INTERACTIONS BETWEEN EXHAUST GAS RECIRCULATION AND EFFICIENCY DURING THE DYNAMIC OPERATION OF TURBOCHARGED SI ENGINES

Dissertation

zur Erlangung des akademischen Grades

Doktoringenieur (Dr.-Ing.)

von

Daniel Hans Langmandel, M.Sc.

geb. am 28.11.1988 in Bad Mergentheim

genehmigt durch die Fakultät für Maschinenbau
der Otto-von-Guericke-Universität Magdeburg

Gutachter: Prof. Dr.-Ing. Hermann Rottengruber
Prof. Dr.-Ing. Peter Eilts

Promotionskolloquium am 15.03.2023

"Man kann ein Auto nicht wie ein menschliches Wesen behandeln. Ein Auto braucht Liebe. Es ist erst dann schnell genug, wenn man morgens davor steht und Angst hat, es aufzuschließen. Beim Beschleunigen müssen die Tränen der Ergriffenheit dann waagerecht zum Ohr hin abfließen."

nach *Walter Röhrl*

Abstract

In the context of increasingly stringent CO_2 restrictions and the growing electrification of the automotive sector, external exhaust gas recirculation offers an as yet largely unused efficiency or CO_2-potential in turbocharged gasoline engines. Among the various configurations, low-pressure exhaust gas recirculation thereby simultaneously provides the greatest steady-state efficiency potential and poses the greatest challenge for the implementation in dynamic engine operation. The main difficulty is the long gas path between the exhaust gas recirculation inlet point and the engine combustion chamber, which is relatively slow compared to the rapid load change of the engine.

This thesis aims to provide a holistic view and analysis of this key problem of low-pressure exhaust gas recirculation and, at the same time, to develop and evaluate proposals for possible solutions to the problem. To this end, the understanding of the problem is first sharpened and the central control levers are discussed. The possible solution approaches are separated into a hardware-based and a software-based one and analyzed individually. The respective advantages and disadvantages are worked out.

Using a simplified container model, a worst-case load step is determined, which serves as a basis for comparison and evaluation of different hardware approaches. In addition to existing systems known from the literature, two self-developed concepts are compared with the configuration on the base engine. The analysis is carried out using a 1D CFD simulation, and the knowledge gained is used to implement a configuration from the system portfolio on the test engine. The fresh air bypass investigated shows considerable advantages compared to the base engine in this regard. Through a comprehensive sensitivity analysis, the remaining shortcomings can be identified. Based on this, recommendations for future hardware configurations are provided.

The implementation of a purely software-based approach can only be achieved with the precise knowledge of the behavior of the engine efficiency and combustion stability depending on the external exhaust gas recirculation rate as well as on the most important operating point parameters. A novel model based on second-degree polynomials is created to depict these relationships. The derivation is performed empirically using nonlinear regression. Physical relationships are analyzed and scrutinized by interpreting the polynomial coefficients. The presented model approach shows consistently positive results in the validation performed on two engines. In the future, the model can be used as an application tool for easy implementation of low-pressure exhaust gas recirculation or directly embedded in an operating strategy. The latter is possible due to the simple basis, the low computational effort, the easy transferability to other engines as well as the coverage of the entire operating range based on simple sensor or calculation variables that are known in operation. A new operating strategy tailored to the model is explained at the end of the thesis by means of a flowchart.

The challenge of implementing low-pressure exhaust gas recirculation in dynamic engine operation is thus analyzed explicitly and from different angles for the first time in the work below. Both the introduced hardware-based and the software-based approach represent an essential added value with novelty character for solving the problem. It is demonstrated for both that they are suitable to realize considerable CO_2-potential.

Kurzfassung

Externe Abgasrückführung bietet im Spannungsfeld immer strengerer CO_2-Beschränkungen und der zunehmenden Elektrifizierung des Automobilsektors ein bisher noch weitgehend ungenutztes Wirkungsgrad- und damit CO_2-Potential bei Turbo-aufgeladenen Ottomotoren. Unter den verschiedenen Konfigurationen stellt Niederdruck-Abgasrückführung dabei gleichzeitig das größte stationäre Wirkungsgradpotential und die größte Herausforderung bei der Umsetzung im dynamischen Motorbetrieb dar. Die wesentliche Schwierigkeit ist die im Vergleich zur schnellen Laständerung des Motors träge, weil lange Gasstrecke zwischen Einleitpunkt der Abgasrückführung und Brennraum des Motors.

Die vorliegende Arbeit hat das Ziel eine ganzheitliche Betrachtung und Analyse dieser zentralen Problematik von Niederdruck-Abgasrückführung zur Verfügung zu stellen und gleichzeitig Vorschläge für mögliche Lösungsansätze zu erarbeiten und bewerten. Hierzu wird zunächst das Problemverständnis geschärft und die zentralen Stellhebel beleuchtet. Die möglichen Lösungsansätze werden in einen Hardware- und einen Software-basierten getrennt und einzeln analysiert. Die jeweiligen Vor- und Nachteile werden erarbeitet.

Anhand eines vereinfachten Behälter-Modells wird ein Worst-Case Lastabwurf ermittelt, der als Vergleichs- und Bewertungsbasis für verschiedene Hardware-Ansätze dient. Neben aus der Literatur bekannten Lösungen werden zwei selbst entwickelte Systeme mit der Konfiguration am Basis-Motor verglichen. Die Analyse erfolgt anhand einer 1D CFD Rechnung, wobei die gewonnenen Erkenntnisse dazu genutzt werden aus dem vorhandenen System-Portfolio eine Konfiguration am Versuchsmotor umzusetzen. Der dabei untersuchte Frischluft-Bypass zeigt im Vergleich zum Basis-Motor erhebliche Vorteile. Durch eine umfassende Sensitiviätsanalyse können die noch bestehenden Schwachpunkte gefunden werden. Darauf basierend erfolgen Handlungsempfehlungen für zukünftige Systemlösungen.

Die Umsetzung eines rein Software-basierten Ansatzes kann nur mit der genauen Kenntnis des Verhaltens des Motor-Wirkungsgrads und der Verbrennungsstabilität abhängig von der externen Abgasrückführungsrate sowie von den wichtigsten Betriebspunkt-Parametern erfolgen. Zur Abbildung dieser Zusammenhänge wird ein neuartiges Modell basierend auf Polynomen zweiten Grades erstellt. Die Ableitung erfolgt empirisch mithilfe Nichtlinearer Regression. Physikalische Zusammenhänge werden durch die Interpretation der Polynomkoeffizienten analysiert und hinterfragt. Der vorgestellte Modellansatz zeigt bei der an zwei Motoren durchgeführten Validierung durchweg positive Ergebnisse. Das Modell kann zukünftig als Applikationswerkzeug zur einfachen Implementierung von Niederdruck-Abgasrückführung oder direkt in einer Betriebsstrategie eingebettet genutzt werden. Letzteres ist durch die einfache Basis, den geringen Rechenaufwand, die einfache Übertragbarkeit auf andere Motoren sowie die Abdeckung des gesamten Betriebsbereichs anhand einfacher und im Betrieb bekannter Sensor- bzw. Rechengrößen. Eine auf das Modell zugeschnittene neue Betriebsstrategie wird am Ende der Arbeit anhand eines Ablaufdiagramms erläutert.

Die Herausforderung der Umsetzung von Niederdruck-Abgasrückführung im dynamischen Motorbetrieb wird in der nachstehenden Arbeit also erstmalig explizit und von verschiedenen Seiten analysiert. Sowohl der vorgestellte Hardware- als auch der Software-basierte Ansatz stellen hierbei einen essentiellen Mehrwert mit Neuheits-Charakter zur Problemlösung dar. Für beide wird nachgewiesen, dass sie dazu geeignet sind erhebliches CO_2-Potential umzusetzen.

Contents

List of Figures

List of Tables

Nomenclature

In addition to the symbols listed here, other symbols are used in the text. As these have only locally limited meaning, they are explained at the appropriate passage.

formula symbol	unit	meaning
$bd10-90\%$	[CAD]	burning duration of the 10 % to the 90 % mass fraction burnt point
c	[J/kgK]	specific heat capacity
ect	[CAD bTDC]	exhaust cam timing
f	[−]	correction factor
ict	[CAD aTDC]	intake cam timing
it	[CAD bTDC]	ignition timing
$imep$	[bar]	indicated mean effective pressure
m	[kg]	mass
$mfb50$	[CAD aTDC]	crank angle of the 50 % mass fraction burnt point
p	[bar]	pressure
t	[s]	time
voa	[mm^2 CAD]	valve overlapping area
x	[−]	proportion
COV	[%]	coefficient of variation
CR	[−]	compression ratio
L	[−]	air demand
LHV	[kJ/kg]	lower heating value
P	[W]	power
Q	[J]	energy
R	[J/kgK]	specific gas constant
T	[K]	temperature
T	[Nm]	torque
V	[m^3]	volume
η	[−]	efficiency
κ	[−]	heat capacity ratio
λ	[−]	fuel-air ratio
ρ	[kg/m^3]	density
σ	[−]	standard deviation
ϑ	[K]	temperature
ω	[rpm]	engine speed
Δ	[−]	difference

Indices

air	fresh air
com	conversion
com	combustion
corr	corrected
diff	difference
e	effective
eg	exhaust gas
f	fuel
gl	global
gp	grid point
h	homogeneous
hyd	hydraulic
i	indicated
imep	indicated mean effective pressure
ind	indicated
is	isentropic
korr	corrected
m	mechanical
max	maximum
meas	measurement
opt	optima
p	isobaric
rel	relative
st	stoichiometric, in connection with formula symbol L
th	theoretical
tol	tolerance
tt	target torque
CP	combustion phasing
Cmp	compression loss
Eng	engine
EGR	exhaust gas recirculated
Exp	expansion loss
FL	friction loss
HP	high pressure
HT	heat transfer
IC	imcomplete combustion
IGE	ideal gas exchange
IM	intake manifold
Mess	measurement
OP	operating point

Sim	simulation
RBR	real burn rate
RC	real composition
RFP	real fluid properties
RGE	real gas exchange
RL	roots loader

η	efficiency

abbreviation	meaning
abs.	absolut
et al.	*latin:* et alia; and others
ect	exhaust cam timing
ict	intake cam timing
imep	indicated mean effective pressure
it	ignition timing
rel.	relative
vl	valve lift
voa	valve overlapping area

0D	zero dimensional
1D	one dimensional
3D	three dimensional
aTDC	after top dead center
BDC	bottom dead center
bTDC	before top dead center
CA	crank angle
CAD	crank angle degree
CFD	computational fluid dynamics
CO	carbon monoxide
CO_2	carbon dioxide
DOE	design of experiment
DP	discharge point
EGR	exhaust gas recirculation
EP	extraction point
H_2	hydrogen
HC	hydrocarbons
HP	high pressure
HP-EGR	High Pressure-EGR
LHV	lower heating value
LP-EGR	Low Pressure-EGR
MLP	measured load point

NO$_x$	nitrogen oxides
nRMSE	normalized root mean squared error
O$_2$	oxygen
OP	operating point
PTA	pressure trace analysis
RMSE	root mean squared error
SI	spark ignition
SSE	sum of squares error
TDC	top dead center
TPA	three pressure analysis
TWC	three-way catalytic converter
VTG	variable turbine geometry
WLTC	Worldwide harmonized Light-Duty vehicles Test Cycle

1 Introduction

In motorcycles, airplanes, ships, trucks or cars - internal combustion engines have been the major propulsion source in the transport sector for more than 100 years. They still represent the most widespread energy conversion device within the mobility sector [71]. Over the course of time and to this day, power-train departments have invested a great deal of effort in continuous improvement and development. Given the shortage of fossil fuels, the increasing environmental impact of pollutants and greenhouse gas emissions, and their associated consequences, further progress in the field of internal combustion engines remains indispensable.

The best known and simultaneously most relevant greenhouse gas is carbon dioxide (CO_2). Global warming constitutes one of the biggest international challenges of the moment. As early as 1979, at the first international climate conference in Geneva, global warming was declared to be a problem that must be taken very seriously and urgently resolved by humanity and nature [62]. Over time, the mitigation of global warming was a recurrent topic, and it was eventually also made the key issue of the Paris Agreement on climate protection in 2015. 195 countries committed to jointly limit the long-term global warming that had occurred since the beginning of industrialization to less than 2 degrees celcius [24]. A central lever for achieving this goal is the reduction of anthropogenic CO_2-emissions. In this respect, a major source is road transport, which accounts for 20 % of all CO_2-emissions [44]. Almost all the vehicles addressed in this context have power trains purely featuring internal combustion engines. In 2017, their market share was close to 99 % [95].

These figures represent an opportunity for the internal combustion engine as they allow the level of CO_2-emissions to be directly influenced by reducing the fuel consumption. Nevertheless, these quotas obligate the automotive industry and place it in a position of responsibility for climate protection. The political legislature emphasizes the bivalent role between opportunity and commitment by regulating the average CO_2 emissions of new cars by law at $95\,g/km$ [25]. Only further improvements in efficiency can help to comply with this ambitious legislation. The development of new technologies and the advancement of existing technologies is therefore indispensable.

One important instrument used in both established internal combustion engine concepts - diesel and gasoline engines - is exhaust gas recirculation (EGR), which describes the dilution of the combustion air with recirculated exhaust gas. For Diesel Engines, EGR reduces the NOx fraction in the raw exhaust gases by decreasing the combustion temperature [100]. In conjunction with Spark Ignition Engines, however, the implementation of EGR provides into the bargain an improved engine efficiency and reduces the fuel consumption. It is conducive not only to the gas exchange but also to the combustion process. In partial-load operation, the engine can be dethrottled by the exhaust gas, which functions as inert gas and does not participate in the combustion. The recirculation of unburned fuel residuals, the reduction of the knocking tendency and of the wall heat losses as well as the increase in the adiabatic coefficient of the combustion gas mixture have a positive impact on the efficiency of the combustion process itself [74].

There are basically two ways of adding exhaust gas or residual gas to the cylinder filling - internal and external EGR. Beside that, for turbocharged engines, a more detailed breakdown of external EGR is made. Among the distinguished and competing configurations of High-Pressure (HP) and Low-Pressure (LP) EGR, LP EGR poses the biggest challenge regarding its dynamic behavior. Only some of the

potential measured in a stationary situation can be transferred to dynamic engine operation. Solving this particular problem and thus implementing an external LP EGR during the dynamic operation of a turbocharged Spark-Ignition Engine in conjunction with the best possible use of the stationary potential provides the motivation for this work. This is the topic of the investigations carried out. The findings of this study offer the possibility of a noticeable reduction in the fuel consumption of turbocharged gasoline engines and contribute to compliance with the CO_2-threshold as well as to the fulfillment of the objective first identified in 1979 in Geneva and finally formulated 2015 in Paris.

1.1 Main topic and objective

In the stationary engine map, both the external EGR rate, which is required to achieve the maximum efficiency, and the EGR tolerance are distributed inhomogeneously. They depend on the engine operating point and thus on the engine speed and above all on the engine load. Consequently, in the event of a change in the operating point during dynamic engine operation, a change in the external EGR rate is necessary to achieve the engineâ€™s best possible efficiency and to stay below the EGR tolerance limit. If the latter is not complied with, unstable combustion with cycle-to-cycle variations can occur. The engine may even misfire.

Compared to the gas dynamics in the intake and the EGR path, the engine load and thus the operating point and the required EGR rate change very quickly, especially in the case of a negative load step (tip-out or declutching). This causes a difference between the set point and actual value of the EGR rate. Over several engine cycles, this causes a lack of control over the external EGR mass that is consumed by the engine.

In this context, the size of the volume between the EGR discharge point or the EGR valve and the combustion chamber is a decisive factor. It is crucial for the speed at which the external EGR rate can be adjusted upon a change in the operating point. The LP EGR system configuration shows particular disadvantages on this issue. As a simplified illustration, Figure 1.1 shows a comparison between the two most important configurations of cooled external EGR - the HP and the LP EGR systems in turbocharged engines:
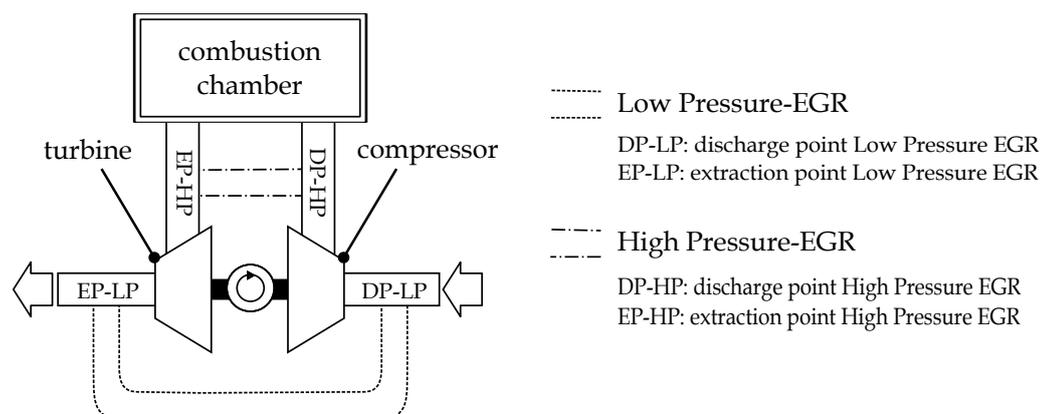
Figure 1.1: Schematic configurations of HP and LP EGR

It is apparent why the LP EGR configuration behaves particularly adversely during dynamic engine operation. It has the largest volume between the EGR discharge point (DP-LP) and the combustion chamber due to the EGR feed upstream of the compressor.

This important disadvantage is counterbalanced by the equally significant advantages of the LP-EGR configuration. Above all, the good mixing of fresh air and exhaust gas, the minor influence on the turbine operation and the gas dynamics during the gas exchange and—last but not least—the applicability in characteristic map ranges at low speeds and high loads (low-end torque) must be mentioned here. In order to implement LP EGR systems in turbocharged gasoline engines on a large scale and to exploit their known advantages, however, minimizing the dynamic disadvantage is key. Within the scope of this work, two possible approaches are presented and analyzed in detail. This makes this study the first to specifically and exclusively deal with the problem of LP EGR during the dynamic operation of Spark-Ignition Engines. It breaks new ground in improving efficiency and implementing as much as possible of the potential feasible in a stationary situation during dynamic operation.

In Chapter 2, the two approaches to solving the problem are outlined and described in detail. In addition, the identification of a dynamic worst case, calculated on the basis of a stationary measurement with simplified models, is explained. Chapter 3 presents the test engines and the measurement setup used for the analysis.

Afterwards, the first of the two potential solutions is further detailed in Chapter 4. Examining the portfolio of hardware setups, well-known concepts from the literature are compared with purpose-built concepts. The analyses are conducted using a detailed 1D CFD software. Based on the system evaluation, one of the configurations is selected for implementation and further investigation on the engine test bench. In conclusion, the results obtained there are presented and discussed in depth. The subsequent Chapter 5 introduces the model-based approach to solving the problem. Physical dependencies are identified and analyzed with the aid of a measurement program specifically developed to set up the model. The mathematical formulation of the model approach serves as the interpretation framework. For the first time, not only is the influence of EGR on the indicated efficiency evaluated on the basis of variations in individual load points, but the interaction of the variations is compared depending only on the starting point of the EGR variation. This behavior is modeled phenomenologically and finally used as the model framework. The chapter ends with the presentation of an operational strategy within which the model could be applied. The premise of the strategy is to minimize fuel consumption knowing the EGR efficiency behavior while maintaining the stability limits of the engine. The thesis ends with a short discourse and a summary of the findings in Chapter 6.

1.2 State of the art

The large number of publications and studies shows that researchers and the industry have already been working on cooled external EGR for a long time. Due to technological developments in the field of internal combustion engines, and especially Gasoline Engines, the function of cooled external EGR and consequently the emphasis of publications has altered over time. The polyvalent impact of exhaust gas on the engine cycle opens up different options for using EGR. Initially, the potential of exhaust gas recirculation in Spark-Ignition naturally aspirated Engines was employed to reduce NOx emissions in particular. This benefit can be obtained due to the temperature-reducing effect of exhaust gas. Studies

dealing with this issue are presented e.g. by Komiyama et al. [51] or by Baruah et al. [7]. Beyond that, there are a several other works adressing the same topic [1, 41, 81, 106].

The invention and avaibility of the three-way catalytic converter (TWC) rendered the additional reduction of NOx raw emissions by means of an in-cylinder measure obsolete. To ensure that exhaust gas aftertreatment functions efficiently and correctly, the TWC nonetheless requires a stoichiometric fuelâ€"air ratio. For this purpose, the air mass flow must be limited in partial-load operation, entailing throttling losses but at the same time justifying the significance and the further use of external EGR. The decisive factor in this respect is the ability of cooled external EGR to reduce gas exchange losses and conversely to assume the role of a fuel saving option. General insights on EGR as an efficiency measure can be found in the works of Wei et al. in [106] and of Jiang in [45].

As an inert gas, the recirculated exhaust gas is not part of the chemical reaction mechanism that takes place during engine combustion. Yet, it is part of the gas volume flow that passes the throttle valve and the inlet valves into the cylinder. Indeed the resulting dethrottling reduces gas exchange losses while simultaneously preserving stoichiometric engine operation. Detailed investigations on this issue were inter alia conducted and explained within the studies by Pischinger et al. [77], Schnittger et al. [84], Kawamoto et al. [48] as well as of Jiang [45].

Further positive effects of external EGR on efficiency are known from the literature apart from those described above. For example, another favorable impact of the inert behavior of recirculated exhaust gas during the combustion process is its function as a heat sink. The heat absorption reduces heat transfer losses during the engine process [74]. The addition of exhaust gas to the combustion mixture not only reduces heat losses but also improves the thermodynamic properties of the gas by decreasing the ratio of specific heats [15, 42, 74] and consequently improving the overall theoretically possible efficiency in the entire engine operating map. Unburned hydrocarbons, H_2 and CO that are components of the exhaust gas can be used in the working process again and enhance the conversion efficiency as well. After all, a major benefit of external cooled EGR is the reduction of the knocking tendency, enabling not only higher compression ratios but also more efficient combustion phasing. Due to its relevance, this topic has received attention from various authors, such as Diana et al., who conduct a general study on knock mitigation [18], Toda et al. who base their research on a natural aspirated SI engine [98] and a number of additional authors focusing on today's most widely used turbocharged SI engine [4, 12, 78].

Due to the multi-layered influences of EGR in turbocharged engines and for the sake of clarity, the following Sankey diagram and the corresponding table in Figure 1.2 summarize the advantages but also the drawbacks of external EGR in turbocharged SI engines on the basis of a loss analysis. A distinction is made between low load operation (naturally aspirated operation) and high load operation (turbocharged operation) as described by Nitschke [74]. The individual influences and losses are evaluated qualitatively, derived from the information found in the above sources. In addition, the loss distribution used in this thesis is presented in the context of the Sankey diagram shown. It becomes important for the analysis of the model's physical background in Chapter 5 and is therefore explained briefly.

The loss distribution is based on a thermodynamically idealised comparison process, used to calculate the theoretical efficiency to be achieved. The efficiency $\eta_{OC,th}$ of this comparison process depends only
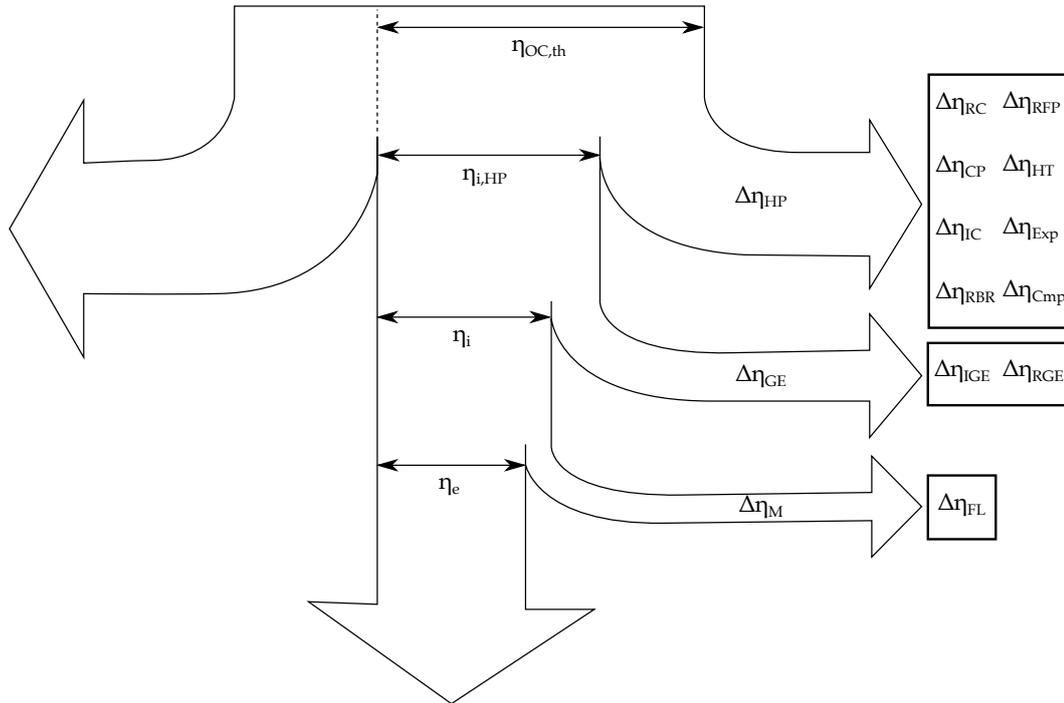
Individual loss	Meaning	Influence of external EGR in low load area	Influence of external EGR in high load area
$\Delta\eta_{RC}$	real composition	++	++
$\Delta\eta_{CP}$	combustion phasing	*o*	++
$\Delta\eta_{IC}$	incomplete combustion	+	+
$\Delta\eta_{RBR}$	real burn rate	−	−
$\Delta\eta_{RFP}$	real fluid properties	+	+
$\Delta\eta_{HT}$	heat transfer	++	++
$\Delta\eta_{Exp}$	expansion loss	*o*	*o*
$\Delta\eta_{Cmp}$	compression loss	*o*	*o*
$\Delta\eta_{IGE}$	ideal gas exchange	++	−
$\Delta\eta_{RGE}$	real gas exchange	++	−
$\Delta\eta_{FL}$	friction loss	−	−

Figure 1.2: Sankey diagram of engine efficiencies and individual losses (see [33] and to [11])

on the compression ratio *CR* and the gas composition as a function of the isentropic exponent κ and is calculated as follows:

$$\eta_{OC,th} = 1 - CR^{1-\kappa} \tag{1.1}$$

The efficiency of the Otto cycle is at the optimum level for the engine process under consideration. The quality of a real engine process can be assessed on the basis of the deviations from this ideal process

[112]. All measures to reduce fuel consumption aim to approximate the real engine process to the ideal comparison process.

In order to calculate a realistic effective engine efficiency based on an idealized Otto process, it is useful to establish an additive chain of individual losses. Due to the complexity of the correlations, it is not possible to accurately measure or calculate all the influences of the individual losses. As a consequence, and since the influences interact, not only their exact distribution but also their defined order is of high relevance [77]. The calculation of losses within the scope of this thesis is carried out using the software *GT Power*. The basics for the loss distribution applied there can be found in [31] and [105]. The approach initially classifies the losses broadly into losses of the high-pressure process $\Delta\eta_{HP}$, losses of the gas exchange $\Delta\eta_{GE}$ and mechanical losses $\Delta\eta_M$. The individual losses are allocated to these main losses as shown in Figure 1.2.

First, all losses attributable to the engine process itself are considered. Starting from the ideal Otto process, the real charge, i.e. deviating from pure air, is taken into account. The properties involved are exclusively functions of the gas composition, which explains the consistently positive influence of external EGR. Within the next proportion, losses are considered that occur when the heat release does not take place at the top dead centre but at the centre of the combustion process. While this loss is still unrelated to external EGR in the low-load range, the knock-reducing effect of externally recirculated exhaust gas means that there is great potential in the high-load range. The centre of combustion can thus be shifted closer towards optimum efficiency. The loss due to incomplete combustion, caused by HC and CO raw emissions must also be evaluated positively in terms of external exhaust gas recirculation, as the unburned emissions are at least partially returned to the combustion process. In contrast to this, the loss of real burning duration, i.e. the deviation from the isochoric heat release, is negatively influenced by external EGR whatever the engine load, since the molecular distances between the reactants are increased by the inert exhaust gas and the reaction is hampered. Similarly to the loss of real composition, the loss of real fluid properties is mitigated by external EGR. The caloric properties of the cylinder filling as a function of pressure and temperature are taken into account. Next, the loss due to heat dissipation to the cylinder walls is calculated. By functioning as a heat sink, external EGR improves this aspect over the entire load range. The reduction in efficiency due to the exhaust valve opening before bottom dead center (expansion loss) and the reduction in efficiency due to the intake valve closing after bottom dead center (compression loss) are both unrelated to external EGR. The losses of the gas exchange are divided into ideal gas exchange losses, which are calculated on the basis of average inlet and outlet pressures, and real gas exchange losses, in which high-frequency pressure oscillations are taken into account. In both cases dethrottling leads to external EGR offering an advantage in the low-load range. Increased exhaust back-pressure and increased compressor work tend to lead to external EGR posing a disadvantage in the high-load range. The mechanical losses are also slightly increased by external EGR, since an increased gas mass must be compressed. For a further detailed description of the individual losses, please see the above literature sources.

The long list of pluses and positive aspects as well as the illustrated figure show how much potential cooled external EGR has and why it is worth putting more effort into its implementation. Nevertheless, cooled EGR is today almost exclusively utilized in conjunction with naturally aspirated gasoline engines [48, 113]. Compared to them, turbocharged gasoline engines can derive even more value from cooled external EGR [15, 54, 87]. Due to downsizing and continuously rising power outputs [56, 100], knocking propensity and exhaust temperatures during high-load operation are increasing. As already mentioned,

external cooled EGR provides an efficient technology to prevent or to compensate for the disadvantages by reducing both, the knocking tendency and the exhaust gas temperatures [35, 54, 104, 106]. The most important publications comparing the key configurations of EGR and analyzing their most important characteristics are discussed subsequently.

Low- and High-Pressure EGR

Low-Pressure EGR (LP-EGR) constitutes one specific configuration of external cooled EGR. The paper by Siokos et al. [87] elaborates the most important benefits and features of LP-EGR in turbocharged SI engines. The configuration has several advantages but also some downsides in comparison to the competing High-Pressure EGR (HP-EGR) system [11]. The principles of both configurations are introduced with the help of Figure 1.1. In [2], [79] and [14], the authors outline some of the main differences between HP EGR and LP EGR in turbocharged SI engines. Below, the focus is on clear differentiation so that the EGR configuration that is chosen in this thesis can be understood and justified.

Several researchers have paid considerable attention to detailed effects and differences in particular. One major distinction made between the two EGR systems is that catalyzed exhaust gas from downstream of the TWC can easily be used in conjunction with a LP EGR configuration. In [79] and [63], the authors compare catalyzed and non-catalyzed exhaust gas in terms of the disposable pressure gradient. Due to the pressure drop caused by the TWC, catalyzed exhaust gas has a lower EGR delivery range. In addition to that, catalyzed exhaust gas lacks hydrocarbons, H_2 and CO. For the combustion process, this means reduced reactivity and conversion efficiency. The smaller proportion of NOx in catalyzed exhaust gas, however, lowers the knocking tendency, as analyzed by several scientists in [43, 47, 94]. Finally exhaust gas that is extracted downstream of the TWC has advantages regarding the fouling of the EGR and intake path, since the proportion of hydrocarbons is significantly lower. This is the topic of many works that study fouling extensively in connection with Diesel Engines [55, 89, 90]. Particle emissions are equally important for the fouling process. Yoo et al. demonstrate their influence on EGR cooler fouling in [114]. Recently, Fischer et al. introduce a coated gasoline particle filter within the EGR path to deliver exhaust gas largely free of hydrocarbons and soot emissions [28]. The System, however, is mainly necessary in conjunction with HP-EGR, since the particle emissions are particularly high in its operation area [68]. In this context, this can also be seen as a disadvantage of HP EGR.

A further beneficial feature of LP EGR is the extraction of the exhaust gas behind the turbine. As Zhong et al. [117] show in their paper, this results in little influence on pressure pulsations during the gas exchange, or on the turbine. Severe disturbances to the turbocharger and the gas exchange in general cause undesired engine performance losses. The consequence of this is, among other things, the limited operating range of the HP EGR and especially the lack of applicability in the Low-End-Torque area at low speeds and high torque on the engine map. In times of downspeeding and downsizing, however, it is precisely these operating ranges in which the engine is frequently operated [101]. This operating area is correspondingly of high relevance for the engine's fuel consumption. The approach taken by Crain et al. [13] emphasises the influence on turbocharger operation on the basis of the compressor. It illuminates the disadvantageous behaviour of the LP-EGR with regard to compressor efficiencies due to increased compressor work. Ultimately, however, his work also comes to the conclusion that the other advantages predominate, especially the lower impact LP EGR has on the turbine.

Another of these benefits is the fact that the mixture of exhaust and fresh gas is worse in combination with HP EGR than with LP EGR, since the gas path for homogenizing the mixture is shorter. The

authors in [57, 88] show that this can provoke unstable combustion and cycle-to-cycle variations. The long gas path of the LP EGR configuration, which is advantageous in this case, however also poses one main downside to its dynamic implementability.

The comparison of the two external EGR configurations in conjunction with turbocharged SI engines offers advantages and disadvantages for both systems. One main drawback that remains unsolved so far is the long gas path of the LP EGR system. During dynamic operation, this results in reduced potential efficiency compared to stationary operation. Nevertheless, as Cloos et al. show in [15], due to its operating area, the positive aspects of LP EGR largely outweigh the advantages of HP EGR. The effort to overcome the dynamic drawback of the LP EGR system can therefore not only be justified but is necessary to keep fulfilling the legal fuel consumption requirements and to further improve the efficiency of SI engines. For this reason, several authors have taken up and investigated the issue of LP EGR during dynamic operation. They propose different solutions and approaches.

Hardware-based approaches

Cunningham, for example, suggests in [17] an additional gas routing to flush the inlet path between the EGR valve and the throttle after a load step as quickly as possible with fresh air. The mixture of EGR and fresh air, which is in the inlet path before the load step, is recirculated to the exhaust pipe downstream of the TWC. This is an important prerequisite for this system. The EGR duct must also be branched off after the TWC to ensure that the exhaust gas aftertreatment system works properly even in the event of a load step. However, Cunningham does not provide evidence that the system functions correctly in his publication.

In [96, 97], Thewes introduces a so-called fresh air bypass. This additionally installed air duct takes fresh air from upstream of the EGR supply position in the standard intake path and returns it directly to the intake manifold downstream of the throttle. This is possible during a negative load step as the pressure conditions prevail. The simulation results presented within the publications conclusively demonstrate that the system functions correctly. Fischer et al. present an almost identical approach in [26] which also includes a fresh air bypass and the associated control valve. The solution with the second fresh air path is also investigated by Siokos in [86]. The 1D CFD simulation performed there also incorporates a criterion for simulative misfire detection. The results show the feasibility of misfire-free transient operation with LP EGR when applying this solution.

In [37], Haas suggests a comparable but enhanced concept that is additionally equipped with a pump in the fresh air bypass. The objective pursued is to accelerate the fresh air delivery through the bypass path. This suggestion, however, requires significantly more hardware. An examination of its specific functions is not part of this publication.

Another way to overcome the dynamic drawbacks of the LP-EGR is proposed by Messing et al. in [72]. Instead of changing the fresh air routing, the load point is to be shifted during a tip-out, i.e. engine load is increased. The load point shift simultaneously improves the EGR tolerance and thus lowers the criticality of the tip-out by means of combustion stability. In practice, a generator is a conceivable option. With this proposal, it is important to bear in mind that the absolute fuel consumption increases due to the load point shift. Moreover, when a generator is used, battery capacity is required that can absorb the excess energy for a short time during the negative load step.

Application-based, model-based and operational-strategy-based approaches

All works cited so far try to solve the problem of implementing LP EGR during dynamic engine operation with the help of hardware modifications. This is often costly and difficult to achieve. For this reason, a similar number of authors have dedicated their work to addressing the issue using the available hardware system. Their approach is to adapt the operational strategy or implement the control of the EGR rate in a smart way. To take a first example, Styles presents a strategy that prescribes the EGR rate solely based on engine speed, but not on engine load [92]. This is to avoid the gap in the adjustment time between the engine load and EGR rate. The EGR rate for each engine speed is always chosen such that the EGR tolerance is not undershot at any load. This method accepts disadvantages in terms of efficiency to ensure that the engine can easily be stably operated.

One approach that takes into account the actual engine operating conditions, including not only the engine speed but also the engine load, is introduced by Yoshioka et al. in [115]. The EGR rate to be set is based exclusively on values stored in the control system. The suggestion to additionally vary the adjustment speed of the EGR valve is considered to be obsolete, especially in connection with LP-EGR, since the adjustment time of the EGR valve is much shorter compared to the delay time of the intake gas path.

Due to the location downstream of the intake duct, adaptations to the variables of the valve train can be used to circumvent the delay time. In this context, Kapus proposes that crankshaft variabilities can be exploited to adjust the internal and thus the total EGR rate during and after a load step [46]. UUsing crankshaft variabilities to control the residual gas mass or the internal EGR rate is not a new concept, and neither is the interaction between internal EGR and external EGR. While the first issue is described by Golloch in [33], among others, the latter is explained in detail by Nitschke in [74]. However, Kapus's idea of linking the topic of crankshaft variabilities with the implementation of LP EGR during dynamic engine operation is novel. Although the suggestion is not supported by simulations or measurement results, various authors, such as Krost in [53] or Böcking in [9] demonstrate the positive influence of a decreasing internal EGR rate on combustion variance and stability, which underlines the potential of the measure. Nevertheless, it should be borne in mind that the adaptation of crankshaft variability can also lead to a disadvantage in terms of efficiency and thus to an increase in fuel consumption. Furthermore, precise knowledge of the behavior of efficiency and combustion variations depending on internal and external EGR and on load point properties is essential for implementation, but normally requires considerable measuring effort.

The paper by Wiese et al. is likewise heading towards a similar direction [108]. The possibility of using variable valve timing to reduce the evacuation time of the intake path is analyzed. In turbocharged operating areas, a high degree of valve overlapping is chosen in order to increase the gas volume flow without changing the cylinder filling or the engine torque, respectively. As the simulation results show, this allows the gas volume in the intake path between the EGR and inlet valves to be flushed approximately 1-3 engine cycles faster. The obvious disadvantages of this approach are that the valve drive must be adjusted before the load point change and that the measure only makes sense in operational areas where scavenging is possible, i.e. turbocharged engine operating points. In addition, efficiency deficits are not considered. In a second publication, Wiese et al. ignore the scavenging approach via the variable valve train but provide an improved model predictive control [109]. A prediction horizon of 0.8 s is used. Still, it is indispensable that the EGR tolerance at lower loads as

well as the corresponding efficiencies are well known, since model predictive control needs these input parameters to work properly.

Sarlashkar et al. [82] present a controller for transient operation with LP EGR that is built around three parts. These consist of the modelling of the gas paths and the associated gas compositions in the respective path, the influence of the actuators in the gas path on the quantities modeled in the first part and finally the influence of fueling and ignition as a measure to ensure that engine operation is as stable and efficient as possible in the potential case of overdilution with EGR. Even though this model refers to Dedicated EGR, which enables increased combustion stability despite dilution due to the reformer gases H_2 and CO, the author claims that the model can be applied to all engines with LP-EGR systems. Though, no information is given on the effort required to use the model for another engine. In addition, the major shortcoming is once more that the model can only be used with detailed information on the efficiency and combustion stability of the respective engine depending on the model input parameters.

Recently, in his thesis, Siokos outlined a holistic approach for the implementation of LP-EGR in turbocharged SI engines [86]. The elaboration is divided into three essential parts. Like many of the authors already mentioned in this chapter, one part deals with the overall potential as well as the individual strengths of EGR in SI engines. Focus is given to the specifics of LP-EGR in downsized turbocharged SI engines. In addition, the approach taken by Siokos attaches particular importance to determining the EGR rate at different locations in the intake path operation. For this purpose, an O_2 sensor after the compressor, together with a feed-forward method including various correction functions, are presented and implemented. Finally, one part of the thesis also focuses on achieving the highest possible potential in transient engine operation using LP EGR by reducing the EGR rate as quickly as possible and increasing the EGR tolerance. In this context, the parameters of the variable valve drive as well as the ignition timing and the throttle valve position during a negative load step are adjusted. The latter is done to keep the volume flow through the intake path artificially at a low level of efficiency and to accelerate the scavenging. Neural networks are used for the implementation, which are trained with the help of measurement data. Even though neural networks are meanwhile frequently used for application and model building on engine control devices, the well-known weaknesses such as black box behavior, high computing times, strong dependence on the quality of the training data and above all the the missing physical basis remain. In fact concerning the handling of the transient operation with LP-EGR, a methodical approach is hardly seen in this thesis. Rather, the most easily practicable possibilities are illuminated in order to generate a benefit for the overall approach.

Conclusion on the state of the art

The long air path between the EGR discharge point and the inlet valves, which can lead to over- dilution, remains one of the central drawbacks when implementing LP EGR. However, the number of studies and publications presented in the previous sections clearly shows that the relevance and potential of LP EGR in turbocharged SI engines has indeed been recognised. Although many of the solutions are already providing major improvements, a a holistic approach which enables the simple, efficiency-optimized conversion of LP EGR into production engines has not yet been found. In this regard, the studies have two major, essential weaknesses in common. Firstly, there is often little physical context, which leads directly to the second problem that the measuring effort required to implement the respective measures, especially for implementation on a new engine, is rather high. Either there is no correlation between the engine behavior with EGR in terms of efficiency and dilution tolerance on the basis of simple engine

parameters known during operation, or the correlation has to be redetermined every time. Beyond these two problems, so far there is no work specifically addressing either the individual proposals, especially regarding the hardware approaches, or an overall assessment of LP EGR during transient engine operation. Moreover, suitable validation is frequently absent.

It is precisely these weaknesses and current shortcomings that this work addresses. In addition to a methodical comparison of different hardware solutions, a model is being developed that considerably reduces the measuring effort and illustrates the relationship between the EGR rate and the indicated efficiency and EGR tolerance on a mathematically simple basis with transparent physics. All insights gained in the simulative environment are generated on the basis of measurement data. Validation, offering proof that the model works, is carried out on the basis of measurements.

2 Implementation of low-pressure exhaust gas recirculation during dynamic engine operation

Parts of this chapter have been previously published in [59] and [60].

The following provides a detailed explanation regarding the problem of LP EGR during dynamic engine operation. This is intended to promote a better understanding of the evaluation process referred to specifically later in this section as well as in the following chapters in general. Based on this, a simplified but efficient methodology for identifying a worst case and its results are outlined. Finally, the possible levers and their general advantages and disadvantages are discussed. However, before addressing the technical problem, some important definitions need to be given. Mentioned frequently in Chapter 1, the EGR rate x_{EGR} defines the exhaust gas mass m_{eg} related to the total cylinder charge that is inside the combustion chamber after the intake valves close. The total cylinder gas mass is thus composed of fresh air m_{air}, fuel m_f and exhaust gas m_{eg} [77]:

$$x_{EGR} = \frac{m_{eg}}{m_{air} + m_f + m_{eg}} \cdot 100\% \tag{2.1}$$

Whenever this work mentions an *EGR rate*, this always refers exclusively to externally recirculated exhaust gas. Exhaust gas that becomes available again through retaining or recirculation from the inlet or outlet channels in the immediately following combustion cycle is denoted with the prefix *internal* or *residual gas*. The sum of the internal and external EGR is referred to as the *total EGR* or *total EGR rate*.

Figure 2.1 initially depicts an engine map containing the EGR rates required to achieve the best possible efficiency during stationary engine operation. The engine map originates from a stationary measurement of Engine 1 equiped with a low-pressure EGR path using extraction before TWC. A detailed engine configuration is presented in Chapter 3. One particularly striking aspect is the inhomogeneous distribution of the EGR rate depending on the engine speed and the engine load. Taking a low speed and medium load range as a benchmark where the highest EGR rates are obtained, the EGR rate decreases at higher loads and speeds, and with lower loads. In the high-speed, high-load region, the non-capable charging system and the lack of a pressure gradient along the EGR path are the main limiting factors. In low-load operation, the increasing cyclical fluctuations and the consequently decreasing combustion stability are crucial. With regard to optimum efficiency, it is also important to note that internal exhaust gas recirculation has advantages over external exhaust gas recirculation in the range of these small loads. This is due to the higher temperature of the internal EGR and the resulting improved thermal dethrottling as well as the improved thermodynamic boundary conditions at the start of combustion which improves the dilution tolerance in comparison to external EGR [74].

Nevertheless, initially independently of the influence of internal EGR, a picture or distribution comparable with Figure 2.1 is also apparent regarding the (external) EGR tolerance. At first, it is also irrelevant whether the engine actuators are still set to make the engine as efficient as possible, or whether all parameters are set to achieve maximum EGR tolerance. The qualitative distribution of the EGR tolerance continues to be almost identical with that of maximum efficiency. It is precisely with regard to this characteristic that the central problem arises.
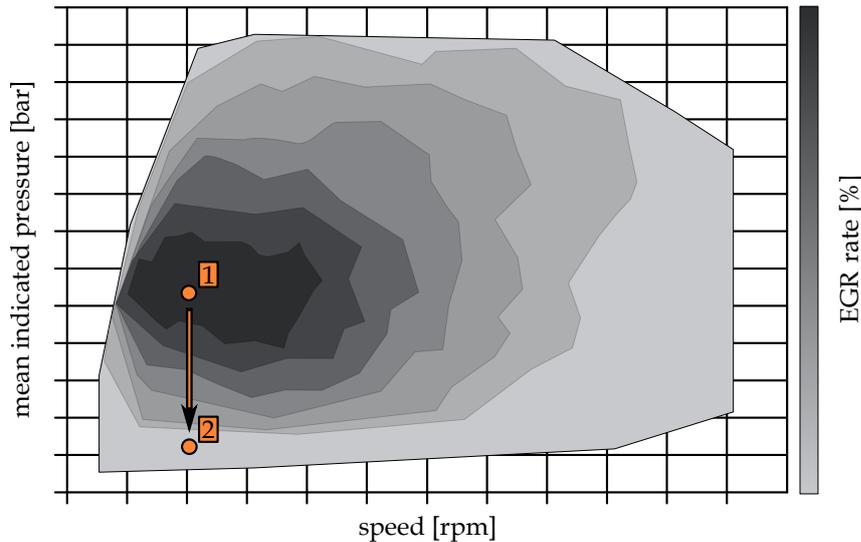
Figure 2.1: Engine map with optimal stationary EGR-rate (qualitative values)

In order to clarify this in further detail, a dynamic operating point change, i.e. a negative load step based on values measured in stationary conditions, is shown in Figure 2.2. Both, the start and the target load point are marked on the stationary engine map with the orange-colored 1 and 2, respectively. The concrete values of the load points are not of importance for the explanation. However, what is of relevance is the transition from an operating point with a high-consumption optimal EGR rate to an operating point with a low-consumption optimal EGR rate. As just explained, the behavior regarding the dilution limits of the load points is analogous. The transient of the two operating points is shown in Figure 2.2 using 1D CFD simulation based as well on the engine configuration presented subsequently in Chapter 3. In addition to the engine torque, which has a decisive effect on drivability and customer awareness, the EGR rate determined in the cylinder and the operating-point-dependent EGR tolerance are shown as a function of time. The EGR tolerance is determined on the basis of the values of 5 load points measured in stationary conditions which are passed during the load transient and a cubic interpolation between these points.

The EGR valve is closed simultaneously to the initiation of the load change and the engine speed is constant. The first notable point is that the decrease in the EGR rate occurs with a considerable time delay after the torque is adjusted. This is due to the dead gas volume between the EGR valve and inlet valves, which must first be supplied for combustion after the load step, i.e. scavenged throughout the engine. The decrease in the EGR tolerance, on the other hand, takes place synchronously with the alteration of the torque. As a result, the course of the in-cylinder EGR rate briefly exceeds the EGR tolerance. Unstable combustion or even misfire events are possible or even very likely. A comparable type of operation in the vehicle would therefore not be possible, because on the one hand the comfort is not sufficiently warranted. On the other hand, due to incomplete combustion and the associated deviations from a stoichiometric air?fuel ratio in the exhaust gas, the faultless functioning of the TWC and the gasoline particle filter (GPF) cannot be guaranteed.
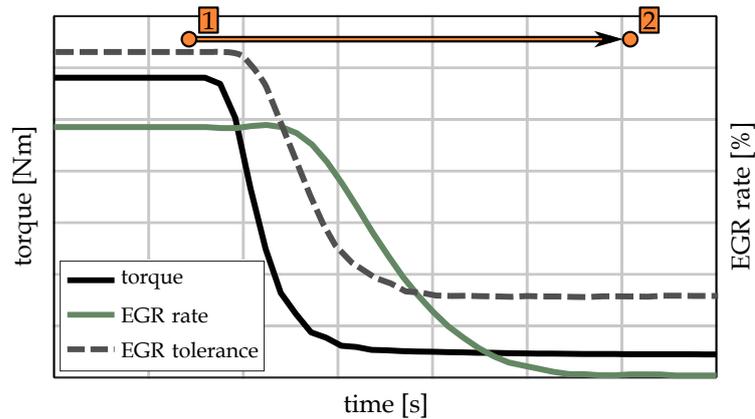
Figure 2.2: Exemplary load step, courses of torque, in-cylinder EGR-rate and EGR-tolerance

The illustration of the three curves during the tip-out in Figure 2.2 appears to be useful, as they can be taken as a basis for deriving measures and identifying the key levers. In this respect, Figure 2.3 shows the contextual correlations between the adjusting levers and possible measures. Adjusting the torque curve directly influences the mass flow through the air path. As a consequence, the reduction in the EGR rate is curtailed, since a high mass flow can be maintained longer after closing the EGR valve. In addition, delaying the reduction in the load slows the decrease in the EGR tolerance. However, it must be borne in mind that adjusting the torque curve has an effect on customer perception and the drivability of the vehicle. It is therefore excluded as a practicable measure. A torque adjustment can only be implemented by a load point shift that is not noticeable for the customer, e.g. by engaging an additional consumer during the transient situation. In addition to the mere reduction of the EGR rate, other possible measures include the torque-neutral adjustment of the load control and a change in the air flow with the help of a hardware measure. An adaptation of the EGR tolerance is also feasible within certain limits. Measures to be considered here include changing the ratio of internal to external exhaust gas recirculation and increasing the dilution limit by altering combustion process parameters such as the ignition timing or air?fuel ratio. Possible disadvantages in terms of efficiency must always be taken into account. In the chapters of this thesis that follow, all measures, except for those that have just been excluded, are examined in more detail. A further differentiation of the particular approaches is initially given in Chapter 2.2.

The following illustrations and explanations will indicate how the problems described above particularly come to bear in downsized turbocharged SI engines, when they are operated in one of the most common driving cycles today: the so-called *Worldwide harmonized Light vehicles Test Procedure (WLTP)*. The velocity profile of the driving cycle is plotted in Figure 2.4. The corresponding load and speed profiles occurring in a compact car with the engine configuration presented in Chapter 3 are illustrated in Figure 2.5. Depending on the engine-vehicle combination, in this driving cycle, fuel savings of approx. 2 - 3 % are possible by identifying and transferring the stationary potentials of LP-EGR. If none of the measures described above are implemented, this potential will be significantly reduced. Thus, remedial measures are necessary in order to be able to realize as much of the stationary potential as possible during real, dynamic driving operation.

adjustment / change of	possible measure
course of the engine torque	- load point shifting - deceleration of torque trajectory
course of the in-cylinder EGR rate	- adaption of the air path - torque-neutral adaption of load control - reduction of the EGR rate
EGR tolerance / dilution limit	- ratio internal / external EGR - adaption of combustion parameters
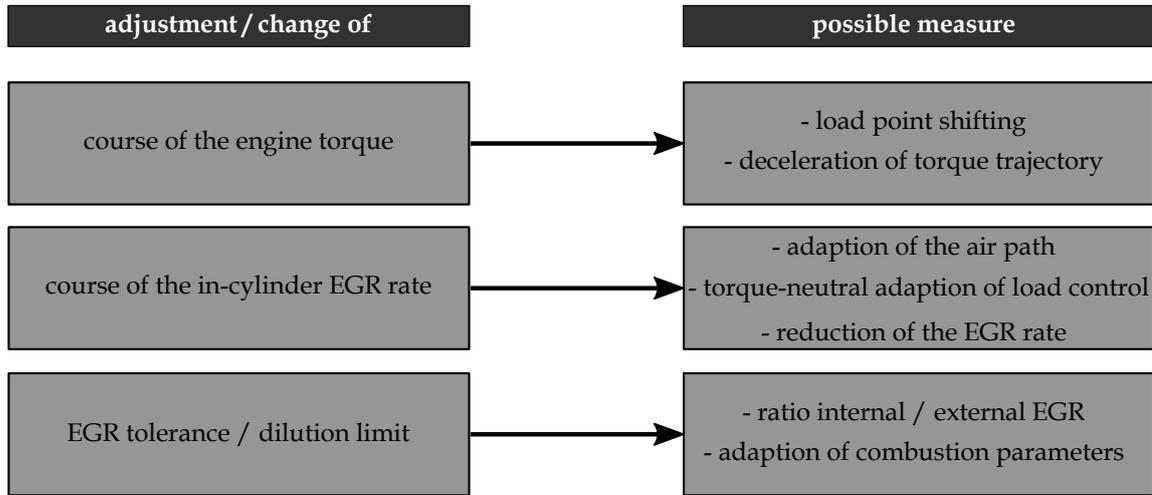
Figure 2.3: Possible control parameters and measures to solve or to improve the dynamic drawback of LP-EGR
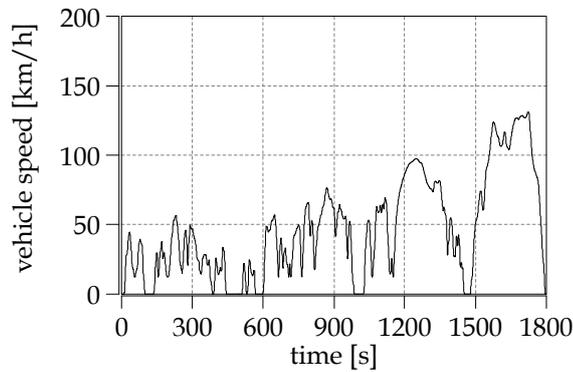
Figure 2.4: WLTP profile: vehicle speed
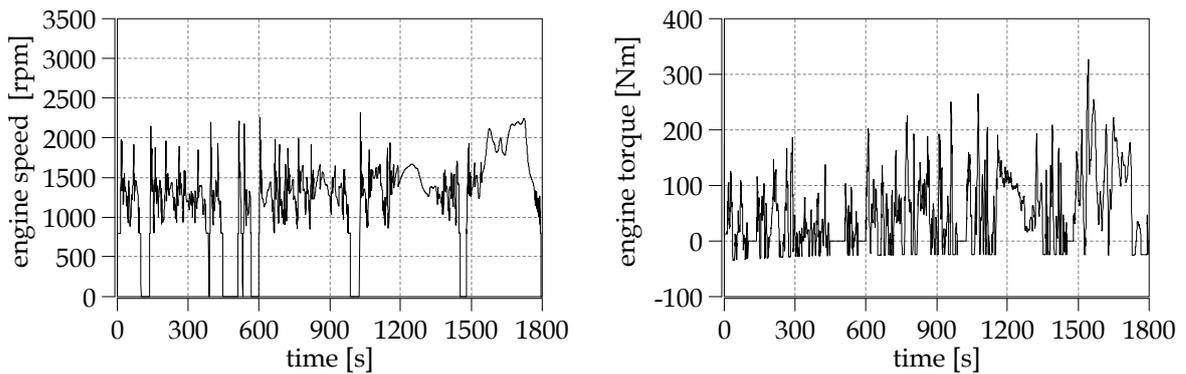
Figure 2.5: WLTP engine speed and engine torque profiles for a compact class vehicle

If the profiles shown in Figure 2.5 are now transferred to the engine map in Figure 2.1 with the help of an evaluation of their percentage dwell time in the individual operating points, the key factors for the fuel savings become striking. Caused by downsizing and downspeeding [101], the WLTP operating points are predominantly located in the area of low speed and low to medium load. Due to the distribution of the EGR rate for maximum efficiency, this results in frequent operation in the range of high EGR rates

and thus, conversely, also in the range of high potential for fuel consumption savings through external EGR. At the same time, it emerges once more that the implementation of a real dynamic driving cycle is very challenging owing to the high proportions in the area of the lowest load associated with low external EGR rates and low EGR tolerances.
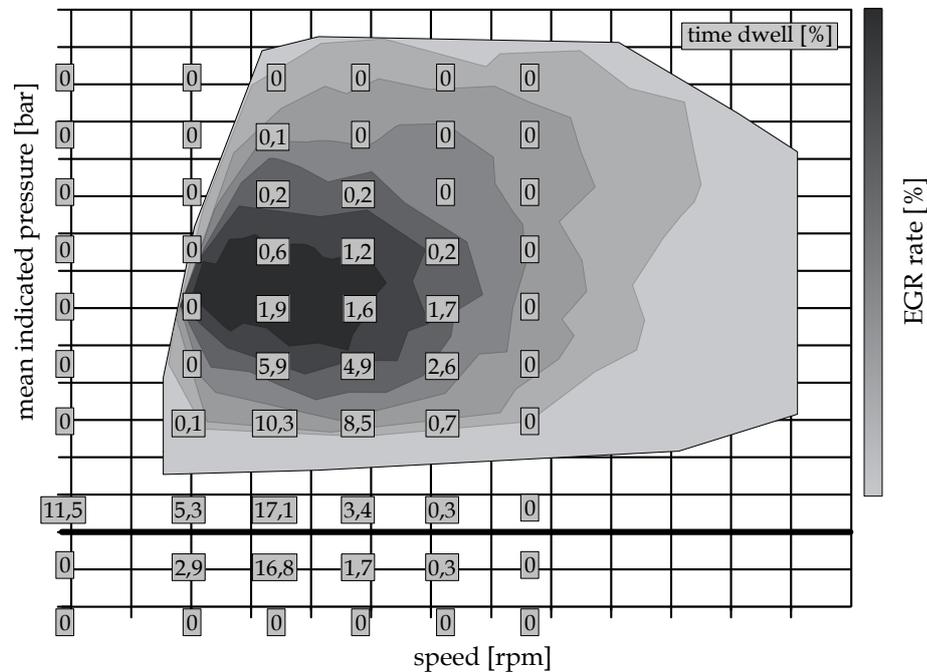
Figure 2.6: WLTP engine speed and engine torque percentage dwell time in EGR rate engine map

This becomes even clearer when a trajectory of the WLTP cycle is included in the same engine map, as can be seen in Figure 2.7. The trajectory marked in yellow demonstrates how many operating point changes take place within the cycle. The vertical lines, which represent a load change, have a decisive effect on the LP EGR system. Load changes and above all the critical tip-outs usually happen extremely fast, since they depend only on the positioning speed of the load actuators, such as throttle valve or valve lift. An engine speed change, on the other hand, is always associated with the acceleration or deceleration of the mass rotating in the motor and is therefore significantly slower than a load change.

In summary, it should be noted that due to the large gas volume between the EGR valve and the inlet valves, there is a delay in the adjustment of the in-cylinder EGR rate that has a decisive influence on the combustion. This dead time is usually longer than the time needed to set the load of the motor. However, the load is a crucial criterion for the EGR tolerance. As a consequence, the EGR rate in the cylinder can exceed the EGR tolerance or the dilution limit during fast negative load steps and provoke combustion instabilities and misfires. It is demonstrated that the dynamic operation and above all the range of high EGR rates in the engine map are crucial for implementing consumption potential, for instance within the WLTP Driving Cycle. Before these crucial enhancement measures as well as their advantages and disadvantages are discussed in more detail, the following section defines once the critical area of tip-outs and how to find it.
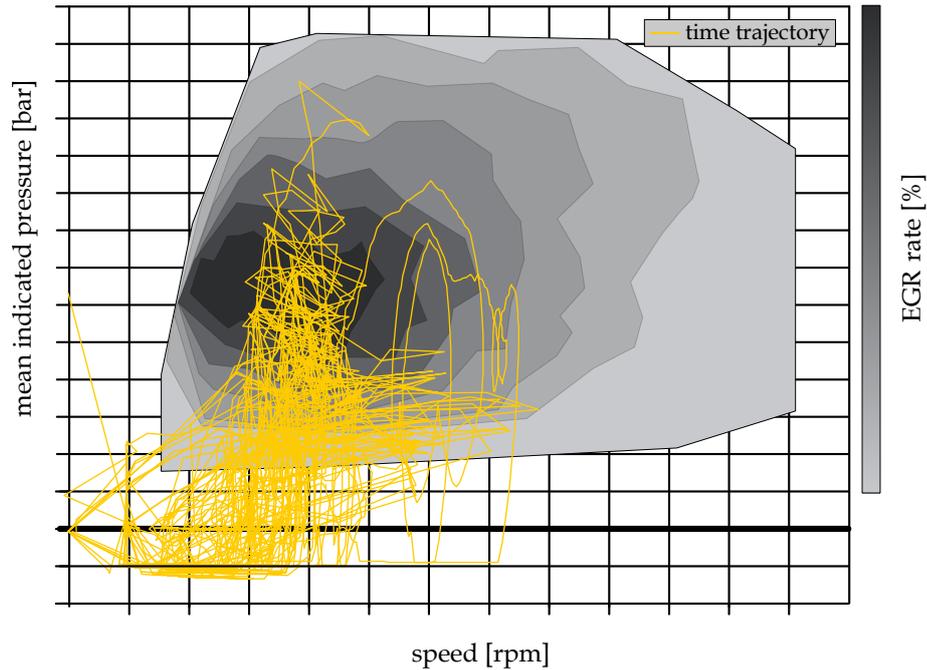
Figure 2.7: WLTP engine speed and engine torque trajectory in EGR rate engine map

2.1 Identification of a dynamic worst case

Systems or components of internal combustion engines are designed to cover a wide operational area, with the ability to function reliably under any circumstances and conditions. There are a large number of independently adjustable actuators and consequently a large number of degrees of freedom. The procedure governing the validation and functional evaluation of such systems is therefore not trivial, but requires either a great deal of effort or a sophisticated methical approach. One such method is analysis based on a worst case scenario. To do this, it is necessary to define the properties as well as the key boundary conditinos of this worst case and to find it on the basis of these guidelines within the degrees of freedom.

If a system can deal with this worst case, it is assured that it will work in all other cases. In addition, when considering and comparing different systems, it must be ensured that the basis for comparison is unique, which applies when working with a worst case. Thus, in this section, the dynamic worst case of negative load steps with external LP EGR is identified. In addition to basic localization and awareness building, the analysis is used as a basis to compare different hardware systems in a simulation and in engine test bench measurements later on in this thesis (Chapter 4).

In the following, a container model developed especially for this study and its derivation as well as its implementation in the *Matlab* software are presented in general. Subsequently, the model is fed with the boundary conditions of the engine presented in Chapter 3 and the relevant results are submitted and discussed. Nevertheless, the model can easily also be used for any other engine with corresponding boundary conditions and any external EGR configuration.

Figure 2.8 initially shows the basic model structure and the model design. On the left-hand side, the engine intake path is visualised schematically. It is modelled as one container and encloses the

inlet path volume from the EGR discharge point to the inlet valves. The container has an incoming mass flow as well as an outgoing mass flow. An ideal gas mixture is assumed in the container. For an engine operated with external EGR, this implies a mixture of exhaust gas and fresh air. This is a considerable simplification, but has no restricting effect on the identification of the worst case. All other thermodynamic properties in the container, such as the pressure and temperature, are also assumed to be homogeneously distributed.

The key question addressed by the model, which is also highlighted in the figure, is how quickly the EGR rate in the container and thus the EGR rate of the outgoing mass flow can be reduced in the event of a negative load step. The course of the EGR rate depending on time thus has to be found.
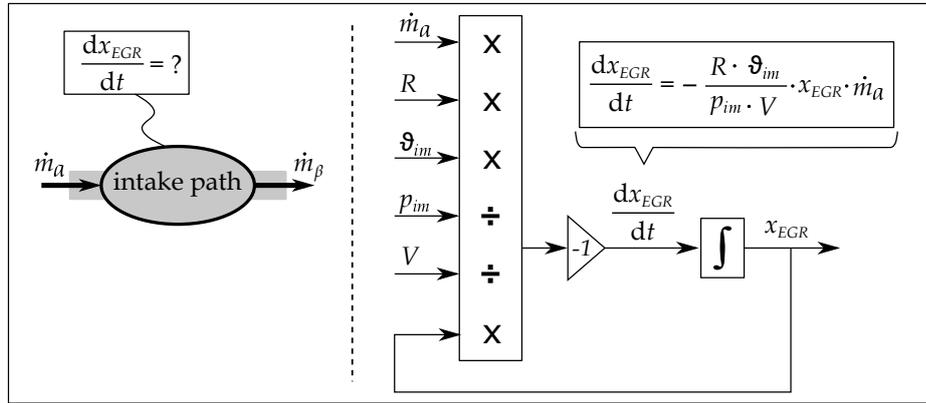
Figure 2.8: Structure of the *Matlab* container model.

In the event of a load transient, the EGR valve is closed in order to reduce the in-cylinder EGR rate as quickly as possible and set the desired EGR rate. As a result, the gas mass flowing into the container m_α consists exclusively of fresh air. The outgoing mass flow, on the other hand, always has the same composition as the ideally mixed gas in the container. The conservation of mass is assumed for the container's content. Its mass balance can therefore be formulated as follows:

$$\frac{dm}{dt} = \dot{m}_\alpha - \dot{m}_\beta \tag{2.2}$$

The exhaust gas or EGR mass flow $\dot{m}_{EGR}$ can be calculated as the product of the total mass flow $\dot{m}$ and the mass fraction of the exhaust gas x_{EGR}:

$$\dot{m}_{EGR} = x_{EGR} \cdot \dot{m} \tag{2.3}$$

Since the incoming mass flow $\dot{m}_\alpha$ consists solely of fresh air ($x_{EGR} = 0$), only the outgoing mass flow $\dot{m}_\beta$ contributes to the change in the exhaust gas mass over time dm_{EGR}/dt in the container. This allows the following formulation to be made:

$$\frac{dm_{EGR}}{dt} = -x_{EGR} \cdot \dot{m}_\beta \tag{2.4}$$

By using the total differential, the change in the EGR mass in the container can also be formulated as shown below:

$$\frac{dm_{EGR}}{dt} = m \cdot \frac{dx_{EGR}}{dt} + x_{EGR} \cdot \frac{dm}{dt} \tag{2.5}$$

As displayed in Figure 2.8, the value of interest is the course of the EGR rate in the container, which is equal to the EGR rate of the outgoing mass flow over time dx_{EGR}/dt. This simultaneously represents the EGR rate that will be provided to the cylinders in the next engine cycle. Equation 2.5 can be resolved according to the value searched for. The term below arises:

$$\frac{dx_{EGR}}{dt} = \frac{1}{m} \left(\frac{dm_{EGR}}{dt} - x_{EGR} \cdot \frac{dm}{dt} \right) \tag{2.6}$$

The so far unknown gas mass in the cylinder m can be calculated on the basis of the ideal gas law [52]:

$$m = \frac{p \cdot V}{R_s \cdot \vartheta} \tag{2.7}$$

All the quantities within this equation are known. The pressure p and temperature ϑ can be measured for the respective operating points. The volume of the container corresponds to that of the inlet path from the EGR discharge point to the inlet valves and is known a priori as well as constant. The gas constant R_s can also be taken as a constant by simplifying it and assuming it to be an ideal gas.

All the values from Equation 2.5 are thus known and the desired quantity dx_{EGR}/dt can be calculated as follows:

$$\frac{dx_{EGR}}{dt} = -\frac{R_s \cdot \vartheta}{p \cdot V} \cdot \dot{m}_\alpha \cdot x_{EGR} \tag{2.8}$$

For a transient negative load point change, the EGR rate in the container can be plotted as a function of time, as shown in Figure 2.9 using the green curve as an example. In addition to the green curve, the diagram also illustrates the engine torque curve in black. Two remarks are required on this subject. First, the torque gradient is assumed to be constant across the entire engine map and during the load step itself. This simplifying assumption has no influence on the search for a worst case. On the other hand, the specific value assumed to be constant is decisive, since it is used both for the evaluation (see below) and for the input variables due to the course of the mass flow. The value used in this analysis is $-750\,Nm/s$:

$$\frac{dT}{dt} = -750\,Nm/s = constant \tag{2.9}$$

This value originates from the evaluation of a large number of vehicle measurements. There is a dependency on the respective driving mode, the driving speed, the load point and the current gear, but there is no predefined value. The figure above is chosen conservatively, i.e. it is a torque gradient in the range of the upper 10 % of the evaluated measurements. It is particularly important to note that no fuel cut-off is permitted for vehicles with GPF. This reduces the torque gradients, since torque reduction is implemented with the aid of poor efficiency and late ignition timing. However, the evaluation only

involved vehicles without GPF. For vehicles with GPF, this means that the selected torque gradient is without doubt also a worst-case value.

For vehicles with GPF, this means that the selected torque gradient is without doubt also a worst-case value. Chapter 4 is the assumption of a constant engine speed during the load transient. The evaluation of the vehicle measurements mentioned proves that the assumption is valid.

For reasons of clarity, all simplifications of the model are summarized again in the following list:

- ideal and homogeneous mixture of gases in the container
- homogeneous temperature and pressure in the container
- constant torque gradient across the entire engine map and during the load step
- constant engine speed during tip-out

What is provided by the model is the course of the EGR rate in the container and thus the course of the in-cylinder EGR rate. In addition, the linearized course of the torque is assumed and consequently known. Figure 2.9 depicts an example of both curves in a diagram and simultaneously serves to illustrate the evaluation methodology. The evaluation criterion is the difference between the current EGR rate in the container when the load point has been set $x_{EGR,tt}$, i.e. the target torque has been reached, and the EGR rate intended for the target load point $x_{EGR,mb}$. The latter, as well as the starting EGR rate for the simulated load steps, is based on the previously shown map of consumption-optimal EGR rates. With the above variables, the evaluation parameter $\Delta x_{EGR,diff}$ can be formulated in the following manner:

$$\Delta x_{EGR,diff} = x_{EGR,tt} - x_{EGR,mb} \tag{2.10}$$

This evaluation is carried out for all possible negative load steps within the measured stationary engine map, which in turn, is again the stationary engine map including the EGR rates for maximum efficiency. It is measured in a grid of $500\,rpm$ in terms of engine speed and in a grid of $2\,bar$ in terms of engine load or mean indicated pressure (*imep*). The consideration of all possible load steps within this grid means that for each measured load point (MLP) all measured operating points at the same speed that have a higher load are considered with respect to a negative load step with the MLP as the target operating point. This can be further illustrated in an example. The load point considered by way of example features an imep of $4\,bar\,imep$ and an engine speed of $2500\,rpm$. In this case, taking all the possible tip-outs into account means taking into consideration the load points at $6\,bar\,imep$, $8\,bar\,imep$, $10\,bar\,imep$, $12\,bar\,imep$ ... up to $24\,bar\,imep$, all at $2500\,rpm$ as starting operating points. This makes a total of 10 possible load steps for this single operating point of $bar\,imep$ and $2500\,rpm$ and shows the large number of cases analysed. Thus, the simplicity of the model as well as the short computation time are of great advantage.

The presentation of this quantity of results is not trivial, but is nevertheless implemented in Figure 2.10. However, the figure requires some further explanation to be fully understood. The map itself depicts the $\Delta x_{EGR,diff}$ worst case values of each single operating point. This means that the displayed values always represent the maximum of $\Delta x_{EGR,diff}$ out of all the simulated load steps for the respective load point. In order not to lose the information about which of the tip-outs examined is the most critical, the numbers are displayed in boxes. The box accordingly contains the most critical starting load point, or its imep, respectively, of all tip-outs examined with the load point shown in the map as a target load point.
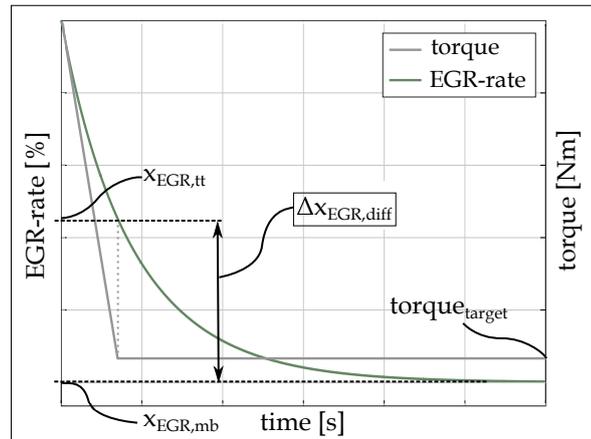
Figure 2.9: Calculated EGR-rate and input torque from the simplified container model.
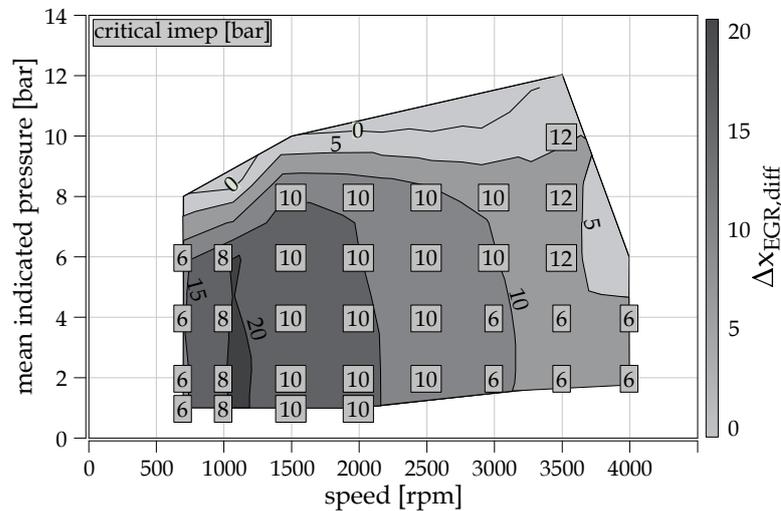
Figure 2.10: Engine map with evaluated results from the simplified model (in accordance with the results presented in [60]).

In order to find the worst-case load step from this information two criteria have to be observed. First, the area with the largest $\Delta x_{EGR,diff}$ must be found, which obviously is the load range at $1000\,rpm$ engine speed. Since all the operating points reveal an almost identical quantity of $\Delta x_{EGR,diff}$, a second criterion must be included, which is the EGR tolerance of the analyzed load area. As this is mainly a function of the engine load, the operating point with the lowest load represents the worst case, which is the point at $1000\,rpm$ and $1\,bar\,imep$. The figure shows that the corresponding starting load point is the point at $1000\,rpm$ and $8\,bar\,imep$. Table 2.1 summarizes the worst case negative load step.

The derivation of worst-case operating points is permissible despite the simplifications made. However, the transmission of the simulated $\Delta x_{EGR,diff}$ values is not permitted. At this point, it needs to be emphasized that the worst case presented here applies specifically to the engine under investigation, in particular to the measured stationary engine map including the EGR-rates for maximum efficiency, the size of the intake path and the applied torque gradient. Nevertheless, the results are representative for

Table 2.1: Worst case load step.

	engine speed	engine load	EGR rate
starting point	$1000\,rpm$	$8\,bar$	$23\,\%$
ending point	$1000\,rpm$	$1\,bar$	$0\,\%$

the typical case of a turbocharged SI engine with downsizing, as these criteria can be compared well with other engines of this class.

The analyses carried out within the scope of this section show that the results are particularly sensitive with regard to:

- characteristics of the engine map with fuel-consumption-optimal EGR rates
- torque gradient
- size of the intake path

All these parameters are considered in and covered by the presented model. They can easily and quickly be transferred to another engine or a sensitivity analysis in general.

2.2 Potential solutions

In the last section, the worst-case tip-out was defined. Generally, a system that can handle this load step can also handle any other load step. Key questions that arise on the basis of the above elaborations are: Is it possible to implement LP EGR and to handle the worst-case load step? What are the possible and available solutions? What are the requirements for a so far unknown solution? What are the specific advantages and disadvantages?

As in Chapter 1.2 and what is currently known from the literature, two fundamentally different approaches are distinguished and examined to answer these questions. The following sections emphasize the concrete differences between the two approaches in terms of levers and impact. In addition, the advantages and disadvantages are explained, as well as the reasons for investigating both paths. The implementation of the two solutions and the detailed presentation will then then follow inChapters 4 and 5.

2.2.1 Hardware-based approach

It is already known that a straightforward implementation of the worst case tip-out with LP-EGR is not possible. This is shown, for example, in Figure 2.2, where a negative load step deviating from the worst case leads to the dilution limit being exceeded. The measurements in Chapter 4 will confirm this fact within a concrete engine application. It is thus reasonable to think about the engine's design and what could be changed to allow easy implementation and, above all, implementation without a loss of efficiency relative to the stationary potential. The first approach pursued in this thesis is therefore to adapt the existing hardware with the aim of modifying the engine such that the worst-case tip-out can be handled. With regard to Figure 2.3, essentially the first two of the three possible control levers are utilized, i.e. an intervention in the torque curve of the engine as well as an intervention in the curve of the in-cylinder EGR rate.

An implementation is to be based on the premise to modify the operating strategy only to the extent necessary that a realization of the hardware modification is possible under all possible operating conditions. Though, the objective is to squander as little as possible of the stationary potential by manipulating the operating strategy. Table 2.2 below is intended to list the basic advantages and disadvantages of a hardware approach.

Table 2.2: Basic advantages and disadvantages of the hardware-based approach.

Advantages	Disadvantages
+ possible use of the full stationary potential	- high cost
+ use of the manipulated hardware for additional purpose	- difficult to package
+ robust implementation possible including back-up against instable motor operation	- increasing variety of engine design variants
+ more accurate than a model-based solution	- additional degrees of freedom
+ less input data required for proper implementation and operation	- input data must be provided by extensive measurements
+ most direct approach (no time delay due to computation)	

Despite the obvious disadvantages of an enormously high effort with the associated problems of packaging and costs, a hardware approach is pursued for two core reasons:

- The thesis strives for a holistic evaluation of the implementation and the associated necessities for transient engine operation with LP-EGR.
- Investigation into the possibility of implementing the full stationary potential of LP-EGR in all operating areas.

The presentation of the various approaches in detail and the associated results are given in Chapter 4.

2.2.2 Model-based approach

The second approach examined in this thesis addresses the problem from another perspective. One central goal is to keep the existing engine hardware and engine design. Instead, the problem of implementing LP EGR during dynamic engine operation is addressed using a model approach. With respect to Figure 2.3, the manipulation is essentially carried out via the last two control levers. On the one hand, the course of the in-cylinder EGR rate, or the rate itself, is adjusted. On the other hand application methods are used to increase EGR tolerance.

It is important to note at this point that the model developed in this paper cannot be used directly for the implementation of LP EGR. However, it provides all the information needed for that purpose. It establishes a basis for implementing LP EGR with the usual effort that would be required, without the additional degree of freedom offered by external EGR. The process itself can then be carried out in two different ways. One is an online use during engine operation with the help of an optimizer embedded in a tailored operating strategy, the other is an offline use to create an application database. In any case, the primary objective of the model is therefore to reduce to a minimum the enormously high measurement effort, which is further increased by the additional degree of freedom of an external exhaust gas recirculation system.

The model developed in this thesis is subject to key requirements. These include short computing times and a correspondingly simple structure. In addition, the model input parameters all have to be known during engine operation, i.e. they have to be measured by the sensors that are usually installed. After all, the model must have a physical basis and must not be based on neural networks, for example.

Again, Table 2.3 below is intended to list the basic advantages and disadvantages of the model approach. The major and most obvious disadvantage of the model solution is the loss of efficiency compared to the stationary potential. However, Chapter 5 will prove that the model is as simple as possible and at the same time optimal based on the existing hardware. Above all the considerable reduction in the measuring effort, as well as the multi-layered application possibilities and the understanding of physical connections favour the model solution.

Table 2.3: Basic advantages and disadvantages of the model-based approach.

Advantages	Disadvantages
+ physical basis conducts to enhanced insight	- models contain simplifications and thus inaccuracies
+ no packaging problems	- efficiency disadvantages are to be expected
+ existing model requires fewer measurements to generate input data	- additional computing effort (possible time delay)
+ no additional costs	
+ easy to implement on other engines	
+ no additional degrees of freedom	

A combination of a hardware and a model solution is also conceivable, but is not analyzed in the framework of this study. The resulting synergies as well as the additional robustness counteract the significantly increased effort and the possible need to extend the model with additional degrees of freedom due to additional hardware components.

3 Experimental analysis

Engine test bench measurements are required for both the hardware-based approach and the model-based approach. This chapter presents all the information required for that purpose. First, the engines investigated are introduced. Subsequently, the engine test bench setup and the particularly relevant measurement systems including the existing restrictions are contemplated. Finally, the methodology for evaluating the measurement data is provided, as well as its limitations.

3.1 Investigated engines

The engine used in the experiments is an inline four-cylinder production engine from the BMW gasoline engine family. Specifically, the unit investigated is the upper power stage of the turbocharged four-cylinder SI engines, labeled *B48B20O0*. Table 3.1 contains the most important engine data. These are equally valid for the 1D CFD simulation model applied in Chapter 4.

Table 3.1: Engine specifications

engine name	BMW B48B20O0
engine type	gasoline, 4-stroke
rated power	$185\,kW$ @ $5200 - 6500\,rpm$
rated torque	$350\,Nm$ @ $1450 - 4800\,rpm$
cylinders	4 inline
displacement	$1998\,cm^3$
compression ratio	10.2
cylinder bore	$82\,mm$
piston stroke	$94.6\,mm$
valves per cylinder	2 intake, 2 exhaust
valve train intake	continuously variable valve lift (0.2 - $9.7\,mm$) and timing (52.5 - $122.5\,CAD\,aTDC$)
valve train exhaust	constant valve lift ($9.9\,mm$) and continuously variable valve timing (60 - $120\,CAD\,bTDC$)
design of intake ports	high tumble ports
volume intake path	$5500\,cm^3$ (from discharge point EGR path to intake valves)
charging system	twin-scroll turbocharger
load control	interaction between throttle, intake valve lift and timing as well as turbocharger wastegate
injection system	direct injection, homogeneous, stoichiometric, timing synchronous to induction

In addition to the first test engine, which is used for the measurements regarding the hardware-based approach as well as for generating measurement data to set up the model, a second engine is indispensable to validate the model in Chapter 5. This motor differs from the one shown in Table 3.1 in terms of the following properties:

- compression ratio of 11.6
- intake ports produce lower tumble: less charge movement

Further deviations from the basic engine could not be realized due to the unavailability of EGR components for another engine configuration. The basic design of both engines is thus identical and is illustrated in Figure 3.1. The engine described in Table 3.1 is supplemented by an LP-EGR path including an EGR cooler as well as an EGR valve.
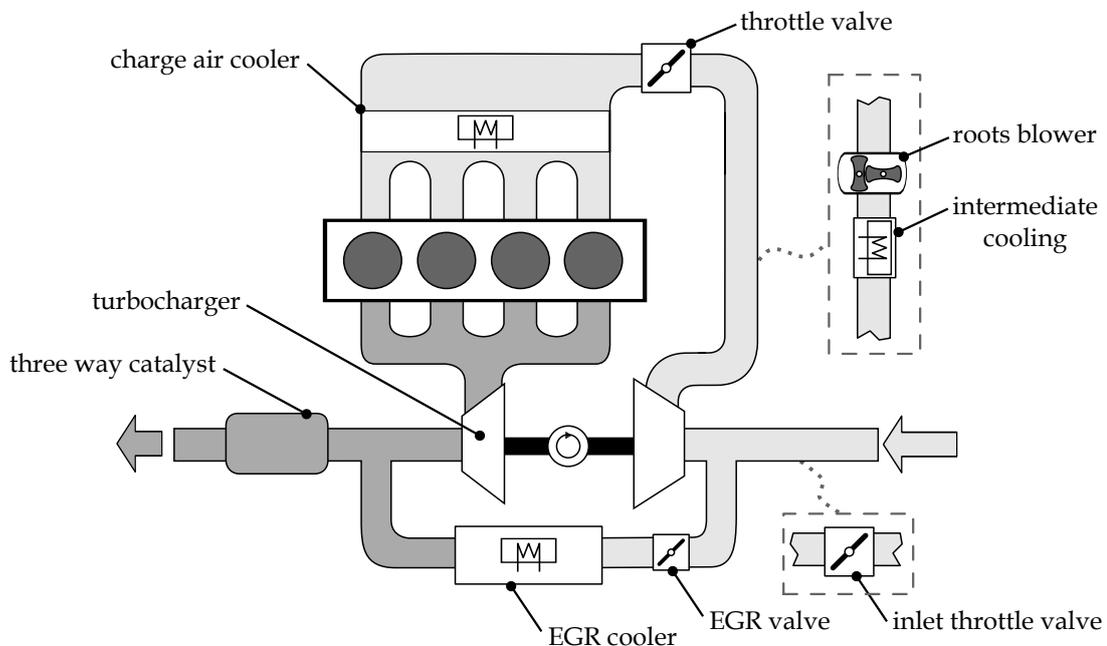
Figure 3.1: Configuration of investigated engine(s)

The illustration requires two more detailed comments. To this end, the extraction point of the EGR path, which is obviously located in front of the TWC, is discussed more precisely by providing proof of the benefits. Moreover, reference is made to the dashed boxes in the graphic. They represent a change in the motor configuration for the measurements in Chapter 5.

EGR extraction point in front of the TWC

As explained in Chapter 1.2, an LP EGR system can involve extraction either in front of or behind the TWC. The particular advantages of extraction in front of the TWC are the increased pressure gradient available, the recirculation of unburned CO, HC and H_2 and the consequently increasing reactivity of the residual gas as well as the reduced loss due to incomplete combustion. In comparison to extraction in front of the TWC, removing the exhaust gas behind the TWC tends to lower the fouling risk. In addition, due to its decreased NOx-fraction, it has a minor propensity for knocking. Finally, the effects on turbine operation are far less severe in the case of extraction behind the TWC. The reason for selecting and implementing an extraction point in front of the TWC, despite being within the scope of this analysis, will be demonstrated by the results of representative measurements below.

First Figure 3.2 depicts a comparison of the indicated fuel consumption of a load variation at an engine speed of $1500\,rpm$ between measurement with EGR extraction in front of and behind the TWC. The measurements were carried out on another test engine. This is not referred to in Section 3.1 as it is only used at this one point. The reason for using this engine is the availability of EGR parts for exhaust gas extraction in front of and behind the TWC. In fact, it is a 3-cylinder inline engine labeled *B38B15M0*,

which is similar to the presented engines in terms of its basic properties regarding the combustion process and load control, but featuring a compression ratio of 11. It becomes very clear that in this area, which is particularly important in customer and driving cycle operation, there are significant advantages regarding fuel consumption and efficiency for the extraction point upstream of the TWC over the entire load variation.
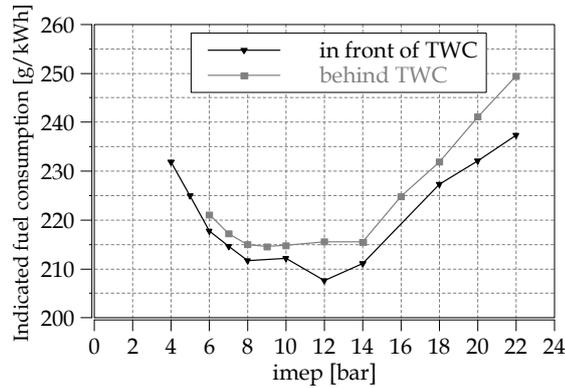
Figure 3.2: Comparison of LP EGR extraction point: indicated fuel consumption

To understand where this big difference originates, Figure 3.3 next provides the essential indications. Based on the identical load variation, the external EGR rate for optimum efficiency and the burning duration from 10 to 90 % mass fraction burned are shown. A major and immediately noticeable difference is the significantly higher EGR rates when the exhaust gas is extracted in front of the TWC. This is only possible as a result of the increased pressure drop available. Of course, the increased EGR rates also entail all the associated benefits in terms of correspondingly increased quantities. Despite the significantly higher EGR rates, the burn duration is only slightly increased. This is due to the higher proportion of unburned residuals mentioned above, which are reincorporated into the combustion process after recirculation.
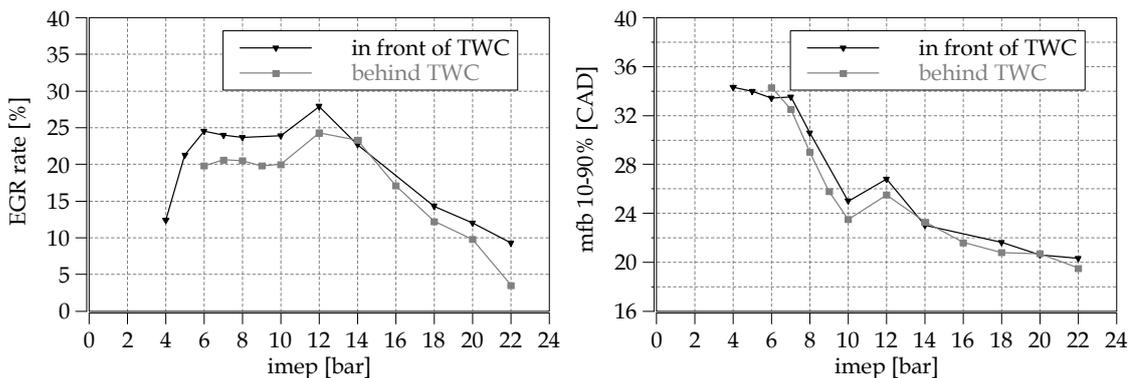
Figure 3.3: Comparison of LP EGR extraction point: EGR rate and burning duration $10 - 90\%$

Figure 3.4 below shows that the configuration where the exhaust gas is extracted behind the TWC does in fact also have advantages. Despite the lower EGR rates, equal or better $mfb50$ values can be implemented over a large load range. This can be explained by the lower proportion of NOx, which is

reflected in a significantly reduced knocking tendency. It becomes particularly clear when the advantage regarding $mfb50$ is related to the EGR rate, i.e. CAD of $mfb50$ per EGR rate percentage point.
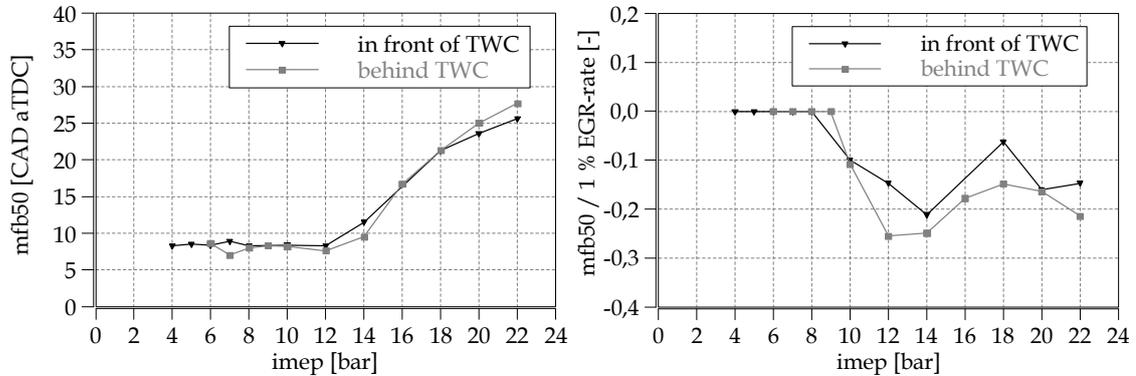
Figure 3.4: Comparison of LP EGR extraction point: mfb50 and mfb50 / 1 % EGR-rate

It should thus be noted that the increased EGR rates and the subsequent efficiency advantage clearly favor placing the extraction point in front of the TWC. Finally, the increased EGR rates in turn again pose a disadvantage during dynamic behavior and are consequently precisely the worst case that will be used for research in this thesis. Problems deviating from the efficiency point of view, such as the fouling risk, are classified as manageable and are not the subject of this investigation.

Modified engine configuration for the model setup

Neither the initial assembly nor the standard configuration of the motor contains the parts framed by dashes in Figure 3.1. On the one hand there is an additional throttle valve in front of the EGR inlet and on the other hand there is a unit consisting of an additional charge air cooler and a roots loader mounted downstream.

The throttle valve is the series throttle valve described in the specifications, as already used in the engine. The control works via an additional slave control unit. The compressor is a roots blower which was developed for the Mini Cooper S. It is belt-driven by a 2-pole electric machine with an electrical power of $7.5\,kW$. The control and regulation works with the aid of the test bench software. The power and efficiency data of the roots blower can be taken from the characteristic maps in Appendix A.1. These values are crucial for correcting the indicated efficiency in operating points where the roots blower is used. The correction procedure is explained in Section 3.3.

The additional components are necessary to generate the measurement data for setting up the model in Chapter 5. This involves on the one hand measuring operating points which have a pressure gradient across the EGR path that is no longer large enough to be able to set the desired EGR rate. If this is the case, the additional throttle valve can be used to decrease the negative pressure at the EGR discharge point and thus to increase the pressure gradient across the EGR path. On the other hand, it is about operating points where the engine can no longer adjust the desired load complying with a high EGR rate. The additional charging via the roots lader provides assistance in this respect. Details on why the measurement of these operating points is required can be found in the aforementioned chapter.

It is important to note that the volume of the inlet tract changes due to the additional installations. However, this is only relevant for dynamic operation. In this configuration, though, the system is only used for stationary measurements.

3.2 Engine test bench

The data from two investigated operating points may only have slight differences within this study. In order to make them apparent, high demands are placed on the experimental setup. Figure 3.5 contains an overview of the engine test bench design as well as of the positioning of the measuring points on the test engine.
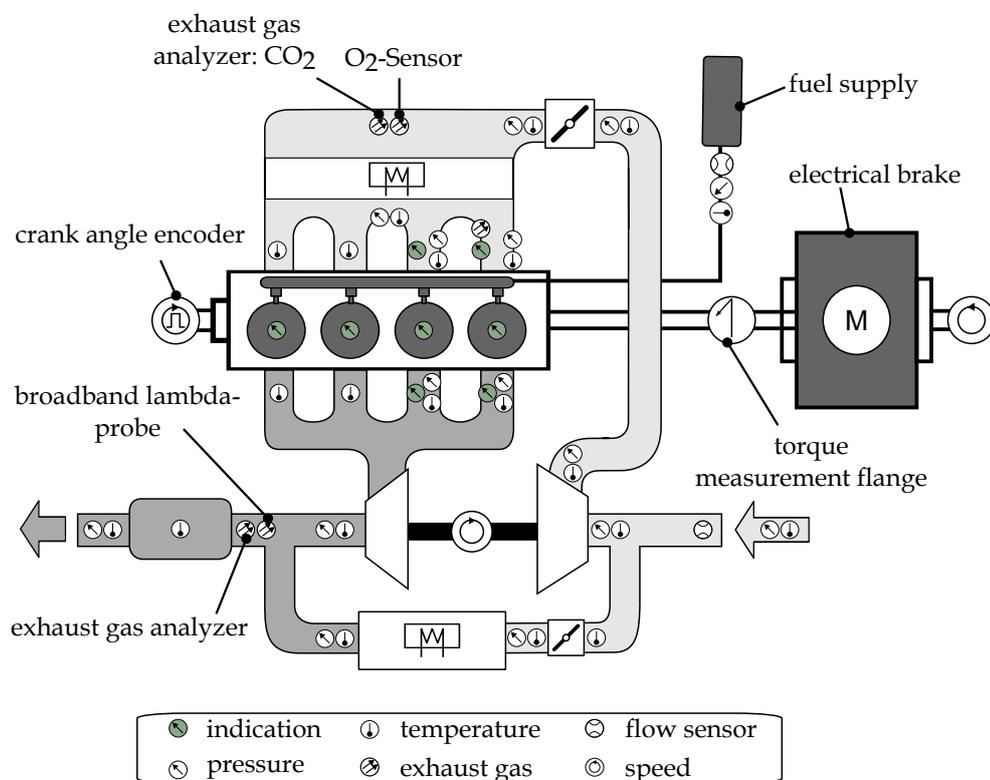
Figure 3.5: Engine test bench (configuration with measuring points)

Fundamentally, on the test bench used and shown, a distinction can be made between two groups of measured variables and similarly between two measuring systems. The criterion for differentiation is whether the measurement system acquires measurements slowly, i.e. based on time-averaged values as well as to a finite time step size, or quickly, i.e. with crank angle resolution. The first of the two systems, known as the automation measurement system, which is *Morphee* from *D2T* in this case, is the slower of the two and comprises several other submeasurement systems. First of all, this includes all temperature and pressure measuring points. They are recorded and read in directly via thermocouples or resistance strain gages, respectively, and are crucial for the description of the gas mixture?s thermodynamic state. The speed and load of the engine, in turn, are controlled and sensed by an asynchronous electric machine. The torque output is additionally measured via a measuring flange mounted on the output

shaft. A combination of a fuel flow meter and a fuel density measurement (*AVL PLU KMA 4000*) determines the amount of fuel absorbed. This is crucial for calculating the indicated efficiency. In the one and same way, exhaust gas measurement technology is particularly important for engines with external exhaust gas recirculation. Within this context, there is a measuring point in front of TWC where the concentrations of CO_2, CO, HC and NOx as well as the fuel?air ratio, are determined using an exhaust gas analyzer (*Pierburg AMA 4000*). In the same exhaust gas analyzer there is a second CO_2 measuring line used to determine the CO_2 concentration from a measuring point in the intake manifold. Deviating from the general formula 2.1 for the determination of the exhaust gas fraction x_{EGR}, the EGR rate is calculated as described in [23] as follows:

$$x_{EGR} = \frac{x_{CO_2,im} - x_{CO_2,air}}{x_{CO_2,eg} - x_{CO_2,air}} \cdot 100\% \tag{3.1}$$

Here, $x_{CO_2,im}$ denotes the CO_2 concentration in the collector, $x_{CO_2,eg}$ the CO_2 concentration in the exhaust gas before TWC and $x_{CO_2,air}$ the CO_2 concentration in the ambient air that is assumed to be known and constant (0.04 %). This exhaust gas or EGR rate measurement method is very accurate, but due to the long gas paths and the associated time delay, it is only suitable for stationary measurement points. For this reason, a second method is applied within the automation measurement system, which detects the EGR rate in an analogous manner to Formula 3.1 but with the respective O_2 concentrations using one broadband lambda probe in front of TWC and another one in the intake manifold. The mathematical formulation is shown below:

$$x_{EGR} = \frac{x_{O_2,im} - x_{O_2,air}}{x_{O_2,eg} - x_{O_2,air}} \cdot 100\% \tag{3.2}$$

The nomenclature corresponds to that of the CO_2 concentrations from Formula 3.1. As a result of the short response time of the lambda probes, this enables a significantly faster and virtually delay-free system response to be implemented and the EGR rate to be determined during dynamic operation. However, this method is significantly less accurate than the measurement with the exhaust gas analyzer. A numerical example will illustrate this in the following. Based on the manufacturer's information on the exhaust gas analyzer, the accuracy in determining the CO_2 concentration is +/-5 % *rel.* regardless of the measuring range [76]. Assuming an EGR rate of 30 % and stoichiometric combustion, the CO_2 concentration in the exhaust gas is 14.5 %, that in the collector is 4.4 % and that in the environment remains constant at 0.04 %. If all these values are now presumed to deviate most unfavourably, a maximum deviation of the EGR rate of +/-0.9 % *abs.* arises. This deviation is classified as quite small since it represents, as described before, the absolute worst case. When this numerical example is now applied analogously for the determination using the O_2 concentration, first of all the significantly higher basic accuracy is noticeable. While the error in the range of a stoichiometric fuel-air ratio at +/-0.3 % *abs.*, is still very low, the possible error in the range of high oxygen rates, such as those present in the intake manifold during the measurement, is significantly higher at up to +/-7 % *rel.* [10]. In contrast to Siokos findings [86], the oxygen content in the ambient air may again be assumed to be constant at 20.95 % within the calculation, since constant ambient conditions and approximately constant humidity are ensured at the engine test bench. The assumptions cause a maximum error of +/-5 % *abs.* regarding the EGR rate. To reduce the error, in a second step, an O_2 sensor specially developed by a supplier for

measuring oxygen in the inlet path is used instead of the second lambda probe in the intake manifold [16]. Thus the initial error in the oxygen determination is reduced from $+/-7\,\%\,rel.$ to $+/-2\,\%\,rel.$ and the total error regarding the EGR rate is decreased to $+/-1\,\%\,abs.$ assuming the worst case. Hence, the error is approximately at the level of the value determined with the exhaust gas analyzer.. Comparable sensors or modified lambda probes are also used in the studies by Sarlashkar [82], Surnilla et al. [93] and Siokos [86].

Finally, the automation measurement system also interfaces with the engine's electronic control unit (ECU) via the *INCA* tool, which makes it possible to adjust all the actuators on the engine separately, such as the ignition angle, valve timing or injection quantity. The actual values of the actuators, i.e. all information available in the ECU during vehicle operation, can in turn be transmitted to the automation measurement system and integrated into the measurement in the same way.

Since dynamic measurements are not only included in this work but a key factor, it is important to point out that the automation measurement system is also able to record each value's course over time instead of the time-averaged value. The maximum resolution in this case is $1000\,Hz$ for a number of measurands. However, this is still slower than the high resolution measurement system that is used. This refers to the indication system (*AVL Indicom*) which measures the cylinder, intake and exhaust pressures with a resolution of $0.1\,CAD$. These values serve as indispensable input variables for the pressure trace analysis described in Section 3.3 below. As can be seen in Figure 3.5, all cylinders are indicated, as are the intake and exhaust ports of the test engine's Cylinders 3 and 4. An additional indication of the intake and exhaust ports of Cylinders 1 and 2 would be desirable. However, for installation space reasons, these could not be implemented in a feasible way.

Piezoelectric pressure transducers are applied to measure the cylinder pressure. This process makes use of a physical property of a piezoelectric body: the fact that a mechanical deformation leads to a proportional change in the electrical polarization [71]. Difficulties are caused by the lack of assignment to the absolute pressure level and the occurrence of thermal drift. Water-cooled pressure transducers are used to reduce the thermal drift. The absolute pressure adjustment is carried out during the compression stroke when the valves are closed using two discrete crank angles. The adjustment is conducted using the calculated polytropic pressure for both discrete moments. Furthermore, the pressure signal must be assigned to the actual crank angle position of the engine. As a reference point, the top dead center is used, which is determined by means of a capacitive sensor. This is necessary since the geometric TDC deviates from the angle of the cylinder pressure maximum due to thermodynamic losses. The concrete correlation with a crank angle is achieved by means of an incremental encoder.

Rather than measuring the cylinder pressures, piezoresistive pressure transducers are applied to capture the crank-angle-resolved pressures in the intake and exhaust ports. This measuring principle uses resistors connected together to form a Wheatstone measuring bridge. The measuring bridge is detuned by the effect of pressure, and a pressure-proportional output signal is generated. The absolute pressure is determined directly by the sensors. Since thermal drift can also occur at least on the outlet side, the sensors installed there are likewise water-cooled. Furthermore, the sensor signals are compared with additional absolute pressure sensors.

For stationary measuring points, the measured values are averaged over a period of 30 s within the automation measurement system, while 256 successive cycles are recorded within the indication system and then averaged for the relevant quantities, such as *imep* or indicated efficiency. For dynamic measurements, the variables are recorded in the automation measurement system with the corresponding

sampling frequency, while the individual engine cycles can be viewed in the indication system. The calculation of the characteristic values from the indication system is then also based on individual cycles. However, it must be borne in mind that the values from the fuel measuring system for calculating the efficiency or the measured values from the exhaust gas analyzer are not reliable during dynamic operation since the systems cause significant delays. These values must therefore be determined for the EGR rate either with the aid of a simulation or using alternative measuring systems, such as the lambda probe and the O_2 sensor.

Summing up, if the measurement data are accurately determined together with the appropriate background knowledge, their informative value is of high quality. Simultaneously, they offer a solid basis for comparison with and validation of simulations and, last but not least, a good basis for creating models which, conversely, are intended to considerably reduce the actual measuring effort and thus the complexity.

3.3 Measurement data analysis

Since the measurement data or the raw data from the test bench are not yet sufficient as a source of information and need to be corrected in some cases, this section covers all topics related to the primary and standardized post-evaluation of the measurement data. First and foremost, so-called pressure trace analysis (PTA) or three pressure analysis (TPA) is to be mentioned, which is essential for generating data for model-building and especially for understanding the model in Chapter 5. In addition, two correction terms are derived and presented. One is required due to the method used to calculate the EGR rate for measurements deviating from the stoichiometric fuel?air ratio; the other is needed due to the test setup deviating from the series engine.

Pressure Trace Analysis

Although some work has already been conducted on a real-time capable PTA on the engine test bench, due to its complexity, the PTA must still be carried out as part of a post-evaluation [66]. Thus, with the data from each operating point measured in stationary conditions, a PTA is carried out using the 0D/1D gas exchange and the *GT Power* combustion tool. Within the *GT Power* software, the PTA procedure is referred to as TPA. Within the scope of this work, however, both terms and designations are to be regarded as equivalent.

The PTA allows the burn rate to be determined based both on the crank-angle-resolved cylinder and on the intake and exhaust pressures. The burn rate, in turn, is essential for quantifying other important values such as the start, duration or end of the combustion process, as well as quite a number of other values which are specified below. To run a PTA, a simulation model is built in which a single cylinder and the associated inlet and outlet ports are truncated from the entire engine at the points where the respective pressure transducers of the real engine are located in the inlet and outlet. An example and simplified model of this kind is shown in Figure 3.6.

The following description is divided into four parts. The necessary input data for a PTA are specified, followed by discussion on the general procedure within the simulation method, the limits of a PTA and the most important results.
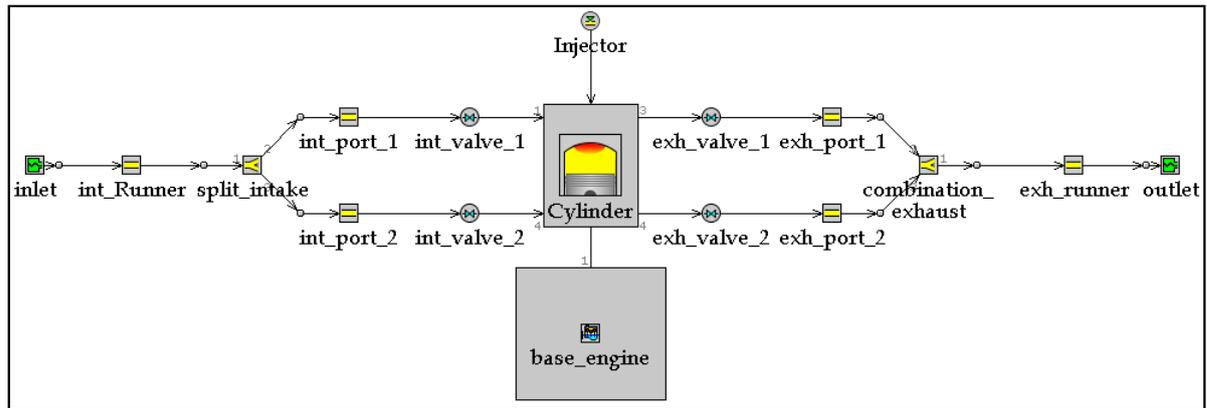
Figure 3.6: PTA model in 1D CFD *GT Power* software

At first, the simulation requires important input data from the measurement, specified within the interfaces which are illustrated by two elements: the *Inlet* and *Outlet*. These are:

- The measured indicated pressure signals in the intake and exhaust ports.
- The average intake and exhaust pressure to adjust the indicated pressure signals, if necessary.
- The average intake temperature as well as the ratio of externally recirculated exhaust gas, i.e. the EGR rate to calculate the charge density.

Additional data required to run a PTA and provided not within the Inlet and Outlet interfaces but within the simulation model in general are listed below:

- The valve lift and timings of intake and exhaust valves. These are crucial to calculate the gas exchange and the cylinder charge correctly.
- The fuel-air ratio or alternatively the fuel mass injected per cycle.
- The fuel properties (lower heating value, density, C/H ratio, evaporation heat).
- The indicated cylinder pressure signal to iteratively calculate and to validate the simulated cylinder pressure.
- The exhaust emissions of HC, CO and H_2. As there is no measured value of H_2, it is calculated as described by Witt [112] based on HC and CO.
- The mechanical engine friction as a function of engine speed.

The calculation of the PTA begins with a gas exchange according to the specified boundary conditions, the resulting cylinder charge and a consecutive combustion cycle with a standard burn rate specified by the software. Once the first simulated combustion has been completed, a further gas exchange occurs, again using the specified boundary conditions. The second combustion takes place after the gas exchange. During this event, the burn rate from the first cycle is adjusted in such way that the simulated cylinder pressure converges with that from the measurement. By repeating the described procedure until the error between the simulated and measured cylinder pressure is below a specified error, the simulated burn rate theoretically corresponds to the real burn rate of the engine.

However, there are still deviations or operating points that do not converge. The discharge coefficients at the valves, for example, are a source of error and thus simultaneously a control lever. They are based on measurements on an idealized test bench and are subject to different boundary conditions within the full engine. Another source of error is the heat transfer at the cylinder walls. It is based on model

values (cf. [41]). The software offers correction factors for both, the valve discharge coefficients and the heat transfer so that the cylinder filling or burning speed can easily be adjusted if necessary. In the case of successful measurements and well-adjusted simulation models, it should after all be possible to calculate a PTA correctly, independently of the engine operating point.

The main results of the PTA that exceed the information content of the raw measurement data and that are of high relevance for this work in general, but primarily for the explanations in Chapter 5, are the information related to the burn rate and its course, such as the start, end and duration of the combustion. Of course, any other discrete point within the burn rate can also be determined. In addition, the course of the temperature inside the cylinder can be determined. One key parameter from the PTA or the associated gas exchange calculation is the internal residual gas rate or internal EGR rate. It determines the total EGR rate, i.e. the sum of the internal and external EGR. These and the ratio between internal and external EGR are crucial in evaluating the possible dilution with exhaust gas (cf. [74]). The PTA and its results are also the prerequisite for a loss analysis. This is calculated automatically when the PTA is processed. Consequently, the results are available for each individual operating point measured in stationary conditions.

From the explanations above, but especially from the listed input data and the calculation procedure, it becomes easily apparent where the limits of the PTA are and where it is particularly significant to ensure accurate processing. The importance of the indicated pressures and the input measurement data in general should be noted here first of all. Any error in the measurement inevitably causes a lack of precision in the simulation. Since a simulation is always a simplification, however, there are also systematic errors, such as those described with the gas exchange and with the wall heat transfer. It is crucial that the adaptation factors are chosen carefully on a physical basis. All results need to be challenged critically at all times. Finally, it must be emphasized that the implementation of a PTA and a loss analysis only applies to stationary measuring points. On the one hand, as described earlier, key input variables cannot be used during dynamic operation. On the other hand, convergence cannot be achieved due to the alternation of the operating point and the associated constant change in important control levers such as the valve lift, lambda or ignition timing.

Correction of the indicated efficiency in operation points using the roots blower

While the PTA is a widely used and well-known tool, the correction of the indicated efficiency for operating points using the roots loader from Figure 3.1 has been specifically developed for this thesis. Therefore, rather than only the most important input and output variables, a short derivation is shown.

Since the roots loader is driven by an electric machine, it is important that the supplementary energy introduced into the process on top of the fuel energy is not mistakenly incorporated into the overall or indicated efficiency. In this context, a detailed consideration of the engine process is necessary. Later it will be seen that the absolute values of the indicated efficiency are less important for the model output. One very relevant aspect, however, is the course of the indicated efficiency and the values as a function of the EGR rate. The indicated efficiency of the motor process is calculated on the basis of measured data in line with [77] using the equation below:

$$\eta_{ind} = \frac{P_i}{\dot{Q}_{fuel}} = \frac{\oint p dV}{m_{fuel} \cdot LHV} \tag{3.3}$$

The term in the denominator represents the supplied energy quantity $\dot{Q}_{fuel}$ as the product of the lower heating value LHV and the fuel mass flow $\dot{m}_{fuel}$. The numerator contains the circular integral of the engine process, as illustrated by the pV diagram in Figure 3.7 below. The result of the circular integral corresponds to the indicated power P_i.
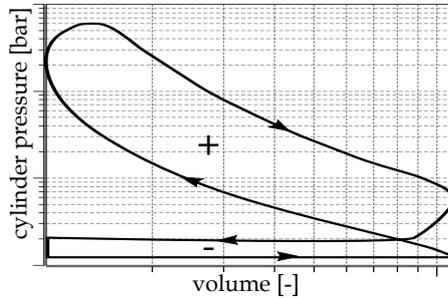
Figure 3.7: pV diagram: illustration of the engine process and the inner work

When calculating a circular integral, clockwise processes are to be considered as positive work and counterclockwise processes as negative work.

There are two ways to consider the work of the compressor in relation to the efficiency, whereby the first of which can itself be solved in two different ways. For this purpose, the hydraulic power input of the roots loader is directly taken into account in Formula 3.3 above in the following form:

$$\eta_{ind,korr,RL} = \frac{P_i - P_{RL,hyd}}{\dot{Q}_{fuel}} \tag{3.4}$$

The hydraulic power of the roots loader is subtracted from the power produced in the engine cycle and is thus not taken into account regarding the indicated efficiency, but is deducted instead. The hydraulic power of the roots loader $P_{RL,hyd}$ can be determined in two different ways. The first possible method is a calculation using the isentropic compressor power equation [116]:

$$P_{RL,hyd} = \frac{\dot{m} \cdot c_p \cdot T_1}{\eta_{is} \cdot \eta_m} \cdot \left(\left(\frac{p_2}{p_1} \right)^{\frac{\kappa-1}{\kappa}} - 1 \right) \tag{3.5}$$

All variables necessary for the calculation are known. The pressure upstream of the roots loader p_1 that downstream p_2 and the temperature upstream of the roots loader T_1 are all captured by measuring points. In the case of the isentropic exponent κ and the isobaric heat capacity c_p, the values of air are assumed as a first approximation. The mechanical efficiency of the roots loader η_m is assumed to be constant ($\eta_m = 0.85$) and provided by the suppliers datasheet [21]. The isentropic efficency, in turn, is taken from a characteristic map and depends on the volume flow and on the pressure ratio. The map is illustrated in Figure A.1 and also given in [21].

The second way to determine the hydraulic power is using the map provided on the left-hand side of Figure A.1. As with the readout of the isentropic efficiency for Equation 3.5, inaccuracies may occur due to the discrete number of sampling points in the map. Because of these possible imprecisions, the results obtained with these correction factors should be outlined first. For the validation of the correction term from Formula 3.4, the results are shown from an EGR variation, which can also be implemented without

using the roots loader. The measurement series (@ 1000 *rpm*, 12 *bar imep*) is performed in three different ways. The first series of measurements uses the wastegate to control the load during turbocharged operation as intended in the engine control unit. In the second measurement series, the load is adjusted using the throttle valve only. Finally, the last measurement series is conducted using the roots loader. The load is controlled in the same way as the second measurement series via the throttle valve, i.e. the power of the roots loader is adjusted via the throttle valve. A correctly implemented correction term thus should value the second and third measurement series with identical indicated efficiencies. The first and second measurement series deviate due to the differences regarding the load control strategy.

In order to show the comparability of the three conducted measurement series, the following Figure 3.8 and Figure 3.9 show the intake manifold and exhaust gas back-pressure, as well as the *mfb*50 and the burning duration, each as a function of the EGR rate.
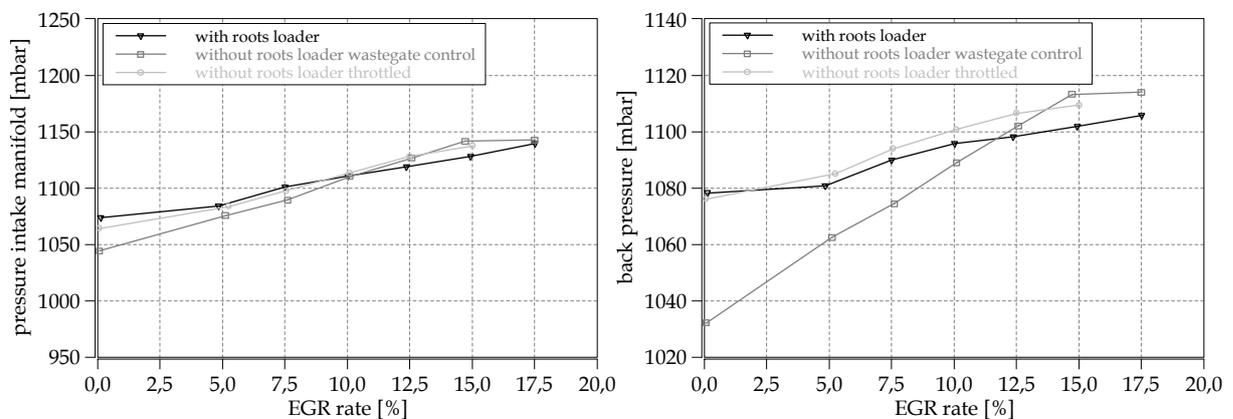
Figure 3.8: pressure intake manifold and back pressure (@ 1000 *rpm*, 12 *bar imep*) for the three measurements series

The pressure in the manifold is approximately the same for all three measurement series. Since all measurement series were also performed with the same valve lift and the same valve timing, they do not differ on the inlet side or in terms of its conditions. In contrast, the exhaust back-pressure in the throttle-valve-controlled measurement series (these two are almost identical) differs considerably from the measurement series controlled by the wastegate. This is due to the continuous motion of the wastegate, which increases the exhaust back pressure. Of course, this has an effect on the gas exchange and its efficiency. Figure 3.9 below indicates whether the high-pressure process is also affected.

Since both the *mfb*50 and the burning duration are approximately the same, almost identical high-pressure processes can be assumed for all three measurement series. Consequently, the differences between the measurement series lie primarily in their gas exchange and its efficiency.

The comparability of the two throttle-valve-controlled measurement series has been proven. The need to compare a throttle-controlled series of measurements with the series using the roots loader to validate the correction term has also been demonstrated. The series of measurements performed with the wastegate based load control mainly differs regarding the exhaust back pressure. The following Figure 3.10 depicts the efficiencies of the three measurement series as well as the curves of the corrected efficiencies of the measurement series, using the roots loader.

A number of aspects are immediately noticeable:
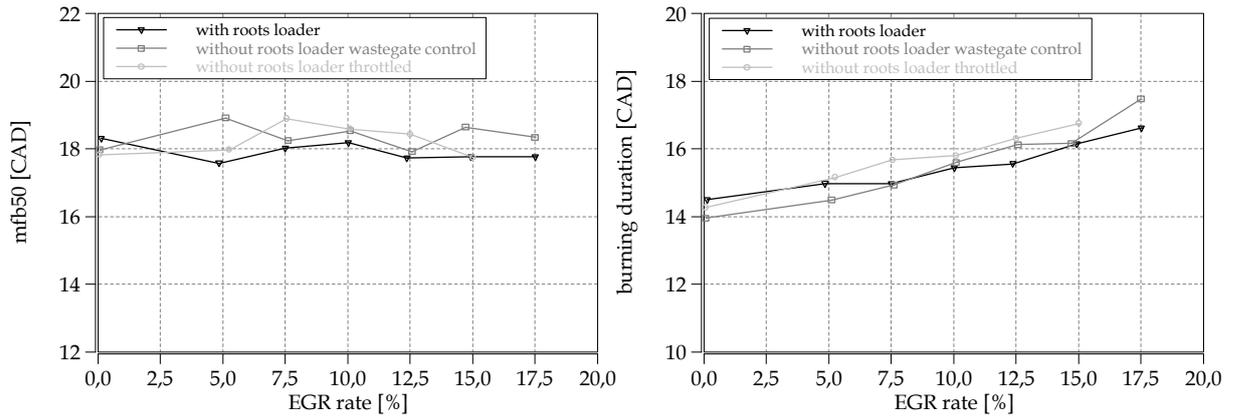
Figure 3.9: mfb50 and burning duration (@ 1000 *rpm*, 12 *bar imep*) for the three measurement series
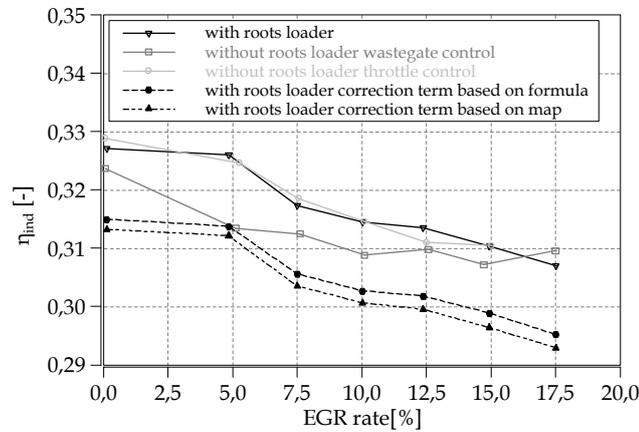
Figure 3.10: Results of an exemplary correction of the indicated efficiency (@ 1000 *rpm*, 12 *bar imep*) using the first possible method for determining the hydraulic power

- The curves of the indicated efficiencies for the two measurement series controlled via the throttle valve appear almost identical, as expected.
- The wastegate-controlled measurement series, on the other hand, differs considerably.
- The curves of the adjusted efficiencies are essentially shifted parallel to the unadjusted curve. They both deviate significantly from the original measurement.
- There is a slight difference between the two adjusted curves themselves.

The parallel shift in the two curves that follow from the correction term is essentially caused by the constant output power of the roots blower across the measurement series (load control by means of the throttle valve). The differences between the two adjusted curves are to be found in the deviation of the real compressor from the description using Formula 3.5.

There are two reasons why this variant of the correction term is not used or valid in either the formula-based or map-based versions. Firstly, the terms adjust the overall efficiency, which in turn affects the efficiency of the high-pressure process. Figure 3.9 shows that this is actually incorrect. In addition and even more crucially, the influence of the compressor is overestimated. This is because the work of the compressor is not fully transferred to the engine process as the correction term supposes,

but is still lossy owing to the roots loaders' position in the intake path. Both disadvantages can be avoided using the method or term finally chosen and introduced below.

The correction term which is ultimately selected is applied exclusively to the gas exchange loop. For the implementation, the high-pressure process and the gas exchange process hence have to be distinguished. To separate the charge exchange from the high-pressure process and be able to evaluate both processes individually, there are essentially three different methods known from the literature: [112], [85] or [103]. Figure 3.11 presents an overview of the three methods depending on their complexity or effort and on their accuracy.
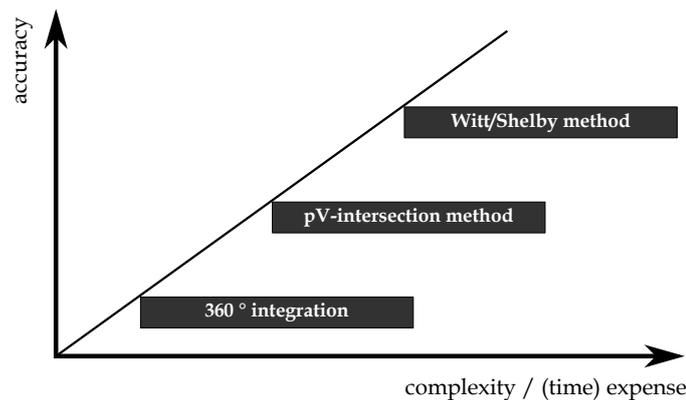
Figure 3.11: Methods for separating the engine process into high-pressure process and gas exchange process, rated by effort and accuracy.

The accuracy of the procedures increases in line with the effort they require. In the case of the 360 ° *integration*, the two parts of the engine cylce are split in a particularly simple way from BDC to BDC. The process is very well suited for both throttled and turbocharged operation and for both positive and negative gas exchange loops. However, it reveals inaccuracies in absolute values when comparing processes with different valve timings. An absolute comparison of two processes with identical valve timings, including with Miller or Atkinson strategies, is nevertheless possible without making a noteworthy error. The main strength of the *pV intersection method* is that it allows operating points with different valve timings to be compared directly. However, one disadvantage of the *pV intersection method* is that the detection of the intersection point between high pressure and charge exchange loop is not trivial and, above all, not unambiguous due to pressure oscillations. In addition, the evaluation of processes with positive gas exchange work proves to be difficult. The best and most accurate method from a thermodynamic point of view is the *Witt/Shelby method*. Again, the calculation is not simple. Furthermore, knowledge of the exact valve timing is indispensable. However, this method can be used to compare all kinds of processes, even with different valve timings and positive gas exchange loops, on the basis of absolute values.

For the correction term in this work, the simplest of the three methods, labeled 360 ? integration, is used. As explained below, the disadvantages of the procedure do not come to bear in this application and the advantages of the simple calculation outweigh the drawbacks. Figure 3.12 is used to explain the procedure in detail.

The separation of the mean indicated pressure and thus also of the indicated efficiency (see Formula 3.3) in line with the 360 ° *integration* corresponds to the distribution in a compression and expansion
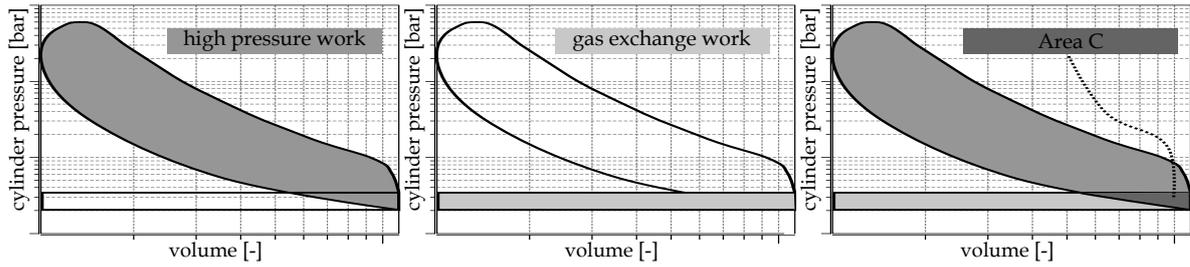
Figure 3.12: From left to right: high-pressure work, gas exchange work and total work including Area C within the 360 ° *integration* (in the manner of [103])

cycle as well as to an exhaust and an intake cycle in the classical sense. The calculation of the two parts - the high-pressure part (*HP*) and the gas exchange part (*GE*) is carried out by definition starting from one bottom dead center and moving via the compression and expansion stroke to the other bottom dead center via the exhaust and intake stroke. The denominator from Equation 3.3 is thus divided as follows [103]:

$$\oint pdV = \int_{HP} pdV + \int_{GE} pdV \tag{3.6}$$

Figure 3.12 illustrates how this division affects the allocation of the engine process on the basis of the pV diagram. The first part of the aforementioned equation represents the *high-pressure work* on the left-hand side of the figure. The second part depicts the *gas exchange work* shown in the middle of the figure. *Area C* on the right-hand side of the figure is part of both the gas exchange and the high pressure work, but once as a negative part with a minus sign and once as a positive part with a plus sign. It is therefore omitted. To derive the correction term, the indicated efficiency can similarly be apportioned:

$$\eta_{ind} = \eta_{ind,HP} + \eta_{ind,GE} = \frac{\int_{HP} pdV + \int_{GE} pdV}{m_{fuel} \cdot LHV} = \frac{\int_{HP} pdV}{m_{fuel} \cdot LHV} + \frac{\int_{GE} pdV}{m_{fuel} \cdot LHV} \tag{3.7}$$

As already mentioned, the use of the roots loader has a substantial effect on the gas exchange. With the separation made above, it is possible to take the gas exchange efficiency into account specifically. At operating points using the roots loader, the numerator of the last term in Equation 3.7 is corrected. For this purpose, the stationary pressure delta present between the roots loader inlet and the inlet channels is subtracted from the absolute pressure level during the intake stroke as a constant value Δp_{RL}. The recorded pressure signal from the indicating pressure transducers in the inlet channels is thus shifted downwards, as can be seen by way of example in Figure 3.13.

The measurement points p_{IM} and $p_{RL,inlet}$ used to calculate the pressure delta Δp_{RL} applied for shifting are shown in Figure 3.14. Both of them are absolute pressure transducers.

The gas exchange loop is calculated with the help of the 360 ° *integration* once without and once with the pressure offset. Both values are then set in relation to each other to obtain the correction factor $f_{\eta_{ind},RL}$:
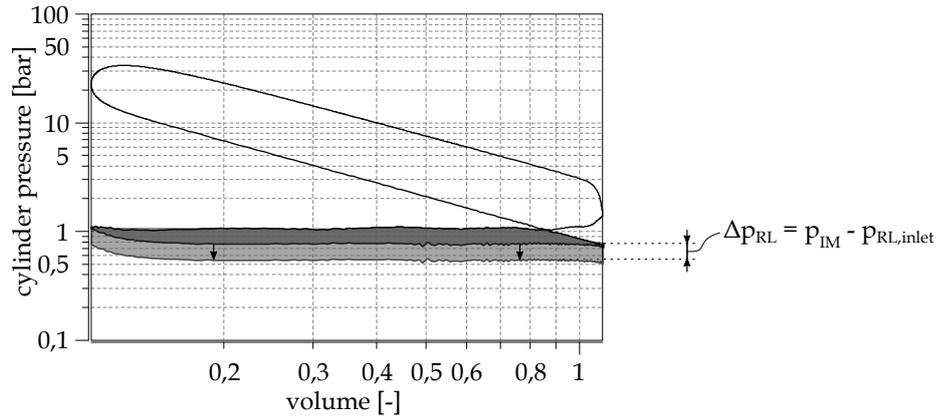
Figure 3.13: Pressure shift in gas exchange loop to correct efficiency when the roots loader is utilized.
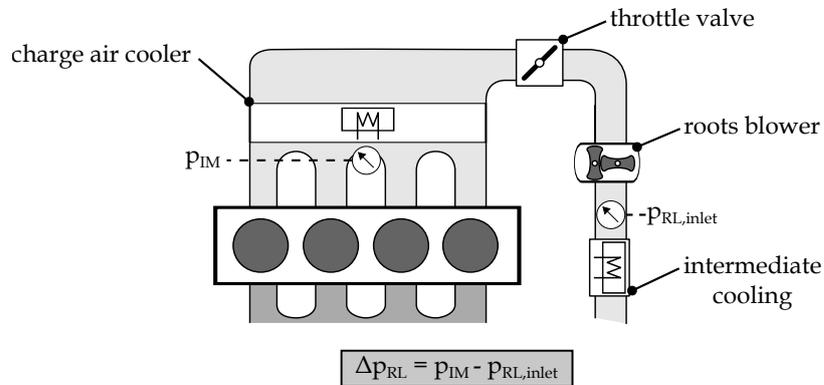
Figure 3.14: Measurement points used to calculate Δp_{RL}.

$$f_{\eta_{ind},RL} = \frac{\int\limits_{GE,360°} p_{shifted}dV}{\int\limits_{GE,360°} pdV} \tag{3.8}$$

The correction factor is subsequently multiplied with the gas exchange efficiency, which is in turn calculated within the software using the *Witt/Shelby method*. This results in a new gas exchange efficiency, which is summed up with the high-pressure efficiency to produce the total indicated efficiency η_{ind} based on Formula 3.7:

$$\eta_{ind} = \eta_{ind,HP} + \eta_{ind,GE,corrected} = \frac{\int\limits_{HP,Witt/Shelby} pdV}{m_{fuel} \cdot LHV} + \frac{\int\limits_{GE,Witt/Shelby} pdV \cdot f_{\eta_{ind},RL}}{m_{fuel} \cdot LHV} \tag{3.9}$$

Using the relative comparison, the weaknesses of the $360°$ integration described in [112], [85] and [103] become irrelevant, since an almost identical error is committed for the numerator and the denominator in Equation 3.8. Nevertheless, the simplicity of this procedure can be exploited in terms of the correction

term. The actual breakdown into the high-pressure loop and gas exchange loop, on the other hand, continues to be carried out using the most accurate method.

The new correction term applied to the series of measurements shown previously is illustrated in Figure 3.15. It clearly demonstrates that when the correction term is used and the flow losses between the roots loader outlet and inlet channels are taken into consideration, the correction is not misinterpreted or overestimated; instead, this produces an approximately identical course for the unadjusted and adjusted values. Due to the boundary conditions provided in Figures 3.8 as well as 3.9, this is the result to be expected.
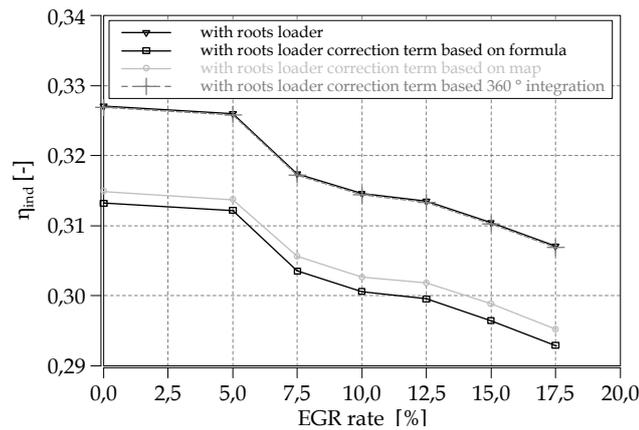
Figure 3.15: Results of the correction term based on the $360°$ *integration* of the indicated efficiency ($@\ 1000\ rpm$, $12\ bar\ imep$) in comparison to the other correction terms

The above findings justify and explain why an exclusive but relative correction of the gas exchange loop or its efficiency is made. The increased influence of the roots loader and hence the increased pressure delta Δp_{RL} are both considered in the resulting indicated efficiency.

As will become clear later in this paper, the correction of the efficiency is only relevant for the model input data, but not for the model output, since it provides information based on relative values. Ultimately, a correctly functioning model thus also indicates that the correction term is working properly.

Recalculation of the EGR rate based on the fuel-air ratio

Since Chapter 5 also includes measurement points and series with a fuel?air ratio deviating from the stoichiometric ratio, it is necessary to clarify the effects such measurement points have on the determination of the EGR rate as presented in Section 3.2. This section provides the correction made to the recorded EGR rate in this context and reveals the derivation of the applied adjustment.

For a stoichiometric fuel-air ratio $\lambda = 1$, the EGR rate can generally be calculated using the mass balance presented in Equation 2.1. The term derived from this in Equation 3.1 likewise applies likewise only to a stoichiometric fuel-air ratio. The issue with $\lambda \neq 1$ is that the recirculated exhaust gas also contains oxygen or fuel. Equation 2.1 therefore needs to be reformulated implementing a correction mass m_{corr}:

$$x_{EGR} = \frac{m_{eg} - m_{corr}}{m_{air} + m_f + m_{eg}} \cdot 100\% \tag{3.10}$$

The correction term can be calculated as follows:

$$m_{corr} = m_{O_2,eg} + m_{f,eg} \tag{3.11}$$

Note that $m_{O_2,eg}$ is the oxygen mass contained in the recirculated exhaust gas and $m_{f,eg}$ is the fuel mass contained in the recirculated exhaust gas. Both masses are negligible during normal operation at $\lambda = 1$. A case distinction can be made. For $\lambda \geq 1$ it applies that $m_{f,eg} = 0$, since it is assumed that, due to the excess oxygen, no fuel residues, or only negligible amounts thereof, are found in the recirculated exhaust gas. The oxygen mass $m_{O_2,eg}$ is then calculated as follows using the O_2 fraction $x_{O_2,eg}$ from the exhaust gas analyzer:

$$m_{O_2,eg} = m_{eg} \cdot x_{O_2,eg} \tag{3.12}$$

where m_{eg} applies:

$$m_{eg} = \frac{(m_{air} + m_f) \cdot x_{EGR}}{1 - x_{EGR}} \tag{3.13}$$

In the case of $\lambda \geq 1$ Equation 3.10 can consequently be reduced to the one unknown x_{EGR}:

$$x_{EGR} = \frac{m_{eg} \cdot (1 - x_{O_2,eg})}{m_{air} + m_f + m_{eg}} \cdot 100\% \tag{3.14}$$

For $\lambda \leq 1$, it is assumed that $m_{O_2,eg} = 0$. This presumption can be met because of the surplus fuel mass. From Equation 3.11 only $m_{f,eg}$ has to be calculated. For this purpose, two empirical approaches for gasoline engines can be found in the literature. One is provided by Bargende et al. in [6], the other by Pischinger et al. in [77]. The approaches both provide an approximation for the conversion efficiency η_c depending on various exhaust gas components. The conversion efficiency η_c is generally definded as the quotient of the burnt fuel mass and the total supplied fuel mass m_f. It can be calculated as follows:

$$\eta_c = \frac{1 - m_{f,eg}}{m_f} \tag{3.15}$$

Both approaches from the literature are implemented on a test basis, but provide analogous results, hence only the approach finally applied is presented hereafter. According to Bargende [6], the conversion efficiency η_c can be calculated using the formula below:

$$\eta_c = 1 - \frac{x_{HC}}{10^6} \cdot (1 + \lambda_{gl} \cdot L_{st}) \cdot \left(1 - 4.5 \cdot \frac{x_{CO_{com}}}{10^6} + 7 \cdot \left(\frac{x_{CO_{com}}}{10^6}\right)^2\right) \tag{3.16}$$

The stoichiometric air consumption L_{st} and the global fuel-air ratio λ_{gl} in the exhaust gas are known, the latter being measured using the exhaust gas analyzer. x_{HC} describes the proportion of unburned hydrocarbons in the exhaust gas in *ppm*, and is also detected by the exhaust gas analyzer. The volume fraction of carbon monoxide present at combustion $x_{CO_{com}}$, on the other hand, is not measured but must be determined using the following equation on the basis of the measured global volume fraction of carbon monoxide in the exhaust gas $x_{CO_{gl}}$. All values must be specified in *ppm*:

$$x_{CO_{com}} = x_{CO_{gl}} \cdot \frac{1 + \lambda_{gl} \cdot L_{st}}{1 + \lambda_{com} \cdot L_{st}} \tag{3.17}$$

For this equation, it is necessary to determine the actual air ratio present at combustion λ_{com}. Bargende proposes the following term:

$$\lambda_{com} = 1 - 3.28 \cdot x_{CO_{gl}} \cdot 10^{-6} + 5.092 \cdot (x_{CO_{gl}} \cdot 10^{-6})^2 \tag{3.18}$$

Thus, all variables from Equation 3.16 can be determined or are known and the conversion efficiency can be calculated. Example results for the correction terms of $\lambda \geq 1$ as well as of $\lambda \leq 1$ are shown in Figure 3.16. The measurement series of an EGR variation at $\lambda = 0.9$, another at $\lambda = 1.0$ and another at $\lambda = 1.1$ are displayed. The uncorrected EGR rate is plotted on the x-axis and the corrected EGR rate is plotted on the y-axis. All measurement series are recorded with identical boundary conditions (@ 2000 *rpm*, 6 *bar imep*).
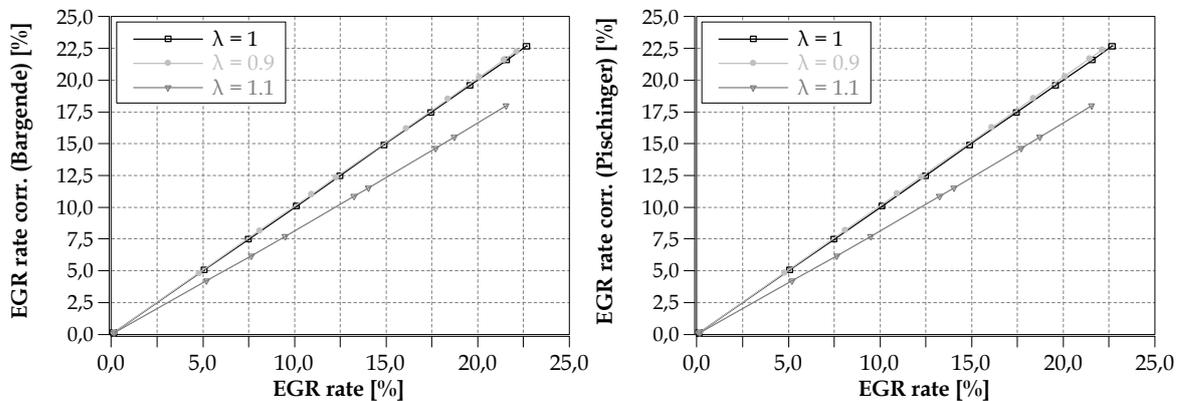
Figure 3.16: Comparison of EGR rate uncorrected and corrected for an fuel-air ratio deviating from the stoichiometric ratio as described by Bargende et al. (left-hand side) and Pischinger et al. (right-hand side)

There is no correction for a stoichiometric fuel-air ratio. The respective measurement series therefore forms an angle bisector in the diagram. The measurement series clearly deviates from this at $\lambda = 1.1$. Due to the additional air mass in the recirculated exhaust gas, the corrected EGR rate is thus lower than the measured or uncorrected EGR rate. The measurement sequence at $\lambda = 0.9$ deviates by only a very small amount from the measurement series at $\lambda = 1.0$. This is due to the small absolute mass of unburned fuel present in the recirculated exhaust gas. Nonetheless, a correction is carried out to eliminate even this small error.

Finally it becomes also clear, that the two correction terms for the operation at $\lambda \leq 1$ and their results differ only marginally from each other and are hence both applicable. For this reason the simpler term used by Bargende is implemented in this work.

4 Investigation into a hardware-based approach

Parts of this chapter have been previously published in [59].

Section 2.2.1 discusses the reasons for investigating of a hardware-based approach, outlining the basic advantages of this solution are outlined. Essentially, a hardware approach theoretically enables the entire stationary efficiency potential of LP EGR to be employed during transient engine operation. Since this thesis provides a holistic perspective of LP EGR during dynamic engine operation, the development and evaluation of the most promising approaches are indispensable. The alternative approach using an unmodified engine hardware is introduced in the upcoming Chapter 5.

At the outset of this chapter, the systems examined are presented one after the other. It is a mixture of approaches known from the literature and self-developed ones. Subsequently, an overall comparison of the six systems is provided, then the particulars of the systems are discussed in individual subsections. The basis for comparison is always the unchanged standard or basic system. An evaluation for each system is conducted in accordance with the worst case load step derived in Section 2.1.

4.1 Analyzed systems

This chapter provides an overview of the systems investigated for handling dynamic engine operation employing an LP EGR configuration. The first step taken is to evaluate all the systems in a simulative setting. The implementation and measurement of one single system on the engine test bench is carried out afterwards using the findings and knowledge gained from the simulation. An overview of all systems is depicted in Figure 4.1 below.

The Basic System, which also forms the starting point of the simulation model, is given in Figure 4.1(a). It concerns the BMW 4 cylinder presented in Chapter 3.1 supplemented by an LP-EGR path with the sampling point before the TWC. All the other systems are implemented on the basis of this engine or system, respectively. To clearly differentiate between and accentuate the variations, deviations from the Basic System in Figures 4.1(b) to 4.1(f) are highlighted in yellow.

Bypass System

The Bypass System that is implemented in the simulation model is derived from an idea portrayed in various literature sources, such as [96, 97], [26] or [86]. It is depicted in Figure 4.1(b). The key feature of this solution is the second air path, which branches off from the actual air path directly after the intake silencer. It is important that the junction point is located upstream of the EGR path discharge point, i.e. it is fed exclusively with fresh air. The second air path is reinserted into the main air path downstream of the main throttle valve. For the implementation within this study, it is reinserted as close as possible to the combustion chamber and directly into the individual intake ports. The second air path is equipped with a throttle valve as a control unit and a non-return valve. The latter serves to prevent the air path from being flushed backwards during turbocharged engine operation.

The second air path is used in the case of a negative load step during operation with EGR. Being filled with a mixture of exhaust gas and fresh air, the main air path is throttled sharply when the load is reduced, while the second air path or bypass path is opened or dethrottled via its throttle valve to hit

(a) Basic System   **(b)** Bypass System   **(c)** High-Pressure Bypass System

(d) Scavenging System   **(e)** BMW Direct Boost   **(f)** Generator System
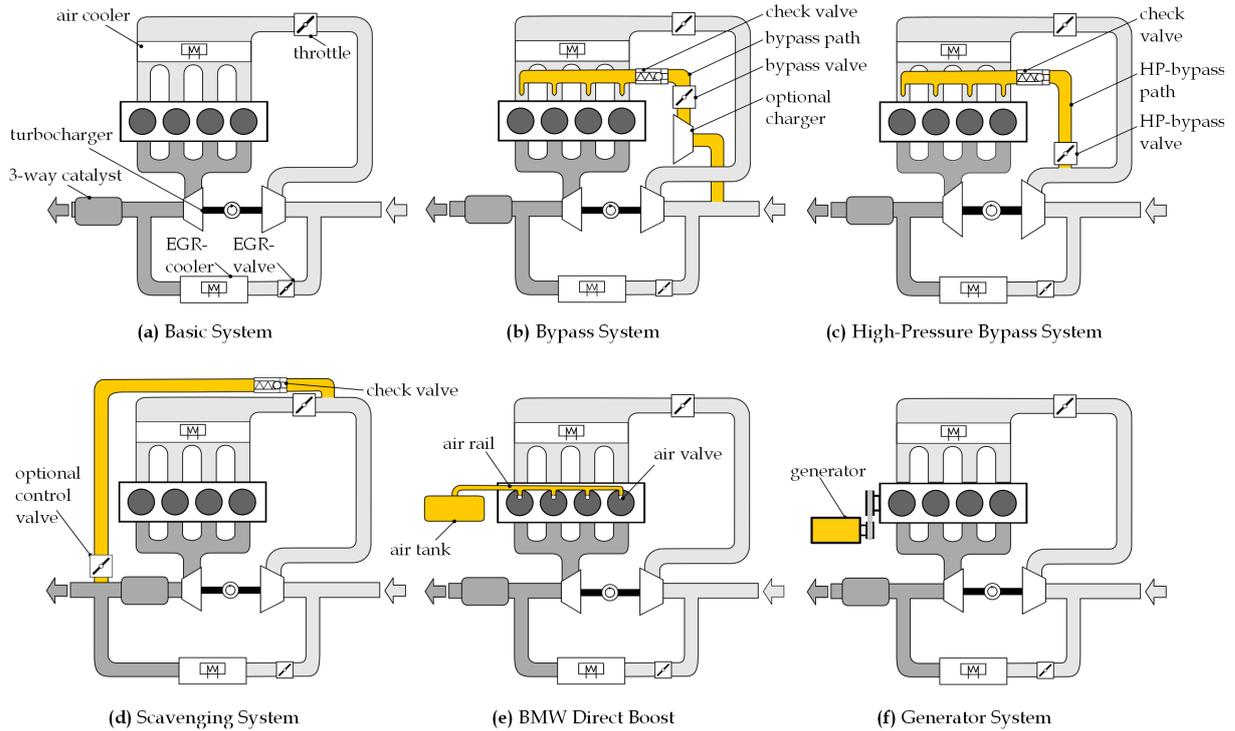
Figure 4.1: Analyzed systems.

the target load point and deliver the air that is required. The second air path contains only fresh air at that moment. The switchover closes and bypasses the actual air path filled with the exhaust gas and air, while the second path quickly supplies the engine with pure air. However, this can only happen if the load is decreased to an natrually aspirated operating point. Due to the negative pressure in the area of the intake manifold, there is a positive pressure gradient across the second air path. As the figure demonstrates, an optional compressor is included in the bypass path. It could be applied whenever demand arises and serves on the one hand to further accelerate the supply of fresh air via the bypass path and on the other hand to extend the operating range into turbocharged areas.

The bypass solution in its configuration without the optional compressor is implemented within the boundary conditions of the engine presented in Table 3.1 including the given package dimensions. Particular attention was paid to the following properties:

- Sufficiently sized cross-section of the second air path (implementation of the maximum load in naturally aspirated engine operation area).
- Small volume between the throttle valve of the main air path and the combustion chambers.
- Discharge point of the second air path as close as possible to the combustion chambers.
- Small volume between the bypass path discharge points and the bypass path throttle valve.

Special emphasis should be placed on these properties not only in the case of the engine presented here, but also when generally designing a system with a bypass path or second air path. Furthermore, the location of the main path's throttle valve is of high importance. It should be placed as close as possible to the combustion chambers, which means keeping the volume downstream of the throttle valve as

small as possible.

High-Pressure Bypass System

This system is a self-developed solution presented in [38]. Its functional principle is identical to that of the Bypass System. It likewise has a second air or bypass path, which is used in the case of a negative load step from EGR operation. This means that the main air path is filled with a mixture of fresh air and exhaust gas. When the load is reduced, the main air path is closed by the throttle valve, while the second air path is opened by actuating its throttle valve. The second air path only contains air at this time. Consequently, the same design rules and specifications apply as mentioned in the last sections for the Bypass System.

The main difference between the two systems is to be found at the point of withdrawal of the bypass duct. In the case of the High-Pressure Bypass System, it is located directly downstream of the compressor outlet and consequently also downstream of the EGR discharge point. There are thus two principle stretegies conceivable in the case of a load step. Either the bypass section is opened immediately and the small proportion of exhaust gas/air mixture located between the EGR discharge point and the bypass branch is passed through the bypass path, or the bypass line remains closed until this small amount has passed the branch point. This means that only fresh air circulates in the bypass path.

When the two bypass systems are compared, the later branching point of the High-Pressure Bypass System is at a disadvantage at first glance. It should be noted, however, that with the High-Pressure Bypass System, the overpressure downstream of the compressor can still be exploited. This can be a significant advantage in terms of scavenging speed. In addition, the target load point does not necessarily have to be a naturally aspirated operating point. Another obvious advantage is the smaller bypass size.

Scavenging System

The Scavenging System as shown in Figure 4.1 (d) is implemented according to the information given in [17]. This solution is based on the idea of flushing the inlet path as quickly as possible during and right after a negative load step from the EGR operating area. The implementation works with the help of a scavenging path which branches off directly in front of the throttle valve of the inlet-air path and is fed to the exhaust system downstream of the TWC. The crucial distinguishing element of this solution can immediately be found here. To ensure the proper functioning of the TWC with a stoichiometric fuel-air ratio, no unregulated air must be supplied upstream of the TWC. This excludes flushing of the inlet path and any supply via the scavenging path upstream of the TWC. However, for the scavenging path to be fed downstream of the TWC, the exhaust gas that is flushed out of the inlet path must already be aftertreated. This can only be guaranteed by an EGR extraction point downstream of the TWC. Despite the drawbacks discussed in Section 3.1, the Scavenging System is employed accordingly. Efficiency disadvantages during stationary operation could be accepted if the implementation of transient operation with EGR has appropriate advantages.

The scavenging line itself has a throttle valve for opening and controlling the flow as well as a check valve to prevent uncontrolled flow from the exhaust tract towards the inlet path. In the event of load

discharge, the throttle valve of the main inlet path is throttled to such an extent that the target operating point's load can be adjusted. In this way, overpressure is created upstream of the throttle valve, i.e. at the branch point of the scavenging path; this can be released towards the exhaust pipe. The flushing of the intake system filled with the exhaust gas/air mixture is thus accelerated without impacting the engine's target load. The function can be further supported by maintaining the compressor's performance at a high level by setting the wastegate. Since the volume downstream of the throttle remains unaffected by the scavenging process, the position of the throttle in the inlet path close to the combustion chambers is also beneficial for this system. The volume between the throttle valve and the intake valves should therefore be as small as possible.

BMW Direct Boost

The *BMW Direct Boost* is introduced in [19] as well as in [49]. Its actual function is to improve the response of turbocharged gasoline engines by injecting compressed air directly into the combustion chamber using an additional intake valve. Within the scope of this work the system is used for the first time to reduce the in-cylinder EGR rate as quickly as possible to the target value during load discharge. For this purpose, the combustion air is at least partly provided by the Direct Boost System instead of by the inlet-air path. As depicted in Figure 4.1 (e), the Direct Boost System essentially consists of a compressed air tank, an air rail and an additional inlet valve connected to the air rail and the tank. The tank contains air at a pressure level between 8 and 14 *bar abs.*, which is stored there over run with the fuel cut off by opening the air valve at the same time as with the compression cycle of each particular cylinder. It is then available virtually for free in all other situations, e.g. negative load steps from the EGR operating range.

In the case of such load steps, the inlet valves are at least partially closed and instead pure air is supplied from the tank to the combustion by opening the additional inlet valve. A key issue is to ensure that the load is set correctly, i.e. when a charged operating point is present, the overpressure is relieved from the intake manifold and compressed air is injected through the Direct Boost System at the same time.

Finally, it should be noted that the system's use as an enabler for transient operation with LP-EGR does not limit its actual functionality as a measure to improve the dynamic engine response. This functionality remains unchanged. The tank capacity, on the other hand, must be assessed sensitively, especially when the system is used as a measure both for load jumps and for load shedding. It is also important to evaluate the extent to which external EGR can be prevented from entering the air tank.

Generator System

The solution using a generator, as shown in Figure 4.1 (f), differs fundamentally from the approaches introduced so far. In contrast to them, the air path is not adjusted or manipulated, but the load or the torque of the engine is directly influenced by interference at the crankshaft. The basic idea of this solution is presented by Messing et al. in [72]. Its concrete application, as shown in Figure 4.1 (f), is elaborated within the framework of this study. A belt-driven starter generator at a voltage of 48 V is coupled to the crankshaft. This 48 V is the voltage level that will be found in the on-board electrical

system of much of the automotive fleet in the future. Compared to the previous voltage level of 12 V, it enables a significantly higher power supply.

When the load is discharged, the generator is switched on instantaneously and demands power. This is converted into electrical energy, stored in a battery and can be reused any time later. The coupling of the generator increases the engine load without increasing the effective torque applied to the output. Though, the enhanced load has two major implications for EGR operation. Firstly, the higher load increases the mass flow through the engine. The exhaust gas-air mixture from the inlet is thus scavenged more quickly. On the other hand, the EGR tolerance is improved by increasing the load point. However, the fact should not be neglected that the absolute consumption of the engine temporarily rises due to the load point shift. Furthermore, the solution can only be implemented as long as the battery capacity is available. A fully charged battery should therefore be avoided by a prudent operating strategy.

As mentioned above, all systems are implemented under the conditions of the existing engine from Table 3.1. As far as possible, the central design criteria mentioned above are also taken into account. When considering the following results, it must nevertheless be borne in mind that the systems could theoretically offer even more potential with a specifically designed engine. If this is the case, however, it is mentioned at the appropriate passages.

4.2 Simulative study

In this chapter, the systems presented above are compared using a detailed engine simulation with the 1D CFD software *GT-Power*. The comparison is likewise based on the worst-case load step described in Section 2.1. The first part of this section outlines the excellent match between the simulation model and the engine test bench measurements. Another system is based on the well tuned base model. Subsequently, the results of the overall comparison are displayed and discussed. The specifics, i.e. the advantages and disadvantages as well as problems and potentials of each system are stated in the following sub-sections.

Figure 4.2 shows the comparison of the behavior of the most important actuators in the motor, obtained from the measurement and the simulation of the worst case load step, which is initiated at 8.9 seconds in the diagrams. The actuators are relevant for the adjustment of the load on the one hand and for the adjustment of the EGR rate (internal and external) on the other hand. A very good match between the measurement and simulation results can generally be found. Nevertheless, a number of comments are required. The throttle valve and the valve lift are seen to differ slightly at the starting load point. This is due to the inaccuracies in the discharge coefficients of the inlet valves and throttle valve described in Section 3.3. Both actuators are calibrated within the simulation model such that the load of the engine, i.e. the cylinder filling, is reproduced properly and the negative pressure in the intake manifold corresponds to that resulting from the measurement ($p_{IM} = -100\,mbar$). In addition, the constant value of the wastegate during the load step is conspicuous. This complies with the application of the engine and is kept constant over a large load range in which the load control is performed using the throttle valve and the valve lift. Its purpose is to increase the turbocharger speed and to improve the

dynamic response of the engine. All actuators that are relevant for the quantity of internal residual gas (intake and exhaust cam timing) and external residual gas (EGR valve) match very well.
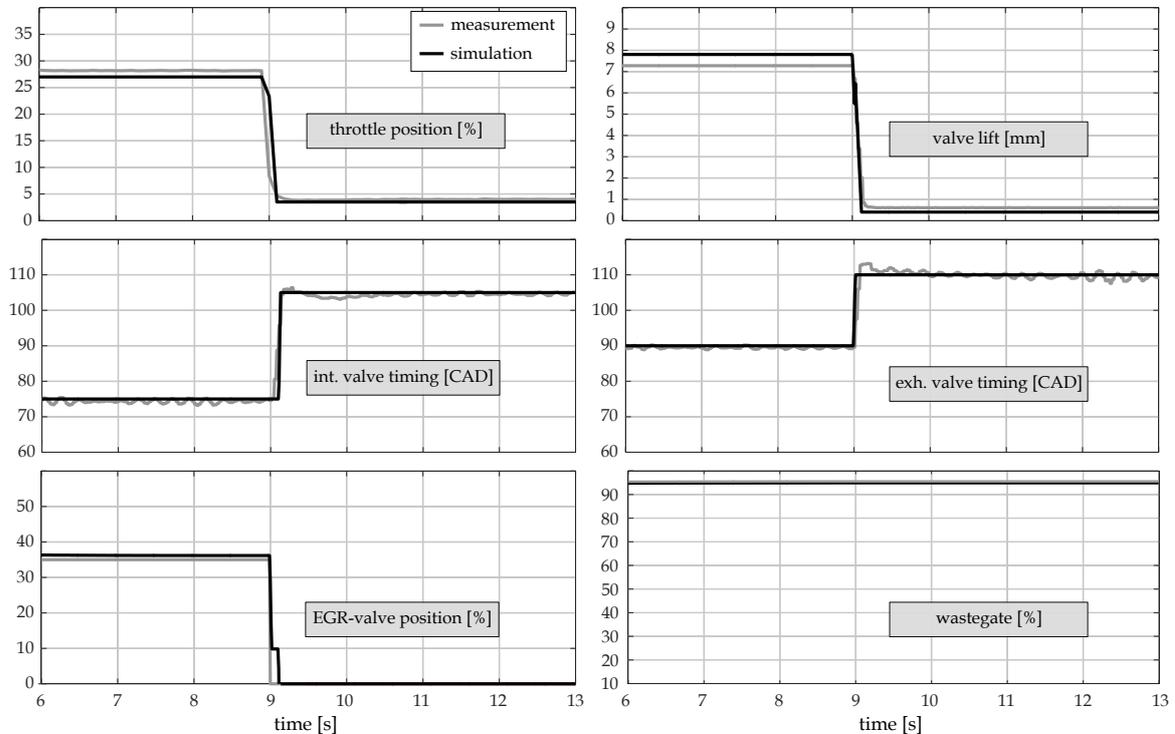
Figure 4.2: Engine actuators signals from simulation and engine test bench measurement.

Figure 4.3 below presents the values which are obtained by means of the actuator setting and which are crucial for the matching, namely the torque characteristic and the external EGR rate. The torque congruence is very satisfactory. The most important aspect in this context is the gradient, which is assumed to be $-750\,Nm/s$, as defined in the worst-case load step identification Section 2.1. The torque gradient is applied for the measurement on the engine test bench; the same gradient can be found in the results of the simulation. It is very important for the individual simulations of the systems that this gradient is fulfilled, otherwise advantages or disadvantages would arise for the respective system. The faster the gradient, the more disadvantageous the behavior regarding the dynamic EGR operation. This demand is not easily met, but is still satisfied by all setups.

Similarly to the torque, the curves of the EGR rate also match very well. They are determined in both the simulation and the measurement at the measuring point of the O_2 intake sensor in the intake manifold. Thus the sensor's signal is used for the comparison. This is mandatory because the EGR rate cannot be measured any closer to the combustion chamber. However, a change in the qualitative course is not to be expected, but a further delay in degradation is very likely.

The concordance between the measurement and the simulation is particularly important because only the Basic System has been implemented in hardware so far. A decided matching of every system is not possible a priori. However, all the other systems are configured on the basis of the standard system.
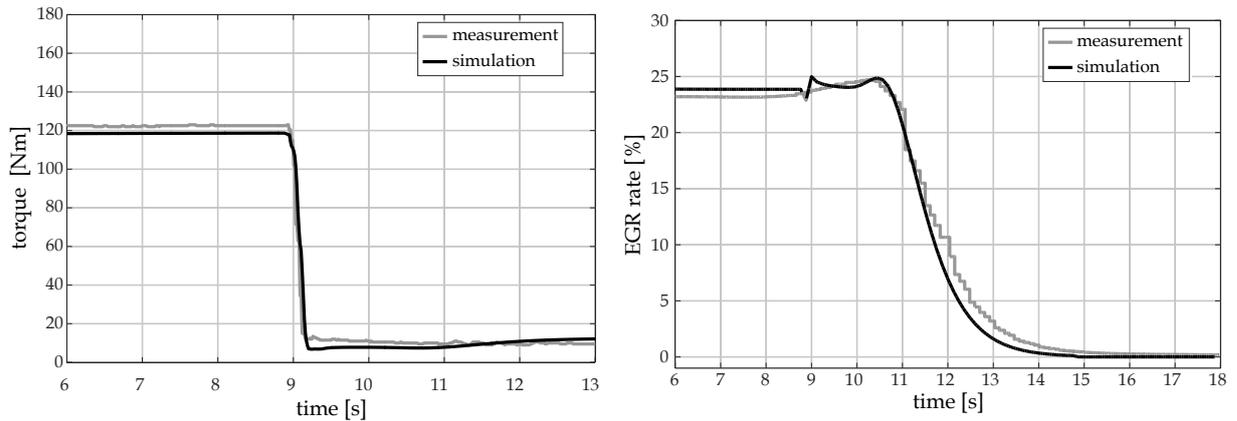
Figure 4.3: EGR rate and engine torque from simulation and engine test bench measurement.

Now that the agreement between the simulation and measurement has been proven, the results of the individual simulation setups can be depicted. For this purpose, the course of the external EGR rate in the combustion chamber of each system over time is illustrated in Figure 4.4.
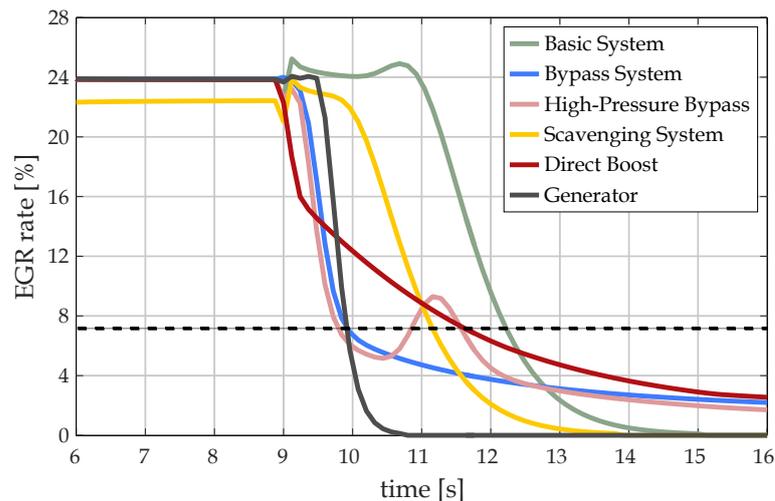
Figure 4.4: EGR rate over time for each system configuration.

Two criteria are used for an initial evaluation of the courses. One is the starting time for the reduction of the EGR rate and the other is the gradient with which the EGR rate is reduced. In all setups that modify the air path, this is essentially a function of the point where the fresh air is introduced. Thus, the two bypass solutions, each of which feeds the fresh air into the individual intake pipes, have a synchronously running gradient. The steepness of the gradient can be explained by the fact that the EGR air mixture from the main path and the fresh air from the bypass path do not mix much due to their introduction close to the combustion chamber. After the short gas column, which is still present between the bypass feeding point and the inlet valves, only fresh air is introduced via the bypass path. The reason for the flattening of the curves after the steep gradient is, that the throttle valve of the main path is not completely closed. This is necessary to prevent the compressor from pumping and also to refill the main path with fresh air. The difference between the two bypass configurations with regard to

the starting point of the EGR rate reduction is the usable overpressure in the case of the High-Pressure Bypass System.

Similarly to the two bypass setups, the curves of the Basic System and the Scavenging System run in parallel. In both systems, the fresh air is supplied directly to the intake-silencer just like in standard engine configurations. The mixture of fresh air and the air/exhaust gas mixture is therefore analogous and runs along the same gas path, which is also, above all, equally sized. As expected, the starting point for reducing the EGR rate of the Scavenging System is much earlier than with the Basic System. This is due to the significantly increased mass flow by scavenging the intake gas path towards the exhaust system. A short note still has to be made regarding the EGR rate at the start load point with the Scavenging System. The desired value can no longer be achieved due to the lower pressure drop across the EGR path using an EGR withdrawal downstream of the TWC; it is therefore 1.3 percentage points lower than in the other systems. However, this does not have a decisive impact on the final evaluation in this section.

In line with expectations, the Generator System should behave similarly, since it has the same feeding point for fresh air, directly at the intake silencer. However, Figure 4.4 shows that the EGR rate reduction has a much steeper curve and an earlier starting point. This is due to the significantly increased total mass flow through the engine caused by the load point shifting.

Finally, the Direct Boost System must be evaluated. The direct air injection into the cylinder causes both an early onset of the EGR rate reduction and a steep gradient in EGR degradation. However, the gradient levels off after a short time, as the required torque gradient cannot be fulfilled otherwise. The air supply must thus be a combination of the main air path and the Direct Boost System.

Although the above evaluations and statements provide a broad and qualitative overview and help to gain a better insight into the specific mechanisms, they do not serve as a conclusive basis for comparison or one that is, above all, quantitative or evaluable. This begs the question of how quickly a particular EGR rate level is reached. In principle, the assessment represents a quantifiable combination of the degradation gradient and the starting point of the reduction in the EGR rate. Although the primary objective is to achieve optimally efficient operation, the criterion for the representability of the load shedding is essentially misfire-free operation combined with smooth running. This means that the EGR tolerance must be complied with at the target load point.

The EGR or dilution tolerance is reached when the combustion variance of the mean indicated pressure COV_{imep} equals its maximum admissible value. In order to guarantee adequate engine smoothness, a maximum tolerated COV_{imep} of 3 % is specified in this study. The definition of the COV_{imep} is provided below, derived from a general definition according to [50]:

$$COV_{imep} = \frac{\sigma_{imep}}{\overline{imep}} \cdot 100\% = \frac{\sqrt{\frac{1}{n}\sum_{i=1}^{n}(imep_i - \overline{imep})^2}}{\overline{imep}} \cdot 100\% \tag{4.1}$$

Here, n describes the number of all observed cycles i. $\overline{imep}$ is the arithmetic mean of the mean indicated pressure of all observed cycles and $imep_i$ is the mean indicated pressure of a single cycle. Thus, the standard deviation of the mean indicated pressure σ_{imep} represents the numerator of the term.

Below, Figure 4.5 presents the course of the indicated efficiency and the COV_{imep} for the target load point (@ 1000 rpm, 1 $bar\ imep$) of the worst case load step, both depending on the EGR rate. Assuming a

maximum permitted $COV_{imep,max}$ of 3 %, the EGR tolerance and consequently the EGR rate for assessing the different systems is 7 %. The relevant line of an EGR rate of 7 % is displayed in Figure 4.4.
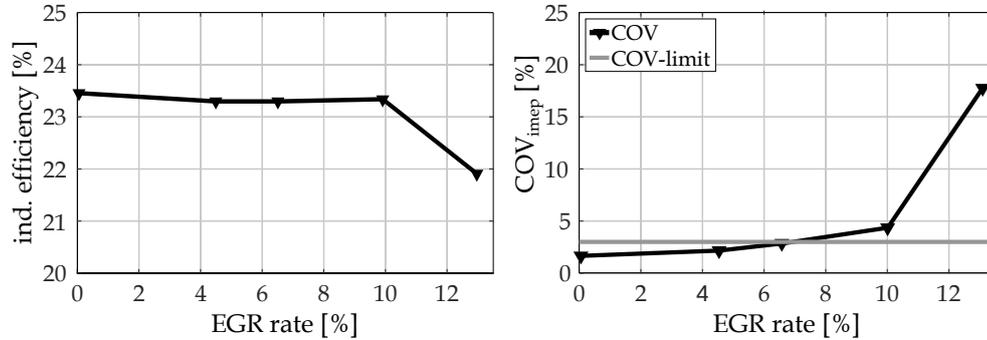
Figure 4.5: Indicated efficiency and COV from an EGR-variation within the target load point of the worst case load step.

A clear illustration of which setup reaches the relevant EGR rate fastest, is shown in Figure 4.6. The bar chart depicts the duration required after the start of the negative load step to undercut the line of 7 % EGR rate.
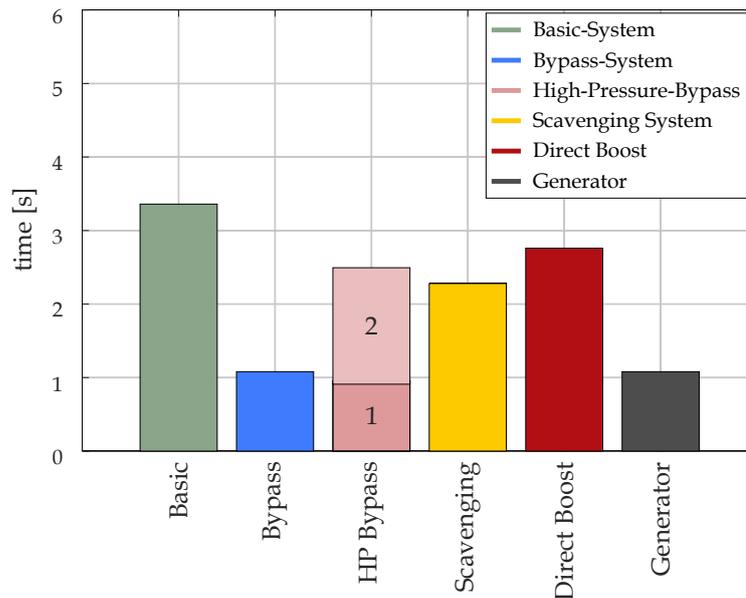
Figure 4.6: Time to decrease EGR-rate beneath EGR-tolerance for each system configuration.

It clearly emerges that all systems achieve improvements compared to the Basic System. With the Direct Boost System, however, the improvement is only marginal. This is due to the previously described mandatory requirement of fulfilling the torque gradient. While an advantage of about 1 s can be realized with the Scavenging System, an improvement of only 0.8 s can be achieved with the High-Pressure Bypass System. As the associated bar is two-piece, this system requires additional explanation. The first part of the bar represents the moment of the first undershooting of the EGR tolerance. Considering only this part, this system is the most advantageous in the available portfolio. However, since the EGR rate

increases again with this system, the EGR tolerance is ultimately only undershot after an additional delay, which is represented by the second part of the bar. This behaviour can be explained by the branch point of the High-Pressure Bypass path, which is located after the EGR discharge point. An optimum balance must be found in terms of the delayed opening of the bypass line and fast opening accompanied by the exhaust air mixture being flushed through the bypass line.

After all, both the Generator and Bypass System are the most beneficial in terms of quickly reducing the in-cylinder EGR rate. They are on precisely the same level and achieve an enhancement of more than 2 s compared to the base system, which means an improvement of more than 65 %.

The evaluation of how long the individual systems need to comply with the EGR tolerance has been completed in principle. This work, however, aims at optimally efficient operation or operation with minimum fuel consumption using LP EGR. The detailed work on this topic is covered by Chapter 5, yet at this point the efficiency behaviour of the individual systems is to be evaluated beyond their mere adherence to EGR tolerance. Figure 4.7 plots the course of the indicated efficiency during the load point change over time.
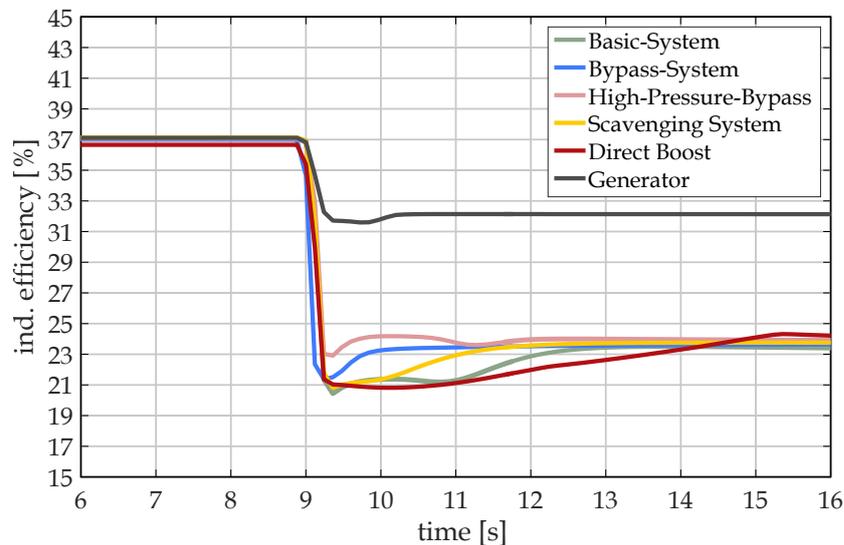
Figure 4.7: Indicated efficiency over time for each system configuration.

The curves displayed represent the efficiencies obtained by the simulation, which of course contain slight uncertainties. The calculation is based on the charge exchange as well as on a multi-dimensional map with burn rates determined with PTA calculatinos depending on the engine load, the engine speed and the EGR rate. Despite the imprecision, the figure gives a clear and qualitative overview of the specific system behaviour.

Each of these curves can be explained very vividly using the diagram on the left of Figure 4.5. Up to the EGR tolerance limit, the indicated efficiency decreases minimally but is very flat. Consequently, at the target load point, until the EGR tolerance is reached, it is not practicable to speak of a genuine disadvantage or advantage in terms of efficiency. This is an essential reason why short-term operation with an EGR rate deviating from the stationary EGR rate with optimal consumption is acceptable as long as it is beneath the EGR tolerance. However, above the maximum allowed EGR rate, there is a noticeable efficiency drawback, which is evident for at least 3 of the 6 setups in Figure 4.7.

The Basic System, the Bypass System as well as the Scavenging System show an efficiency shortfall as long as the EGR tolerance has not yet been reached, i.e. as long as it is on the right-hand side of the EGR tolerance on the efficiency curve of Figure 4.5 and is therefore subject to the noticeable efficiency disadvantage. The Generator System behaves similarly, but at a significantly higher efficiency level. Although the efficiency here is improved by the load point shift, this still leads to a higher fuel consumption overall. Further details will be given later.

Additional effects can be seen with the Direct Boost and the High-Pressure Bypass System. The former requires considerable throttling of the gas exchange due to the injection of additional compressed air during combustion. The losses increase, and do not decrease until the amount of compressed air injected is reduced. The undershoot after load point change is therefore particularly pronounced here and significantly longer than would be expected from the combination of the 'efficiency over EGR rate' and 'EGR rate over time' curves. The High-Pressure Bypass System behaves in the opposite way. Compared to the Bypass System, it has the significant benefit that the throttle valve of the main path can be closed almost completely and the entire mass flow can be made available via the Bypass Path. This is shorter and does not have to pass through a charge air cooler, which significantly reduces gas exchange losses. The efficiency after the load point change thus reaches a higher level much faster than would have been expected.

In the following subsections, further specifics of the individual systems are highlighted along with amendments made to them, and their key influencing variables are examined before a final overall assessment is made with regard to the implementation of one of the investigated systems in hardware on the engine test bench.

Basic System

The negative load step using the Basic System has been discussed at several points in this thesis. Nevertheless, some amendments remain to be incorporated at this point. Before the following text focuses on elementary design rules, a variation in the valve timing in the target load point is pointed out. The valve timing of the start load point is not changed. The valve timing variation is intended to demonstrate two aspects: on the one hand, whether it is fundamentally feasible to significantly reduce the residual gas rate after the load step and on the other hand, whether it is possible to accelerate EGR reduction by changing the valve timing. Deviating from the valve timing applied under stationary conditions, as previously investigated, four further combinations of intake and exhaust cam timings are analyzed.

Based on the findings from [30], three different strategies are elaborated for controlling and obtaining the internal residual gas. The first is so-called *exhaust channel recirculation*. Both the closing time of the exhaust valves as well as the opening time of the intake valves are subsequent to the TDC. However, there is a valve overlapping. The residual gas is obtained by returning the exhaust gas from the exhaust channels. With the *intake port recirculation*, the two valve events described are ahead of TDC. Again, there is a distinctive valve overlapping and the residual gas is drawn in again during the consecutive working cycle having first been discharged into the intake ports. Finally, it is also possible to close the exhaust valves a long way before TDC and to open the inlet valves long after TDC without any valve overlapping to directly trap the residual gas in the cylinder. The strategy is called *combustion chamber trapping*. All three strategies from [30] are presented in Figure 4.8. For the general definition including the relevant reference points of the valve timing, see Figure 5.8.
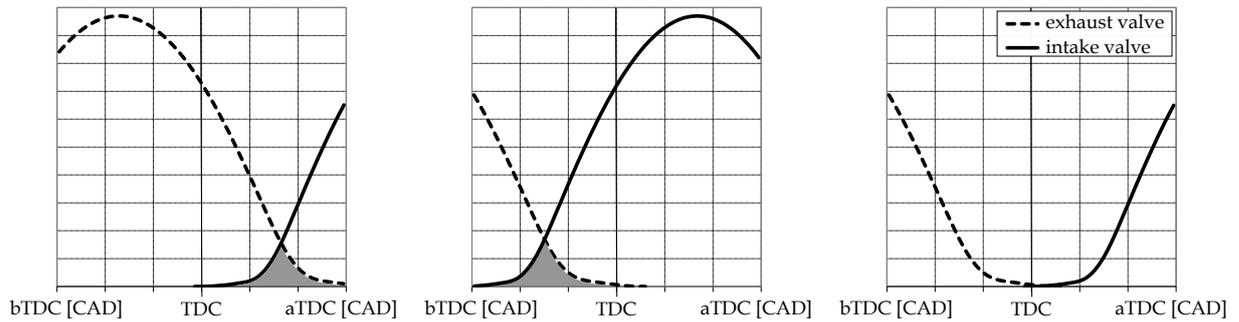
Figure 4.8: Valve timing strategies for internal EGR: exhaust channel recirculation (left), intake port recirculation (middle), combustion chamber trapping (right).

In addition to those described, the valve timings for a maximum valve overlap are investigated. Out of the strategies in the portfolio, the series application, which is illustrated for comparison, can be classified as combustion chamber recirculation. The influences that all analyzed valve timings have on the internal and external EGR rates during the load point change can be seen in Figure 4.9.
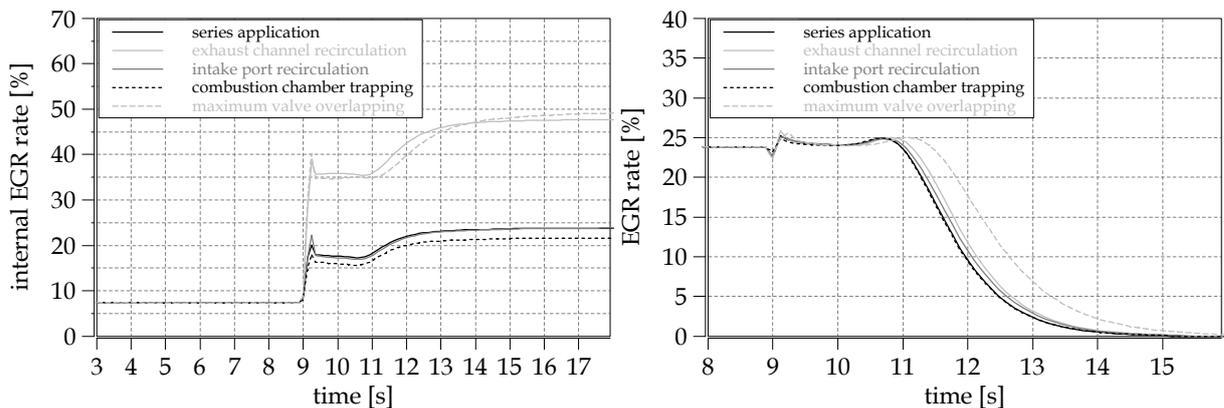
Figure 4.9: Internal and external EGR rate using different valve timing strategies for target load point of negative load step.

The basic course of the curve is essentially the same for all setups. The internal EGR rate increases significantly after the load step, and even further in a second step when the pressures in the inlet and outlet have stabilized. All curves differ only by a parallel offset. Both the *exhaust channel recirculation* and the configuration with the highest level of valve overlapping show significantly increased levels of internal EGR. Thus the exhaust channels can be identified as the main source for internal EGR at this operating point. This can be substantiated by the existing pressure conditions as well as the valve lift. Compared to the inlet valves, which have only a small stroke of 0.5 *mm* at the target load point due to the variable valve train, the exhaust valves are fully open (9.9 *mm*). Accordingly, the *intake port recirculation* features significantly lower internal EGR rates and is comparable to the basic application in this regard. The internal EGR rate using the strategy of *combustion chamber trapping* is slightly lower and therefore advantageous with regard to the total dilution with exhaust gas. Due to the low exhaust back-pressure in the partial-load area, a large portion of the residual gas is already pushed out despite the early exhaust valve closure and is therefore not kept in the combustion chamber. In order to achieve a low dilution with internal EGR and thus a low total dilution, the *combustion chamber trapping*

strategy is consequently the most advantageous, although the existing series application also features comparatively low internal residual gas rates. However, it must be noted that this results likewise in lower indicated efficiencies.

An overall assessment, of course, can only be made after considering the influence of valve timing on the external EGR rate. Due to the fast adjustment of the camshafts of $0.2\,s$, the valve timing is of relevance during virtually the entire operating point change. In this case, the clear tendency can be observed that a higher valve overlap causes a slower reduction in the EGR rate. The reason for this is straightforward: instead of scavenging in the direction of the exhaust tract, the naturally aspirated engine operation after load point change leads to reverse scavenging towards the inlet and consequently to a reduction in the mass flow and in the EGR rate degradation. Scavenging with a high valve overlap, as suggested, for example, in the work of Siokos [86], is therefore not feasible in this case.

For the entire evaluation, it should be noted that no quantitative conclusions can be drawn about the indicated efficiency and the practicability of these configurations. It is therefore quite possible that individual setups significantly exceed the dilution tolerance and therefore exhibit high COV_{imep} values or misfires. Rather, this is a fundamental examination which highlights the influence of different valve timings in comparison to the basic application. In the conflict between the best possible indicated efficiency, maintaining the dilution tolerance and at the same time achieving a rapid reduction in the EGR rate, it becomes clear though that the currently applied valve timing represents a very good compromise.

Two basic investigations will finally illustrate the essential issues when designing an engine with LP EGR as well as the limitations of the simulation model. In this context, Figures 4.10 and 4.11 show sensitivity analyses on the enlargement of the intake path in regard on pipe length and of the intake manifold volume, analyzing its simple expansion. The insights gained apply independently of the system.
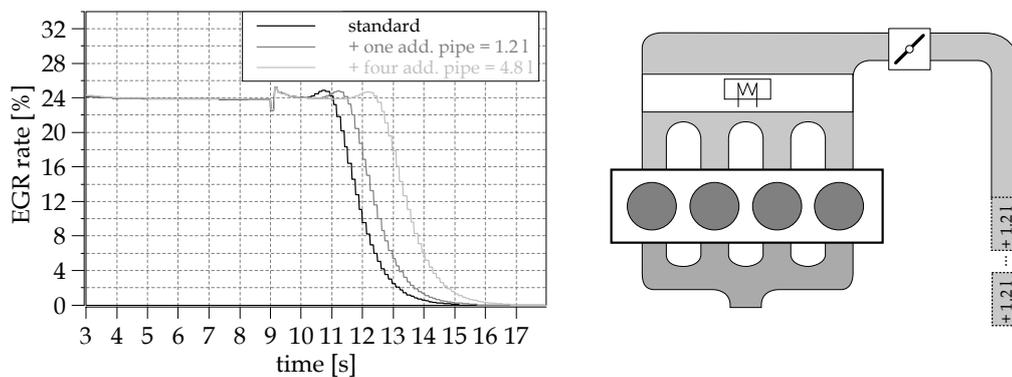
Figure 4.10: Sensitivity analysis: enlargement intake path - additional pipes.

As shown schematically on the right-hand side of Figure 4.10, for this investigation the inlet path between the compressor outlet and the throttle valve is extended by installing additional pipe segments. The pipe segments used correspond to a straight pipe segment modelled as one part in the 1D CFD software with a volume of $1.2\,l$. Compared to the basic configuration, one and four additional pipe segments are installed, which corresponds to an additional inlet volume of $1.2\,l$ or $4.8\,l$, respectively. Accordingly, this also appears as an additional dead volume between the EGR valve and the inlet valves, as evidenced by the diagram on the left. Again, the EGR rate is plotted over time during the load point

change. When the inlet path is extended, the curves are virtually shifted in parallel corresponding to the additional volume, thus significantly delaying the decrease in the EGR rate. The almost pure parallel shift is essentially a consequence of the assumption that the simulation only allows a mixture within so-called discretization elements (10 *mm* length), but not between the individual discretization elements. The (additional) pipes are also modelled with the help of such discretization elements. For long, thin elements, however, this modeling is reasonable and realistic. In this way, interaction between the gas columns of the fresh gas and the EGR air mixture is largely prevented. The design rule that emerges from this examination is that the inlet path should be minimized to as small a volume and as short a length as possible. The engine under investigation, for example, is also available with an outsourced air-to-air charge air cooler. In this configuration, the intake volume doubles by 6 *l* to 12 *l* and, of course, the transient operation with LP EGR also deteriorates noticeably.

Since the gas-carrying elements in the inlet are not only modeled similarly to the pipes, but also as coherent volumes, this influence also needs to be evaluated. As shown on the right-hand side of Figure 4.11, the collector volume directly downstream of the throttle valve is increased by 2, 4 and 6 *l* compared to the base volume. The resulting EGR rate curve is shown on the left-hand side of the figure.
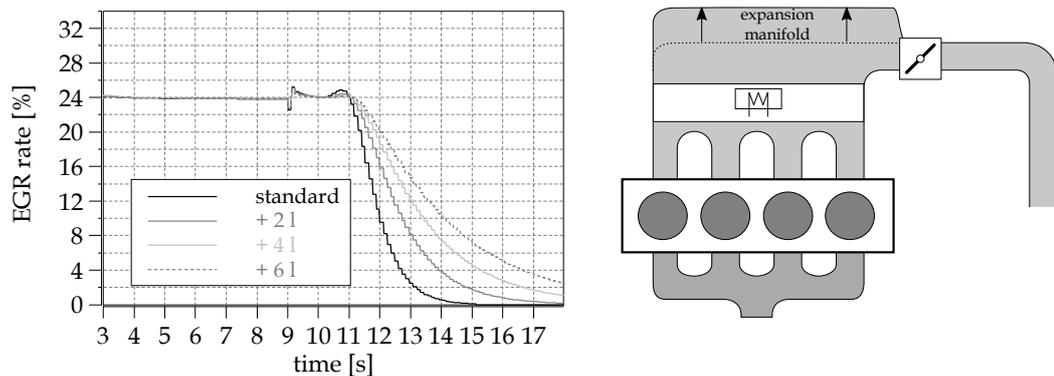
Figure 4.11: Sensitivity analysis: volume intake manifold.

The volume is expanded as a whole and not divided into discretization elements. This results in a completely contradictory behaviour of the EGR rate. Instead of a parallel shift, which changes the start time of the EGR rate reduction, only the gradient of the EGR rate reduction is affected. The larger the volume, the smaller the gradient. The behaviour seen is a result of the mixture between fresh air and the EGR air mixture in the modelled collector. The behavior resembles the simplified container model from Section 2.1. As a design rule, it can be stated once more that such volumes must be kept as small as possible for dynamic operation with LP EGR.

In addition to the basic design rules and the sensitivity to large inlet volumes, the two considerations mainly demonstrate the limitations and specifics of the 1D CFD simulation, which are important in connection with the calculation of dynamic operating points using LP EGR. A way has to be found to correctly model large cumulative volumes and small volumes built by discretization elements. The modelling performed within this thesis can clearly be evaluated as such, since the course of the EGR rate from the simulation corresponds almost exactly to that from the measurement (see Figure 4.3). Nevertheless, these findings will become relevant again in a later section when evaluating the test bench measurements of the worst-case load step.

Bypass System

With regard to the Bypass System, two additional evaluations are discussed. The first aspect analyzed is the influence of the bypass path discharging into the intake manifold instead of directly into the individual intake pipes. This is suggested for example in [97] and in [26]. On the other hand, the influence of an additional pump in the bypass path is investigated as shown in Figure 4.1 (b). An electric charger with a voltage level of 12 *V* is used for this purpose.

At first, on the right-hand side of Figure 4.12, the discharge to the intake manifold is schematically illustrated. The discharge is directly located downstream of the throttle valve, i.e. directly upstream of the inlet of the intake manifold. Its purpose is to use the negative pressure generated by the throttle valve of the main path during a load step to accelerate the mass flow across the bypass. The system can be compared with the schematic depicted in Figure 4.1. The left-hand side of Figure 4.12 shows the EGR rate over time for the Basic System as well as the Bypass System with the discharge location into the intake pipes and the intake manifold.
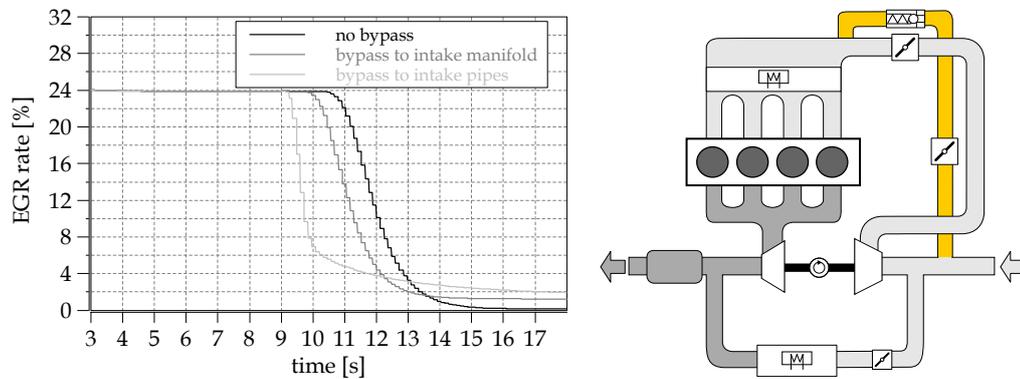
Figure 4.12: Comparison of a Bypass System using the intake manifold as discharge point and one using the inlet channels as a discharge point.

What clearly emerges is the much slower degradation of the EGR rate when the discharge location is not in the intake pipes but in the intake manifold. The reason for this is simple but obvious. The discharging further upstream bypasses less of the intake path. The gas volume that is located between the feeding point and the intake valves must therefore first be delivered to and used in the combustion chambers. The parallel displacement of the EGR curves, similar to that shown in Figure 4.10, is due to the introduction at the inlet of the intake manifold. Even though this has advantages in exploiting the negative pressure generated by the throttle valve, it is not possible to mix the fresh air with the EGR?air mixture in a sufficient way. In contrast, the curve of the bypass being fed into the intake pipes does not exhibit a purely parallel displacement: mixing is assumed to occur. However, this takes place upstream, directly in the intake manifold's volume, as explained in Figure 4.11. At the discharge point, there is hence a backflow in the direction of the intake manifold.

Figure 4.1 indicates an optional charging unit in the bypass line. A unit of this kind is implemented in the simulation model a an electric charger with a voltage level of 12 *V*, operated at maximum power directly after the load step. The course of the EGR rate for the Basic System, for the Bypass System with introduction into the intake pipes and for the Bypass System with introduction into the intake pipes and the electric charger switched on can be seen in Figure 4.13.
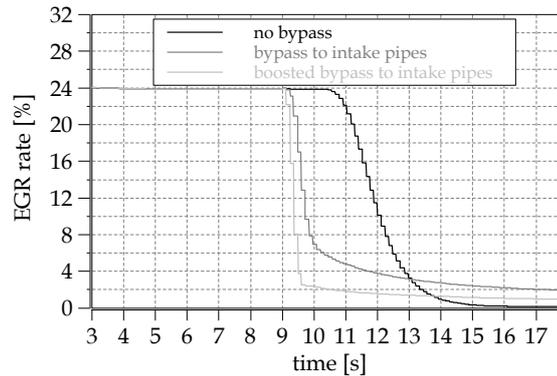
Figure 4.13: Comparision of a Bypass System with and without additional boosting.

With regard to undershooting the EGR tolerance, the additional charging achieves an advantage of approx. $0.6\,s$ in comparison to the system without it. Thus, the boosting accelerates the system by 55 %. Both the starting time of the EGR reduction and the gradient are improved. The investigation reveals the potential of the existing system and how easily a further improvement of the system could be implemented. However, it should be noted that the installation of an additional charging unit increases the package requirement enormously. For the vehicle-engine combination considered, this would not be practicable.

High-Pressure Bypass System

Since the overpressure from the compressor of the turbocharger can be used in case of a load step, the High-Pressure Bypass System has the boosting or charging unit incorporated. However, due to the resulting dead-end section between the EGR discharge point and the bypass branch, the opening moment of the bypass valve is crucial for its function. A sensitivity analysis of the opening moment of the bypass valve after the load step is illustrated in Figure 4.14 by plotting the EGR rate during the load point change as a function of the opening moment over time. The curves displayed include a range from the bypass valve opening in synchrony with the load reduction to it opening 3 working cycles after the load change.
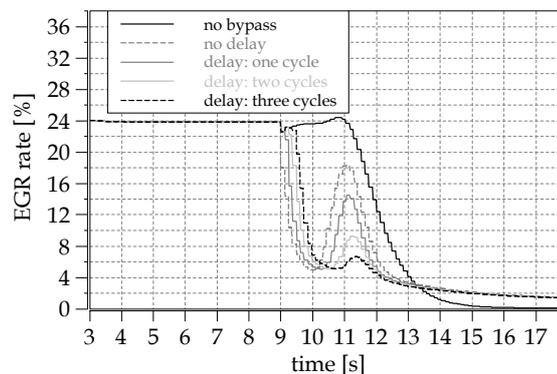
Figure 4.14: Sensitivty analysis: opening time of the bypass path.

The main dilemma of the system becomes apparent. The earlier the bypass valve is opened, the faster the EGR rate is initially reduced. This means that the EGR rate reduction both starts earlier and follows

a steeper gradient. However, it is also true that the second overshoot described above is likewise more pronounced. For the solution illustrated in the overall comparison, a delay of two working cycles is chosen. The overshoot here is only marginally higher than the EGR tolerance and illustrates how well the system functions when applied correctly. This system therefore has further potential for improvement if the EGR discharge point is as close as possible to the compressor inlet and the bypass branch point is as close as possible to the compressor outlet, i.e. the dead-end section is kept accordingly small. In the special case of the application here, it would be reasonable to slightly increase the permissible external EGR rate at the target load point using appropriate measures, such as changing the valve timing.

Scavenging System

With regard to the Scavenging System, both the basic functionality and the reasons for the suboptimal operation will be examined in more detail. For this purpose, in Figure 4.15 the mass flows over the scavenging path and over the main path are plotted
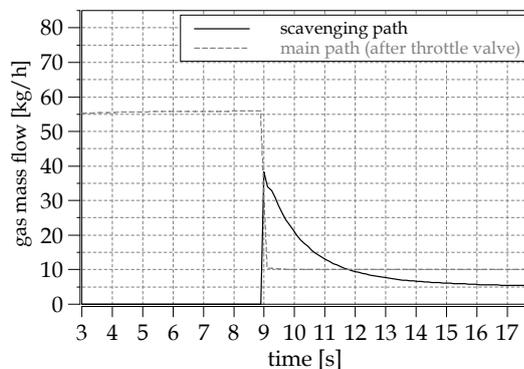
Figure 4.15: Air mass flows over time for Scavenging System.

Immediately after the load point change, the mass flow across the scavenging path increases to a comparable extent as the mass flow across the main path downstream of the throttle decreases. However, while the main path mass flow can be adjusted very quickly and remains constant, the mass flow via the scavenging drops again and the scavenging effect diminishes. Figure 4.16 provides an explanation. The behaviour of the turbine speed and the pressure downstream of the throttle valve for the Basic System and the Scavenging System are compared.

Although the Scavenging System, unlike the Basic System, closes the wastegate as soon as the load is adjusted, the turbocharger speed drops significantly faster than for the Basic System. The actual flushing function by the compressor can therefore not be maintained. The reason for this is two-part. On the one hand, the exhaust gas back pressure is increased by the discharge downstream of the turbine and thus the operation of the turbocharger turbine is negatively influenced. On the other hand, the boost pressure in the intake manifold decreases considerably faster, since it can quickly be reduced via the scavenging path, as shown on the right-hand side of the figure. In order to increase the pressure again, however, the throttle valve must be opened further, which leads to a reduction in the stagnation effect at the branch point of the scavenging pipe and is therefore just as obsolete.

The Scavenging System can only function if the intake path until the branching point is correspondingly small or the starting point of the load step is particularly high, so that a great deal of intake pressure can be relieved via the branching pipe and the entire intake path can be scavenged.
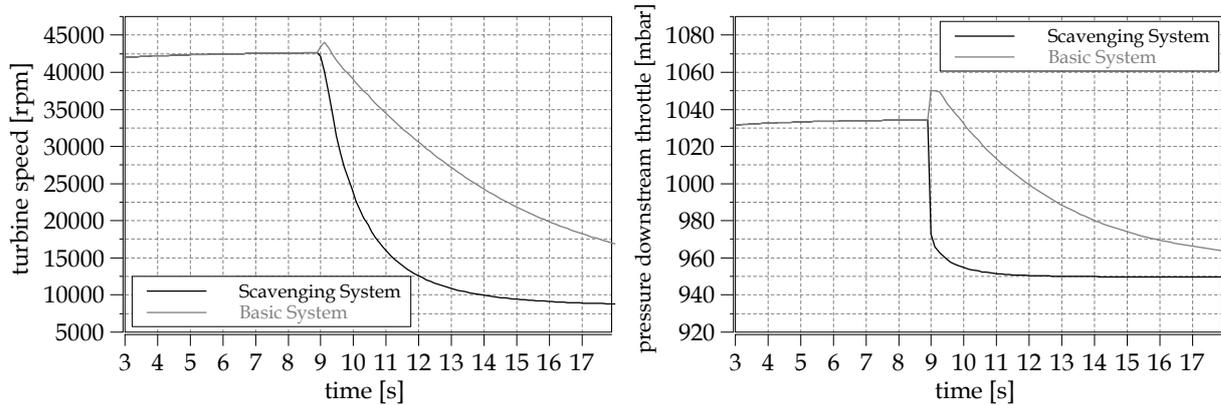
Figure 4.16: Turbine speed and pressure upsream throttle valve over time for Scavenging System and Basic System.

Direct Boost System

With the Direct Boost System, the first decisive factor is how the mass flows are distributed via the main path and via the Direct Boost System. Figure 4.17 provides the respective curves during the load point change over time.
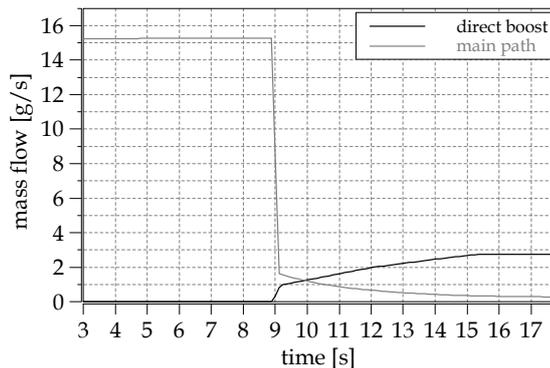
Figure 4.17: Mass flow over time through Direct Boost System and through main path.

The immediate increase in the mass flow via the third inlet-valve of the Direct Boost System is noteworthy. It is accompanied by a rapid drop in mass flow via the standard air path. The mass flow via the Direct Boost System subsequently increases gradually, just as the mass flow via the inlet path decreases slowly. This is where the central problem originates. The mass flow from the Direct Boost, consisting of fresh air, can only be built up and the EGR rate reduced to the extent that the mass flow in the main path can be reduced. However, as the turbocharger continues to rotate, this happens only at a very sluggish pace. Figure 4.18, which shows the turbine speed for the two systems, shows that this problem is even worse for the Direct Boost System compared to the Basic System.

The BMW Direct Boost System was originally conceived as a dynamic measure in the low-end torque area of the engine map. This feature comes to bear here as well. Although the wastegate of the turbocharger is fully opened simultaneously with the load step, the turbocharger speed initially rises slightly and subsequently declines at a very sluggish rate. The air blown in accelerates the turbine by enthalpy induction. In contrast, the speed of the turbocharger of the Basic System undergoes a
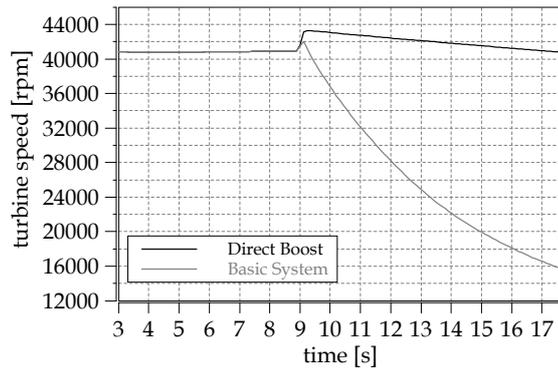
Figure 4.18: Turbine speed over time for Direct Boost System and for Basic System.

degressive decrease after the load point change. The ensuing contradiction and its consequences for the Direct Boost System are illustrated in Figure 4.19, which shows the pressures in the inlet directly upstream and downstream of the throttle valve for the Basic System and the Direct Boost System.
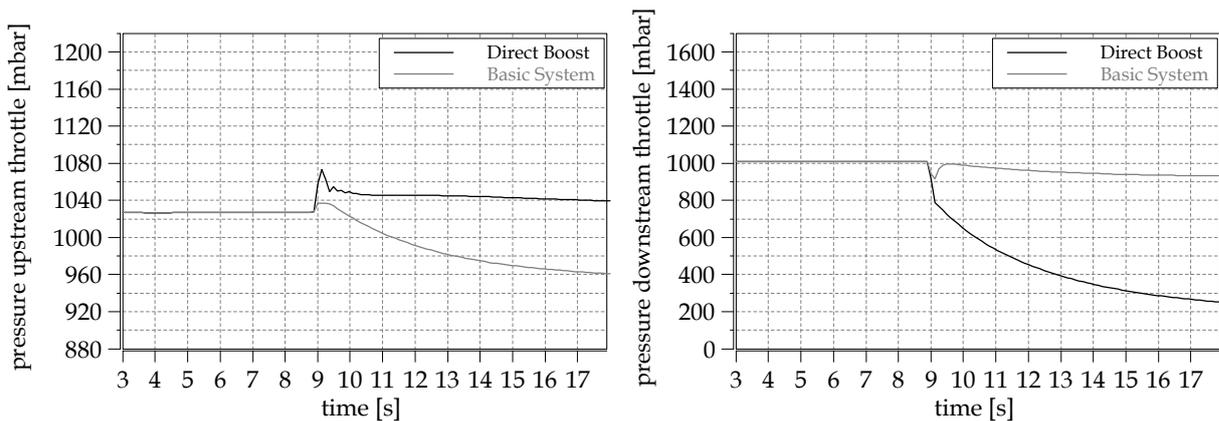
Figure 4.19: Absolute pressure upstream and downstream of the throttle valve over time for Direct Boost System and for Basic System.

While the pressure decreases faster and significantly stronger for the Direct Boost System than for the Basic System due to throttling downstream of the throttle valve, the pressure upstream of the throttle valve increases as a result of the accelerated turbocharger after the negative load step. The system is very sensitive in this respect. If further throttling is applied, additional air must be provided by the Direct Boost System, which increases the enthalpy transfer into the turbocharger's turbine. At the same time, the compressor delivers higher pressure which means correspondingly increased throttling is needed to comply with the torque requirement. If the pressure between the compressor and throttle valve rises above a critical level, the compressor might pump, causing an unwanted backflow.

This clearly reveals that the mass flow across the main path can only be restricted to a certain degree. Either the pressure is reduced via the throttle valve so that the pressure upstream of the throttle valve increases to such an extent that the compressor starts to pump or the pressure is not reduced sufficiently and the torque requirement can no longer be met. As stated at the beginning of this chapter, all systems are evaluated within the boundary conditions of the engine used. Remedies, away from these boundary conditions, would include increasing the cross-section of the wastegate to effectively prevent

the turbocharger from accelerating and increasing the volume of the inlet path until the throttle valve to better suppress compressor pumping. However, this would again prevent the fast flushing of the inlet path in the event of a load step from the EGR operating range. Nevertheless, the Direct Boost system can become a very effective measure to implement dynamic engine operation with LP EGR if the engine as a whole is adapted accordingly.

Generator System

The system with a generator fundamentally behaves like the Basic System with regard to the air mass flows and the turbocharger, only at an increased level due to the load point shift. The key issue with this system is instead the impact of the increased load point on the overall system. Figure 4.20 presents both the indicated work during the load point change and the absolute fuel mass flow for the Basic and the Generator System.
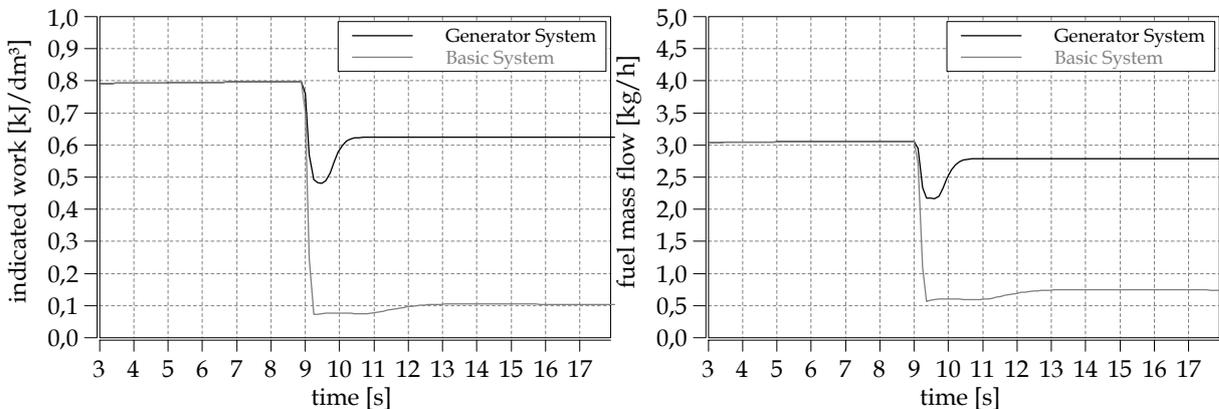
Figure 4.20: Indicated work and fuel mass flow over time for Generator System and for Basic System.

Figure 4.7 reveals that the inidated efficiency of the Generator System is significantly improved in line with the indicated work shown here. This is mainly due to the significantly reduced throttle losses. The fuel mass flow, on the other hand, is increased by a factor of 3, which has to be considered negative during normal operation. The additional energy introduced by the fuel, though, is stored by the generator in a battery and can be reused later, for example to start the engine. However, the conversion losses should be carefully evaluated in any case. In addition, the operating strategy and battery size must be coordinated so that the generator can be switched on at any time and energy can be stored in the battery, otherwise the load point shift will not work and the system will operate at the same level as the Basic System.

Evaluation of key aspects

All systems are evaluated in terms of their functionality of quickly reducing the EGR rate in response to the worst-case load step. The Generator System and the Bypass System have proven to be equally advantageous in this respect. In the previous subsections, important specifics of the individual systems are discussed and background information for the particular behavior is elucidated. However, the serial implementation of a system is usually not only based on one singular function, but also requires the evaluation of other aspects. Thus, before one of the systems can be implemented in hardware and analyzed on the test bench, the systems are assessed with regard to other criteria such as cost,

package, applicative effort, series maturity and further benefit using Table 4.1. The latter of the listed aspects represents the possible customer benefit extending beyond the function of implementing LP EGR operation and thus saving fuel. Since the individual criteria are of different relevance, they are weighted from 1, which means 'less important' to 3 for 'particularly important'. The weighting factors are given in the rightmost column in the table. A similar classification has been chosen for the evaluation values with 0 standing for 'insufficient' or 'high effort' and 3 meaning 'excellent' or 'low effort'.

Instead of explaining all of the values individually below, only those deviating from expectations based on what has been explained so far are addressed, along with those that are regarded as particularly crucial. The first thing that should be mentioned in this context is the application effort in connection with the Basic System. Since the Basic System needs to be applied with LP EGR in order to run it reasonably for the customer, a 2 instead of a 3 is assigned here. The application effort for the Bypass and the High Pressure Bypass System is even greater, since the additional bypass valve and the mixing of gases with different EGR rates allow further degrees of freedom. Accordingly, the respective value is 1.

While the Scavenging System is rated at 1 for functionality, the High-Pressure Bypass System is rated at 2. This is because the High Pressure Bypass System is the best system in terms of the initial undercutting of the EGR tolerance. Reducing the dead section between the EGR inlet point and the branch of the bypass path or a minimally increasing of the EGR tolerance would turn this system into the most advantageous one. A similar argumentation is chosen for the Direct Boost system, and a 2 is given. The Direct Boost System can reduce the EGR rate very quickly when modifications are made, such as increasing the wastegate cross section or the inlet volume.

In terms of cost, the Bypass System is considered to be very competitive, since only simple components have to be modified. Meanwhile the Direct Boost System and the Generator System are considered to be very expensive and therefore valued at 0, the former due to modifications at the cylinder head and the latter, due to modifications at the belt drive and because of the additional $48\,V$ system including all necessary components. The same applies to the package, where the cylinder head and belt drive are considered critically difficult to modify. The High Pressure Bypass and the Scavenging System are rated worse than the Bypass System, as modifications have to be made at the turbocharger outlet and for the Scavenging System at the TWC outlet, respectively, and these are particularly space critical components.

The application effort is evaluated similarly for all systems. Only the Scaevnging System stands out positively: unlike the other air path solutions, no mixing of two gases with different EGR rates is required.

Only the Direct Boost and the Generator System offer additional benefits. While the former can act as a dynamic measure in its actual function, the generator can also boost in dynamic mode or provide additional comfort functions through the $48\,V$ system.

Finally, all systems are ready for series production, although the Direct Boost, the Scavenging, the High-Pressure Bypass and the Generator System are rated slightly lower, mainly because of sensitive hardware modifications.

After the overall evaluation of the systems, it can be concluded that the Bypass System has emerged as the favorite. It combines the most advantageous function with simple, cost-effective implementation. The system is therefore implemented in hardware and analyzed in the following section on the basis of test bench measurements including sensitivity analyses and full-factorial DoEs.

Table 4.1: Evaluation matrix for the different systems.

Criteria:	Basic System	Bypass System	High-Pressure Bypass	Scavenging System	Direct Boost	Generator	weighting
functionality	0	3	2	1	2	3	3
cost	3	3	2	2	0	0	2
package	3	3	2	2	1	1	1
applicative effort	2	1	1	2	1	1	1
further benefit	0	0	0	0	3	2	2
series maturity	3	3	2	2	2	2	1
sum	**15**	**19**	**15**	**17**	**16**	**17**	**-**

4.3 Engine test bench measurements

If the bar chart from Figure 4.6 is translated into working cycles, the values for the individual systems are as given in Table 4.2. It becomes clear that the Bypass System remains a total of 9 working cycles above the EGR tolerance, while the torque reaches the target value within only one working cycle. All other systems are at the same level or even worse or slower. On the basis of the simulation and the criteria assumed so far, it can be assumed that the system will not function straight off and without adjustments or cutbacks.

Table 4.2: Cycles until EGR tolerance in target load point of the worst-case load step is undershot.

	Basic System	Bypass System	High-Pressure Bypass	Scavenging System	Direct Boost	Generator
Cycles to undershoot EGR tolerance	28	9	20 (8)	18	23	9

Merely evaluating the EGR rate would add little value to the measurement compared to the simulation, as it has already been shown that the measurement and simulation match very well in this respect. It is much more important to generate information about the combustion and its properties, such as the COV_{imep} and the misfire events. These values are not sufficiently covered by the 1D CFD simulation. Although semi-empirical approaches ([102], [107]) as well as physical approaches ([20], [53]) approaches are suggested in the literature on how to implement cycle-to-cycle variations and misfires in 0D/1D CFD simulation, these methods all relate to steady-state operating points. The simulation cannot yet provide any information on the behavior if there is a transient change in the operating point and all the boundary conditions and input parameters thus change rapidly. However, the COV_{imep} cannot be obtained from dynamic measurements either, since both values from Equation 4.1, the standard deviation of the mean indicated pressure σ_{imep} as well as the mean of the indicated pressure $\overline{imep}$ change

rapidly and considerably, but also disproportionately during the load step. A new value is therefore needed to be able to interpret the results obtained from the engine test bench.

One simple, persuasive criterion that can be implemented when evaluating the test bench measurements is the determination of the number of combustion misfires. The speed at which the load, and thus the mean indicated pressure, changes during the load point change means that the driver does not notice cyclical fluctuations in detail. However, when the engine produces combustion misfires, this is clearly perceptible. Furthermore, these must be avoided to ensure that the exhaust aftertreatment functions properly, as the fuel?air ratio in the TWC may differ significantly from the stoichiometric one. In order to obtain a characteristic value, the percentage misfire rate per two working cycles or per four engine revolutions, respectively, is integrated as illustrated in Figure 4.21. In addition to the integral, a possible course taken by the misfire rate over time is shown as an example. In concrete numerical values, 12.5 % means that one cylinder per 2 cycles misfires, since a total of eight combustions take place within these two cycles within the 4-cylinder engine investigated. The starting point of the integration is the starting point of the load step, while the ending point of the integration is the end of the measurement, which is set 10 s after the load step starts. The calculated value is not a physically interpretable value, but one that can be used to quantitatively evaluate and compare the configurations analyzed below.
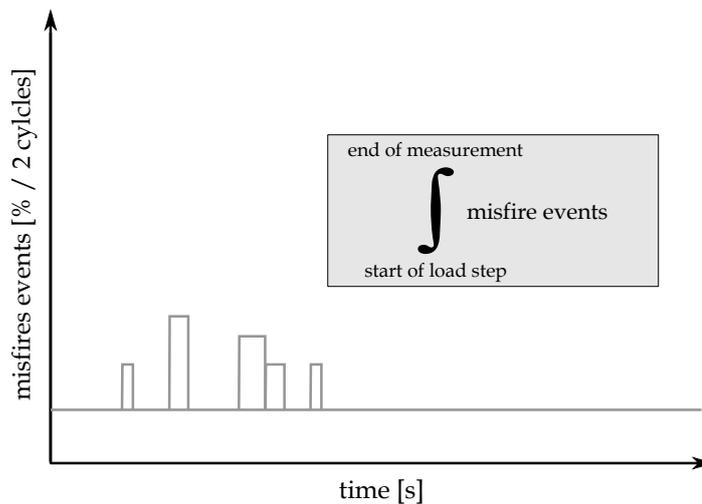
Figure 4.21: Integrated level of misfire.

In the first part of the measurements, a sensitivity analysis is performed with respect to the throttle and bypass valve angle. The measurements of the worst case load step are based on a full factorial *Design of Experiments* (DoE), adjusting both valves within a physically reasonable interval. The full-factorial DoE is used to capture detailed effects of the low mass flows in the area of the target load point. The setting of the values of the throttle and the bypass valve is pre-controlled by the engine control unit within the limits given in Table 4.3.

When both valves are changed simultaneously, the target load as well as the specified torque gradient will of course change. To ensure that all configurations can still be displayed correctly in accordance with the specifications, the valve stroke is used to make the necessary adjustments. In addition to the values from Table 4.3, the configuration with a closed bypass valve is examined. This corresponds to the standard configuration without Bypass System. The results of the sensitivity analysis are shown in

Table 4.3: Sensitivity analysis: variation limits and step size of the throttle and bypass valve angle in the target load point.

	minimum	maximum	increment
throttle valve angle	0°	4°	0.5°
bypass valve angle	5°	20°	2.5°

Figure 4.22. The value of the integrated misfires for the respective combination of bypass and throttle valve angle is provided. The value is calculated as the arithmetic mean of three identical load steps. Statistical influences can thus be excluded.
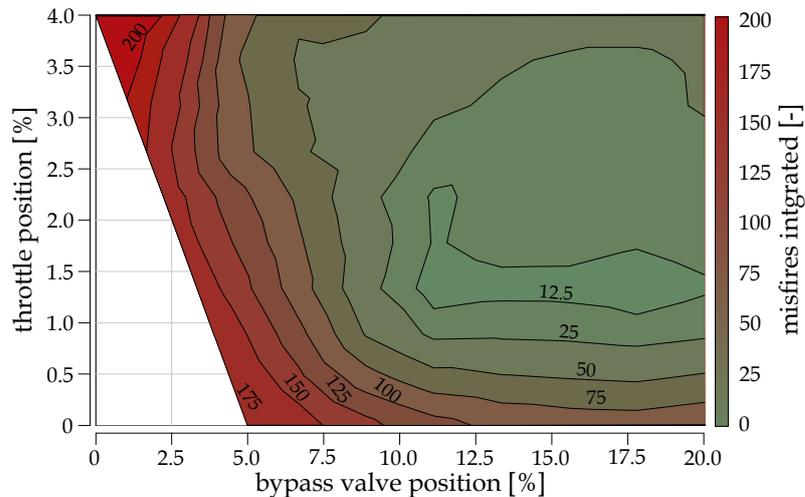
Figure 4.22: Sensitivity study: throttle and bypass control valve position.

The first thing that stands out is the point in the upper left corner of the map, which represents the Basic System and produces by far the most misfires during the load point change. This again proves that LP EGR using the steady-state values cannot be implemented by the Basic System during transient operation. For the remaining part of the map, starting from the map area containing the lowest values of about 12.5, three areas with different influences can be identified. The first of these is located to the left of the area with the lowest values. The isolines run approximately parallel to the y-axis. The bypass valve angle is crucial or too small in this particular area. When the angle is increased, the fresh air mass flow via the bypass path increases and the misfire rate declines. A second area can be found on top of the shell containing the most favourable values. Again, the values of the misfire rate increase, with the isolines parallel to the x-axis. This implies an influence of the main path's throttle angle, which is too large in this map area, i.e. the mass flow of the EGR air mixture from the main path is not sufficiently throttled. The area below the zone with the lowest misfire rate is of particular interest. Again, the isolines run parallel to the x-axis. Minor throttle valve angles lead to a striking deterioration of the misfire rate, which contradicts the previous observation regarding the blockage of the main path. If throttling is too sharp, the EGR air mixture is pumped back across the compressor in the direction of the branch-off point of the bypass-path. Instead of pure fresh air, the EGR air mixture is supplied to the

combustion by the bypass path. A schematic view of the mechanism described here is shown in Figure 4.23.
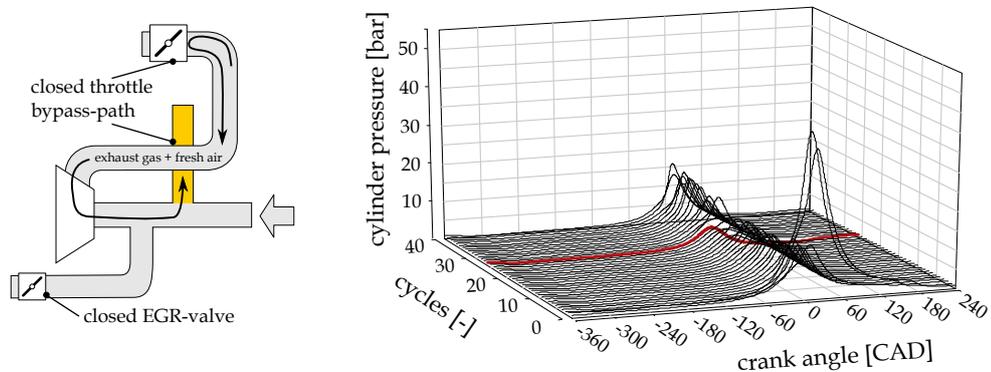
Figure 4.23: Backflow from the main path and corresponding flow chart pressure over crank angle.

Proof of this assumption is shown on the right-hand side of the figure by the flow chart of the consecutive cylinder pressure curves over the crank angle (Cylinder 2) during the load point change. In this particular case involving a completely closed throttle valve, there is a misfiring combustion 23 cycles after the target load is set. The misfire event is emphasized in red on the flow chart. This misfire event, occurring long after the actual load step, strongly indicates that the EGR rate is only exceeded with a considerable delay.

Within the simulation, this behavior is not recognized by means of the EGR rate, since it is only possible to simulate compressor pumping qualitatively in the software. However, Figure 4.24 below illustrates that a backflow occurs in the 1D CFD simulation, even if it is relatively small. The moment the backflow arises is crucial for the analysis.
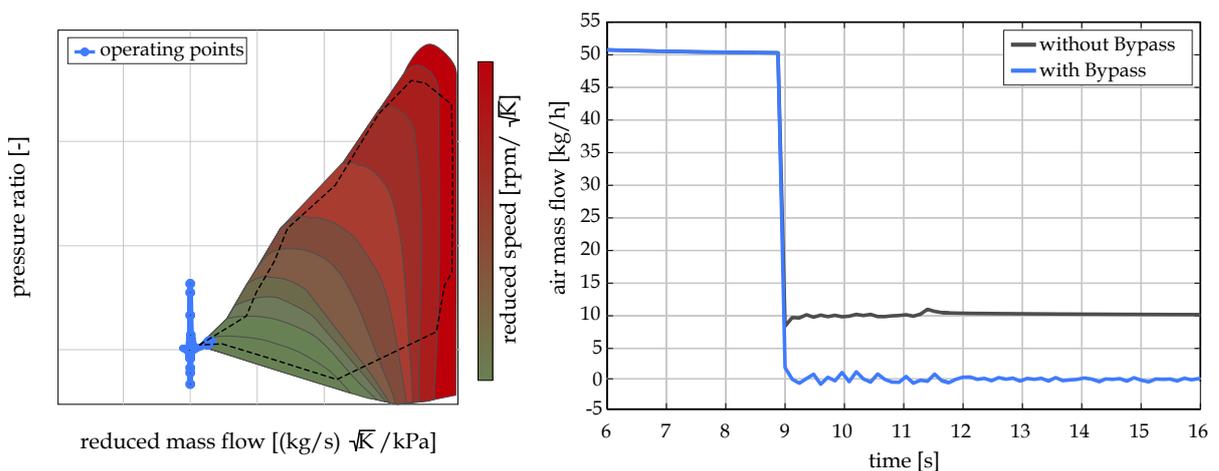
Figure 4.24: Compressor map including operating points and mass flows of the bypass configuration from the simulation.

On the left-hand side, the operating points during the load step are depicted on the compressor map. Pressure ratios less than 1 and mass flows less than 0 both appear, indicating backflow. This behavior can also be seen for the mass flow in the main path, shown on the right-hand side of the figure once with

a closed throttle valve and open bypass valve (blue line) and once in the Basic System without a bypass and normally open throttle valve. When the throttle valve is closed, the mass flow oscillates around the zero position between a negative and a positive mass flow. The first significant reverse flow occurs after about $0.4\,s$. The moment of the misfire is estimated to coincide well with the occurring mass flows and delays. The flow rate over the bypass is approximately $2.2\,l/s$. Based on a cycle duration of $0.12\,s$ and 23 cycles, $6.1\,l$ flows along the bypass path. The capacity from the compressor outlet to the inlet valves is $4.7\,l$, opposite to the actual direction of flow through the bypass path, as shown schematically in Figure 4.23. Following a delay of $0.4\,s$, this simplified calculation results in a flow volume of about $5.5\,l$. This coincides very well with the $6.1\,l$ volume passed after 23 cycles. The small difference is attributed to statistical effects and mixing.

Given that the worst-case load step under the given boundary conditions cannot be implemented entirely free of misfires even with the Bypass System at its optimum setting, the following sensitivity analysis is intended to demonstrate the adjustments to the operation points required to achieve misfire-free operation. The EGR rate of the starting load point and the target load are defined as the main levers. The latter strongly influences the EGR tolerance of the target load point. Similarly to Figure 4.22, Figure 4.25 displays the integrated misfire rate depending on the starting EGR rate and the target load. The setting is the one previously analyzed as appropriate at a misfire rate of 12.5, a bypass valve opening of 15 % and a throttle valve position of 1 %.
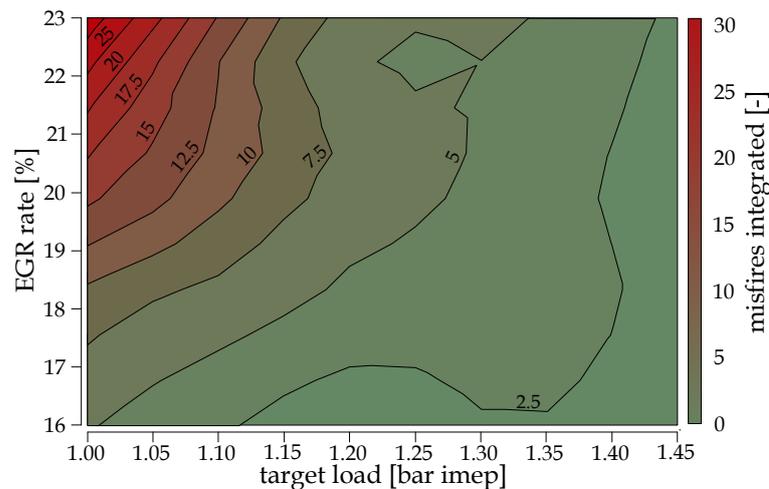
Figure 4.25: Sensitivity study: starting EGR-rate and target load.

This map can again be broken down into two parts. The separation line runs at an EGR rate of about 18 %. Above that, the isolines run approximately parallel to the y-axis, so the impact comes especially from the target load and thus the EGR tolerance. An increased dilution tolerance reduces the misfire rate. Below an EGR rate of 18 % the isolines tend to be parallel to the x-axis. A positive influence can therefore be generated by reducing the starting EGR rate.

There is a map range that contains values below 2.5, which means that for the 3 load steps measured per setup, one working cycle still suffers one misfire. The area is cut off diagonally in the map and lies between a target load of $1.1\,bar\,imep$ with a starting EGR rate of 16 % and a target load of $1.4\,bar\,imep$ with a starting EGR rate of 23 %.

In addition to the sensitivity analyses carried out here, others were conducted for the fuel?air ratio along with the target load and the throttle position. The results are in turn presented in the form of a characteristic map in Figures A.2 and A.3, enclosed in the Appendix to this thesis. Deviations in the results obtained for a stoichiometric fuel?air ratio compared with those presented in this chapter are due to a slight remaining statistical variation. Finally, no significant improvements can be achieved by modifying the fuel?air ratio. The results achieved when the bypass valve is regulated with a pilot control are likewise given in the Appendix (Figure A.4). The faster opening of the bypass path does not lead to further advantages either.

Even with a further reduction in the EGR rate at the starting point, which implies an efficiency disadvantage of $0.5\,\%\,points$, as well as with a slight increase in the target load, which impairs both the efficiency and the absolute fuel consumption, a completely misfire-free operation is not feasible using the Bypass System. The changes detailed in the Annex are similarly not target-oriented. In principle, this corresponds to the simulation results and can thus be expected when considered in isolation. Due to the introduction of the dedicated bypass path into the intake pipes, however, there are grounds to expect that a charge stratification between the fresh air and the EGR?air mixture can be accomplished in the cylinder. The excitability of the mixture directly at the centrally located spark plug and thus the combustion stability might therefore be increased. A schematic illustration of this mechanism and a further explanation are given in the Appendix under Figure A.5.

If a more detailed evaluation of the individual cylinders is carried out, it becomes clear why this does not work either. It is clearly noticeable that the closer the cylinders are located to the throttle valve, i.e. to the point of introduction of the main path, the worse the combustion stability becomes. As an example to prove this point, the previously described flow chart of the cylinder pressure over the crank angle for the individual consecutive cycles of Cylinders 1 and 4 is plotted in Figure 4.26. Cylinder 1 is located the closest to the inlet point of the main path, Cylinder 4 the furthest away.
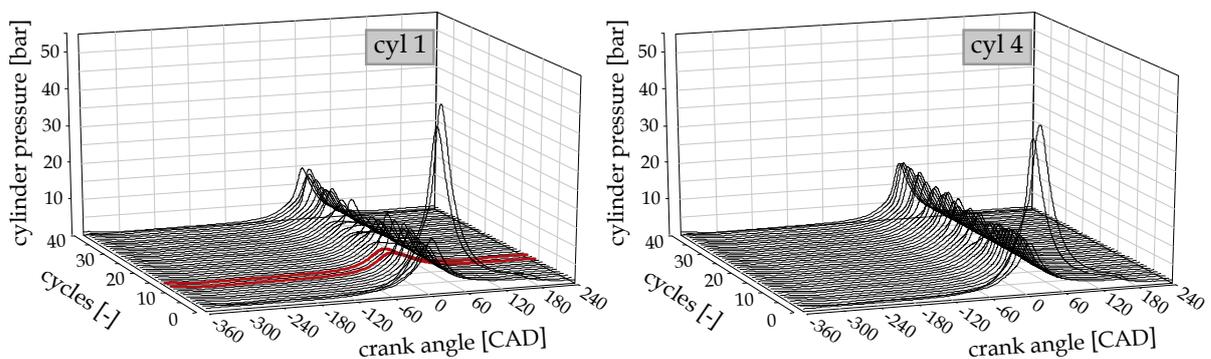
Figure 4.26: Flow chart pressure over crank angle Cylinder 1 and Cylinder 4.

The differences between the two cylinders are clearly visible. Apparently, the COV_{imep} of Cylinder 4 is much lower than it is for Cylinder 1. While the latter has two working cycles after the load point change, where misfires occur (highlighted in red), Cylinder 4 shows only slight differences in the cylinder pressure curves between the individual cycles and consequently high combustion stability. Based on the isolated consideration of a single cylinder, it is reasonable to assume that the system functions.

It is necessary to answer the question of why this characteristic behavior occurs depending on the cylinder order, which can be observed to a greater or lesser extent in all possible settings. For this

purpose it is advisable to use the simulation again. All analyses made so far assume that the EGR rate averaged throughout all cylinders. This assumption applies to LP EGR, as the mixing of the gases is very good due to the available length of pipe and a homogeneous composition can be presupposed. However, this assumption has to be verified anew due to the bypass path feeding close to the inlet valves. For this purpose, the simulation model is further detailed in the area of the discharge point. The inlet point, which is modelled as a single volume so far, is divided into four individual pipes after the check valve. The results obtained for the single cylinder EGR rates are shown on the left side of Figure 4.27. On the right-hand side, the sequence of the cylinders is shown for clarity.
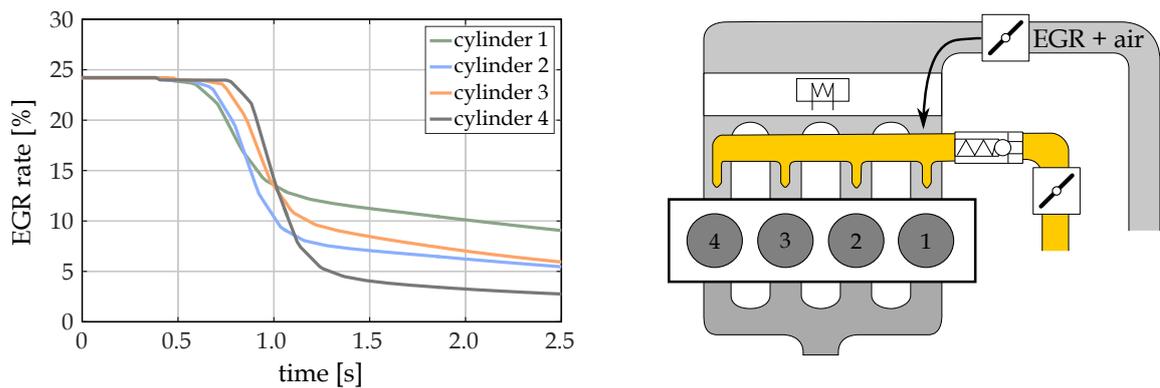
Figure 4.27: Cylinder separated EGR-rate and unequal distribution of the EGR-air mixture.

Clearly, the single cylinder EGR rates differ. Although the reduction in the EGR rate starts first in Cylinder 1, which is closest to the main path or throttle, that cylinder has the highest EGR rate immediately after the load step. This EGR rate exceeds the EGR tolerance; hence, significant cycle-to-cycle variations are found after the load point change. This behavior reverses the further away a cylinder is from the throttle valve. Thus, Cylinder 4 is the last cylinder in which the EGR rate starts reducing, but has the lowest EGR rate after the load step.

With the help of the diagram on the right-hand side of the figure and knowledge of the throttle position, a straightforward explanation can be given for this behavior. Both the main path and the bypass path are aligned asymmetrically. In the plan view shown here, both paths begin from the right-hand side and hence supply Cylinder 1 first. Thus, when the bypass valve is opened, the fresh air reaches Cylinder 1 initially and the EGR rate is reduced first. However, since the throttle valve cannot be closed completely due to the backflow via the compressor, there is a residual leakage via the throttle valve. As the mass flow is small, the entire intake manifold is no longer evenly flooded with it. Instead, most of this mass flow is fed to Cylinder 1. The EGR rate is accordingly the highest of all cylinders. One central problem of the bypass solution when implemented in practice is the apparently asymmetric feed of both the main path and the bypass path. Due to the current package requirements, however, this is absolutely necessary and cannot be avoided. Using a symmetrical air path system is expected to be further beneficial and to improve the ability to achieve a misfire-free implementation.

4.4 Conclusion

This chapter discusses various hardware options for improving the reduction in the EGR rate during load steps. The basis for the evaluation is the worst-case load step presented in Chapter 2.1. The basis for comparison is the standard engine from Chapter 3.1, named Basic System. The simulation results indicate that there is a reduction in the critical time span until the EGR tolerance is undercut by almost 70 % for two of the systems investigated. After an overall evaluation, the Bypass System is implemented in hardware on the engine test bench. The results achieved are in fact better than expected based on the simulation. Various sensitivity analyses reveal the optimal setting between the throttle and bypass valve, as well as possible improvements by reducing the EGR rate at the start load point or by increasing the target load. Finally, it remains to be stated that the bypass system is limited particularly by the assyemtric supply, which is dictated by the package requirements.

A combination of the different systems may also be conceivable. For example, the Scavenging System could be used in combination with the Direct Boost or the Bypass System to discharge the mass flow across the throttle valve. However, it is important to note that each combination of these systems increases the requirements in terms of costs and package, but also in terms of complexity. This is a key argument for not considering combinations of individual systems, although the probability of implementing the entire stationary efficiency potential appears to be within the realm of possibility. The decision to implement these measures or even a combination is a trade-off between the benefit in terms of efficiency and the necessary expenditure. Yet not only costs and packaging must be taken into account but also the application effort of such systems, as they all bring additional degrees of freedom into an engine configuration that already features the additional degree of freedom because of the LP-EGR configuration itself.

This is the major impetus for the investigations and explanations in the next chapter. While this chapter demonstrates the difficulty of understanding the interrelationships of a single load shedding operation and of optimizing the system in this way, the findings in the next chapter offer the possibility to optimize it with minimal effort. The model presented minimizes the application effort and maximizes the system understanding to achieve the best possible efficiency level for engines with a LP EGR system. Although all results are developed and presented using the Basic System, the model developed could subsequently be implemented for the system solutions presented in this chapter by means of extensions and adaptations.

5 Investigation into a model-based approach

This section constitutes a central point of this thesis and makes a new contribution to the field. Parts of it have been previously published in [60].

The increasing complexity of combustion engines due to the growing number of variable engine parameters is leading to a significantly higher demand for engine test bench measurements. The results obtained are not only needed for the application but also for the fundamental investigation and evaluation of engines. LP EGR further increases the level of sophistication in turbocharged SI engines by introducing an additional degree of freedom.

The last chapter demonstrates that dynamic operation in particular - either in general or incorporating LP EGR - shows signs of influences that cannot be fully evaluated with stationary measurements a priori. The measurement of discrete steady-state values is therefore only conditionally reliable during transient operation, even in the case of an extremely fine grid of measurement points, considerably increasing the required effort. On this account, an applicative level of operational safety always needs to be provided, meaning that part of the efficiency potential remains unused.

Although models likewise have inaccuracies that prevent the efficiency potential from being fully exploited, it is conceivable to calculate and determine all necessary and possible intermediate operation states during dynamic operation without significant effort. Assuming the same advantage in terms of efficiency could be achieved during operation with LP EGR, either with the help of a model or with the help of an extensive and complex measuring program which would also have to be made anew for each engine configuration, it is the model that has clear advantages, since the effort required for multiple implementation is considerably lower. The key prerequisites of a model fulfilling this assumption are of course both a sufficiently short computing time to be able to carry out online calculations as well as a feasible model structure that is determined using a smart measuring program to set the model up. In addition, the model must be simple to apply to other engines.

When classifying models, important criteria include the computing time and the modeling depth or equivalently the accuracy of the model's reproduction of reality. Figure 5.1 shows a classification of different models in accordance with this framework. The computing time usually increases in line with the accuracy of the model, represented by the bisector in the diagram. In the ideal case, a model can be classified as deviating from the bisector, thus reducing the computing time without losing accuracy. This is possible if models are not completely self-sufficient from the start, but have to be calibrated beforehand. Hence they represent a mixture of measurement and modeling. Using these models, the computing time can be reduced, but the accuracy remains high.

Furthermore, it is important that the model is based on physics, or at least that it can be interpreted on a physically logical basis. The model that is introduced in this chapter is initially set up empirically with the help of a custom-designed measurement program. It is thus not initially based on purely physical relationships or equations. However, the individual influence equations of the respective model input parameters can and must be interpreted and implemented on the basis of physics. That means that a physically evaluable model, which can be transferred to other applications, is built on an empirical measurement database for one engine. For this reason the model is classified as semi-empirical or physically based. Semi-empirical models have several advantages. They represent complex mechanisms
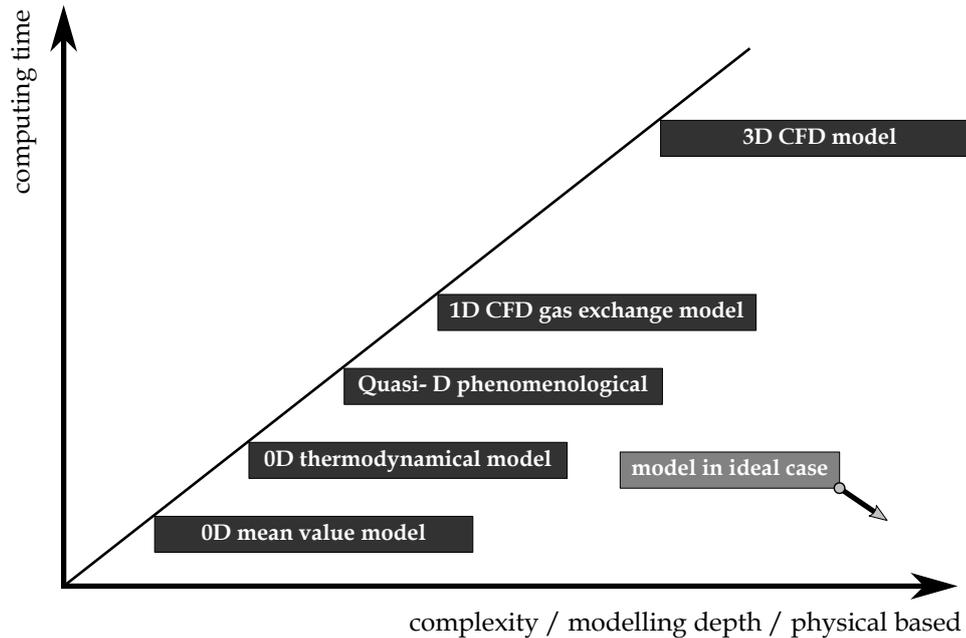
Figure 5.1: Comparison and classification of different models, according to [70, 100].

in a simplified way and adopt certain assumptions. The parameters of semi empirical models, however, can be adapted relatively quickly to existing problems. This is also the case in this study, where the model can be transferred to other engines with different characteristics by a simplified calibration.

A descriptive classification of the model from this work follows Goodman [34]. In the theory of empiricism and physical basis, a model is created as shown on the right-hand side of the schema in Figure 5.2.
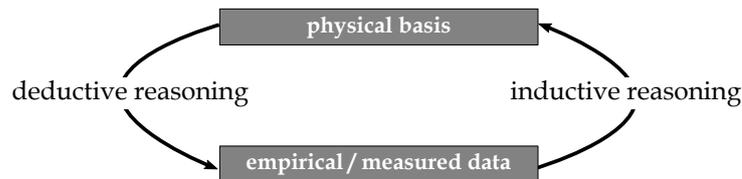
Figure 5.2: Descriptive classification of the model, according to [34].

Using *inductive reasoning*, a conclusion about the physical basis is made on the basis of the measured data. The method of *deductive reasoning* is represented by the initial application of the model, when data are derived from the physical basis known at the time, but these data are no longer measured; they are replaced with data from the model.

In the following sections, the model development, the structure and the mathematical formulation are explained, the the measurement program required for the initial setup is presented. Subsequently, the constraint equations of the individual model input parameters are discussed, then the engine-specific calibration is presented along with the reduced measurement program required in that regard, and the model data are validated with measurement data. The results obtained are discussed under the aspect of how well the model can be used for the virtual calibration and online adjustment of spark-ignition

engines so that ultimately, a possible operational strategy for implementing LP EGR during dynamic engine operation can be proposed on the basis of the model.

5.1 Model development and formulation

The model developed in this thesis is subject to key requirements. These include short computing times and a correspondingly simple structure. In addition, the model input parameters all have to be known during engine operation, i.e. they have to be measurable by the standard installed sensors. Significant effort has been made to ensure that the model has a physical basis that can be interpreted and after all comprehended. This clearly distinguishes it from models based on neural networks, such as the one presented by Siokos in [86].

Even though neural networks are now frequently used in applications and in modeling on engine control units, a major disadvantage of that method is the lack of information about how the model behaves and works. Instead, the black box approach is applied where input variables are provided to the model and output values are generated. Even if a well calibrated model can indeed work properly using this approach, it is not appropriate to create a simple, understandable model. Rather, a mathematically derivable approach is required. To identify this type of approach, it is first necessary to break down the data or information required for operation with EGR. This involves:

- $\eta_{ind} = f(x_{EGR}, \text{operating point})$
- $COV_{imep} = f(x_{EGR}, \text{operating point})$

Both the indicated efficiency η_{ind} and the coefficient of variance of the mean indicated pressure COV_{imep} are needed as a function of the external EGR rate x_{EGR} and the operating point's characteristics. The former information serves to ensure that operation is as efficient as possible, the latter is an essential condition for stable and misfire-free engine operation. The definition of the variables raises two key questions:

- How can the characteristics of the two values be described by a simple mathematical approach?
- What defines the operating point and how is it determined?

A simple but distinct mathematical approach from elementary algebra that serves as a possible answer to the first question is given by a polynomial of degree k. In its general formulation and depending solely on the external EGR rate, the equation is as follows:

$$f(x_{EGR}) = \sum_{i=0}^{k} n_i x_{EGR}{}^i \tag{5.1}$$

Even though no polynomial approach for modeling the effects of the EGR rate is known so far, this type of modeling is used widely in literature for various other applications. In that regard Winsel [111] investigates the fundamental approximation properties of such approaches. Edwards, though, uses polynomials to optimize fuel efficiency [22], it is utilzed by Lumpp [67], Urano et al. [99] and Gärtner [32] use them to model nitrogen oxides in diesel engine exhaust. Furthermore, Hathaway [39] and Wilhelm [110] employ their polynomial methods to model the mechanical friction in combustion engines and achieve good results.

In order to show that a polynomial approach is suitable as the basis for the model in this thesis, Figure 5.3 first shows an EGR variation for two different operating points. The plotted values are the measured indicated efficiency and the measured COV_{imep} depending on the EGR rate. In addition, a second degree polynomial, which is adapted to the course of the indicated efficiency, is inserted into the diagram. It can be formulated as follows:

$$f(x_{EGR}) = \eta_{ind} = n_2 x_{EGR}{}^2 + n_1 x_{EGR} + n_0 \tag{5.2}$$

Since there is only a dependence on the EGR rate, Equation 5.2 can be used in this case. All other adjustable engine parameters are kept constant while the EGR rate is varied. The left-hand diagram depicts an operating point at $3000\,rpm$ and $8\,bar\,imep$, the right-hand diagram shows an operating point at $2000\,rpm$ and $2\,bar\,imep$. The exhaust and inlet valve timing and the $50\,\%$ mass fraction burned are constant throughout.
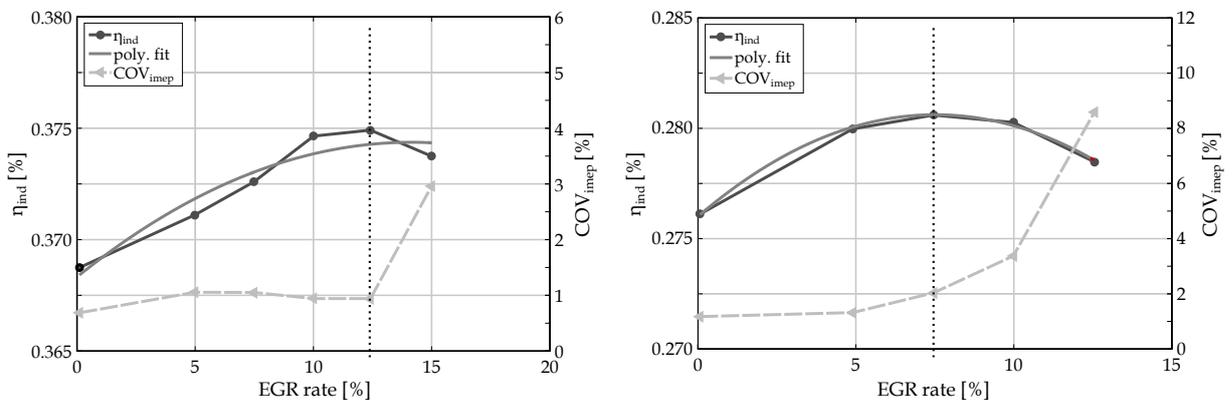
Figure 5.3: Measurement data of indicated efficiency and COV as well as polynomial fit.

Even if these are only example curves, it is already obvious how well the polynomial can be used to reproduce the course of the measured data. The maximum of the measured data course is particularly well mapped by the polynomial. This point is not only of major relevance with regard to the maximum efficiency, but also needs to be mentioned with regard to the cyclical fluctuations. In both cases, the curve of the COV_{imep} rises sharply up above the EGR rate, beyond which the level of the indicated efficiency drops. On the basis of the two examples, there are hence two theses to be noted:

- The course of the indicated efficiency as a function of the EGR rate at constant operating conditions throughout the variation can be reproduced with sufficient accuracy by a second degree polynomial.
- The COV_{imep} increases strongly at the same EGR rate at which the indicated efficiency decreases. The reason for this is considered to be the disproportionately increasing losses due to incomplete combustion and the real burn rate.

It has to be proven whether these theses are not only valid for the examples but also generally. For this purpose, a broad basis of measurement data is used. About 3500 EGR rate variations from a wide range of operating areas are evaluated. Analogous to those shown in Figure 5.3, the EGR rate is varied under constant boundary conditions until the COV_{imep} takes on inadmissibly high values ($> 3\,\%$ in

analogy to Chapter 4). The evaluation criterion is the quality with which the polynomial depicts the x and y values of the maximum of the indicated efficiency curve. The x value represents the EGR rate $x_{EGR,max,tol}$ corresponding to the maximum efficiency $\eta_{ind,EGR,max}$, represented by the y value. The comparison consequently focuses on the crucial aspect of the curve. The results are presented using the two diagrams in Figure 5.4. The x-axis of the diagrams shows the value from the measurement, the y-axis the corresponding value obtained by the respective fitted second-degree polynomial.
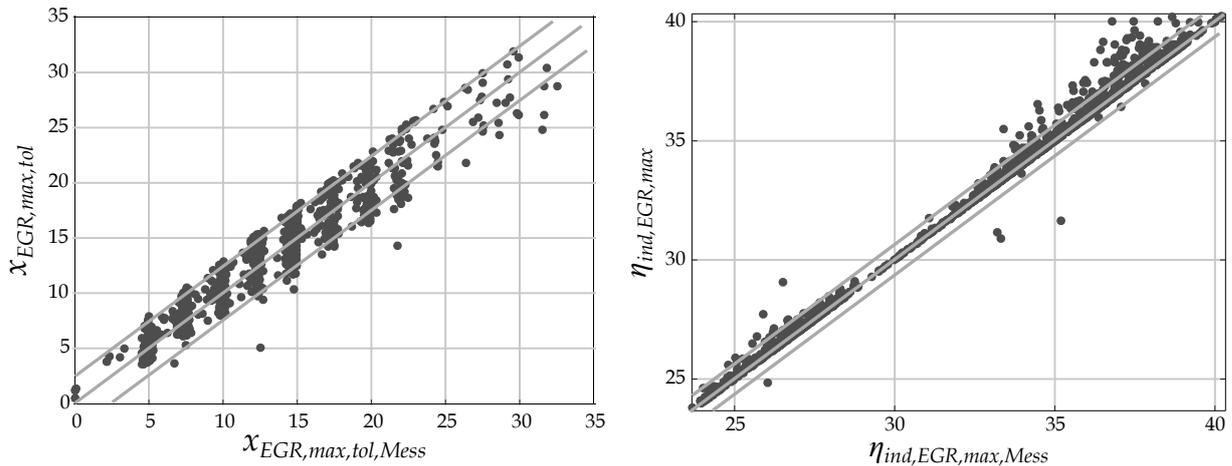
Figure 5.4: Reproduction quality of the polynomial fitted EGR-rate $x_{EGR,max,tol}$ that corresponds to the maximum efficiency within an EGR-variation to the measured value $x_{EGR,max,tol,Mess}$ and of the polynomial fitted maximum efficiency $\eta_{ind,EGR,max}$ within an EGR-variation to the measured value $\eta_{ind,EGR,max,Mess}$.

Regarding the EGR rate at the maximum efficiency, it can be stated that the values from the polynomial fit are virtually all within the scatter band of $+/-2.5\%$ and the projection quality is therefore very good. The tolerance bandwidth is assumed as the EGR variation is carried out in steps of 2.5 %. Greater accuracy can therefore not be achieved. The fact that there are fewer points outside the set tolerance indicates that the values attained by measurement are larger than those found through the polynomial fit. With regard to stable engine operation, this is the conservative and reliable side, since the maximum permissible COV_{imep} would not be exceeded using a polynomial fit.

Similarly to the x value of the maximum efficiency, the y value, i.e. the maximum indicated efficiency itself, is reproduced well by the polynomial fit. The scatter band used for this purpose is narrowed to $+/-1\%$. In the comparison between measured values and those found via the polynomial fit, almost all the values lie within the tolerance band around the bisector in the diagram. Some values from the measurement are slightly overestimated by the polynomial fit, but the overall mapping quality is nevertheless still very good.

It remains to be verified whether the efficiency maximum coincides with the increase in the cyclical combustion fluctuations quantified by COV_{imep} values above the allowed limit. Again, the 3500 EGR variations mentioned above are evaluated in this respect. For more than 99.3 % of these series of measurements the claim is valid. This value corresponds to the numbers found in the work by Siokos [86], where the assumption even applies to 99.5 % of the EGR variations for a series of measurements evaluated in a similar way. Instead of the evaluation on the basis of COV_{imep}, Siokos uses the fuel fraction burned as the limit value for his evaluation, which is comparable to the approach used here.

However, this quantity cannot be determined online at the test bench, but only from the post-evaluation of the measured data. The method used in this study is therefore easier to implement.

As the two claims above can be considered as substantiated, another crucial advantage of the polynomial approach, especially when second-degree polynomials are used, should be emphasized. Each of the polynomial parameters can be interpreted in terms of its influence on the course of the polynomial curve. Figure 5.5 below shows how a change in the polynomial coefficients n_k with $k = \{0,1,2\}$ affects the polynomial course as formulated in Equation 5.2. For clarity, the individual influences are additionally captured as descriptions in Table 5.1. A column that contains the meaning of the mathematical influence on the behaviour of indicated efficiency over the EGR rate is also included.

Table 5.1: Impact of polynomial coefficients on the curve and their meaning in terms of EGR variation.

polynomial coefficient	influence on the course of the curve	meaning regarding EGR efficiency curve
n_0	y-axis intercept	indicator of efficiency level regardless of the EGR rate
n_1	rotation around the vertex	influence on the EGR sensitivity and the EGR tolerance but no definite physical meaning
n_2	curvature of the polynomial curve	indicator of the EGR tolerance and sensitivity of the efficiency on EGR

While the n_0 parameter only shifts the curve along the y-axis, i.e. at the level of the indicated efficiency throughout the entire EGR variation range, the n_2 coefficient is crucial for the curvature of the polynomial and, transfered to the efficiency curve, thus for the EGR rate of the maximum efficiency or the EGR tolerance. The n_1 parameter rotates the curve around its vertex. There is no unambiguous interpretation in terms of the efficiency curve. Considered separately, the n_1 parameter does not directly influence the position of the efficiency maximum in x and y directions, but it does have an indirect impact on it. To understand this, the formulation from Equation 5.2 must be transformed by simple algebraic conversion into the so-called point vertex shape shown below:

$$\eta_{ind} = n_2 \left(x_{EGR} + \frac{n_1}{2n_2} \right)^2 - \frac{n_1{}^2}{4n_2} + n_0 \tag{5.3}$$

The diagram in the lower right-hand corner of Figure 5.5 once again highlights the key point, i. e. the maximum of the EGR efficiency curve. To be precise, this can easily be read off or calculated from the point-vertex shape of the second-degree polynomial derived above, based on the polynomial coefficients. The x-value of the maximum is described by the second summand in squared brackets:

$$x_{EGR,\eta_{ind,opt}} = \frac{n_1}{2n_2} \tag{5.4}$$

The y value of the maximum, i.e. the maximum efficiency, on the other hand, is characterized by the last two summands in Equation 5.3;

$$\eta_{ind,opt} = -\frac{n_1{}^2}{4n_2} + n_0 \tag{5.5}$$

Each operating point serving as the starting point of an EGR variation is characterized by a unique combination of properties and parameters. If the course of the indicated efficiency depending on the EGR rate is fitted by a polynomial, this combination of operating point parameters can be assigned to the resulting polynomial coefficients. It is now known that these, in turn, allow conclusions to be drawn about the shape of the curve and contain information about the maximum efficiency achievable with external EGR. The combination of operating point parameters and polynomial coefficients is a key element of the modeling approach taken in this study. The following summary and conclusions can be made:

- By assigning an operating point to the polynomial coefficients resulting from the fitting of the EGR variation's efficiency curve, influences of the operating point parameters on the polynomial coefficients and thus on the behavior of the EGR efficiency behavior can be derived.
- If a relationship is found that allows individual operational parameters to be used to draw conclusions about polynomial coefficients, it is possible to form a polynomial from the set of parameters of an operating point that simulates or models and predicts its EGR efficiency curve.
- Since it has already been proven that the EGR efficiency curve also contains information on the maximum EGR tolerance, this prediction or modeling can be used for the application and conceptual specification of combustion engines.

The mathematical approach has been specified, but the second question asked at the start of this section, about the operating point, and the parameters used to describe it clearly and unambiguously, remains unanswered. Particular emphasis is placed on using sensor values and variables that are readily available in conventional production engines. In addition, it is essential that the parameter set of the operating point is clearly defined and that all values relevant for operation with LP EGR are taken into account. Thus, a verifiable, physical basis for the model can be established. This is of great importance when the model is transferred and applied to other engines.

For the description of an operating point, the engine load (*imep*) and the engine speed (ω) are of crucial significance. Both parameters are known during engine operation and consequently contributed as input variables. In the framework of this thesis, the engine load is described by the mean indicated pressure, but during real driving operation the relative filling or the engine torque is usually applied. However, the quantities can easily be converted back and forth. The influence of load and speed on the EGR efficiency and the EGR tolerance behaviour is described throughout various literature sources, such as Nitschke [74] or Chao et. al. [14].

The relationship between the internal and external EGR rate is equally important for the characterization and behavior of an operating point subjected to an EGR variation [74]. The model takes this ratio into account using the exhaust cam timing (*ect*), the valve overlapping area (*voa*), and the relative pressure in the intake manifold (p_{IM}). Collectively, these parameters clearly specify both the amount and the sort of the internal residual gas (see Figure 4.8).
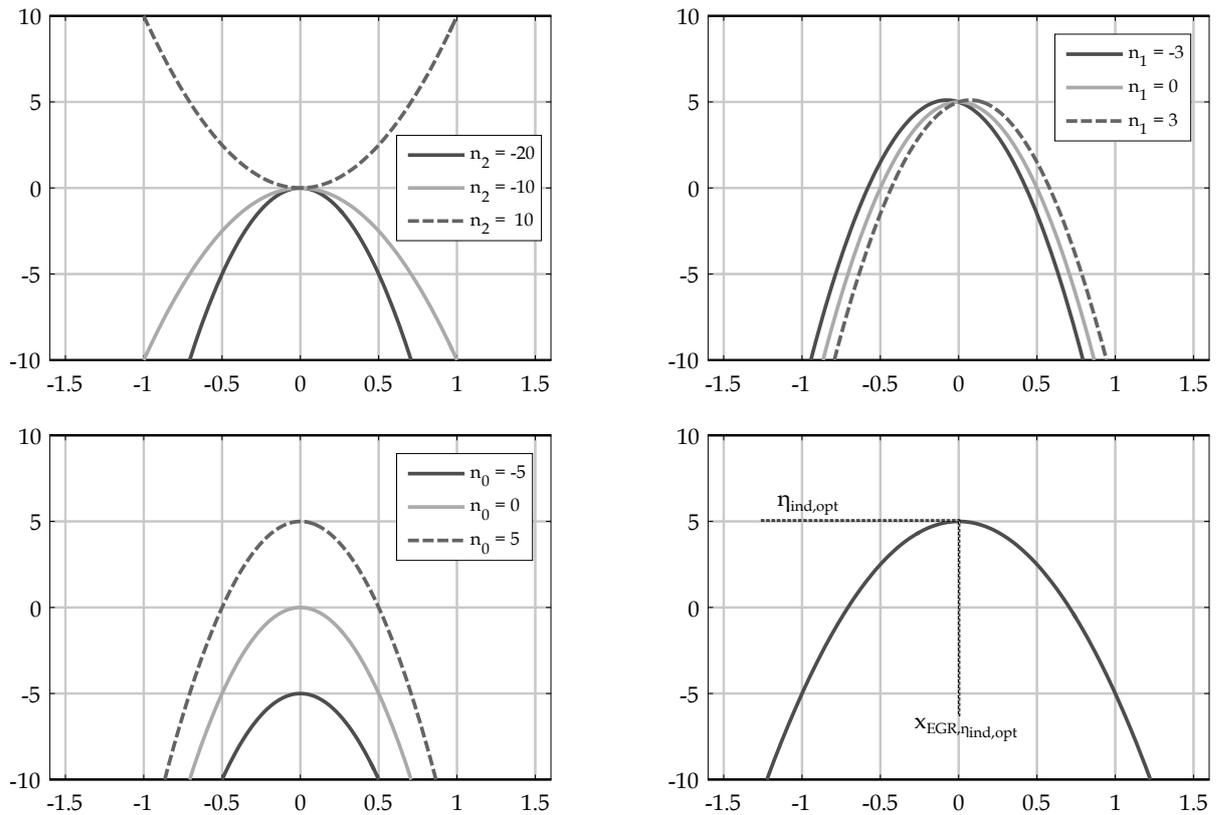
Figure 5.5: Influence and forms of interpretation of the coefficients of a second degree polynomial.

In their work Lauer et. al. [61] as well as Bunsen [11] depict the importance of the ignition timing (*it*) regarding the dilution tolerance and the efficiency. The ignition timing determines the thermodynamic boundary conditions in the first phase of the combustion, which in turn are crucial for the laminar burning rate. This is directly incorporated into the overall burning rate. Its slowdown leads to increased cyclical combustion fluctuations and causes the engine process to deviate further from the ideal equal space Otto Cycle, thus reducing the achievalbe efficiency. Finally, the portfolio of input variables also implies the fuel air ratio (λ). The research carried out by Fricke [29] and Heikes [40] proves that the fuel air ratio influences the burning rate in a similar way as the ignition timing does with the above mentioned consequences.

In total, this results in seven parameters which are used to describe the operating point explicitly and which function as model input variables. For the sake of clarity and for the assignment of the numbering, the parameters, a_j with $j = \{1,2,3,4,5,6,7\}$, are re-listed below:

- exhaust cam timing ($a_1 = ect$)
- valve overlapping area ($a_2 = voa$)
- engine speed ($a_3 = \omega$)
- engine load ($a_4 = imep$)

- ignition timing ($a_5 = it$)
- fuel-air ratio ($a_6 = \lambda$)
- pressure intake manifold ($a_7 = p_{IM}$)

The model output consists of the coefficients of a second degree polynomial: n_0, n_1 and n_2. Each of the seven model input variables listed features an equation for each of the three coefficients according to following formula:

$$n_{k,a_j} = f_k(a_j) \tag{5.6}$$

Assuming that all other boundary conditions are kept constant, this equation allows the respective input parameter to influence the polynomial coefficients n_k for $k = \{0,1,2\}$. All equations for the seven model input variables resulting from Term 5.6 can be summarized in Matrix A as follows:

$$A = \begin{pmatrix} f_2(a_1) & \dots & f_2(a_7) \\ \vdots & \ddots & \vdots \\ f_0(a_1) & \dots & f_0(a_7) \end{pmatrix} \tag{5.7}$$

The individual equations are defined and established on the basis of empirical measurement data. unlike from the term they are supposed to model later, they may, but do not have to be described by a second-degree polynomial. The equation type applied is allowed to differ from one input variable to another. It is rather essential that both the error between the $f_k(a_j)$ model term and the measured data is kept to a minimum and that there is a physically sound and explainable or transparent correlation regarding the course of the polynomial coefficient depending on the particular input variable. Especially for the latter of the two criteria, information can be found in the literature. One method that can fulfill both of these requirements is regression analysis. It is conducted based on the measurement data, whose derivation and extraction are initially described in greater detail in Section 5.2.

Regression analysis is a statistical procedure for modeling relationships between different variables: dependent ones, which are the polynomial coefficients n_k and independent ones, which are the model input parameters a_j. It is generally used to describe and analyze relationships in data. On the other hand, regression analysis can also be used to make predictions, which is particularly relevant for this thesis and generally so for technical appliactions [75].

The errors that occur for each model input variable as well as a more extensive description of the procedure are presented in the sections 5.3.1 through 5.3.7 along with the correlations or mapping equation used. The choice of the mapping equation, e.g. polynomial or exponential function, is based on the above criteria of a minimal error and a physically meaningful dependency. In this section, however, the basic model structure is discussed in further detail.

The matrix described in Equation 5.7, together with the equations it contains, represents the physical layer of the model. It can be evaluated for each engine operating point (*OP*) by inserting its parameter set a_j into the correlations contained in A. The resulting Matrix A_{OP} is specific to the operating point and formulated as follows:

$$A_{OP} = \begin{pmatrix} n_{2,a_1} & \cdots & n_{2,a_7} \\ \vdots & \ddots & \vdots \\ n_{0,a_1} & \cdots & n_{0,a_7} \end{pmatrix} \tag{5.8}$$

The entries in the matrix are no longer dependencies or functions but specific values for a particular operating point. If the model's applicability to only one motor had to be proven, it would be sufficient to use j-weighted values for the 3 polynomial coefficients and define this weighting once. However, for the model to be applied to other motors, a second matrix, Matrix B, is introduced. It contains a motor-specific parameterization and is defined as follows:

$$B = \begin{pmatrix} \beta_{2,a_1} & \cdots & \beta_{0,a_1} \\ \vdots & \ddots & \vdots \\ \beta_{2,a_7} & \cdots & \beta_{0,a_7} \end{pmatrix} \tag{5.9}$$

The size of Matrix B is (jxk), which is $(7x3)$ in this case, and is determined only once per engine. Thus, unlike Matrix A, it is only dependent on the engine, but not on the operating point. To specify B, a reduced measurement program, introduced later in Section 5.2, is used. When B is implemented, engine-specific properties such as the compression ratio and charge movement can be taken into account. B represents a motor-specific reweighting of the particular input parameters on the EGR efficiency curve.

To take into account the weighting factors from B in the constraint equations from A, both matrices are scalar-multiplied. The resulting Matrix C is formulated below:

$$C = \begin{pmatrix} \sum\limits_{j=1}^{7} \beta_{2,a_j} \cdot n_{2,a_j} & \cdots & \sum\limits_{j=1}^{7} \beta_{0,a_j} \cdot n_{2,a_j} \\ \vdots & \ddots & \vdots \\ \sum\limits_{j=1}^{7} \beta_{2,a_j} \cdot n_{0,a_j} & \cdots & \sum\limits_{j=1}^{7} \beta_{0,a_j} \cdot n_{0,a_j} \end{pmatrix} \tag{5.10}$$

The size of Matrix C is (kxk), which in this case equals $(3x3)$. Since the second-degree polynomial is uniquely defined by the three coefficients n_2, n_1 and n_0, a cross-effect of the weighting factors and polynomial coefficients is excluded. Accordingly, the assumption of linear independence can be applied. The scalar product is simplified so that only the diagonal entries of Matrix C have to be considered. This results in:

$$\vec{N} = \begin{pmatrix} n_{2,OP} \\ n_{1,OP} \\ n_{0,OP} \end{pmatrix} = \begin{pmatrix} \sum\limits_{j=1}^{7} \beta_{2,a_j} \cdot n_{2,a_j} \\ \sum\limits_{j=1}^{7} \beta_{1,a_j} \cdot n_{1,a_j} \\ \sum\limits_{j=1}^{7} \beta_{0,a_j} \cdot n_{0,a_j} \end{pmatrix} \tag{5.11}$$

The vector $\vec{N}$ contains the polynomial coefficients of the operating point OP ($n_{2,OP}, n_{1,OP}, n_{0,OP}$) and can be calculated by summing the polynomial coefficients n_{k,a_j} belonging to the model input variables a_j multiplied by the corresponding weighting factors β_{k,a_j}.

Finally, for any given operating point, the second-degree polynomial, which maps the efficiency behavior when adding external EGR and the associated EGR tolerance, can be calculated. The calculation is solely based on the knowledge of the model input variables a_j. No prior knowledge or measurement is necessary other than the model calibration regarding the specific engine.

Below, the measuring program for setting up the two matrices A and B is defined. Section 5.3 then goes on to discuss the influence of the individual model input variables and the derivation of the respective influence equations by means of nonlinear regression analysis.

5.2 Measuring program for initial model setup and calibration

In the following, two different measuring programs are presented. While one is designed for deriving the influence equations in Matrix A to understand and deduce important physical relationships, the other is designed for the engine-specific calibration of the model and thus for deriving Matrix B. The effort is proportionate. The measurement program for Matrix A is extensive and covers all relevant engine map ranges. In contrast, the measuring program for Matrix B is reduced to the minimum required.

In order to understand the following explanations clearly, it is important to know that for each combination of model input variables a_j according to the system explained below, an EGR variation is carried out until the EGR tolerance is exceeded. In that regard, the first two variation steps are applied at EGR rates of 0 and 5 %. Starting from there, the EGR rate is further increased in steps of 2.5 % to ensure that the limit of the EGR tolernace is covered with sufficient accuracy. Based on the EGR variation, the polynomial can be fitted and the associated coefficients can be assigned to the respective combination of model input variables a_j that is defined by the starting operation point of each EGR variation.

The development of the measurement program for Matrix A follows three steps using different methods for the design of experiments (DoE). In the third step, an additional distinction is made as to whether the operating point is turbocharged or naturally aspirated.

Conventional methods of experimental design involve considerable effort in terms of both the execution and evaluation of the experiments. Both in practice and in this specific case, the question posed in the study is often limited, i.e. the task is to reduce the effort required. This is achieved by employing statistical experimental designs or DoE. The time and effort required from experimental

design to execution of the experiment and evaluation can be significantly reduced [83] by that. The following is a list of the steps or methods used:

    a. full factorial design with asymmetrical distribution

    b. fractional factorial design with central point

    c.1 full factorial design (*ect* and *voa*) - naturally aspirated operation

    c.2 full factorial design (*ect* and *ict*) - turbocharged operation

a. Full factorial design with asymmetrical distribution

At the outset of the thesis, in Figure 2.1, the relevant engine map areas are illustrated based on the qualitative EGR rate. These are taken into consideration when developing the measuring program for Matrix *A*. The measuring points that are shown subsequently in Figure 5.6, and which are distributed along the engine load and speed, are mainly derived from these areas. However, the areas that prove to be particularly important for dynamic operation when analyzing the worst-case scenario (cf. Figure 2.10) are taken into account and also included.
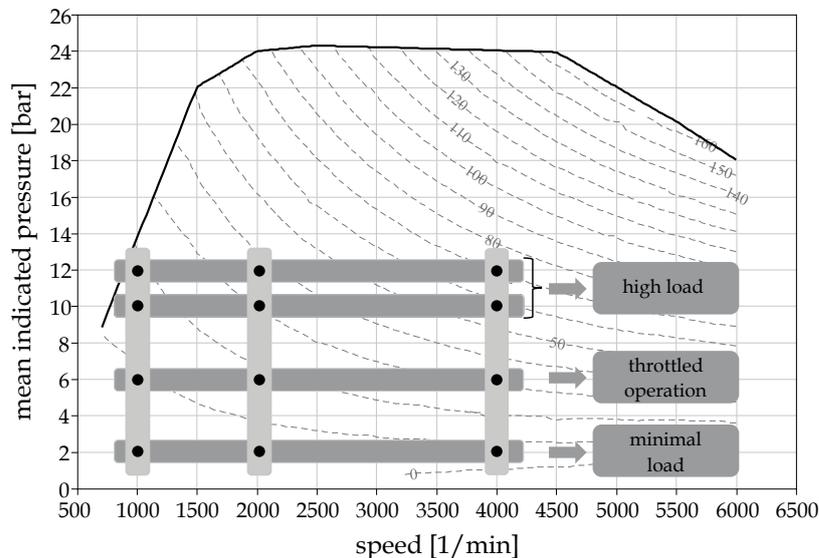
Figure 5.6: Measured load points for data acquisition to set up matrix *A*.

In total, the program comprises 16 engine speed and load combinations, which are subdivided into three speeds (@ 1000, 2000 and 4000 *rpm*) and four load variations (@ 2, 6, 10 and 12 *bar*). This covers the relevant speed and load range very accurately. Special emphasis is placed on the upper load area of the EGR operation by means of a narrow grid. Here, particularly high advantages can be achieved in terms of efficiency, and correspondingly high EGR rates can be obtained. Hence a narrow grid is crucial. Of course, the medium and minimum load ranges are also covered. As with the load, the measured operating points are asymmetrically distributed in terms of the engine speed, the main focus being on low speeds. This pays tribute to their critical nature during dynamic operation due to the lower air flows in the area of low engine speeds.

b. Fractional factorial design with central point

For each of the operating points depicted in the engine map, the influencing parameters fuel air ratio λ, pressure in the intake manifold p_{IM} as well as ignition timing *it* are varied. The latter value, though, is not varied directly, but rather via the detour of the 50 % mass fraction burned $mfb50$, which is kept constant throughout the EGR variation. In doing so, the $mfb50$ is specified and the ignition angle is adjusted such that $mfb50$ remains unchanged. This is necessary since the $mfb50$ is not known during vehicle operation. To achieve a stable measurement of an EGR variation, however, a constant $mfb50$ that is not shifted towards late values is indispensable. Otherwise the COV_{imep} would unintentionally be falsified and thus invalidate the entire model concept. The ranges of variation in the parameters applied are listed in Table 5.2.

Table 5.2: Varied influencing parameters and corresponding intervals.

fuel air ratio λ	$\{0.9, 1.0, 1.1\}$ [-]
pressure intake manifold p_{im}	$\{50, 100, 150\}$ [*mbar neg. pressure*]
50 % mass fraction burned $mfb50$ (ignition timing *it*)	$\{8, 12, 16\}$ [*CADaTDC*]

All values contained in the table are varied for all load points, with only one parameter being varied at a time. The standard configuration or central point is represented by $\lambda = 1$, $p_{IM} = 100\,mbar$ and $mfb50 = 8\,CADaTDC$. Thus, for example, if λ varies within a load point, i.e. $\lambda = 0.9$ or $\lambda = 1.1$, only the default values of p_{IM} and $mfb50$ are allowed. Figure 5.7 provides a three-dimensional schema to illustrate the principle graphically for better understanding. The central point is black, while the variation points are gray.
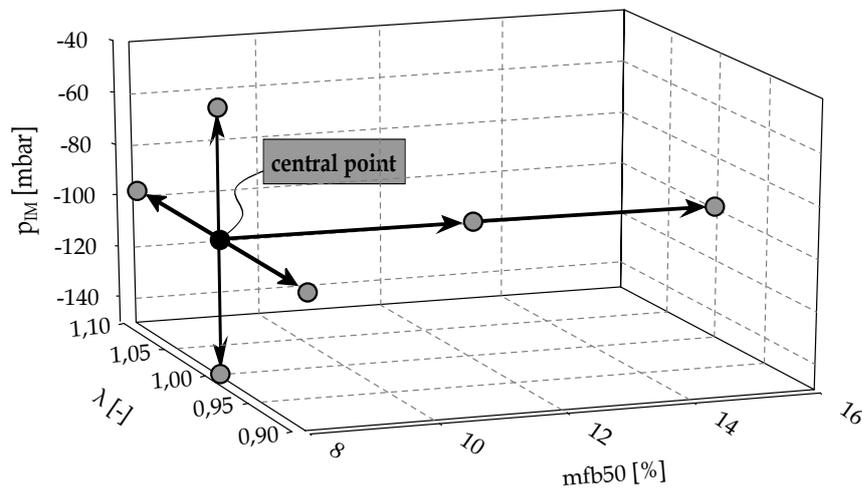
Figure 5.7: Fractional factorial design with central point.

Note that for turbocharged operation (@ 10 and 12 *bar*) the pressure intake manifold loses its significance or purpose and can no longer be varied as the load cannot be achieved otherwise.

c. Full factorial design for naturally aspirated and turbocharged operation

Due to the particular relevance of the internal EGR rate and the ratio of the internal and external EGR rates, the parameters of the exhaust cam timing *ect* and valve overlapping area *voa* or intake cam timing *ict*, respectively, are handled differently. Instead of a unique and isolated variation, independently of the other influencing parameters, these are varied fully factorially for each combination of *imep*, ω, λ, p_{IM} and *mfb*50. Before the details are elaborated upon, Figure 5.8 and the depicted valve lift curves show the definition of the quantities used. While the *ect* and *ict* are measured in [*CAD*], *voa* denotes the integral or area in the range of valve overlap and is given in [$mm^2 CAD$]. The limits for the integral calculation relate to a valve lift of 0.1 *mm*.
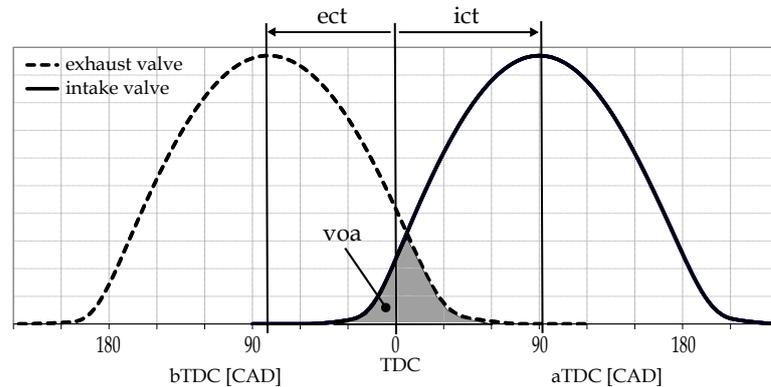
Figure 5.8: Illustration of the exhaust cam timing *ect*, the intake cam timing *ict* and the valve overlapping area *voa*.

By varying the parameters of the valve train, both the internal residual gas rates and the type of residual gas are altered as shown in Figure 4.8. In this context, the size and position of the valve overlapping area are crucial. Due to the engine's special feature of the variable valve lift, which also comes to bear with the EGR variation, the experimental design needs to differentiate between charged and non-charged operation. During partial-load or uncharged operation, as explained in Table 3.1, the load control is implemented via an interaction between the throttle, the intake valve lift and the turbocharger wastegate. Due to the dethrottling when increasing the EGR rate the valve lift, inter alia, must be increased to maintain a constant engine load during EGR variation. However, this conflicts with the requirement to keep the critical boundary conditions as well as the model input parameters unchanged while the external EGR rate alters. Increasing the valve lift means increasing the valve overlapping area, which in turn has a direct influence on the internal EGR rate. A differentiated evaluation of the influence of the external EGR rate is consequently no longer easily possible.

For this reason, during partial-load operation, i.e. when the load is still controlled by the valve lift and there is no full valve stroke yet, it is not satisfactory to simply set a fixed intake and exhaust cam timing. Though, it is necessary to specify the exhaust cam timing and the valve overlapping area while keeping them steady throughout the EGR variation. Both, the type of internal residual gas, which is defined by the position of the valve overlapping area and thus by the exhaust cam timing, and the quantity, which is mainly dictated by the size of the valve overlapping area and thus by itself, can be kept unchanged across the entire EGR variation range in this way. While the exhaust cam timing can be set to a constant value, the valve overlapping area has to be defined by a correlation between the valve intake cam timing *ict* and the valve lift *vl*, i.e. *ict* = $f(vl)$. For *ect = const.* and *voa = const.*, these

dependency curves can be described analytically. They are specified within the test bench software during the EGR variation so that increasing values of the *vl* cause increasing values of the *ict* and *voa* remains constant.

Finally, the question arises of which ranges of values are reasonable with regard to the valve overlapping area. An evaluation of stationary engine maps, both with and without external EGR, indicates values in the range between 0 and maximum $140 \, mm^2 \, CAD$. In theory, however, values up to $300 \, mm^2 \, CAD$ can be achieved with the given valve train. However, this only applies to the maximum valve lift, which is not relevant for the two load stages considered in the partial load range (@$2, 6 \, bar \, imep$). At a theoretical EGR rate of 30% and thus the maximum dethrottling or valve lift that can occur in the low load area, a value of $220 \, mm^2 \, CAD$ is achieved. Accordingly, the discrete values listed in Table 5.3 are used for *voa*. The narrower grid for smaller values is chosen because smaller values mean lower internal EGR rates and therefore higher external EGR rates are possible. This in turn increases the sensitivity of the polynomial coefficients. It is important to note that not all of the valve overlapping values can be applied for each exhaust cam timing. Since, among other things, the dependency $voa = f(ect)$ exists, it is geometrically impossible to realize large values of *voa* for large exhaust cam timings. For example, with $ect = 120 \, CAD \, bTDC$ only a maximum of $voa = 5 \, mm^2 \, CAD$ can be realized. Which *voa* values are varied for the individual *ect* values is shown on the left-hand side of Figure 5.9 by means of the points in the characteristic diagram.
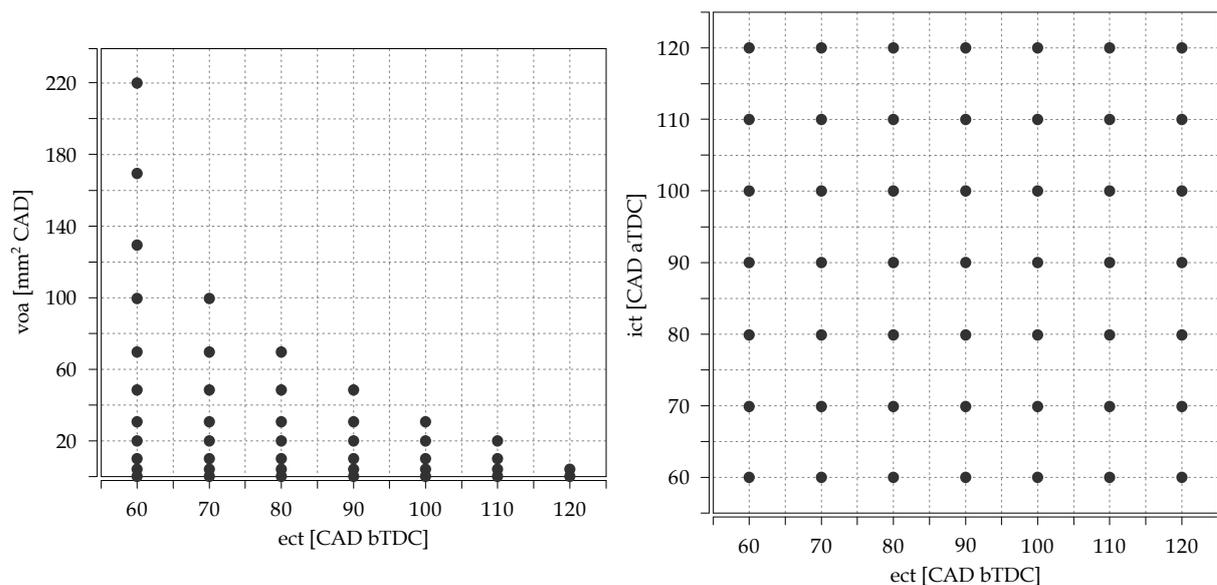
Figure 5.9: Distribution of the measuring points in the partial load and in the high load area.

The values listed in the table for the *ect* apply to both the partial load range just described and the two load stages at 10 and $12 \, bar \, imep$. However, as a full valve stroke is already present, a simpler strategy can be adopted and the intake and exhaust cam timing can be adjusted independently of each other using fully factorial steps. The values of the inlet valve timing, used for the measurement points of the two upper load levels, are also listed in Table 5.3 and depicted in the diagram on the right-hand side of Figure 5.9.

Table 5.3: Cam phasing variables including their variation range.

exhaust cam timing *ect*	$\{60, 70, 80, 90, 100, 110, 120\}$ $[CAD\,bTDC]$
intake cam timing *ict*	$\{60, 70, 80, 90, 100, 110, 120\}$ $[CAD\,aTDC]$
valve overlapping area *voa*	$\{0, 5, 10, 20, 30, 50, 70, 100, 130, 170, 220\}$ $[mm^2\,CAD]$

This results in a total of about 3600 operating points. For each of them, the EGR rate is varied according to the system described above until the maximum COV_{imep} or EGR tolerance is reached. There are, however, operating points that cannot be implemented even without external EGR due to excessive cyclical fluctuations. When designing the measuring program, it is not possible to predict which operating points will be affected. No approximation polynomial is calculated for these and they are excluded in the following model setup.

For each of the measured operating and EGR variation points, a TPA including a complete loss analysis is conducted subsequently to the measurement. Although this involves a great deal of effort, it is at the same time indispensable for understanding the impact mechanisms and for interpreting the results as well as the model equations later on.

To determine Calibration Matrix *B* two things are essential. The first is the measuring program, as with Matrix *A*, and the second is an understanding of how Matrix *B* is derived on the basis of the measuring program. Both issues are described in the following two subsections.

Measuring program for Matrix *B*

The sequence as well as the basic structure of the measurement procedure for the engine-specific Matrix *B* are similar to that for Matrix *A*, though including modifications and reductions to step *a* as well as reductions to step *b*. While step *a* now implies a fractional factorial design with a central point, instead of the full factorial design, step *b* is only conducted for the central point of step *a*. The other four variation points are direclty subjected to step *c*. This step remains unchanged throughout the full range of its variation, as with Matrix *A*. However, the effort is considerably reduced. Figure 5.10 illustrates the three steps *a* to *c* for Matrix *B* including the considered load points. The arrows illustrate how three load points are treated directly in step *c*. Only the central point is first treated in step *b* and afterwards also in step *c*.

All in all, this results in a total of approx. 380 measuring points, again with the proviso that not all of these points can be measured during operation without external EGR due to high cycle-to-cycle variations because of the apparent internal residual gas. Conversely, this means that the mere measuring effort for Matrix *B* is only about 10 % of that for Matrix *A*.

Derivation of Matrix *B*

Unlike Matrix *A*, *B* does not contain influence equations but instead scalars. They are calculated based on a multiple step iteration. The basic idea is to use the data available from *A* for each measured EGR variation in each of the approx. 380 operating points to compute the polynomial using homogeneous weighting. Consequently, for each operating point, a curve of the indicated efficiency over the EGR rate from the measurement and one from the unadjusted simulation model is available. Concerning Matrix *B*, this means that all entries are assumed to be equal, meaning 1/7, at first. In doing so, there is
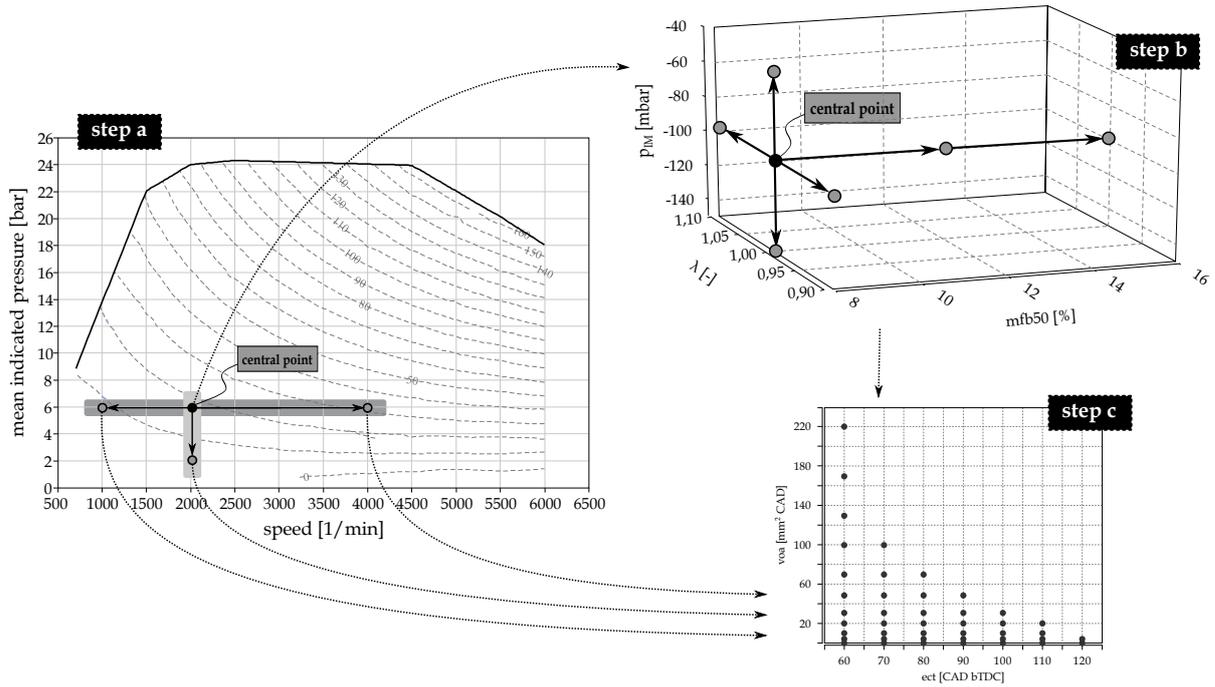
Figure 5.10: Measuring steps for Matrix *B*.

likely to be some deviation between the approximated polynomial efficiency curve derived from *A* and the measured data to a greater or lesser extent, since some of the model input variables have a greater influence on the shape of the efficiency curve than others including engine specifics.

Though, the deviation obtained by the equal weighting needs to be minimized for each polynomial. The method applied in this context and described below is firstly related to the individual measured data as well as its fitted polynomial curve. It is subsequently further specified in terms of the general setup of Matrix *B*. For optimization, the point vertex shape from Equation 5.3 is used. To check how well a modelled polynomial maps the course of the measured data, the vertex or the extreme point of the course is used as a criterion, i.e. $\eta_{ind,opt}$ according to Equation 5.5 and the associated EGR rate $x_{EGR,\eta_{ind,opt}}$ according to Equation 5.4. The polynomial parameters are manipulated or factorized for each measured data progression such that the vertex of the course from the model matches the one from the measurement as closely as possible. However, when adjusting $\eta_{ind,opt}$ it is only allowed to adjust the multiplicant including n_1 in Equation 5.5 and for the adjustment of $x_{EGR,\eta_{ind,opt}}$ only the multiplicant with n_2 may be adjusted in Equation 5.4. The strict separation is necessary to exclude cross effects and above all to let the optimization converge faster.

The optimizer *fminsearch* from the comercial software *Matlab* is employed as an optimization tool. It generally enables the minimum of unconstrained multivariable functions to be found [69]. The first input equation including $\eta_{ind,opt}$ can be formulated for $j = \{1,2,3,4,5,6,7\}$ as follows:

$$(\eta_{ind,opt})_{meas} - \frac{4n_{2,h}n_{0,h} - \boldsymbol{\beta_{1,a_j}} \cdot n_{1,h}{}^2}{4n_{2,h}} = f_{\beta_1} \tag{5.12}$$

The scalar β_{1,a_j} has to be matched such a way that f_{β_1} assumes minimal values. The values $n_{2,h}$, $n_{1,h}$ and $n_{0,h}$ used in the equation are the values for the respective operating point derived from the equations of Matrix A by homogeneous weighting. They are treated as fixed values and may not be altered during optimization. By analogy, the optimizer input equation for $x_{EGR,\eta_{ind,opt}}$ can be stated using the following formula:

$$(x_{EGR,\eta_{ind,opt}})_{meas} - \frac{n_{1,h}}{2 \cdot \boldsymbol{\beta_{2,a_j}} \cdot n_{2,h}} = f_{\beta_2} \tag{5.13}$$

Here too, β_{2,a_j} is adjusted so that f_{β_2} takes on minimal values. Finally, the parameter β_{0,a_j} for weighting $n_{0,h}$ has to be found. Only two equations, along with the corresponding measured values, can be derived as criteria from the point vertex form. Thus an additional criterion and input equation for the optimizer has to be identified. For this purpose, the integrated area of the polynomial from the model within the boundaries between an EGR rate of 0 % and the EGR rate at maximum efficiency $x_{EGR,\eta_{ind,opt}}$ is compared to the area under the efficiency curve from the measurement. The latter cannot be integrated in the standard way due to the discrete number of data points, but is instead calculated in simplified terms using the trapezoidal rule. The resulting equation is:

$$(area)_{meas} - (area)_{model} = \frac{(\eta_{ind,x_{EGR,r+1}} - \eta_{ind,x_{EGR,r}}) \cdot (x_{EGR,r+1} - x_{EGR,r})}{2} - \ldots \tag{5.14}$$

$$\ldots \int_{x_{EGR}=0}^{(x_{EGR,\eta_{ind,opt}})_{meas}} (\beta_{2,a_j} \cdot n_{2,h} x_{EGR}^2 + \beta_{1,a_j} \cdot n_{1,h} x_{EGR} + \cdot \boldsymbol{\beta_{0,a_j}} \cdot n_{0,h}) \, dx_{EGR} = f_{\beta_0}$$

The indices r represent the individual measuring points of the EGR variation belonging to the respective operating point. Likewise, in this case, the value of $f_{\beta_{j,0}}$ needs to be minimized by optimizing β_{0,a_j}. The values β_{2,a_j} and β_{1,a_j}, which are also taken into consideration in the equation, are not modified in this optimization step but are taken over from the previous steps.

By comparing the areas under the two curves, the main object seems to be the adjustment of the indicated efficiency's absolute level between simulation and measurement. That is perfectly correct. However, in the first model applications it is found that the model's absolute level of the indicated efficiency does not reach a satisfactory level, meaning that this last optimization step fails to converge in many cases. For this reason the model has to be simplified at this point. It is assumed that the efficiency level of the base operating point, i.e. the operating point without LP EGR, is known. The n_0 polynomial parameter in the model is thus taken for granted. This is a valid and feasible simplification, since this kind of information is usually obtained in a standard engine development phase. The model continues to provide crucial information about the behavior of the engine and of the indicated efficiency when LP EGR is incorporated. The exact consequences of this facilitation will be clarified again in Chapter 5.4.

Equation 5.2, however, and the calculation of the weighting factor β_{0,a_j} are still also needed and used, since they influence not only the absolute level, but also the curve's shape. In connection with the extreme value of the curve adapted within the previous equations, the polynomial from the simulation is thus further adjusted in all points. This is crucial for later application as it contributes to the accurate reproduction of the measured values, and above all, to the prediction of as yet unknown operational conditions.

The abort criterion of the optimization is either convergence or a maximum number of 200 iteration loops, with one loop including the three steps presented. However, if the calculation still does not converge for an operating point, its factors are excluded from further calculations so that the results are not distorted or negatively influenced.

Finally, all 380 operating points of the measuring program for Matrix *B* are optimized in the manner described above. Provided there is convergence, the three parameters β_{2,a_j}, β_{1,a_j} and β_{0,a_j} are this obtained, meaning 380 times 3 β-parameters. Each set of these three paramters can be assigned to a distinct and unique set of model input variables a_j. To obtain all 21 entries of the (7x3)-Marix *B*, an averaging method is necessary. For that purpose, the β-values are clustered according to the measuring program.

The values are then weighted according to the grid points (*gp*, i.e. the nominal values of the input parameters aj specified in the measurement. The grid points are listed in Table 5.2 and 5.3 and furthermore illustrated in Figure 5.10. The averaging is first carried out individually for each grid point, i.e. all values of a grid point are arithmetically averaged. The resulting mean values of the grid points are then averaged again with an equal weighting to determine the final parameter β_{k,a_j}. The procedure is described below in Equation 5.16 :

$$\beta_{k,a_j} = \frac{1}{m} \cdot \sum_{u=1}^{m} \left(\frac{1}{m_{gp}} \cdot \sum_{o=1}^{m_{gp}} \beta_{k,a_j,gp_o} \right)_u = \frac{1}{m} \cdot \sum_{u=1}^{m} \left(\overline{\beta_{k,a_j,gp_o}} \right)_u = \overline{\left(\overline{\beta_{k,a_j,gp_o}} \right)_u} \tag{5.15}$$

In this context, *o* describes the index of the respective operating point, gp_o the grid point belonging to the operating point and m_{gp} the number of all operating points per grid point. The index of a single grid point is defined by *u* and the total number of grid points by *m*.

This procedure ensures that grid points with a higher number of measured values, which are inevitably present due to the fractional factorial desgin with a central point, are not weighted more strongly than those with fewer measured values. Finally, the goal of Matrix *B* is to map not the measuring program, but the respective parameter influence. This is ensured by the described procedure. Furthermore and in the first place, this approach results in different values for the different model input variables as described later in Chapter 5.4. If a simple arithmetic mean were applied across all operating points, an identical β_k would result for all a_j.

One exception to the described method is the ignition angle *it*, because it is determined according to the *mfb*50 specification and therefore has no grid points in the conventional way. The arithmetic mean of the ignition angle *it* is determined from the β_k parameters of all operating points and is not specifically weighted in terms of its grid points. This is feasible since the ignition angle is in any case widely scattered due to the great variation in the remaining model input variables a_j and, as can be clearly seen in the further course in Figure 5.38, has an even and approximately equidistant distribution of individual values between its minimum and maximum value. Making the assumption that the number of operating points per grid point m_{gp} is one for all grid points, it finally follows for the ignition angle that:

$$\beta_{k,a_5} = \frac{1}{m} \cdot \sum_{u=1}^{m} \left(\frac{1}{m_{gp}} \cdot \sum_{o=1}^{m_{gp}} \beta_{k,a_j,gp_o} \right)_u = \frac{1}{m} \cdot \sum_{u=1}^{m} \left(\beta_{k,a_j,gp_o} \right)_u = \overline{\left(\beta_{k,a_j,gp_o} \right)}_u \qquad (5.16)$$

This Chapter is only intended to explain the basic procedure for deriving Matrix *B* and the associated measuring program. The concrete β_k-values and the associated diagrams are presented in Chapter 5.4. There, the *B* matrices resulting for the two analyzed engines are also presented by comparison, and the significance of the *B* matrix for the model's accuracy is explored further.

5.3 Influence of the operational parameters on the polynomial coefficients

While the last chapter explained the basic model structure, the measuring program, the model derivation and the formation of Matrix *B*, this section serves to deal with the actual model equations and with their physical interpretation. Before the individual subchapters are discussed, the structure used in each of them and the method for determining which is the nonlinear regression will be briefly explained.

In a global sense, a regression investigates the relationship between a target value and one or multiple input variables. Within the scope of this thesis, the input variables relate to the model input variables a_j described above. A mathematical relationship between these and the polynomial parameters n_k shall be found. For this purpose the polynomial parameters, which result from the second degree polynomial fit of the individual EGR variation measurements, are evaluated depending on the model input variables and the regression is carried out.

A fundamental distinction is made between linear and nonlinear regression. Unlike standard linear regression, which is limited to the approximation of linear models, nonlinear regression analyzes functions that cannot be described as linear. Frequently, such functions are derived from theory or concern well-known mathematical model approaches. There are virtually unlimited possibilities to apply the deterministic part of the model function, and nonlinear regression allows models to be estimated with arbitrary relationships between the dependent and independent variables [80].

In the context of this thesis, the software *Matlab* or its fitting toolbox is used to compute the nonlinear regression. A number of mathematical model functions are available in the software to generate a mathematical relationship or a mathematical formulation from the measured data depending on the model inputs a_j The respective model function must basically be selected anew and freely for each input variable and for each relationship examined. However, as can be seen in Sections 5.3.1 to 5.3.7, there are models or equations which are not suitable and consequently others that prove to be favorite and most suitable types.

In the standard statistics literature, such as [27] or [91], the following three framework criteria are often chosen to evaluate the goodness or quality of a model:

- The measurement data could also originate from the model. This is based on the assumption of least-square fitting.
- The model coefficients can be assessed with a high degree of reliability.

- The model describes much of the data variance and can predict new and deviating data feasibly and with high quality.

These criteria are still formulated very generally and are not yet fixed. For this reason, the actual criteria for selecting the model equation in this thesis are listed below:

- The graphical evaluation of the model data: The measurement data and the associated model course are plotted together in a diagram, meaning that the deviation of a large number of individual points can be viewed easily and quickly. The disadvantage of this method is that it does not offer objective comparability.
- The numerical evaluation of the model data: For this purpose, a number of characteristic values are calculated, which will be described later in this section. If the characteristic values are used and interpreted correctly, they represent an objective possible means of comparison. The disadvantage, however, is that all data are summarized in one single value.
- The physical meaning: The model behavior, like the measurement data, should have a physical basis and be interpreted on that basis. Of course, it is necessary to have not only a precise knowledge of the physical correlations but also a reasonable measurement program, so that the behavior is not misinterpreted. If no physical meaning can be found, it must be evaluated whether the model or the underlying data is wrong.

Under these boundary conditions, both the model terms for the polynomial coefficients n_k and the relationship between other important influencing variables must be described. In principle, the model or description term can be selected in an open-ended manner. Thus, no fixed model equations are prescribed. It is crucial that all three criteria are implied, but the graphical evaluation of the results together with the physical interpretation is the most important part. The evaluation based on the numerical parameters is postulated.

Those numerical parameters have yet to be defined at this point. The software *Matlab* has already implemented the most important ones. Some of them will be used in this work and are described below. The relevant information for this comes from [69] as well as [27].

First of all, the so called *sum of squares error - SSE* is to be mentioned. This is the sum of all deviations between a measured and predicted value. It can be formulated as follows:

$$SSE = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 \tag{5.17}$$

Here, y_i denotes the particular measured data value, $\hat{y}_i$ the modeled value belonging to the same grid point and n the number of sample points. The smaller SSE is, the better the model maps the measured data. It is important to keep in mind that SSE also includes the absolute value of the measurement data.

Another significant value for the model evaluation is the Root Mean Squared Error ($RMSE$). It is also known as the standard error of the fits and gives a reference value for the standard deviation of a random component in the data. It is defined as follows:

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n} (y_i - \hat{y}_i)^2}{n}} = \sqrt{\frac{SSE}{n}} \tag{5.18}$$

The smaller the *RMSE*, the better the model equation works. However, it still depends on the absolute values of the modeled quantities. For this reason, the normalized *RMSE* is also presented:

$$nRMSE = \frac{RMSE}{(y_{max} - y_{min})} \tag{5.19}$$

The introduction of the normalized *nRMSE* allows even different dimensions of values to be compared easily. In addition to the definition shown here, it can also be normalized using the mean value of the measured data $\bar{y}$. The results are equivalent in terms of information, though it is important that the definition is stringent and consistent.

The following subsections contain the visualization of the three polynomial parameters over the respective model input variable a_j. The model equations used are explained and the corresponding numerical values are shown. Subsequently, the physics behind the process is examined and explained. This is a special advantage of the procedure of this thesis. Because all the operating points examined are extensively analyzed during post-processing, such as a complete loss analysis, physical correlations can be explained clearly and substantiated. For this purpose, further important values used for explanation are presented in the individual subsections depending on the respective model input. The dependencies shown here are not discussed according to the criteria of regression analysis presented above, but are chosen in a similar manner.

5.3.1 Engine speed

This subsection deals with the model influence of the engine speed. As described in the last chapter, the three polynomial parameters n_k are initially depicted in Figure 5.11. Before the interpretation of the results is continued as shown in the figure, explicit reference is made to Table 5.1, which lists the polynomial parameters and their meaning with respect to an EGR variation.
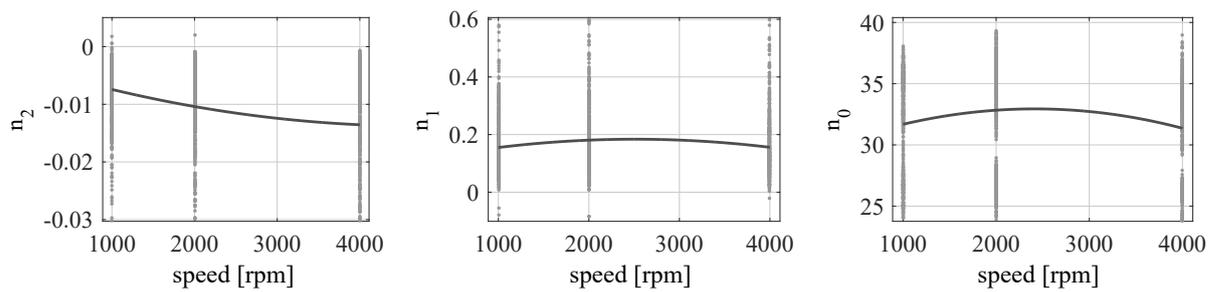
Figure 5.11: Polynomial parameters n_{2,a_3}, n_{1,a_3} and n_{0,a_3} depending on the engine speed (ω).

The gray scatter-plot shows the polynomial coefficients resulting from the measurements produced by the test program introduced in Figure 5.6. The course that results from the nonlinear regression is shown in black. In this case a second-degree polynomial is used as a model equation for all three coefficients. It is presumed as a matter of course that the model equation is selected in the clash between graphical and numerical evaluation as well as physical basis. To ensure that there is sufficiently good

comparability, the y-axis scaling introduced for the engine speed here is maintained uniformly for the polynomial coefficients in the other subsections.

Even if the deviations of the individual values from the fitted course are in part clear, the first thing to be noted is that the regression analysis and the resulting course can in principle reflect the overall behavior. The deviations addressed can be explained by the small number of grid points regarding the engine speed, where only 3 are concerned. In conclusion, the numerical values of the coefficients n_2 and n_1 are also sufficiently accurate. These values as well as those of the n_0-coefficient, indicating the assessment of the model term from Section 5.1, are shown in Table 5.4.

Table 5.4: Characteristic values to evaluate the model quality in relation to the engine speed n_{mot}.

polynomial coefficient	basis for fit	*SSE*	*RMSE*	*nRMSE*
n_2	second degree polynomial	0.334	0.0143	0.0617
n_1	second degree polynomial	37.8	0.152	0.0973
n_0	second degree polynomial	$2.88 \cdot 10^4$	4.19	0.247

The two values, *SSE* and *RMSE*, which depend directly on the absolute level of the coefficients, show a very satisfactory level for n_2 and conditionally also for n_1. Conversely, these values deviate to some extent for n_0. It also becomes clear in the normalized value *nRMSE* that the regression of n_0 works least well. The *nRMSE* value allows all coefficients to be compared directly. The error level of n_0 is about three times as high as that of the modeling of n_2 and more than twice as high as that of the modeling of n_1. Finally, the *nRMSE* of n_0 is still much closer to 0 than to 1. Thus, a physical interpretation is nonetheless still possible for the values of n_0.

One key factor, however, is that the behaviour can be understood physically. Ultimately, the principal behavior above the engine speed must be consistent. The curve of n_2 indicates that the EGR tolerance decreases for increasing engine speed while the sensitivity of the indicated efficiency to the supply of external LP EGR increases, i.e. the curve becomes steeper. The first thesis can be substantiated by the graph on the left-hand side of Figure 5.12. Here, the measured values of $x_{EGR,max,tol}$, which is equivalent to the EGR tolerance, are plotted above the engine speed. It can be clearly seen that the EGR tolerance decreases at increasing speeds. This in turn can be explained by the further two diagrams given in the figure. At first, the middle part shows the increasing burning duration in *CAD* over the engine speed. Next to it, the ignition timing *it* is depicted, which shifts to earlier crank angles at high engine speeds.
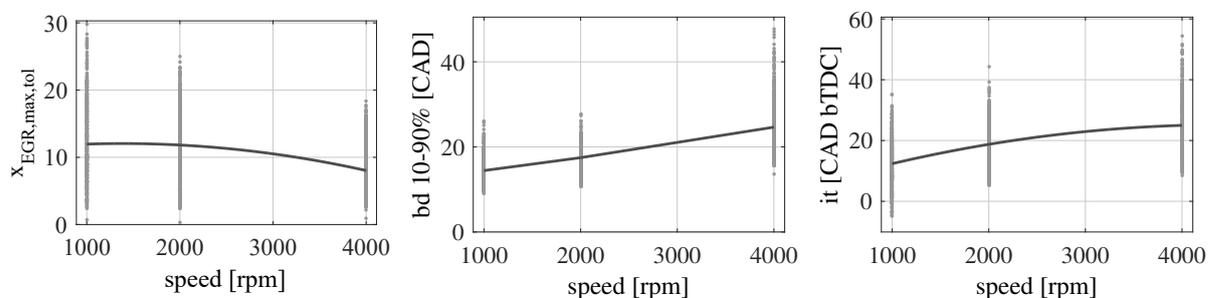
Figure 5.12: $x_{EGR,max,tol}$, burning duration *bd* 10-90 % and ignition timing *it* depending on the engine speed (ω).

The available absolute time span for the combustion is noticeably lower at higher engine speeds. This is due to an extension on the cyclic level, even under the assumption of a constant burning duration on the time level. To ensure a constant $mfb50$, the ignition timing is thus shifted forward within the engine cycle. Though this provokes a combustion start under poorer thermodynamic boundary conditions, i.e. cylinder pressure and temperature. In [36], Güldner shows that the laminar combustion speed is strongly dependent on the conditions prevailing during the combustion start. In the contribution by Alger et al. [3] it is proven that the laminar combustion speed has a significant impact not only on the ignition delay but also on the combustion duration and thus on the cycle-to-cycle variations of the engine. In this context, the experimental results in the study by Wenig [107] likewise show an increase in the COV_{imep} with increasing engine speed and thus support the observations of a decreasing EGR tolerance at rising engine speeds in this work. By adding additional external EGR, the buffer available at higher speeds to further slow down the combustion speed and thus lengthen the burning duration is significantly lower than at lower speeds. Consequently, the EGR tolerance level is lower. A decrease in n_2 is both logical and consistent.

Besides that, the n_2 curve also indicates a steeper slope of the indicated efficiency at higher engine speeds. The sensitivity increases and in spite of a lower EGR tolerance, there is still great efficiency potential to be exploited. As can be seen in the left-hand diagram in Figure 5.13, this is mainly explained by the growth in ideal gas exchange losses at increasing speeds. As shown in Figure 1.2 and the corresponding table, the reduction in the gas exchange losses is a central influencing factor of external EGR which also comes to bear in this context. Due to the increased gas exchange losses at higher speeds, the potential to be exploited is significantly higher. One aspect that should not be concealed and is therefore also shown in the middle diagram in Figure 5.13 is that contrary to this, the losses due to real heat transfer decrease at higher engine speeds. In absolute numbers, however, the decline in the losses due to real heat transfer is clearly less significant than the increase in the losses due to ideal gas exchange. This is why the existing efficiency potential tends to increase with the engine speed.

Finally, the right-hand graph in Figure 5.13 illustrates the maximum efficiency of the individual EGR variations from the measurements $\eta_{ind,EGR,max}$. These confirm the basically correct course of n_0, representing almost a perfect parallel shift of this curve. At the same time, however, they exhibit an essential sheariness. Due to the broad spread of the operating point parameters overall, there are widely varying levels of efficiency. These are very hard to model by simply modeling the speed or, as will be shown later, by modeling just one other model input parameter. This always results in a comparatively bad characteristic value in the corresponding numerical assessment values for n_0. For the actual application of the model, a proposal is made in Section 5.4 to circumvent this problem.

The physical justification for the course of n_0 can be found in the gas exchange losses, which increase significantly at higher engine speeds. As they account for a large part of the losses, the general efficiency level and thus n_0 inevitably decrease.

5.3.2 Engine load

After the course depending on the engine speed is explained, this section studies how the engine load influences the course of the individual polynomial coefficients n_k. Figure 5.14 depicts the course of the three coefficients depending on the engine load *imep*.
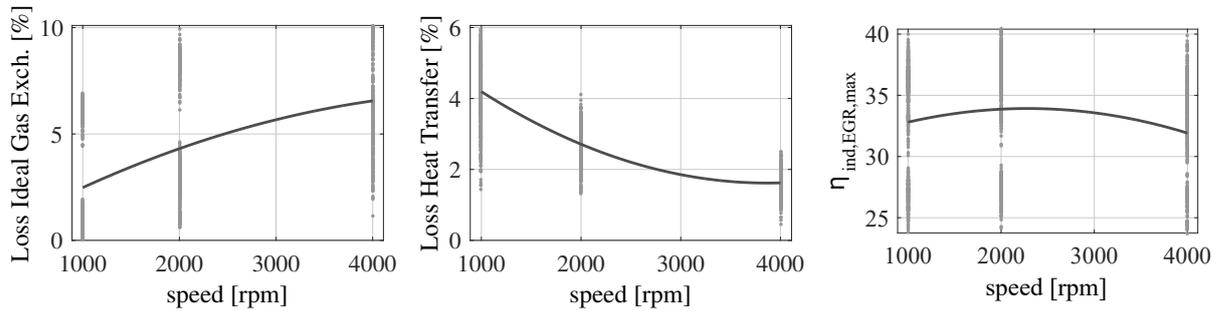
Figure 5.13: *Loss ideal gas exchange, loss heat transfer and $\eta_{ind,EGR,max}$ depending on the engine speed (ω).*
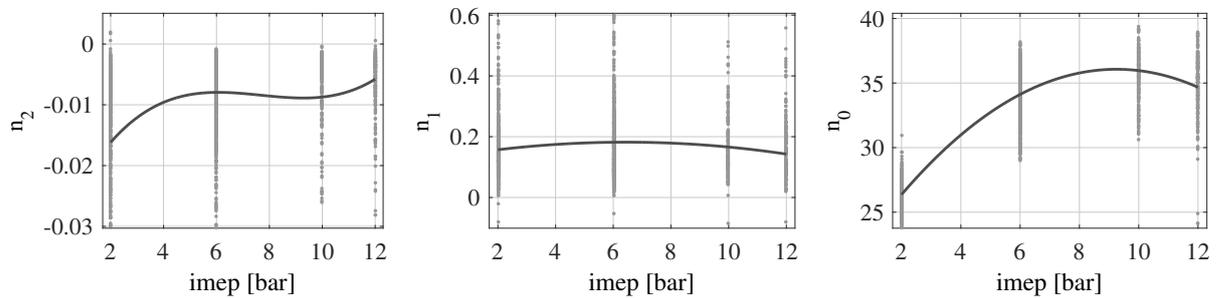
Figure 5.14: Polynomial parameters n_{2,a_4}, n_{1,a_4} and n_{0,a_4} depending on the engine load (*imep*).

The associated numerical values for the evaluation of the regression analysis are given in Table 5.5. These are all comparable with the corresponding values for the engine speed from Table 5.4, both in terms of the ratios between each other and the absolute level. This is mainly due to the same database be used in both cases. In addition, a similar number of grid points is used for the two model input parameters introduced so far. The model basis in this case is a third degree polynomial for n_2 and a second degree polynomial for n_1 and n_0.

Table 5.5: Characteristic values to evaluate the model quality in relation to the engine load *imep*.

polynomial coefficient	basis for fit	*SSE*	*RMSE*	*nRMSE*
n_2	third degree polynomial	0.33	0.0139	0.0611
n_1	second degree polynomial	37.2	0.145	0.097
n_0	second degree polynomial	$2.22 \cdot 10^4$	3.68	0.217

With regard to the courses of the individual polynomial coefficients, however, a clearly differentiated picture emerges in comparison to the equivalent courses depending on the engine speed. First, the n_2 curve stands out in this context. It is modeled by a third-degree polynomial and increases above the engine load. Between the 6 *bar* and the 10 *bar* load point, the values do not increase, and the 10 *bar* load point an upward trend follows again. At first a statement of principle is deduced from this, which is that the EGR tolerance increases noticeably with a rising load. On the left-hand side of Figure 5.15, $x_{EGR,max,tol}$ is once again illustrated but as a function of *imep*. The basic course of the curve is a mirror of the course of the n_2 curve and thus confirms the stated claim.

The reason for this behavior is inherently twofold. On the one hand, the internal residual gas rates that are present in this load range are significantly lower, as can be seen in the middle of Figure 5.15. Consequently, the total EGR rate decreases. The charge exchange in this operating range does not allow the high internal residual gas rates. On the other hand, the filling ratio and the charge density in the higher load range are advantageous for an increased EGR tolerance. In fact, the distances between the reactants are significantly increased by the inert exhaust gas fraction. This effect is more pronounced in the range of small loads, since the reactant distances are already high without EGR. Due to an increased charge density in the high-load range, the maximum reactant distances are reached later on. In addition, the conditions for ignition are fundamentally more advantageous at higher loads due to an increased charge temperature and pressure at the ignition time. The effects and influences described above are supported, for example, by Birkigt and his studies in [8].
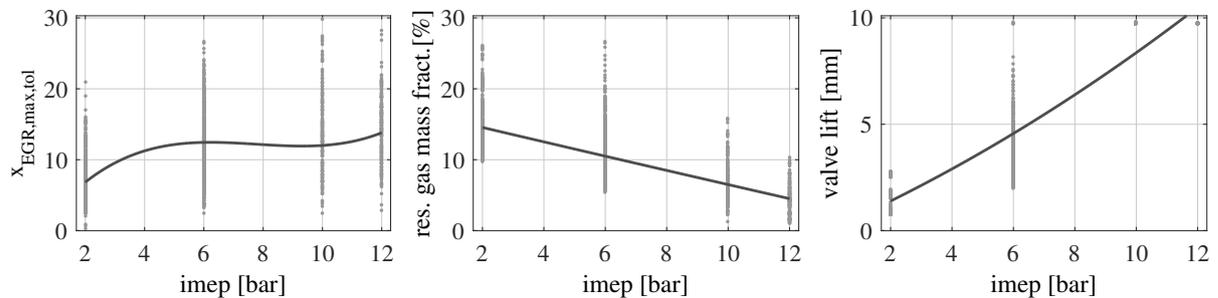
Figure 5.15: $x_{EGR,max,tol}$, residual gas mas fraction and valve lift depending on the engine load (*imep*).

The second assertion made above is that the EGR sensitivity or the efficiency potential that can be implemented by the use of external EGR decreases at elevated engine loads. This can be understood from the first two diagrams in Figure 5.16. Towards higher loads, the gas exchange losses decrease, as can be seen in the left-hand diagram. Thus, a central efficiency control lever when external EGR is applied shrinks. It is particularly noteworthy that the losses decline despite the decreasing residual gas mass fraction. This can be attributed to the strong dethrottling at high loads, as can be seen in the course of the valve stroke in the right-hand diagram in Figure 5.15.

An essential lever in the upper load range is the $mfb50$. The losses assigned to this typically increase according to the load, as can also be recognized in this case in the middle diagram in Figure 5.16. However, the increase is noticeably smaller than the decrease in the gas exchange losses. The potential from EGR consequently decreases overall, especially since other significant losses associated with EGR, such as the losses due to the heat transfer and the losses due to the gas composition, also decrease.

As the losses decrease, the efficiency level increases, as can be seen from the n_0 curve and the $\eta_{ind,EGR,max}$ curve, which both increase towards higher loads. At the junction from 10 to 12 *bar*, however, the efficiency level decreases slightly. This effect is mainly caused by the steeper curve of losses due to combustion phasing at 10 *bar* as well as losses due to the real burn rate and incomplete combustion. Both curves are shown in Figure 5.17 below.

At the end the interpretation of n_1 is outstanding. The course of n_1 is very flat and hardly shows any dependence on the engine load. A more in-depth discussion is therefore not necessary.
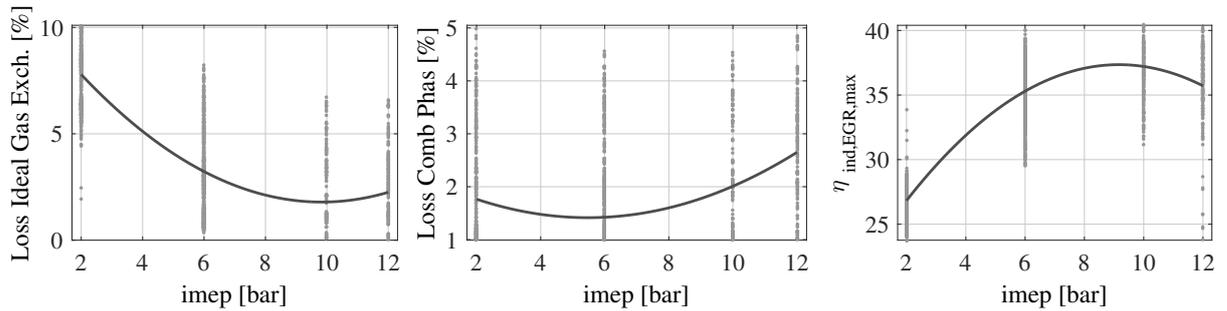
Figure 5.16: *Loss ideal gas exchange, loss combustion phasing* and $\eta_{ind,EGR,max}$ depending on the engine load (*imep*).
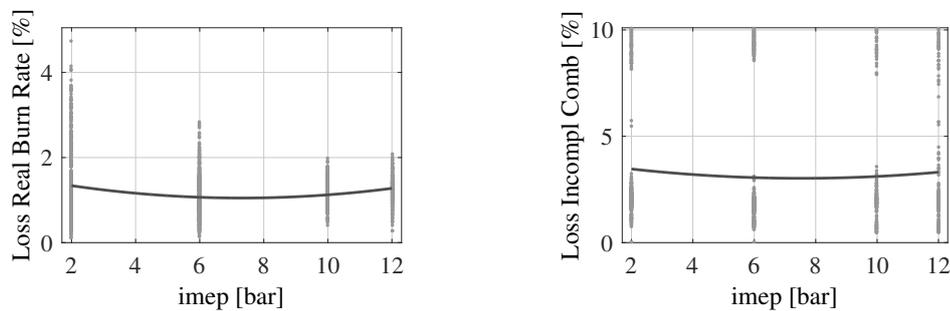
Figure 5.17: *Loss ideal gas exchange* and *loss real burn rate* depending on the engine load (*imep*).

5.3.3 Exhaust cam timing

This chapter goes into further detail on the influence of the exhaust cam timing *ect*: a central parameter affecting the internal residual gas mass fraction and thus the total residual gas mass fraction. Again, it should be considered that the exhaust cam timing, as shown in Figure 4.8, is used to influence the type of residual gas within the measuring program of this thesis. Figure 5.18 shows the three polynomial coefficients depending on the *ect*.
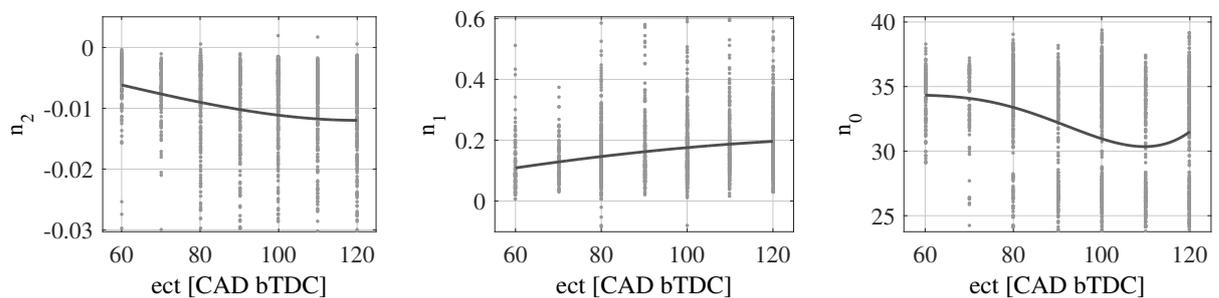
Figure 5.18: Polynomial parameters n_{2,a_1}, n_{1,a_1} and n_{0,a_1} depending on the exhaust cam timing (*ect*).

Table 5.5 completes the corresponding numerical characteristic values. Now, there are more grid points and the numerical values differ slightly more than before from those of the engine load and the engine

speed. The numerical values and their ratios nevertheless remain comparable with the previous ones. Thus no further discussion is necessary.

Table 5.6: Characteristic values to evaluate the model quality in relation to the exhaust valve timing *ect*.

polynomial coefficient	basis for fit	*SSE*	*RMSE*	*nRMSE*
n_2	third degree polynomial	0.337	0.0144	0.0646
n_1	third degree polynomial	37.7	0.152	0.0972
n_0	fourth degree polynomial	$2.84 \cdot 10^4$	4.15	0.245

Examining the course of n_2, a slightly degressive decrease for larger exhaust cam timings becomes apparent. This indicates a decrease in EGR tolerance when exhaust cam timings increase. Besides that, the EGR sensitivity and with it the additional efficiency potential, which can be obtained by adding EGR, presumably increases. The former of the two statements is contrary to the initial expectation and must therefore be more thoroughly understood and questioned. Assuming that there is a small valve overlapping when large exhaust cam timings are implemented, a tendency towards smaller internal residual gas rates should improve the external EGR tolerance. However, it is important to bear in mind that the *ect* is, inter alia, used to control the different residual gas strategies according to Figure 4.8, while the residual gas level is adjusted and assigned via the valve overlapping area *voa*. Though, the two effects cannot be easily separated in the evaluation. The diagrams in Figure 5.19 help to resolve the contradictions. The EGR tolerance $x_{EGR,max,tol}$, the internal residual gas rate and the valve lift are displayed as functions of the exhaust cam timing.
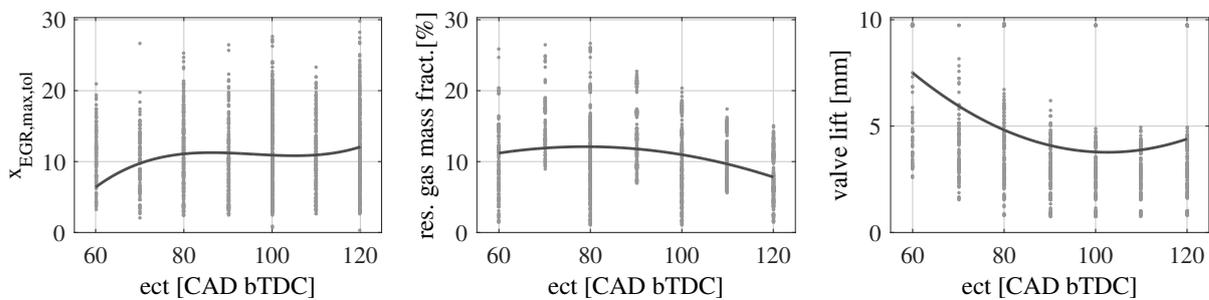
Figure 5.19: $x_{EGR,max,tol}$, residual gas mass fraction and valve lift depending on the exhaust cam timing (*ect*).

First of all, it is obvious that the interpretation of n_2 in this exceptional case does not correspond to the EGR tolerance $x_{EGR,max,tol}$ from the measurements. The EGR tolerance increases slightly at higher exhaust cam timings, in line with the original expectation. However, this course of n_2 is not the result of a wrong regression analysis: the effect of increasing EGR sensitivity simply outweighs the influence of the EGR tolerance. This can be explained by a decreasing internal residual gas rate and a resulting smaller total EGR rate at high exhaust cam timings. As a result, there is still considerable potential for the supply of residual gas. However, a decisive factor for the lower residual gas rates is not only the exhaust cam timing, but also the valve stroke, which is noticeably smaller within the measurement series at higher exhaust cam timings. This is also due to the strategy of controlling the residual gas level using the valve overlapping area.

One thing that remains to be clarified is how the different residual gas strategies affect n_2 and the associated phenomena. As a rough rule of thumb, large *ect* values are necessary for combustion chamber trapping, while small values of *ect* are related to intake port or exhaust channel recirculation. Whereas the first strategy provides rather hot residual gas for combustion, the intake port as well as the exhaust channel recirculation lead to a comparatively stronger cooling of the internal residual gas. This is a decisive factor in the behavior of n_2 depending on the *ect* as well as in the contradictory restraint of n_2 and $x_{EGR,max,tol}$. As previously described, hot internal residual gas stabilizes the combustion, as it improves the boundary conditions at the start of combustion by increasing the combustion chamber temperature. This means that not only the mere lowering of the internal residual gas leads to an increase in the EGR tolerance, but also the alternative residual gas strategy. The EGR tolerance benefits disproportionately from this double effect of the lowered residual gas mass fraction and favorable residual gas strategy. Though, in the course of n_2 also the effects of an increased EGR sensitivity need to be depicted. The non-congruent curves of the internal residual gas rate and the EGR tolerance $x_{EGR,max,tol}$ above *ect* are further evidence of the decisive influence of the residual gas strategy.

The influence of the *ect* on the polynomial coefficient n_1 is the most significant among the parameter influences on n_1 considered so far. The absolute value of n_1 at the maximum exhaust cam timing is only half that at the minimum exhaust cam timing. This means that the curves of the indicated efficiency above the EGR rate become steeper before reaching the EGR tolerance, i.e. the EGR sensitivity increases, which is totally consistent with the behavior of n_2. At the same time, this means that the curves of the indicated efficiency decline less sharply after reaching EGR tolerance, i.e. there is a less rapid increase in cycle-to-cycle variations. This statement can be clearly proven by the diagrams in Figure 5.20.
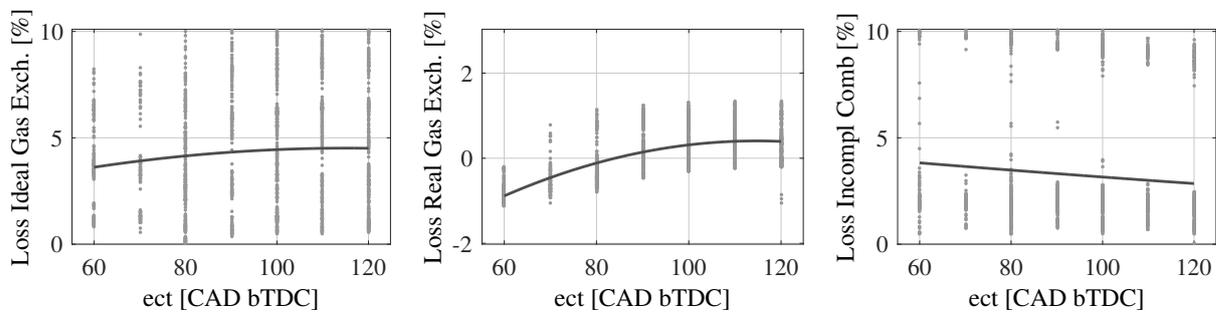
Figure 5.20: *Loss ideal gas exchange*, *loss real gas exchange* and *loss incomplete combustion* depending on the exhaust cam timing (*ect*).

Both the losses due to ideal gas exchange and the losses due to real gas exchange very clearly show the curve flattening out towards high values of the exhaust cam timing. This can be transferred one-to-one to the curves of the indicated efficiency above the EGR rate and thus to the curve of n_1. The flattening curves are caused by higher EGR tolerances and thus higher EGR rates in these areas, which make the gas exchange losses increase less strongly.

In addition, the losses due to incomplete combustion in the right hand diagram in the figure are also significantly lower for large *ect* values. These losses are of particular importance with regard to cycle-to-cycle variations and decrease because of two phenomena. One is the change in the residual gas

strategy towards hotter residual gas, the other is the basic decrease in internal residual gas rates at high *ect* values.

In conclusion, all three losses depicted exhibit a flattening or a decrease at high exhaust cam timings. This leads to a less steep decrease in the indicated efficiency when the EGR tolerance is exceeded. The three losses shown account for a high percentage of the total loss and have a correspondingly high influence on the course of the indicated efficiency depending on the external EGR rate.

Finally, the course of n_0 must be analyzed again. The course of this coefficient is first progressive, then flattens out and even rises again at high *ect* values. As Figure 5.21 illustrates, this behavior coincides with the maximum efficiency of the EGR variations $\eta_{ind,EGR,max}$. Besides the losses just described, however, the losses due to the gas composition are mainly responsible. These are shown in the right-hand diagram in Figure 5.21.
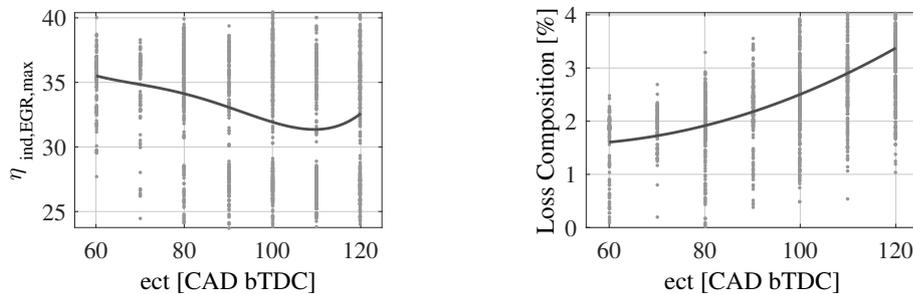
Figure 5.21: $\eta_{ind,EGR,max}$ and *loss composition* depending on the exhaust cam timing (*ect*).

The losses due to gas composition increase exponentially. This can be explained by the decreasing internal residual gas rates. In sum, the illustrated losses, which, as already described, make up a major part of the total loss, lead to the behavior of the EGR variations' maximum indicated efficiency and the behavior of n_0 over the exhaust cam timing.

5.3.4 Valve overlapping area

Following the last subsection dealing with the exhaust cam timing, , a key lever for setting the type of internal residual gas, this chapter will discuss the influence of the valve overlapping area, *voa*. This is the model input variable that has the greatest influence on the level of the internal residual gas rate. Figure 5.22 introduces the three coefficients of the second-degree polynomial depending on the valve overlapping area.

The corresponding numerical values can be found in Table 5.7. They are at a similar level to the previous analyses. Table 5.7 also reveals the use of an exponential equation as the model basis. It is particularly relevant in the present application, since it is the only way to capture the effects of the finer measurement point grid in the range of lower *voa*-values.

The first polynomial coefficient n_2 shows a degressive increase towards larger valve overlapping areas up to a virtually complete curve flattening for the maximum *voa* values. This means an increase in EGR tolerance to external EGR and a decrease in sensitivity to external EGR. The courses of the indicated efficiency depending on the EGR rate become flatter. The flattening of the n_2 curve itself or the asymptotic approach to $n_2 = 0$ shows that the curve changes into a linear course for very large
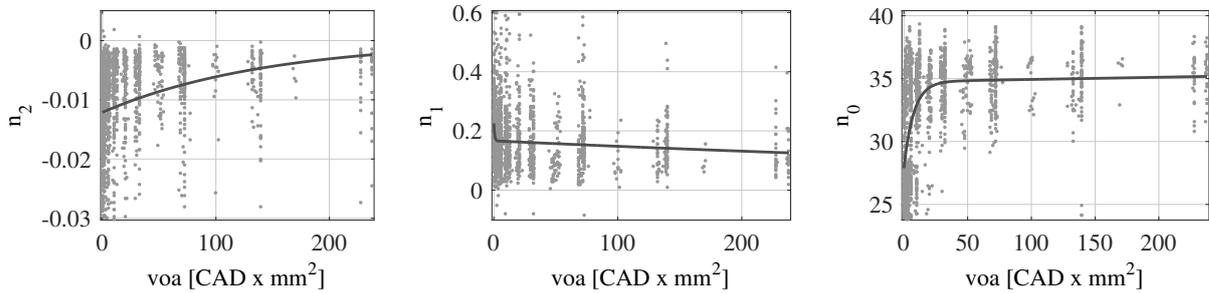
Figure 5.22: Polynomial parameters n_{2,a_2}, n_{1,a_2} and n_{0,a_2} depending on the valve overlapping area (*voa*).

Table 5.7: Characteristic values to evaluate the model quality in relation to the valve overlapping area *voa*.

polynomial coefficient	basis for fit	*SSE*	*RMSE*	*nRMSE*
n_2	exponential	0.34	0.0141	0.0622
n_1	exponential	37.8	0.157	0.0976
n_0	exponential	$2.5 \cdot 10^4$	3.91	0.23

voa. If these large valve overlapping areas are equated with higher internal residual gas rates, this behavior initially contradicts the logic. Figure 5.23 below shows that this is not the case, that high valve overlapping areas are not associated with high internal residual gas rates, and that higher internal residual gas rates do not lead to an increase in tolerance to external EGR.
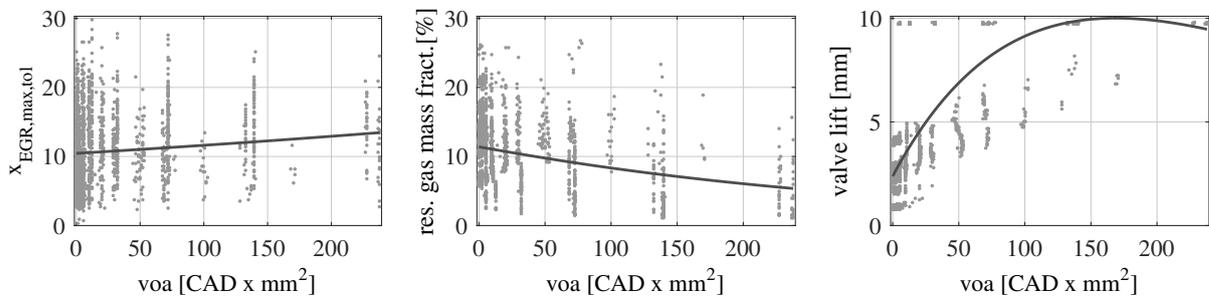
Figure 5.23: $x_{EGR,max,tol}$, residual gas mass fraction and valve lift depending on the valve overlapping area (*voa*).

First, the course of the EGR rate at maximum efficiency, synonymous with the EGR tolerance $x_{EGR,max,tol}$, confirms that the EGR tolerance does tend to increase towards a larger *voa* and thus corresponds to the conclusion of n_2. The middle diagram shows that this statement is fully justified. It shows a decreasing residual gas mass fraction. Of course, this raises the question as to why this behavior occurs. That is answered in the right-hand diagram in the figure. Large *voa* values can only be applied for large valve strokes, which are directly related to the engine load. This does not contradict the physical independence of the single model input parameters, but only implies that large values of *voa* can only be realized with less internal residual gas and at higher loads. The curves that initially appear to run contrary to the actual expectation and the explanation that can be given by the corresponding parameters once again show the importance of discussing the model on a physical level.

The course of n_1 is quite flat over almost the entire value range. However, it shows a strong increase in the area where the valve overlap is small or non-existent. This indicates that without valve overlapping, the drop in the indicated efficiency curve flattens significantly after the maximum efficiency is exceeded, and at the same time, corresponding to the n_2 curve, has a steeper slope to the left of the efficiency maximum. The reason for this trend can be found in many operating points with low load in the smaller *voa* range, where the lever for supplying external EGR and the resulting decrease in gas exchange losses is significant. This contention is supported by the diagrams on the left and in the middle in Figure 5.24, where the two loss components of the gas exchange are depicted. These both increase strongly for small valve overlapping areas and thus offer a large potential for an increase in efficiency. This is also accompanied by the smaller valve strokes in this area, which provide a higher efficiency potential for dethrottling using an inert gas.

Finally, on the right-hand side of the figure, the losses due to the gas composition are displayed. Their behavior is comparable to the gas exchange losses, although the course for a high degree of valve overlapping is likewise overestimated. However, the losses due to the gas composition confirm the behavior of n_1. The reason for the increase in the composition losses in the case of a small *voa* can be considered to be that although there is more internal EGR, there is less external EGR due to the reduced EGR tolerance.
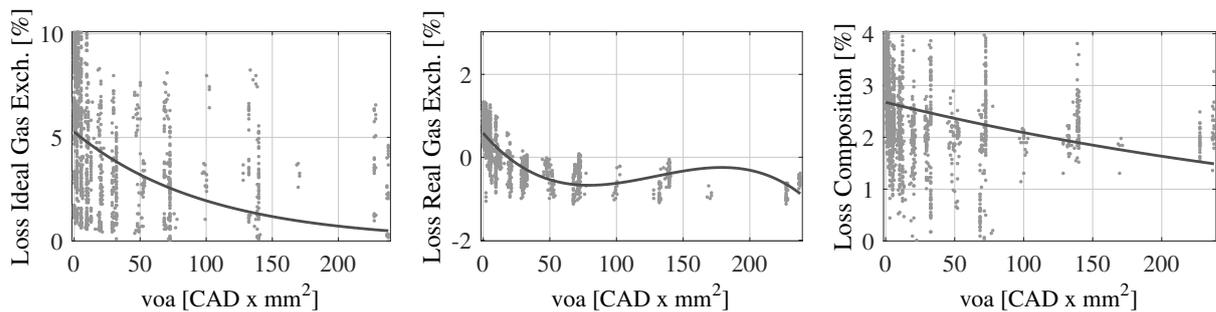
Figure 5.24: *Loss ideal gas exchange*, *loss real gas exchange* and *loss composition* depending on the valve overlapping area (*voa*).

Eventually, n_0 in the right-hand diagram in Figure 5.22, just like n_1, shows a flat course over almost the entire value range of *voa*. For small values, however, there is a steep drop in n_0 and thus in the indicated efficiency in the base point of the EGR variation. This behavior is analogously reflected in the behavior of the maximum efficiency of the EGR variation $\eta_{ind,EGR,max}$, depicted in Figure 5.25. The reasons underlying this behavior follow from the losses that have already been illustrated and explained. These losses increase strongly in unison when the valve overlapping is small and together represent a large portion of the total loss.

5.3.5 Pressure intake manifold

Another variable which influences the type and the level of the internal residual gas rate is the pressure in the intake manifold p_{IM}. Depending on p_{IM} the three polynomial coefficients of the second degree polynomial are shown in Figure 5.26. It is important to mention that the algebraic signs on the x-axis are contrary to the real values. This is due to the implementation of the test program via an automation
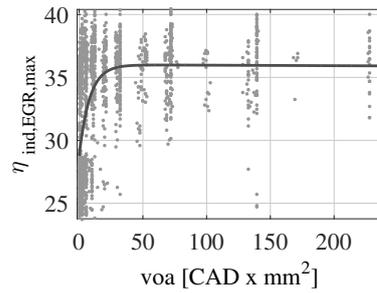
Figure 5.25: $\eta_{ind,EGR,max}$ depending on the valve overlapping area (*voa*).

software. Positive values therefore signify a negative pressure in the intake manifold, while negative values denote charged operation.
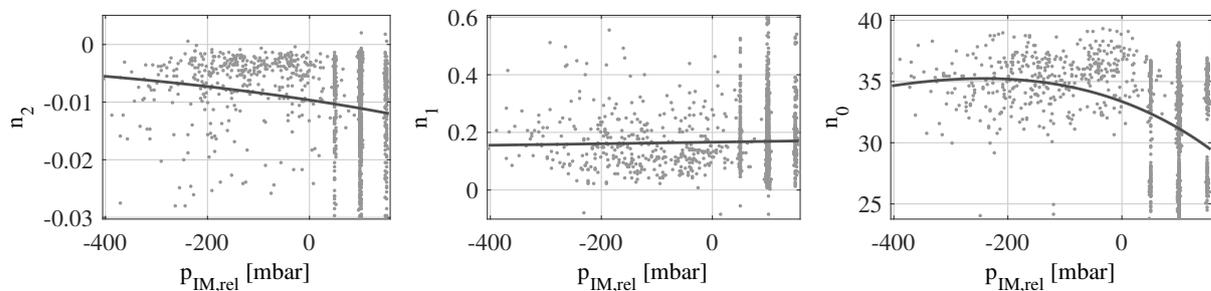
Figure 5.26: Polynomial parameters n_{2,a_7}, n_{1,a_7} and n_{0,a_7} depending on the intake manifold pressure (p_{IM}).

The corresponding numerical parameters, which are listed in Table 5.8, deviate slightly more than those for the previous model input parameters. One key issue, which becomes apparent is, that p_{IM} can only be freely adjusted in the partial-load range, i.e. in the naturally aspirated load points. In all charged operating points, the pressure in the intake manifold automatically results from the load point. This provokes some outliers in these value ranges, which are also visible in the graphic evaluation in Figure 5.26. However, the basic usability and interpretability of the polynomial coefficients depending on the intake manifold pressure remains unrestricted. An exponential function is used to model all three polynomial coefficients.

Table 5.8: Characteristic values to evaluate the model quality in relation to the intake maniflod pressure p_{IM}.

polynomial coefficient	basis for fit	SSE	RMSE	nRMSE
n_2	exponential	0.319	0.014	0.0603
n_1	exponential	48.4	0.171	0.11
n_0	exponential	$4.35 \cdot 10^4$	5.16	0.304

The course of n_2 shows a decrease towards lower intake manifold pressures. This means that the EGR tolerance decreases and the EGR sensitivity increases simultaneously. For the practical implementation in a driving situation, this means that an increase in the intake manifold pressure can be used to increase the EGR tolerance. The behavior of n_2 matches the EGR rate at the maximum indicated efficiency $x_{EGR,max,tol}$ very well, as the left-hand diagram in Figure 5.27 demonstrates. In the range of negative

manifold pressures, the underpressure and thus the internal residual gas rate are controlled by the throttle valve. The key residual gas strategy is the intake port recirculation. The correlation between the throttle position and internal residual gas rate can be seen in the middle and right-hand diagrams of Figure 5.27. To achieve a lower intake manifold pressure, stronger throttling is required. The resulting increased internal residual gas rate is e.g. demonstrated in the research by Nitschke in [74] and Löbbert in [65].
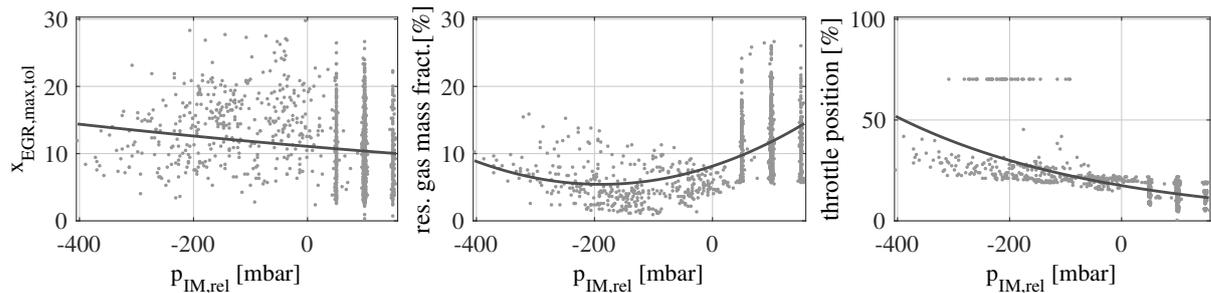
Figure 5.27: $x_{EGR,max,tol}$, residual gas mass fraction and throttle position depending on the intake manifold pressure (p_{IM}).

This correlation does not apply to the turbocharged range in either diagram. Nevertheless, the internal residual gas rate significantly rises during turbocharged operation. Yet the question arises as to where the increase in the internal residual gas originates under these boundary conditions. Essentially, the sharply increasing exhaust back-pressure is responsible. This is illustrated by the left-hand diagram in Figure 5.28. The increase in the internal residual gas rate and the exhaust back-pressure is accompanied by an extended burning duration, as shown in the right-hand diagram in Figure 5.28. These connections are also established in the known literature and illustrated, for example, by Löbbert in [65].
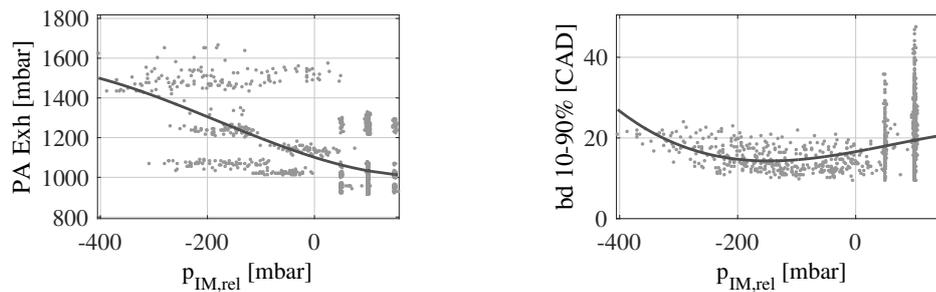
Figure 5.28: Exhaust back pressure *PA Exh* and burning duration depending on the intake manifold pressure (p_{IM}).

The coefficient preceding the linear part of the polynomial, n_1 shows no signs of an influence dependent on p_{IM} and therefore needs no further discussion. In contrast, the constant component n_0 indicates a considerable effect regarding p_{IM}. Especially in the area of high negative pressure, there is a strong reduction in n_0. This behavior is also found at the maximum efficiency of the EGR variation $\eta_{ind,EGR,max}$, as the left-hand diagram in Figure 5.29 shows. The reasoning is illustrated with the losses due to ideal gas exchange in the right-hand diagram in the following figure.
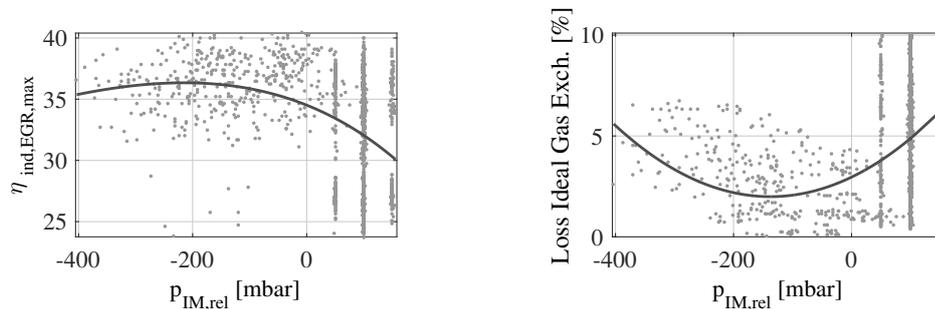
Figure 5.29: $\eta_{ind,EGR,max}$ and *Loss ideal gas exchange* depending on the intake manifold pressure (p_{IM}).

The losses due to ideal gas exchange increase at both edges of the value range. The curve almost mirrors the course of n_0, already indicating the main reason for the course of n_0. The reasons for the course of the gas exchange losses are likewise given above. While the exhaust back-pressure accounts for this loss during turbocharged operation, during throttled operation it is the increased throttling losses due to the higher intake manifold vacuum. This conclusion is reinforced by the work of Nitschke [74], which provides a detailed analysis based on single operating points. For the majority of the other losses the evaluation of this thesis shows only a minor influence with respect to the intake manifold pressure, which is why their cross-effect on n_0 depending on p_{IM} is negligible.

5.3.6 Fuel-air ratio

One parameter which quickly influences the combustion and thus the indicated efficiency and the EGR tolerance is the fuel-air ratio λ. It is, of course, important not to interfere with the proper operation of the TWC. However, a short-term deviation from the stoichiometric fuel?air ratio is possible depending on the extent and duration of that deviation. Figure 5.30 illustrates the behavior of the polynomial parameters n_2, n_1 and n_0 which result from the EGR variations depending on the λ of the base operating point.
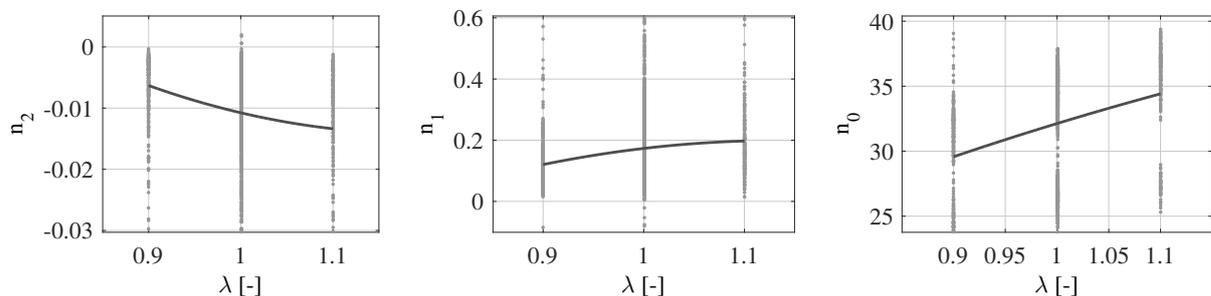
Figure 5.30: Polynomial parameters n_{2,a_6}, n_{1,a_6} and n_{0,a_6} depending on the fuel-air ratio (λ).

The numerical values are specified in Table 5.9. Due to the small number of grid points, it is sufficient to use a second-degree polynomial as the basis equation for the model. The quality of the regression

analysis is consistently satisfactory and at the same or even a better level than those of the other model input variables.

Table 5.9: Characteristic values to evaluate the model quality in relation to the fuel-air ratio λ.

polynomial coefficient	basis for fit	SSE	RMSE	nRMSE
n_2	second degree polynomial	0.316	0.0139	0.0599
n_1	second degree polynomial	37.8	0.152	0.0973
n_0	second degree polynomial	$2.28 \cdot 10^4$	3.73	0.22

The influence of λ is unique for all three polynomial coefficients. Regarding n_2, an approximately linear decrease from the fat ($\lambda < 1$) to the lean ($\lambda > 1$) range can be observed. The corresponding interpretation is a decreasing EGR tolerance along with an increase in how sensitively the efficiency reacts to external EGR. The latter statement can also be deduced from the course of n_1, which increases slightly degressively from the rich to the lean range. Besides the increased sensitivity and thus steeper curve on the left-hand side of the efficiency maximum, this also implies that the slope of the curve is less pronounced after the maximum efficiency is reached.

Figure 5.31 below will be used to start with the discussion and evaluation of the first thesis, namely a reduction of the EGR tolerance. For this purpose, the EGR tolerance from the measurement $x_{EGR,max,tol}$ in the left-hand diagram and the internal residual gas rate in the right-hand diagram are each depicted depending on λ.
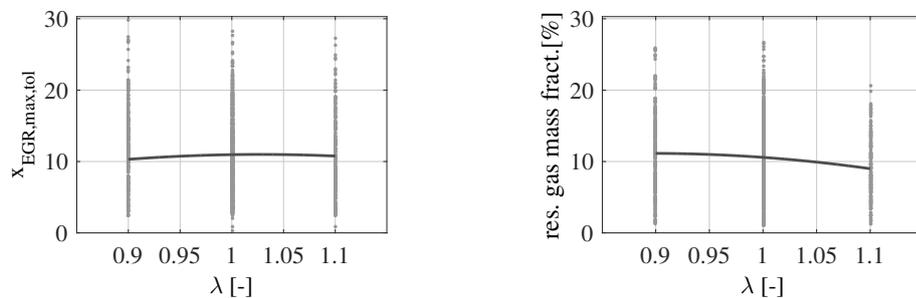
Figure 5.31: $x_{EGR,max,tol}$ and residual gas mass fraction depending on the fuel-air ratio (λ).

Contrary to initial expectations and interpretations, the EGR rate, which is part of the maximum indicated efficiency $x_{EGR,max,tol}$, does not exhibit any perceptible dependence on λ. The influence of λ on the internal residual gas rate is similarly small. This suggests that the decrease in n_2 is not initiated by the reduced EGR tolerance, but by the increased sensitivity towards external EGR. This thesis is also supported by the n1 curve, which is steeper on the left of the maximum indicated efficiency for the lean operating range. This can be proven with the plots in Figure 5.32. The losses due to fluid properties and incomplete combustion as well as the exhaust back pressure are shown. The first mentioned are the losses due to the temperature- and pressure-dependent thermodynamic properties of the cylinder charge mixture. They are not to be confused with the losses due to the real gas composition, where the real cylinder charge is considered to be different from pure air under standard thermodynamic conditions. The observed increase towards the lean operating range corresponds to the findings of Nitschke in [74]

and Heikes in [40]. Due to the lean operation, the cylinder charge can only be cooled down to a small extent, which leads to an increase in temperature and thus to an increase in losses compared to the rich operation, where cooling is provided by the additional fuel. Heywood demonstrates in [41] that the isentropic coefficient rises as the temperature sinks and is thus closer to the ideal process. Heikes and Nitschke also show that the cooling effect works much better using external EGR. The large proportion of losses due to real fluid properties can thus be reduced again, which has an enhancing effect on the EGR sensitivity.

The losses due to incomplete combustion shown next to it reveal a strong decrease towards lean operation. Although this is contrary to the increased efficiency potential in the first step, it also implies that during rich operation no further external EGR can actually be added since external EGR has an additional negative effect on the losses due to incomplete combustion. A buffer to increase these losses and in return reduce the share of losses, on which external EGR has a positive effect, therefore only exists in the rich operating range.

The exhaust back pressure also displays a slight rise due to the increased filling in the lean operating range. Similarly, this is positive with respect to external LP EGR and implies improved sensitivity, since the pressure in the exhaust tract can be lowered by extracting exhaust gas.
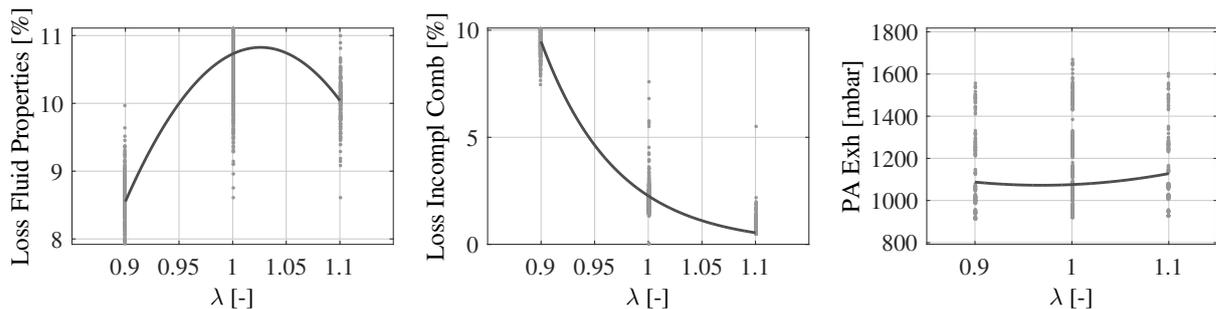
Figure 5.32: *loss fluid properties*, *loss incomplete combustion* and exhaust back pressure *PA Exh* depending on the fuel-air ratio (λ).

One issue remains to be clarified regarding the λ variation, which is why the n_0 curve rises from the rich to the lean range. The left-hand diagram in Figure 5.33 depicts the maximum indicated efficiency of the EGR variation $\eta_{ind,EGR,max}$. The diagram in the middle of Figure 5.32 shows the losses due to incomplete combustion. First of all, it is apparent that the course of $\eta_{ind,EGR,max}$ is in turn a virtually parallel shift in the course of n_0. The basic behavior of n_0 is thus plausible and logical. Furthermore, it is obvious that the losses due to incomplete combustion dominate in the rich operating range. The reason for this can be seen in the surplus fuel. The significantly reduced efficiencies in the rich operating range are based on these losses and are feasible.

Finally, the losses due to gas exchange often used in the previous subsection are shown in the two diagrams in the middle and on the right-hand side of the figure. It becomes clear that they exhibit virtually no dependence on the fuel?air ratio. This is mainly caused by the fact that both naturally aspirated and turbocharged operating points are considered. While throttling in the lean range is advantageous for naturally aspirated operation, it is especially disadvantageous in the charged
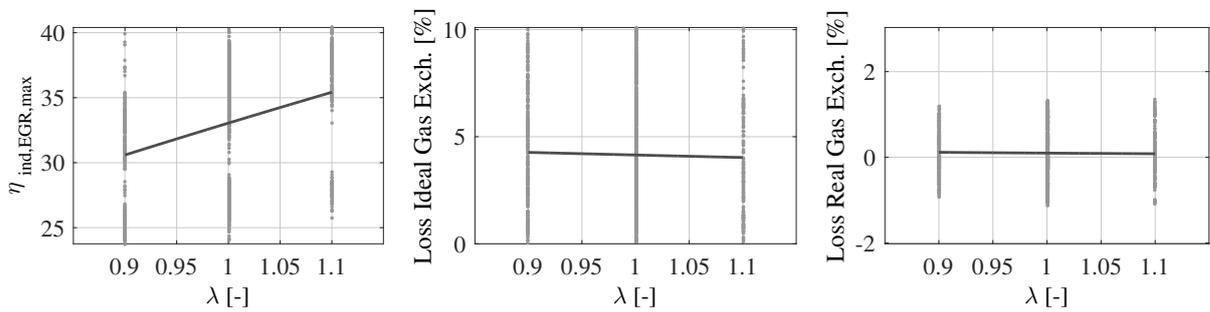
Figure 5.33: $\eta_{ind,EGR,max}$, *loss ideal gas exchange* and *loss real gas exchange* depending on the fuel-air ratio (λ).

operating range, as additional air must be compressed. The influences of the different load conditions, however, are taken into account by using the engine load as a model input parameter.

5.3.7 Ignition timing

Finally, the influence of the ignition timing *it* or ignition angle on the polynomial coefficients is to be examined. The ignition angle occupies a special role in two respects. On the one hand, the ignition angle is not directly dictated in the series of measurements but via the detour of the $mfb50$. This is essential to be able to compare the individual measurement series with each other. In the absence of the same $mfb50$, there would be differences in the combustion process, which would make a variation in the other model input parameters obsolete. It would no longer be reasonable to compare the single measurement series. On the other hand, the ignition angle represents a very fast control path and can be changed from one cycle to the next. During vehicle operation, only the ignition timing is known. For this reason, too, it is indispensable to use the same $mfb50$ as a model input variable, not an unknown one. The polynomial coefficients depending on the ignition angle *it* are plotted in Figure 5.34.
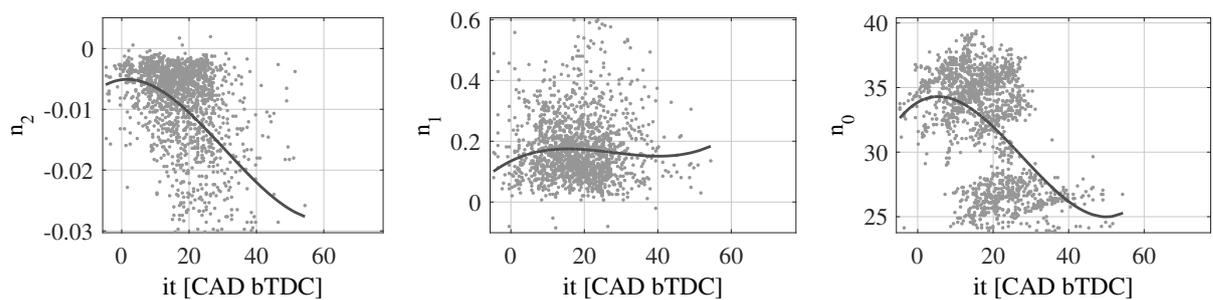
Figure 5.34: Polynomial parameters n_{2,a_5}, n_{1,a_5} and n_{0,a_5} depending on the ignition timing (*it*).

The corresponding values for evaluating the regression are given in Table 5.10. Their level is still adequate and satisfactory.

The ignition timing is not only a fast control path, it is also a very effective and efficient one. This becomes clear from the course of n_2. It drops sharply for early ignition timings, which are located at the right-hand end of the x-axis. In turn, this indicates a decreasing EGR tolerance, but an increasing

Table 5.10: Characteristic values to evaluate the model quality in relation to the ignition timing *it*.

polynomial coefficient	basis for fit	*SSE*	*RMSE*	*nRMSE*
n_2	third degree polynomial	0.332	0.0139	0.062
n_1	third degree polynomial	37.0	0.15	0.0963
n_0	third degree polynomial	$2.64 \cdot 10^4$	4.01	0.237

sensitivity towards external EGR. The thermodynamic reason for this lies in the significantly more favorable boundary conditions with respect to the combustion chamber pressure and combustion chamber temperature, in case the combustion is initiated closer to the TDC. As a result, the laminar burning velocity and with it the initial combustion are improved and the EGR tolerance is increased while cycle-to-cycle variations are reduced. The relationship between the initial combustion and laminar burning velocity is described by Alger et. al. in [3]. The interdependence of the laminar combustion velocity, temperature and pressure are outlined by Gülder in [36] as well as by Metghalchi and Keck in [73]. A confirmation of the decreasing EGR tolerance is given by the left-hand diagram in Figure 5.35, where the EGR tolerance from the measurement $x_{EGR,max,tol}$ is shown.
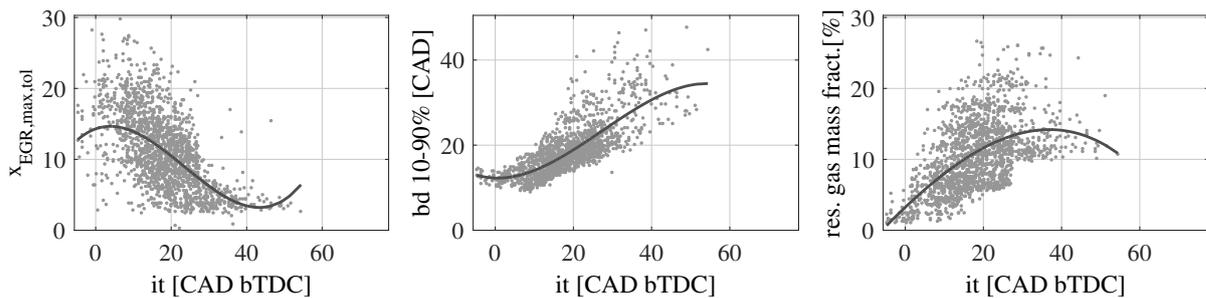
Figure 5.35: $x_{EGR,max,tol}$, burning duration and residual gas mass fraction depending on the ignition timing (*it*).

In addition to the EGR tolerance $x_{EGR,max,tol}$, the burning duration and the internal residual gas rate are shown in the middle and left-hand diagram in the figure. Both values and their curves confirm the decrease in the EGR tolerance. The progression of the burning duration also shows its extension at earlier ignition angles, i.e. under worse thermodynamic boundary conditions. Of course, the reduced tolerance regarding external EGR also goes along with the increased internal residual gas rates.

Although the internal residual gas rates in the measurement series increase with earlier ignition angles, the sensitivity to external EGR grows. Basically, the increase is logical, since rising residual gas rates cause earlier ignition angles and vice versa, while earlier ignition timings go hand in hand with longer burning durations. Nevertheless, the question arises of why the internal residual gas rate increases simultaneously to the sensitivity to external EGR, which is also confirmed by the slightly rising curve of n_1 towards earlier ignition angles.

For explanation, the diagram on the left-hand side of Figure 5.36 is used. There the losses due to ideal gas exchange are plotted. These increase strongly for early ignition angles. This means that despite the increasing internal residual gas rates, the operating points with earlier ignition angles tend to have an increased gas exchange loss and thus an increased potential for the efficiency being improved by

external EGR. This can be attributed to the engine load, which is shown in the middle diagram. It is clear that the particularly early ignition angles are associated with the low-load operating points. As smaller valve strokes and throttle valve angles are found there, the gas exchange losses increase, as shown in the left-hand diagram.
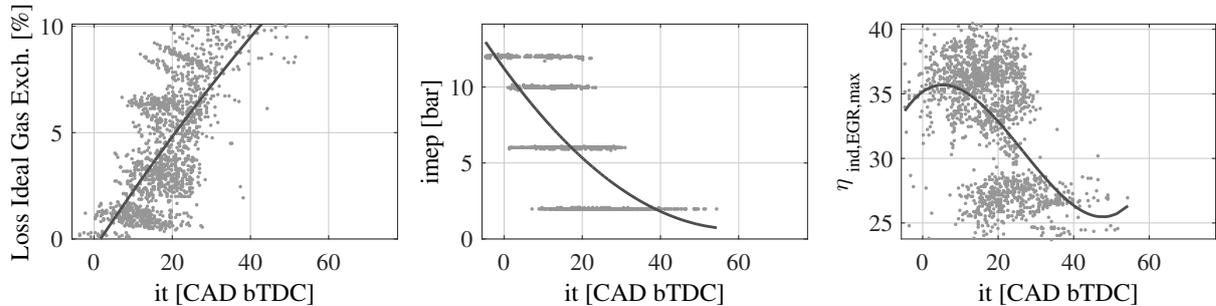
Figure 5.36: *Loss ideal gas exchange*, engine load *imep* and $\eta_{ind,EGR,max}$ depending on the ignition timing (*it*).

Lastly, the course of n_0 needs to be verified. In this context, the right-hand diagram in Figure 5.36 shows the maximum efficiency of the associated EGR variations and therewith initially confirms the basic behavior of n_0. The question about the course of the n_0 curve can be answered with the help of the left-hand plot in Figure 5.36. Increasing losses due to ideal gas exchange are displayed. Since they account for a large proportion of the total losses, this is also associated with poorer efficiency at earlier ignition angles.

In Subsections 5.3.1 to 5.3.7, the influences of the model input parameters on the polynomial coefficients are shown and discussed, or the necessary quantities are assigned. However, this procedure is very abstract with respect to the model's application. The next chapter therefore deals with the concrete application in order to generate the indicated efficiency curves over EGR rate depending on the model input parameters, before the use of the model within a function structure is pointed out in Section 5.5.

5.4 Validation of the model

This chapter can be thematically divided into several parts. Initially, the results of the two weighting matrices are shown and compared with each other. The subsequent model validation follows the principle of moving from the global to the specific. An evaluation of the overall validation results is given by analogy with Figure 5.4. By way of example, these results are also illustrated with interchanged or missing *B* matrices, demonstrating the relevance of the weighting matrix. The accuracy of the model is shown with respect to an individual, unique modification of individual model input parameters at different EGR variations, and on the basis of the maximum indicated efficiency and the associated EGR rate. Finally, the accuracy of the model is depicted by means of individually indicated efficiency curves.

Section 5.2 explains the procedure for calculating the *B* matrices. The actual results are introduced in this chapter and discussed based on the knowledge gained in the previous section on the influence equations for the polynomial coefficients. As becomes apparent, this is necessary for the model?s

implementation and subsequent validation. Below, Figures 5.37 and 5.38 show the individual weighting factors β_{k,a_3} for the engine speed ω and β_{k,a_5} for the ignition angle it, both from Engine 1, as examples.
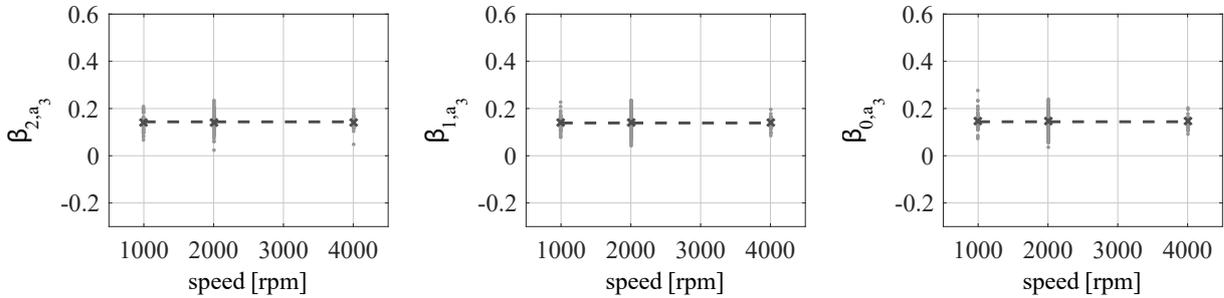
Figure 5.37: Parameters β_{2,a_3}, β_{1,a_3} and β_{0,a_3} depending on engine speed (ω).

The determination of the single values displayed in the scatter plot is accomplished according to the presented procedure. The calculation of the overall values β_{k,a_j} is done by a weighted average of the single values. The overall values are indicated by the black dotted trend line. All other diagrams as well as the diagrams for Engine 2, are given in Appendix A.3, as they do not provide any additional knowledge or added value at this point.
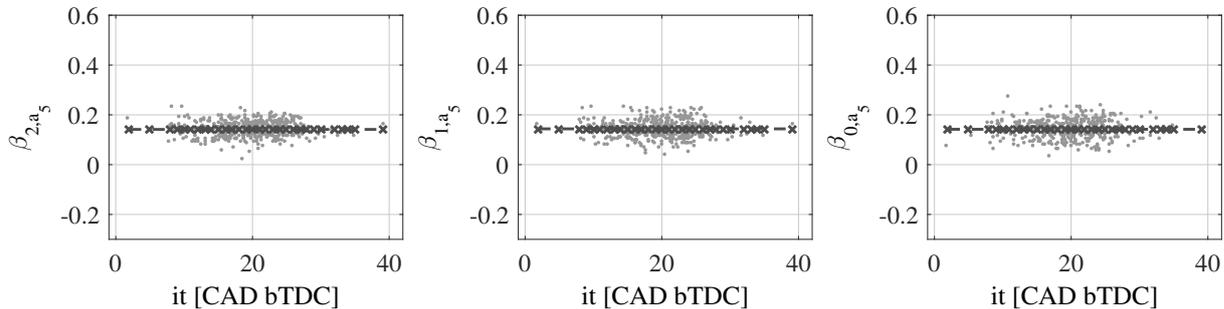
Figure 5.38: Parameters β_{2,a_5}, β_{1,a_5} and β_{0,a_5} depending on ignition timing (it).

The schema below shows the B matrices for both investigated engines. While the left matrix depicts Engine 1, the right matrix shows Engine 2, which is used to validate the model. All values are rounded to three decimal places and reveal a relatively large deviation from the respective start value of the optimization loops 0.1429. The sum of each column must equal 1. According to the definition of B, the left column contains the β_2 values, the middle column contains the β_1 values and the right column contains the β_0 values. The row assignment corresponds to the numbering of the model input parameters with $j = \{1,2,3,4,5,6,7\}$. To illustrate the differences between the two engine, Figure 5.39 indicates the percentage by which the individual factors deviate from one engine to the other with respect to the mean value, for all model input parameters.

$$B_{Eng1} = \begin{pmatrix} 0.1446 & 0.1447 & 0.1424 \\ 0.1455 & 0.1487 & 0.1471 \\ 0.1436 & 0.1388 & 0.1440 \\ 0.1460 & 0.1449 & 0.1452 \\ 0.1414 & 0.1437 & 0.1419 \\ 0.1487 & 0.1440 & 0.1436 \\ 0.1469 & 0.1343 & 0.1344 \end{pmatrix} \quad ; \quad B_{Eng2} = \begin{pmatrix} 0.1451 & 0.1446 & 0.1420 \\ 0.1416 & 0.1449 & 0.1394 \\ 0.1504 & 0.1401 & 0.1442 \\ 0.1483 & 0.1464 & 0.1467 \\ 0.1508 & 0.1459 & 0.1403 \\ 0.1460 & 0.1475 & 0.1433 \\ 0.1452 & 0.1416 & 0.1435 \end{pmatrix} \qquad (5.20)$$

Since the weighting factors of the *B* matrices are iteratively determined using an optimizer, and lack a physically analyzable basis, only the most important differences between the two engines or their *B* matrices are to be discussed. For the weighting factors β_2 and β_1 Matrix B_{Eng1} shows significantly higher values for *voa* and significantly lower values for the pressure in the intake manifold p_{IM}. Both values show how internal residual gas affects the maximum of the efficiency curve. However, it is not possible to infer whether higher or lower values exert a higher or lower influence, respectively. This deviation or difference is nevertheless logical, since Engine 2 is more robust with regard to residual gas influences due to the higher turbulence level and the increased compression ratio. Regarding the deviations at β_0, the engine speed and the fuel?air ratio deserve particular emphasis.
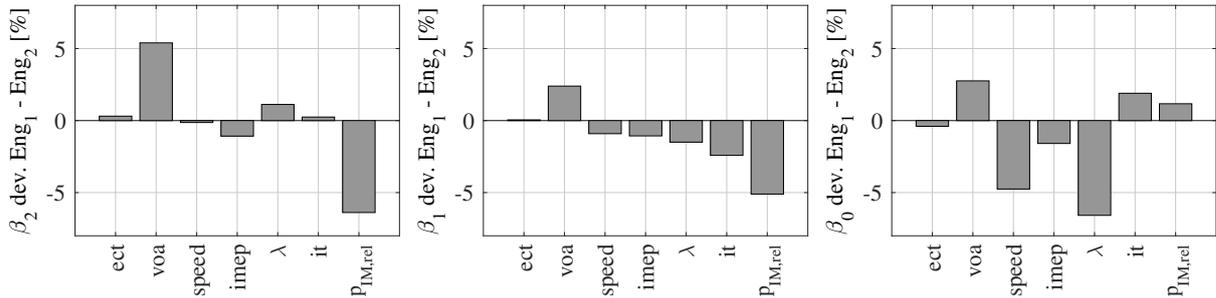
Figure 5.39: Percentage deviation between Engine 1 and Engine 2 of β_2, β_1 and β_0 for all model input parameter a_j.

In the following sections, differently specified results are used to evaluate model?s quality and whether it functions properly. Similarly to Figure 5.4, results from Engine 1 are presented in Figure 5.40. Measurement points from approximately 1000 EGR variations within the parameter space of the model are used as a basis for comparison. In each case, the decisive points of the efficiency curve are considered, i.e. the maximum efficiency $\eta_{ind,EGR,max}$ in the right-hand diagram and the associated EGR rate $x_{EGR,max,tol}$ in the left-hand diagram. In both diagrams, the simulated value is placed on the x-axis while the corresponding value from the measurement is placed on the y-axis. The perfect simulation result is therefore located at the bisector of the diagrams. In addition, scatter bands of +/-5 % are included.

With very few exceptions, the simulated values of the maximum efficiency η_{ind} lie within the marked scatter band or even within an imaginary much narrower scatter band of +/-2 %. Despite the performed normalization of the efficiency without external EGR, these results are very satisfactory and prove the model?s validity very reliably. The values for the EGR tolerance $x_{EGR,max,tol}$, i.e. the EGR rate
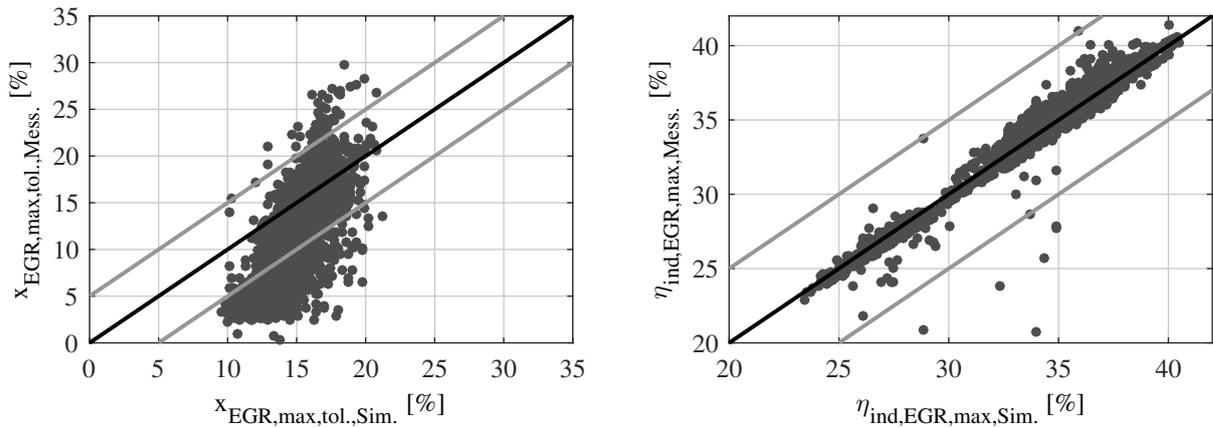
Figure 5.40: Reproduction quality of the modeled EGR-rate $x_{EGR,max,tol,Sim.}$ in comparison to the measured value $x_{EGR,max,tol,Mess.}$ and reproduction quality of the modeled maximum indicated efficiency $\eta_{ind,EGR,max,Sim.}$ in comparison to the measured value $\eta_{ind,EGR,max,Mess.}$ for Engine 1.

associated with the maximum efficiency, are less well calculated. Although a large proportion of the measurement points still lie within the scatter band, the model particularly overestimates those values which are measured with EGR tolerances of 10 % and below. It must be kept in mind that as the EGR variations in the measurement are gridded in steps of 2.5 % external EGR rate, the accuracy cannot be higher in the simulation yet from the beginning. Moreover, the EGR tolerance, which is identified with a COV_{imep} over 3 % and consequently with an increase in cyclic fluctuations in the measurement, is subject to a certain level of statistical scatter and inaccuracy, as described, for example, by Aydin in [5]. In this context, the accuracy of the model is considered to be adequate. Though, some potential for improvement is seen especially in the mapping width and the sensitivity, i.e. that the range of the calculated values becomes wider, in the way that the measured values do.

Figure 5.41 below is intended to illustrate whether the deviations in the EGR tolerance can be assigned to a specific model equation or the associated model input variable. For this purpose, all model input variables are plotted normalized over the corresponding parameter space on the x-axis and the relative deviation between the measurement and the simulation is plotted on the y-axis. Thus, all model input variables or their influence on the calculated values can be directly compared with each other. The curves represent trend lines. The diagram on the left again deals with the EGR tolerance, while the one on the right deals with the maximum efficiency.

Both diagrams confirm the findings from Figure 5.40. While the curves for the maximum tolerable EGR rate show considerably larger deviations depending on the model input variables and vary strongly over the value range, they exhibit only small deviations over the entire value range and are virtually independent of the model input variable for the indicated efficiency $\eta_{ind,EGR,max}$. Deviations that are constant over the value range are less severe and may usually be remedied by a constant shift, while deviations that vary over the value range to a greater or lesser extent indicate that the model has shortcomings. In the lower value range, the valve overlapping area *voa*, the fuel-air ratio λ and the engine load *imep* are particularly striking, as is the ignition timing *it* in the upper range. The finding that the calculated EGR tolerance mainly deviates at measured values below 10 % emphasizes the relevance of the engine load and the ignition timing, given that such comparatively small EGR tolerances are
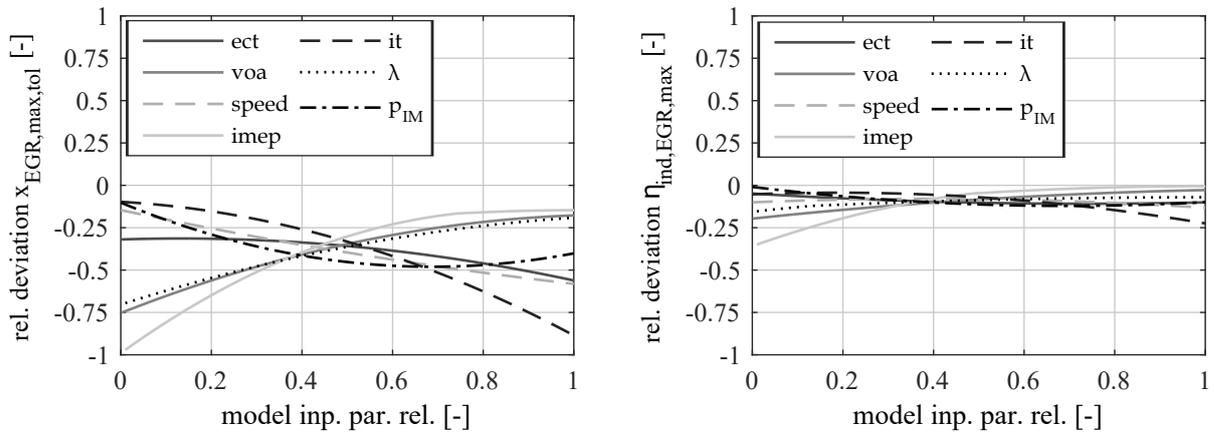
Figure 5.41: Relative change in model outputs $x_{EGR,max,tol,Sim.}$ and $\eta_{ind,EGR,max,Sim.}$ standardized over the relative parameter space of each model input variable.

particularly associated with the range of small loads, and these in turn with early ignition angles. The latter can be found close to 1 within the normalized value range.

The range in which the model reveals certain weaknesses can thus be narrowed down. However, considering the measurement program used, the EGR tolerance in conjunction with the increase in cycle-to-cycle fluctuations, and the chosen polynomial approach, improving the model in these areas would require considerable effort and a more complicated approach. Accordingly, the question arises of whether the given model quality, while adequate under these circumstances , is still sufficient for a feasible application. Before this issue is discussed further, the results for Engine 2, demonstrating the highly important model property of applicability to other engines, are presented below.

Similarly to Figure 5.40, Figure 5.42 shows the measured values of the EGR tolerance and the maximum indicated efficiency plotted against the values calculated in the model. In regard to the indicated efficiency η_{ind}, the model is confirmed as being very accurate for Engine 1, with only slightly overestimated calculated values across the entire value range. The calculated EGR tolerances also show promising results. The deviations from the measured values are even slightly smaller than for Engine 1. This confirms the tendency for larger deviations for small EGR tolerances and a smaller scatter in the case of the calculated values compared to the measured values.

The question of the *B* matrices' relevance and influence is still unresolved. For this purpose, below, based on Engine 2, Figures 5.43 and 5.44 show the calculated model results when Matrix *B* is omitted, and the results when Engine 1's Matrix *B* of is used for Engine 2 instead of its own Matrix *B*. In each case, the results with Engine 2's Matrix *B* are included as a dark scatter plot in the figure as a basis for comparison. The results with the matrices swapped or without a Matrix *B* are displayed as a light scatter plot.

If the efficiency curves are calculated unweighted, i.e. without Matrix *B*, there are significant deviations between the measurement and the model, especially with regard to the EGR tolerance. The calculated results show hardly any scatter. This is due to the fact that individual influence equations partially neutralize each other in the absence of any weighting. Likewise, the maximum indicated efficiencies from the model show an increased deviation from the measurement. Assuming that the
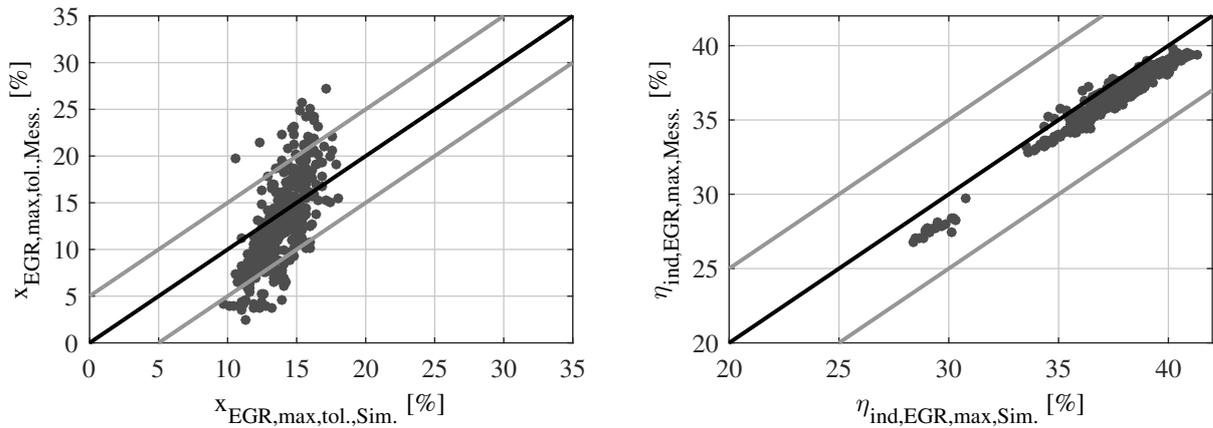
Figure 5.42: Reproduction quality of the modeled EGR-rate $x_{EGR,max,tol,Sim.}$ in comparison to the measured value $x_{EGR,max,tol,Mess.}$ and reproduction quality of the modeled maximum indicated efficiency $\eta_{ind,EGR,max,Sim.}$ in comparison to the measured value $\eta_{ind,EGR,max,Mess.}$ for Engine 2.

indicated efficiency of the operating point without external EGR is known, except for other minor deviations there is essentially a parallel shift compared to the values without Matrix *B*.

The deviation of the calculated values from the measured values with respect to the EGR tolerance is significantly smaller when using the wrong Matrix *B* (that of Engine 1) than without using any Matrix *B*, but it is still noticeable. For the maximum efficiency, a parallel shift can again be observed owing to the above-mentioned reasons. In fact, the deviation is slightly larger than without any weighting. The use of a non-motor Matrix *B* therefore instead leads to a parallel shift in the wrong direction.

The use of an engine-specific *B* matrix is undoubtedly necessary. Since the formation of the *B* matrix is relatively simple and the associated measurement program is significantly less complex than that for model generation and especially than the measurement of all EGR varitions, the effort can be justified without further ado.

Now that the excellent and virtually unrestricted applicability of the model to a new engine with different characteristics has been demonstrated, the question posed above is to be addressed once again as to whether the model quality is sufficient for a reasonable application despite the deviations in EGR tolerance, especially in areas with small EGR tolerances. Apart from the compelling applicability to other engines, two main reasons argue in favor of this. First, the quality shown above is only achieved on the basis of seven model input variables that are constantly available during engine operation. The computation time, complexity and effort are all at a minimum. Second, and much more crucially, it is not so much the absolute values as the relative values that are significant. How do the model and the measurement relate to each other? The following figures serve as an answer.

The results of the model and the measured values of the EGR tolerance (on the left) and the maximum efficiency of the corresponding EGR variation (on the right) are depicted when one single model input parameter is varied. All other model input parameters are constant and unchanged throughout the parameter variation. Variations in the engine speed ω, the valve overlapping area *voa* and of the crank angle of the 50 % mass fraction burnt point *mfb*50 are shown as examples. The remaining variables
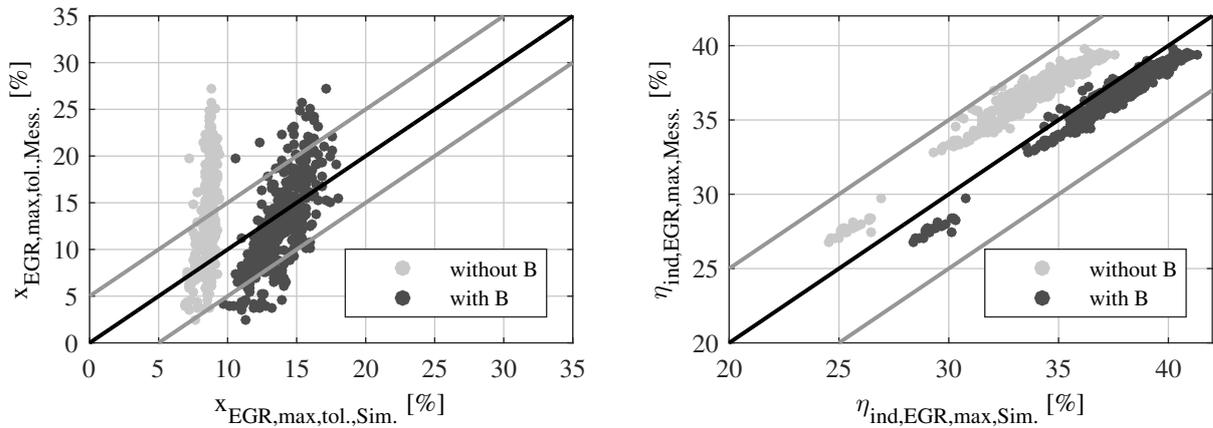
Figure 5.43: Reproduction quality of the modeled EGR-rate $x_{EGR,max,tol,Sim.}$ in comparison to the measured value $x_{EGR,max,tol,Mess.}$ and reproduction quality of the modeled maximum indicated efficiency $\eta_{ind,EGR,max,Sim.}$ in comparison to the measured value $\eta_{ind,EGR,max,Mess.}$ for Engine 2 with and without implementation of Matrix B_{Eng2}.

reveal analogous results and findings. They can be found in Appendix A.3. All variations are assigned to Engine 2 to emphasize again the model's transferability to other engines.

In Figure 5.45, the above-mentioned values are plotted as a function of the engine speed. It is again emphasized that only the speed is changed for all three measurement points depicted. The calculated efficiency maxima reflect the course of the measurement very well with a slight offset. The offset is within the specified accuracy range. The course of the EGR tolerance is also predicted well by the model. Although there is an offset again, the qualitative curves agree very well with the measurement. The latter is the key information in the case of a change in the vehicle speed, for example.

When the valve overlapping area is varied, the good agreement of the efficiency curves is likewise striking. Although the values have an offset, their relative course is almost identical. In contrast, the measurement and calculation of the EGR tolerance behave differently. While the measurement jumps between low and high EGR tolerances despite the monotonic slope of the prescribed valve overlapping area, the model shows hardly any signs of an influence. The values from the model lie approximately between the jumping measured values, as a mean value. This highlights a crucial difficulty in modeling the EGR tolerance. Since it is subject to a certain statistical scatter and measured values are of course always subject to errors, it is not possible to achieve 100 percent model accuracy. Rather, the model needs to function in a way that continues to map the trend of the measurement as accurately as possible. The graph proves that this works well. The model behavior even has advantages during operation or in the application in the vehicle, because a customer-friendly application always needs some kind of wear and tear and cannot follow the non-rounded measurement trend. This issue will be discussed in more detail in the Section 5.5 below. Though, the fact that the calculation and measurement do not completely coincide here does not have any serious disadvantages. Rather, it at least partly explains the deviations between the calculated EGR tolerance and the measured EGR tolerance which were observed and discussed above.

Finally, Figure 5.47 shows the results for the crank angle of the 50 % mass fraction burnt point $mfb50$. The maximum efficiency is mapped very well. The EGR tolerance is also calculated well by the model.
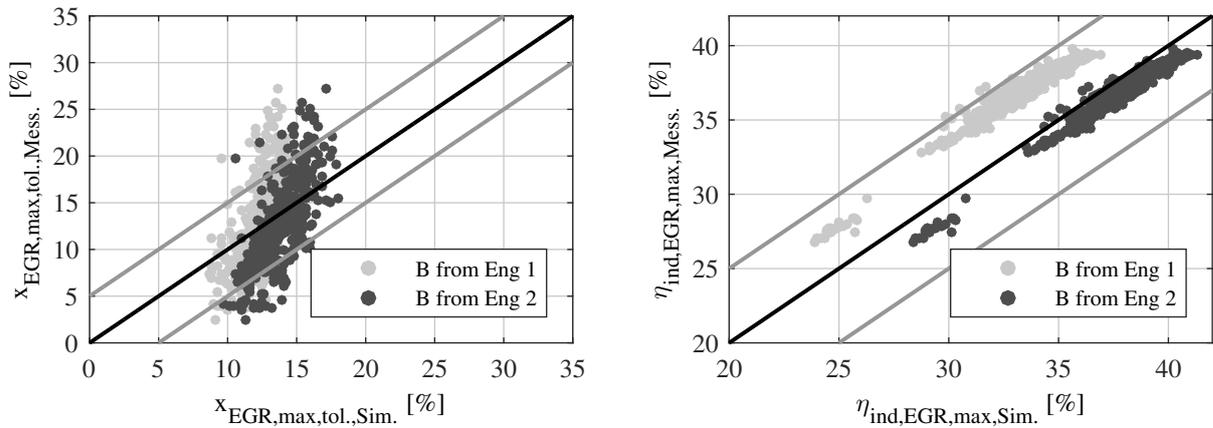
Figure 5.44: Reproduction quality of the modeled EGR-rate $x_{EGR,max,tol,Sim.}$ in comparison to the measured value $x_{EGR,max,tol,Mess.}$ and reproduction quality of the modeled maximum indicated efficiency $\eta_{ind,EGR,max,Sim.}$ in comparison to the measured value $\eta_{ind,EGR,max,Mess.}$ for Engine 2 with implementation of Matrices B_{Eng2} and B_{Eng1}.
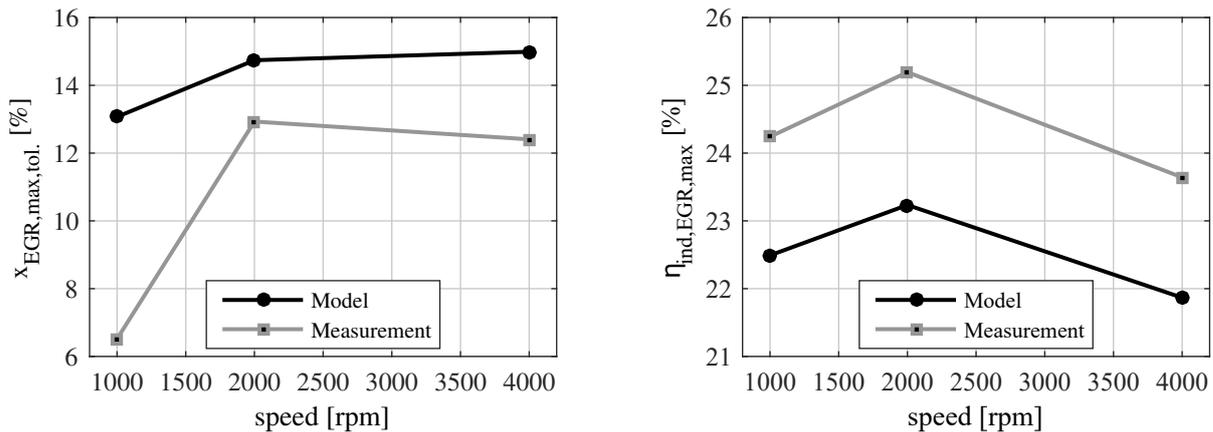
Figure 5.45: $x_{EGR,max,tol}$ and $\eta_{ind,EGR,max}$ depending on the engine speed @2 *bar* engine load, all other model input parameters also constant.

The underestimation at $mfb50$=8 *CAD aTDC* is a deviation to the safe side during operation and is also within the targeted and specified range.

Figures 5.48 to 5.51, providedetailed diagrams of EGR variations between the measurement and model. The principle within the individual diagrams is the same as in Figures 5.45 to 5.47. Except for the EGR rate, only one model input parameter is varied, while all other parameters are kept constant. Since all the information is contained in one diagram in this case, i.e. the exact curve as well as the EGR tolerance and the maximum efficiency, the results for Engine 1 are shown on the left-hand side of the diagrams and the results for Engine 2 on the right-hand side. To further improve the understanding of the model accuracy, the left and right diagrams illustrate different operating points. In order to have all model input parameters illustrated once in this framework, the results of the parameters that are omitted above are shown. The remaining figures can be found in Appendix A.3.
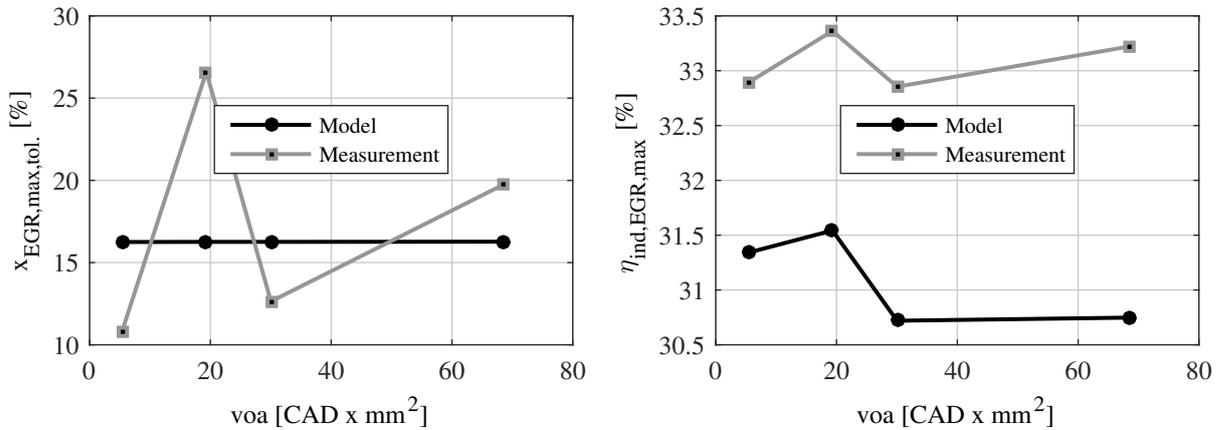
Figure 5.46: $x_{EGR,max,tol}$ and $\eta_{ind,EGR,max}$ depending on the valve overlapping area @6 *bar* engine load and @1000 *rpm* engine speed, all other model input parameters also constant.
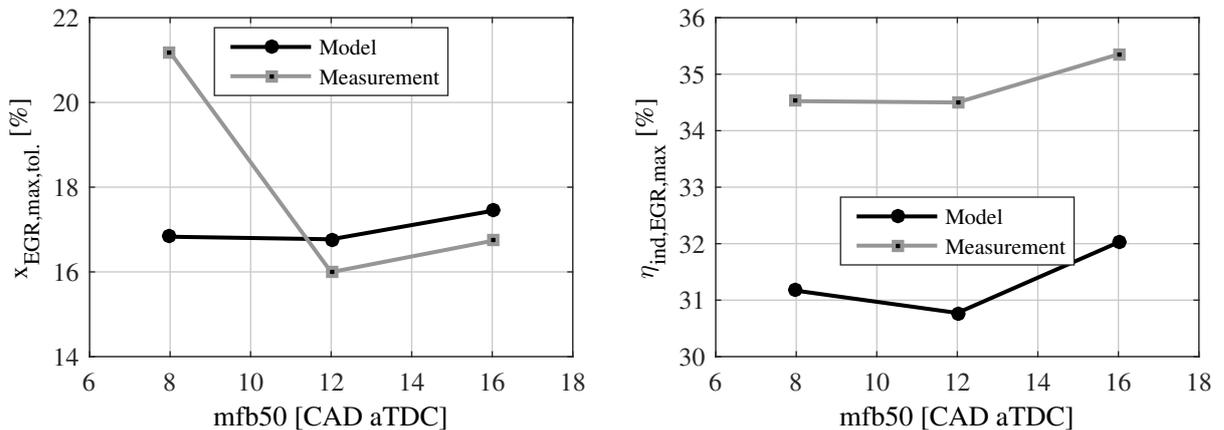
Figure 5.47: $x_{EGR,max,tol}$ and $\eta_{ind,EGR,max}$ depending on the *mfb*50 @6 *bar* engine load and @1000 *rpm* engine speed, all other model input parameters also constant.

All diagrams contain the course from the measurement as a solid line and the course from the model as a dashed line. The assignment to the operating point is made on the basis of the line color, indicated in the legend. In addition, the maximum efficiency value is highlighted by a triangle for the measurement and by a circle for the model.

Figure 5.48 shows a variation in the engine load, with both diagrams containing three different operating points and the EGR variations carried out at each of them. In all cases, the efficiency curves are very well reproduced by the computed polynomial. Furthermore, and much more importantly, the trend of the maximum efficiency is also correctly predicted by the model in all cases.

A variation in the exhaust cam timing is presented in Figure 5.49. For Engine 1, three operating points are plotted, and for Engine 2, two are plotted. The mapping accuracy is very good for Engine 2, both in terms of the maximum of the curve, and of the curve leading up to it. For Engine 1, the curves themselves are slightly overestimated by the model progression. Nevertheless, the efficiency maximum and the trend between the operating points are predicted very well.
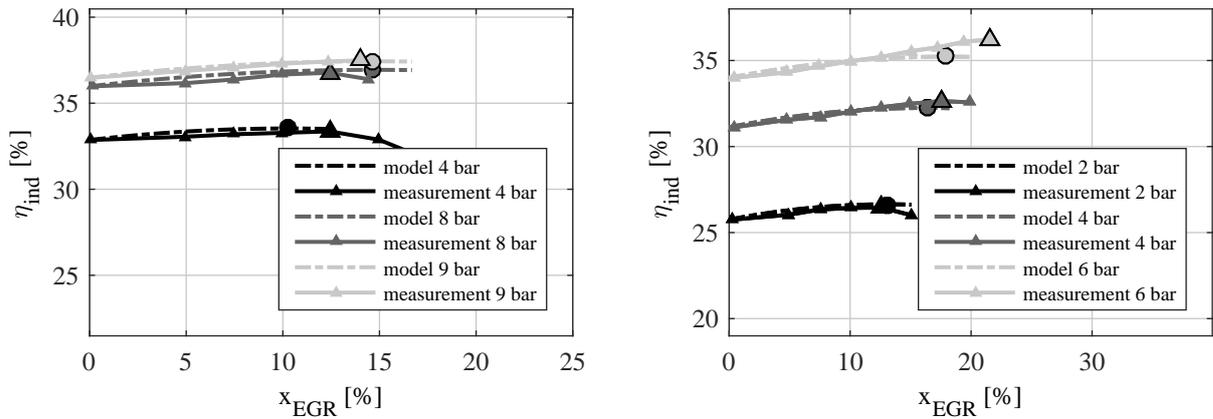
Figure 5.48: Comparison of measured and modeled data - *imep* variation for Engine 1 @3500 *rpm* (left-hand side) and for Engine 2 @1500 *rpm* (right-hand side, all other model input parameters also constant).
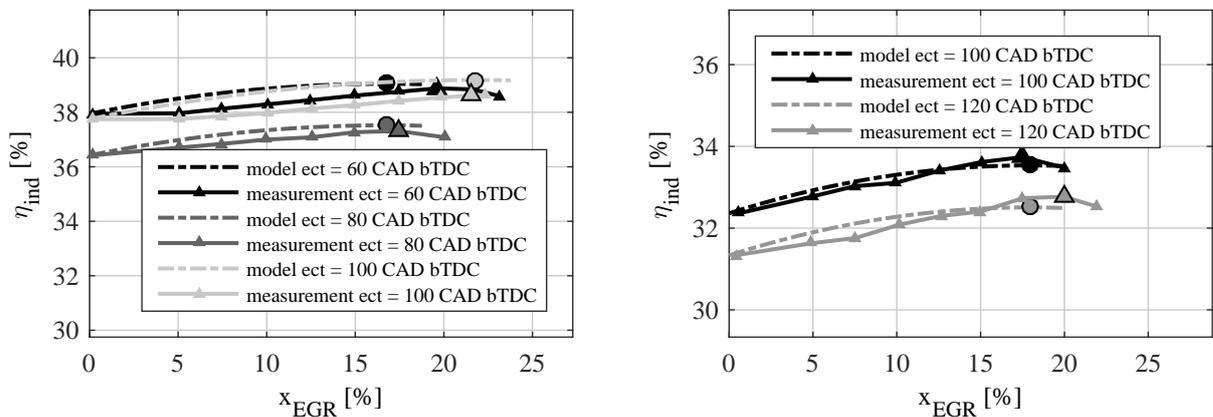
Figure 5.49: Comparison of measured and modeled data - *ect* variation for Engine 1 @2000 *rpm*, 10 *bar* (left-hand side) and for Engine 2 @1500 *rpm*, 4 *bar* (right-hand side, all other model input parameters also constant).

For the variation in the intake manifold pressure, Figure 5.50 shows two operating points and their efficiency curves depending on the EGR rate for both engines. While the maxima are represented very well for Engine 2 and the curve is slightly overestimated by the model, the agreement for Engine 1 is very good with regard to both criteria.

Finally, Figure 5.51 presents a variation in the fuel-air ratio, including three different λ values for both engines. For Engine 2, the mapping accuracy is excellent yet again. As mentioned in [59], an overestimation of the EGR tolerance can occur during lean engine operation, as shown in the left-hand side of the figure for Engine 1. Nevertheless, even there, the efficiency curves and the trend of the associated maxima are predicted very well by the model.

This in turn shows how well the model predicts the trends from the measurement in particular, but also how closely the trajectories match up to the maximum of the curve, and consequently how suitable the polynomial approach is. Before a suggestion for a model-based operating strategy is provided, Table 5.11 finally summarizes the strengths and drawbacks of the model. This is important to better
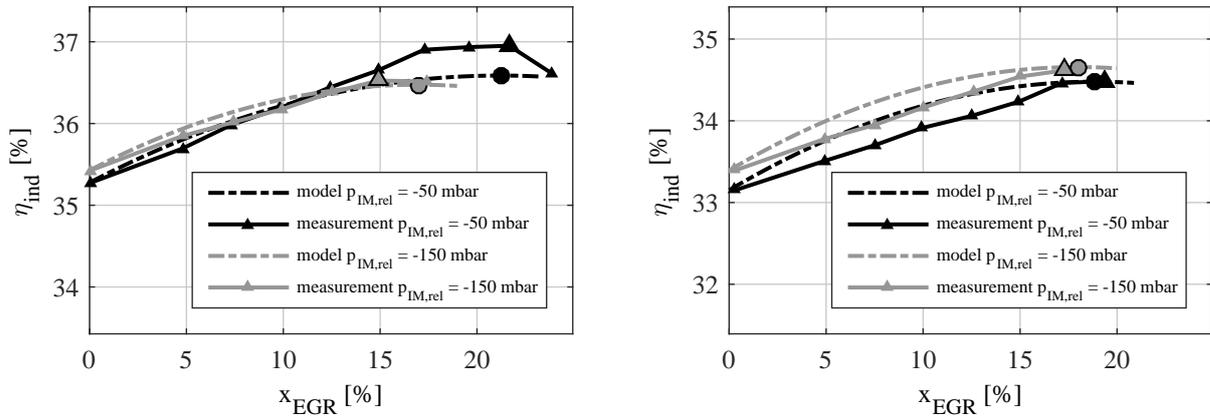
Figure 5.50: Comparison of measured and modeled data - p_{IM} variation for Engine 1 @1000 *rpm*, 6 *bar* (left-hand side) and for Engine 2 @3000 *rpm*, 5 *bar* (right-hand side, all other model input parameters also constant).
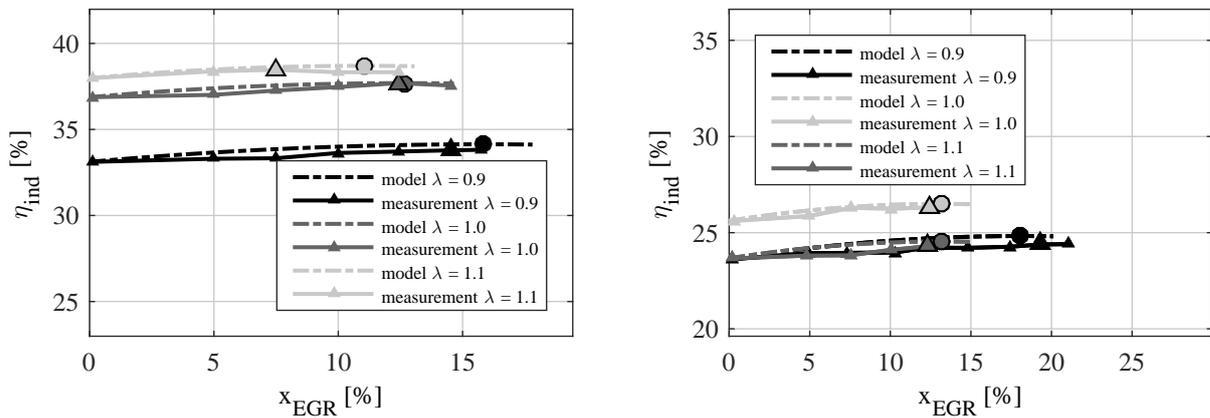
Figure 5.51: Comparison of measured and modeled data - λ variation for Engine 1 @3000 *rpm*, 8 *bar* (left-hand side) and for Engine 2 @1500 *rpm*, 2 *bar* (right-hand side, all other model input parameters also constant).

understand and justify the proposed model application or operating strategy in the following chapter. It also concludes by answering the question of whether the model is adequate for the application.

From the table, the following issues should be highlighted again. Even if the EGR tolerance might be calculated more accurately, the advantages of a good prediction of the relative behavior together with a very accurate calculation of the maximum efficiency are obvious. Thus, it undoubtedly offers the potential to implement significant improvements over previous models, such as those proposed by Styles et. al. in [92] or the neural network-based one by Siokos [86]. As a model, it is always superior to measurement with the advantages in this respect mentioned in Table 2.3. Especially in combination with a smart operational strategy, the advantages can come to the fore. Such a strategy will be provided in the following section.

Table 5.11: Model strengths and shortcomings.

Model strengths	Model shortcomings
+ very high accuracy in calculating the maximum efficiency	- slight deviations in calculating the stand alone EGR tolerance
+ fairly accurate modeling of the relative behavior depending on the model input variables, both for the maximum efficiency and for the EGR tolerance	- normalization of the indicated efficiency without EGR required
+ the model levels out outliers $\rightarrow$ less application effort	
+ easy and reliable implementation, also for other engines	
+ very low computation times, using simple mathematical operations such as addition and multiplication within the polynomials	
+ physical basis to improve the understanding of how the model functions	

5.5 Suggestion for a model-based operating strategy

This section provides a proposal for implementing the model as part of a particular operating strategy. Parts of the section have previously been published by Langmandel et al. in [58], which further emphasizes the novelty character.

Figure 5.52 introduces a schematic diagram of the operating strategy. The legend for the individual function blocks is provided at the bottom left of the figure. The operating strategy is divided into a total of four different function blocks. The *query* function block is marked in the figure with a black border if it has a temporal component. The model developed in this work is required and used in all *action* and *query* blocks. The flow chart of the operational strategy is thus almost exclusively based on the model developed here.

Starting from the current operating conditions defined by the model input parameters, the first step is to evaluate whether the currently present cyclic fluctuations are below the permissible limit, i.e. whether COV is less than or equal to the permissible value COV_{tol}. Two things are assumed in this process. First, it is assumed that the cyclic fluctuations are known during normal operation, i.e. operation without external EGR. Second, when operating with external EGR, the statement again applies that the EGR tolerance is exceeded beyond the point of maximum efficiency in an EGR variation. With these two statements, the first *query* can be answered either with *yes* or *no*.

If the *query* is answered with *no*, the *action* block, which is used several times in the flow chart, is processed. This block is not explained in each case again but instead an explanation is given once at this point. In the *action* block, the model input variables and/or the EGR rate are adjusted such that the requested condition $COV \leq COV_{tol}$ is fulfilled. One prerequisite is, of course, that this is not noticeable to the driver. The engine load and the engine speed are thus excluded as adjustment variables. However, if the model is used in the context of the proposed operating strategy, the load and speed are essential in

Figure 5.52: Flow chart of the model strategy.

determining the actual conditions, alongside other possible conditions. The adaptation of the remaining model input variables, or the EGR rate, is done under two main premises in this block. As already mentioned, one is that the *COV* decreases below the allowable value, which is the trigger premise for this *action* block. The second is that the efficiency is optimized.

A factor of time and efficiency gain, which is still to be defined, is used to select the one state out of all the possible operating states that best fulfills the trigger premise. The function of the optimizer is thus designed such that the trigger premise is always fulfilled and a maximum gain or a minimum loss of efficiency is implemented in the shortest possible time. The output of the action block is a new set of operating point parameters consisting of the model input parameters and the EGR rate. All information necessary for the *action* block and for the optimizer is provided by the model developed in this work.

Just as the *action* block generally needs to access information about the duration of the operating state change, so does the next *query* block, which asks whether possible operating states are covered that are caused by the driver?s request or by the driving conditions. It is a rather abstract *query* block. The *query* is used to ensure that no unforeseen operating states occur, for example due to a fast tip-out as in the case of the worst case load step presented in Section 2.1. Similarly to the derivation of this worst-case load step, the possible trajectories for the individual operating parameters are deposited depending on the operating state. They are known from measurements. The trajectories of the EGR rate can be generated via a simplified filling model, as illustrated in Section 2.1. To align the EGR rate from the filling model to actual conditions, it is recommended to install an O_2 sensor in the intake manifold in addition to the standard sensors, as presented by Liu et al. in [64] or by Siokos in [86]. Even if the

first step within this *query* block is a temporal question, the model information is still required to know how the efficiency and the EGR tolerance react in the possible and queried operating states.

If this *query* is answered with *no*, then an *action* block is again run through, which besides the optimization of the efficiency, this time involves the trigger premise of the excluded operating states. The operating state must therefore be adapted such that future operating states exceeding COV_{tol} are avoided. Again, this is done aiming at the best possible efficiency.

Though, if this *query* is answered with *yes*, the next and last *query* in this flow chart branch is processed, which is the question of whether the efficiency in the current operating state is optimal. If that question can be answered with *yes*, the current operating state corresponds to the target operating state and the flow chart is not run again until the driver changes the actual operating state. If the efficiency is not optimal, a final *action* block is executed with the sole purpose of optimizing the efficiency by changing the EGR rate or changing the other operating point parameters. The flow chart is run through again with the resulting state.

The branch of the flow chart that follows the first two *action* blocks remains to be discussed. It is a backup branch and ensures that the engine operates without fault in all situations. The operating state calculated from the respective *action* block is first checked in a *query* to see whether there is currently enough time to set the desired operating state. It thus involves a double *query*, as the time component is already considered in the *action* block.

A *yes* answer to the *query* block sets the new operating state and the flow chart is run through again from the beginning. If, on the other hand, it is answered with *no*, the operating state is adjusted by an immediate reduction of the EGR rate in an *action* block. If there is not enough time for this either, the subsequent *query* block leads to a restriction of the operation for the driver: this is the worst-case scenario. The restriction is chosen such that, for example, speed or load gradients are limited unless it is absolutely necessary for the driving situation. Misfire-free operation can thus be ensured. If the time required to reduce the EGR rate to the target corridor is sufficient, the flow chart is again computed from the beginning.

The flow chart and the individual blocks are calculated continuously and repeatedly in real time. This is only possible due to the model's very simple structure, which requires neither a high computational effort nor a high computational time. Besides that, the focus is on ensuring that the model, and thus the operating strategy, are viable with the sensor data and values that are readily obtainable and known in the case of conventional production engines.

An implementation of the model is beyond the scope of this work. However, it provides an excellent basis for further work on the topic and for implementing the model in a test vehicle. The implementation of an increased efficiency potential would be possible directly in the vehicle without any further effort. It demonstrates distinct advantages and can only be implemented with the help of a model such as the one developed in this work. A holistic implementation of LP EGR even during dynamic vehicle operation can thus be ensured without major effort but enables much of the potential efficiency from stationary operation to be tapped.

6 Summary and outlook

This thesis deals with the problem of an efficiency-optimized implementation of LP-EGR in turbocharged spark-ignition engines. The ongoing discussions regarding ever stricter CO_2 limits and the increasing electrification of the automotive sector demand a further reduction of both the fuel consumption and the associated CO_2 emissions. Exhaust gas recirculation technology, which has been largely untapped in the gasoline engine sector, still offers usable potential in this regard. During stationary operation, the advantages are well known and have been demonstrated and widely discussed by a wide range of literature sources. LP-EGR has proven to be particularly advantageous for turbocharged spark-ignited engines. However, dynamic operation is problematic when it comes to the actual implementation, since the long gas path between the discharge point and the combustion chamber results in a long dead time. While this issue has been addressed in various literature sources with individual solutions, no work is known so far that addresses the implementation of dynamic operation using LP-EGR explicitly. Precisely this is the subject of this study.

To this end, the problem of dynamic operation is first explained further in order to better understand the reasons for the difficulty and the underlying core issues. A central concern is seen in the inhomogeneous distribution of the EGR rate in the efficiency-optimized stationary engine map. However, the inhomogeneous distribution of the EGR tolerance in the steady-state map is at least as essential, with load dependence being identified as particularly problematic. This is mainly due to the fact that the load in the combustion engine can be changed with very high gradients. It becomes clear that the EGR tolerance essentially alters along with the load, i.e. rapidly, and the EGR rate alters very slowly due to the long gas path. The crucial, main control levers identified to improve dynamic operation including LP EGR are reducing the in-cylinder EGR rate faster, and influencing the EGR tolerance using an in-depth knowledge of that rate . The criticality and frequency of passing through different EGR tolerance ranges is further exemplified by the trajectories of a WLTC driving cycle visualized in the steady-state engine map.

The thesis introduces two approaches to solve the problem: a hardware-based and a software-based one. The advantages and disadvantages of these two approaches are elaborated and discussed. Among other things, the software approach is characterized by a physical basis for a better understanding of the system. No further hardware costs arise and the packaging does not cause any further difficulties. However, software- and model-based solutions always contain a certain degree of inaccuracy, and the stationary potential cannot be fully exploited with a software solution alone. In contrast, a hardware-based solution can exploit the full steady-state potential. If the system is chosen wisely, there may be other advantages for the customer besides the mere ability to rapidly reduce the in-cylinder EGR rate. The additional costs, the increase in the system's degrees of freedom and the packaging issue are the main disadvantages of a hardware-based solution.

Before the two approaches are presented in depth, a detailed presentation of the test engine and the chosen test configuration is provided. Special attention is paid to two aspects. On the one hand, it is explained why the exhaust gas is taken upstream of the TWC and not downstream of it. In the comparative measurement of the two configurations, the possibility of converting higher EGR rates and, even more importantly, the higher efficiency potential due to unburned fuel residues that have not

yet been converted in the TWC are shown to be advantageous for the extraction in front of the TWC. On the other hand, attention is paid to the EGR measurement technique and an error analysis of the systems used is introduced. As well as assessing the EGR rate based on the CO_2 ratios in the manifold and the exhaust gas, an O_2 sensor is applied and it is also assessed based on the O_2 ratios. Although the measurement error is slightly higher in this case, the sensor is suitable for determining the EGR rate in the manifold, i.e. close to the combustion chamber, virtually without delay.

In addition to the measurement setup, the most important evaluation tools are explained. While the TPA described is still a standard procedure, there is a more detailed explanation of the correction terms for compensating the work of the roots loader, which is included in the measurement setup for implementing comprehensive EGR variations even at higher loads, and for adjusting the EGR rate if the fuel?air ratio deviates from the stoichiometric one. The approach for correcting the roots loader uses approaches known from the literature for dividing the engine process into a gas exchange part and a high-pressure part. An approach based on the method of $360°$ *integration* is chosen to separate the gas exchange and to compensate for the supplied compressor work before calculating the indicated efficiency. Evidence that this correction term works properly is provided by a series of measurements. When correcting the EGR rate at $\lambda \neq 1$, a distinction must be made between a lean and a rich mixture. The correction term for a lean mixture can easily be derived analytically. For rich mixtures, on the other hand, an approach from the literature must be used to estimate the amount of unburned fuel still present in the recirculated exhaust gas. For this purpose, an equation is selected as described by Bargende. It can be seen that the amount of correction during lean operation is significantly higher than during rich operation.

A methodology that is as simple as possible is needed to evaluate the hardware-based approach. Since a virtually infinite number of load steps can occur based on the data of a stationary engine map, and since it is extremely time-consuming to evaluate each one individually with regard to possible measures, a simplified container model is developed to address this issue. The container model does not allow absolute conclusions to be drawn but supports relative conclusions and, in addition to a very short calculation time, is also easy to apply to other engines and characteristic maps.

The defined worst-case load step is tested using various hardware configurations in addition to the Basic System. These configurations include systems which are known from the literature and others developed by the author. The portfolio includes various bypass solutions, a new so-called *Direct Boost System* with direct injection of compressed air into the combustion chamber, and a system featuring a generator for load point boosting. First, all systems are examined in a 1D CFD environment. The evaluation criterion is the time to undershoot the EGR tolerance at the target load point. For all systems, a detailed evaluation of the main weaknesses and the remaining potentials is shown. The presented Fresh Air Bypass and the Generator System are identified as advantageous with respect to this property and perform 70 % better than the Basic System. The decision to implement a hardware system for measurement on the engine test bench is based on a holistic evaluation with further various criteria, such as other customer benefits, costs and implementation effort, and a weighting of these criteria. The Fresh Air Bypass achieves the best score, is implemented in hardware and is measured accordingly.

A new evaluation methodology based on the misfire events which occur is used to analyze the measurements. This is necessary because during dynamic operation, the known characteristic values such as COV_{imep} cannot be used. The first main finding of the sensitivity analysis based on the worst-

case load step is that the Bypass System is also very advantageous in comparison to the Basic System, but still features occasional misfire events. An optimum combination of the throttle angle of the main path and that of the bypass path is identified. One difficulty lies in the need to avoid the complete closure of the main path, as this results in backflow via the compressor and the entry of an EGR air mixture into the fresh-air bypass. However, another more serious concern is the asymmetric design of the bypass and intake manifold. Improving the distribution of the fresh air and the EGR?air mixture could contribute to serious improvements. An intake system designed for the bypass solution offers further potential in this respect and leads to justified hope that the bypass will then allow almost all the stationary potential to be achieved.

Apart from an intake duct specifically designed for the bypass, two guidelines can be given for tapping further potential efficiency with the help of a hardware adaptation, and thus for solving the problem of the slow EGR path with a negative load step. First, it would be important to design the entire engine specifically for an LP EGR system at the layout stage. This means an intake path that is as small and symmetrical as possible and, if applicable, a VTG turbine or a large wastegate diameter for fast scavenging of the intake path. In addition, a compressor that can withstand pumping events is advantageous. The engine used in this thesis is a production engine complemented by an LP EGR system, so it is anything but specifically designed. On the other hand, the combination of the systems presented in this work offers further potential. Examples include the *Direct Boost System* with a scavenging bypass or the Fresh Air Bypass, likewise with a scavenging bypass, in which case the discharge point would have to be located upstream of the main throttle in order to completely isolate the main path. Furthermore, high-voltage hybrid systems offer even greater potential compared to the Generator System, which would further increase the advantages already proven.

The crucial aspect in developing a software-based approach for achieving optimally efficient dynamic operation with LP EGR is exact knowledge of the correlation of the most important operating point parameters and the associated EGR efficiency behavior, plus knowledge of the EGR tolerance. The approaches used so far are either based on simplified assumptions or require a great deal of measurement effort to obtain this information. One exception is the approach described by Siokos in [86]. It partially provides the above information, but for one it does not provide the context information and is based on a neural network. This makes the approach a black box. It does not provide the necessary insight into correlations, nor can errors and inadequacies be effectively detected and, most importantly, understood.

In this thesis, for the first time, a model is presented that combines the above-mentioned properties. With the help of a large number of EGR variations, it can be shown that the modeling of the course of the indicated efficiency over the EGR rate can be very well approximated by a second-degree polynomial. This has the major advantage that each change in one of the three polynomial coefficients can be used to physically interpret the influence on the shape of the polynomial and thus on the shape of the EGR efficiency curve. In this way, the relationship between the parameters of an operating point and the polynomial coefficients can be established and, in turn, the influence exerted by the parameters of an operating point on the shape of the EGR efficiency curve can be interpreted and analyzed.

Furthermore, it is demonstrated that a decrease in the efficiency curve of an EGR variation can be associated with a sharp increase in the COV_{imep} in more than 99 % of all measurements. This means that the EGR tolerance can be derived based on the maximum of the EGR efficiency curve. Yet again, the second-degree polynomial shows clear advantages, since it can be transformed into the so-called

point-vertex form via simple mathematical operations. In this form, the x and y values of the maximum of the curve can be easily calculated from the polynomial coefficients. Using the polynomial coefficients, it is thus possible to derive not only the curve, but also information on the maximum possible efficiency and the associated EGR rate, which is interpreted as EGR tolerance.

A total of seven variables are defined to describe the operating point. They are selected on the premise that they are known continuously during operation and that new sensors are not required. Furthermore, they are intended to describe an operating point and its efficiency as comprehensively as possible. Thus, in addition to the engine load and speed, variables are selected for describing the mixture, such as the fuel-air ratio λ, in addition to variables influencing the internal residual gas rate, such as the exhaust cam timing, and variables relating to the combustion, such as ignition timing.

The influence of the selected operating point parameters on the polynomial coefficients of the EGR efficiency curves is determined empirically. Using the method of non-linear regression, influence equations are determined for each operating point parameter on each of the three polynomial coefficients of the EGR efficiency curve. For this purpose, a measurement program is set up using DoE methods. The result is a matrix of size 3x7, named Matrix A. The curves of the equations in A are shown in separate sections for each of the operating point parameters and all of them are checked for plausibility. For this purpose, any other variable from the measurement or the post-evaluation is used; these can be plotted depending on the respective operating point parameter. A complete loss analysis is performed for each single operating point of each EGR variation. This breadth of information is novel and gives the model equations in A a physically justifiable basis.

As not every operating point parameter has an equal influence on the polynomial coefficients of the EGR efficiency curve and engine specifics must also be taken into account, a second matrix, the weighting Matrix B, is introduced. It has a size of 7x3 and, in contrast to Matrix AA, does not contain any influence equations but only scalar weighting factors. The factors are determined for each specific engine with a significantly reduced measurement program in comparison to the uniquely determined Matrix A and with the help of optimization equations. Using the scalar product of the two matrices A and B, the corresponding EGR efficiency curve can be calculated for each operating point or its set of operating point parameters, enabling information on the maximum possible efficiency and the EGR tolerance to be provided.

The validation of the EGR efficiency curves calculated within the model against measurement data from EGR variations demonstrates very satisfactory conformity. Notably, the maximum efficiency is in excellent agreement: both the absolute and the relative values are predicted very well. The relative behavior is also reproduced very well with respect to the EGR tolerance. The absolute values are also well mapped to a large extent, although there are individual values which deviate by more than 5 %. However, this is not a crucial drawback for the later application of the model, since the relative values are what matters. This is due to the assumption made that the engine behavior without external EGR is well-known. These statements apply both to the engine used to create the model and to another engine used just for validation. In this context, the importance of the weighting Matrix B is also demonstrated.

The validation chapter of this thesis therefore proves how well the model works despite its simple basis, the short computation times as well as the low computational effort. Interrelationships have been scrutinized on the basis of the model equations and can be interpreted physically using other relevant quantities. In its simplicity, its information content and its applicability throughout the entire operating range, the model thus offers the ideal basis as an application tool or for devising an operating strategy

to achieve the optimum efficiency of LP EGR during dynamic operation. A proposal for an operating strategy of this kind, albeit without concrete implementation, is presented in the final chapter of this thesis.

For the further development of the model, the practical implementation of an operating strategy of this kind in a test vehicle is recommended. Further potential for improving the model could lie in the implementation of models for evaluating cyclic fluctuations, such as that described by Krost in [53] or by Wenig in [107]. However, the hurdle to overcome is to shorten the computation time of these models. The combination of one of the presented hardware approaches with the model or the operation strategy is also conceivable. For an additional improvement of the EGR tolerance in the low-load range, in [11], for example, Bunsen shows the suitability of multiple ignitions and an increase in the ignition energy. If these measures were to be applied, they would have to be integrated into the model as additional operating point parameters.

Bibliography

[1] Abd-Alla, G.H. (2002): *Using exhaust gas recirculation in internal combustion engines: a review*. Energy Conversion and Management, volume 99, pp. 1027–1042.

[2] Alger, Terry, Chauvet, Thierry and Dimitrova, Zlatina (2008): *Synergies between High EGR Operation and GDI Systems*. SAE Int. J. Engines, **1**2008-01-0134, pp. 101–114.

[3] Alger, Terence, Gingrich, Jess and Mangold, Barrett (2007): *The Effect of Hydrogen Enrichment on EGR Tolerance in Spark Ignited Engines*. SAE Technical Paper Series, 2007-01-0475.

[4] Alger, Terence, Gingrich, Jess and Mangold, Barrett (2011): *Cooled exhaust-gas recirculation for fuel economy and emissions improvement in gasoline engines*. International Journal of Engine Research - INT J ENGINE RES, **12**, pp. 252–264.

[5] Aydın, Kadir (2011): *Effect of engine parameters on cyclic variations in spark ignition engines*. 6th International Advanced Technologies Symposium (IATSÂ'11), pp. 57–63.

[6] Bargende, Michael, Burkhardt, Christine and Frommelt, Alfred (2001): *Besonderheiten der thermodynamischen Analyse von DE-Ottomotoren*. MTZ, Motortechnische Zeitschrift, 62, pp. 56–68.

[7] Baruah, P. C., Benson, R. S. and Balouch, S. K. (1978): *Performance and Emission Predictions of a Multi-Cylinder Spark Ignition Engine with Exhaust Gas Recirculation*. SAE Technical Paper Series, 780663.

[8] Birkigt, Andreas (2012): *Analyse von Vorentflammungsphänomenen an hoch aufgeladenen Ottomotoren mit Direkteinspritzung*. Ph.D. thesis.

[9] Boecking, Tobias (2015): *Grenzpotentiale von ottomotorischen Magerbrennverfahren hubraumverkleinerter aufgeladender Motoren mit Direkteinspritzung*. Ph.D. thesis.

[10] Bosch (2009): *Technical Customer Information: Planar Wide Band Lambda Sensor with Pumped O2 Reference*.

[11] Bunsen, Eike-Philippe (2010): *Beitrag zur Arbeitsprozessoptimierung hochaufgeladener Ottomotoren*. Ph.D. thesis.

[12] Cairns, Alasdair, Blaxill, Hugh and Irlam, Graham (2006): *Exhaust Gas Recirculation for Improved Part and Full Load Fuel Economy in a Turbocharged Gasoline Engine*. SAE Technical Paper, 2006-01-0047.

[13] Cairns, Alasdair, Fraser, Neil and Blaxill, Hugh (2008): *Pre Versus Post Compressor Supply of Cooled EGR for Full Load Fuel Economy in Turbocharged Gasoline Engines*. SAE Technical Paper Series, 2008-01-0425.

[14] Chao, Yuedong, Lu, Haifeng, Hu, Zongjie, Deng, Jun, Wu, Zhijun and Li, Liguang (2017): *Comparison of Fuel Economy Improvement by High and Low Pressure EGR System on a Downsized Boosted Gasoline Engine*. SAE Technical Paper Series, 2017-01-0682.

[15] Cloos, Lutz Kilian, Glahn, Claus, Königstein, Achim and Shin, Steve (2015): *Externe Abgasrückführung am aufgeladenen Ottomotor Â– Eine Technologiebewertung im Fahrzeug.* Wiener Motorensymposium, 36.

[16] Continental (2015): *Application Note for Intake-O2 Sensor*.

[17] Cunningham, Ralph Wayne Mich., Mich, Ross Dykstra, Pursifull, Russell, John David Oreg., Surnilla, Gopichandra Mich. and van der Wege, Brad Alan Mich. (2010): *Doppel-Drossel zur verbesserten Tip-out-Stabilität in einem aufgeladenen Motorsystem*.

[18] Diana, S., Giglio, V., Iorio, B. and Police, G. (1998): *Evaluation of the Effect of EGR on Engine Knock*. SAE Technical Paper Series, 982479.

[19] Dönitz, Christian, Wabbals, Din, Kempny, Rafael, Röhr, Carsten and Brehm, Norbert (2016): *Responseverbesserung durch Drucklufteinblasung bei aufgeladenen Ottomotoren.* Ladungswechsel im Verbrennungsmotor, 9. MTZ-Fachtagung, Stuttgart.

[20] Dulbecco, Alessio, Richard, Stephane and Angelberger, Christian (2015): *Investigation on the Potential of Quantitatively Predicting CCV in DI-SI Engines by Using a One-Dimensional CFD Physical Modeling Approach: Focus on Charge Dilution and In-Cylinder Aerodynamics Intensity..* SAE Technical Paper Series, 2015-24-2401.

[21] Eaton Corporation (2001): *Eaton M45 Datasheet*.

[22] Edwards, S.P., Grove, D.M. and Wynn, H.P. (2000): *Statistics for Engine Optimisation*. London and Bury St. Edmonds: Professional Engineering Publishing Limited.

[23] Eichlseder, H., Klütting, M. and Piock, W.F. (2008): *Grundlagen und Technologien des Ottomotors*. Springer-Verlag, Berlin Heidelberg New York, ISBN 978-3-211-47104-3.

[24] European Commission (2015): *The Paris Protocol - A blueprint for tackling global climate change beyond 2020*.

[25] European Parliament (2009): *Regulation (EG) No. 443/2009*, 04/11/2018.

[26] Fischer, Hubert and Gutzer, Ulrich (2015): *Verbrennungskraftmaschine*.

[27] Fischer, Klaus (1988): *Grundlagen der Statistik*. Gabler Verlag, 3rd edition, ISBN 978-3-322-85682-1.

[28] Fischer, Michael, Kreutziger, Philipp, Sun, Yong and Kotrba, Adam (2017): *Clean EGR for Gasoline Engines Â– Innovative Approach to Efficiency Improvement and Emissions Reduction Simultaneously*. SAE Technical Paper Series, 2017-01-0683.

[29] Fricke, F. (2007): *Untersuchungen zu aufgeladenen ottomotorischen Magerbrennverfahren*. Ph.D. thesis.

[30] Fuchs, Thorsten (2012): *Analyse unterschiedlicher Ladungswechsel-Phasing-Varianten an einem Turbo-Ottomotor mit vollvariablem mechanischem Einlassventiltrieb*. Ph.D. thesis.

[31] Gamma Technologies (2015): *GT-SUITE Engine Performance Application Manual Version 7.5*.

[32] Gärtner, U. (2001): *Die Simulation der Stickoxid-Bildung in Nutzfahrzeug-Dieselmotoren*. Ph.D. thesis.

[33] Golloch, R. (2005): *Downsizing bei Verbrennungsmotoren: Ein wirkungsvolles Konzept zur Kraftstoffverbrauchssenkung*. VDI-Buch, Springer-Verlag, Berlin Heidelberg New York, ISBN 978-3-540-27490-2.

[34] Goodman, Nelson (1979): *Fact, Fiction and Forecast*. Harvard University Press Cambridge, Massachusetts and London, ISBN 0-674-29071-2.

[35] Grandin, Börje, Ã...ngström, Hans-Erik, StÃ¥lhammar, Per and Olofsson, Eric (1998): *Knock Suppression in a Turbocharged SI Engine by Using Cooled EGR*. SAE Technical Paper Series, 982476.

[36] Gülder, Ömer L. (1984): *Correlations of Laminar Combustion Data for Alternative S.I. Engine Fuels*. SAE Technical Paper Series, 841000.

[37] Haas, Daniel (2016): *Brennkraftmaschine und Verfahren zum Betreiben der Brennkraftmaschine*.

[38] Haas, Georg (2015): *Potentiale alternativer Zündsysteme in Verbindung mit Ladungsverdünnung für aufgeladene TVDI-Brennverfahren*. Master's thesis.

[39] Hathaway, R. (2000): *Internal Combustion Engines - ENGINE MODELING, ME 468 Engine Design - Forschungsbericht*.

[40] Heikes, Henning (2014): *System- und Komponentenanalyse für hohen thermodynamischen Wirkungsgrad beim Ottomotor*. Ph.D. thesis.

[41] Heywood, J. B. (1988): *Internal Combustion Engine Fundamentals*. McGraw-Hill, ISBN 0-07-028637-X.

[42] Hoepke, Bjoern, Jannsen, Stefan, Kasseris, Emmanuel and Cheng, Wai K. (2012): *EGR Effects on Boosted SI Engine Operation and Knock Integral Correlation*. SAE Int. J. Engines, **5**2012-01-0707, pp. 547–559.

[43] Hoffmeyer, Henrik, Montefrancesco, Emanuela, Beck, Linda, Willand, Jürgen, Ziebart, Florian and Mauss, Fabian (2009): *CARE Â– CAtalytic Reformated Exhaust Gases in Turbocharged DISI-Engines*. SAE Technical Paper Series, 2009-01-0503.

[44] International Energy Agency (2016): *CO2 emissions from fuel combustion by sector in 2014*, 12/19/2017.

[45] Jiang, N., Liu, J., Zhang, X. and Cheng, X. et al. (2013): *Study on Engine Performance Influenced by External Cooled EGR*. Proceedings of the FISITA 2012 World Automotive Congress: Volume 1: Advanced Internal Combustion Engines (I), pp. 587–598, Springer-Verlag, Berlin Heidelberg, ISBN 978-3-642-33841-0.

[46] Kapus, P. and Glanz, R. (2007): *Brennkraftmaschine*, wO Patent App. PCT/AT2006/000,468.

[47] Kawabata, Y., Sakonji, T. and Amano, T. (1999): *The Effect of NOx on Knock in Spark-Ignition Engines*. SAE Technical Paper Series, 1999-01-0572.

[48] Kawamoto, N., Naiki, K., Kawai, T. and Shikida, T. et al. (2009): *Development of New 1.8-Liter Engine for Hybrid Vehicles*. SAE Technical Paper Series, 2009-01-1061.

[49] Kempny, Rafael and Rottengruber, Hermann (2016): *Direct Air Boost. Analysis of the combustion process of in-cylinder boosted SI engines*. WKM Symposium Graz.

[50] Kohn, Wolfgang (2005): *Statisitk: Datenanlyse und Wahrscheinlichkeitsrechnung*. Springer-Verlag Berlin Heidelberg, ISBN 978-3-540-26768-3.

[51] Komiyama, K. and Heywood, J. B. (1973): *Predicting NOx Emissions and Effects of Exhaust Gas Recirculation in Spark-Ignition Engines*. SAE Technical Paper Series, 730475.

[52] König, A. (1856): *Grundzüge einer Theorie der Gase*, volume 175. 10th edition.

[53] Krost, Philipp (2018): *Experimental Characterization and Quasi-Dimensional Modeling of Cyclic Combustion Variations in Spark Ignition Engines*. Ph.D. thesis.

[54] Kumano, Kengo and Yamaoka, Shiro (2014): *Analysis of Knocking Suppression Effect of Cooled EGR in Turbo-Charged Gasoline Engine*. SAE Technical Paper Series, 2014-01-1217.

[55] Kumar, N. Pradeep, Krishnan, S. Suseel Jai and Thasan, N. Sakthi (2013): *Effects of fouling in EGR Coolers in Automobiles - a Review Study*. International Journal of Innovative Research in Science, Engineering and Technology, **2**.

[56] Lake, Tim, Stokes, John, Murphy, Richard, Osborne, Richard and Schamel, Andreas (2004): *Turbocharging Concepts for Downsized DI Gasoline Engines*. SAE Technical Paper Series, 2004-01-0036.

[57] Lakhlani, Hardik, Barman, Jyotirmoy, Rajput, Karan and Goswami, Angshuman (2013): *Experimental Study of EGR Mixture Design and its Influence on EGR Distribution Across the Cylinder for NOx - PM Tradeoff*. SAE Technical Paper Series, 2013-01-2743.

[58] Langmandel, Daniel, Haas, Daniel and Orlick, Hannes (2018): *Verfahren zum Steuern einer externen Abgasrückführung eines Motors*.

[59] Langmandel, Daniel, Rottengruber, Hermann, Haas, Daniel, Orlick, Hannes and Brehm, Norbert (2018): *Different hardware approaches implementing low-pressure exhaust gas recirculation during the dynamic operation of turbocharged gasoline engines*. Automotive Engine Technologies.

[60] Langmandel, Daniel, Rottengruber, Hermann, Orlick, Hannes, Haas, Daniel and Riegger, Franziska (2018): *A physical-based approach for modeling the influence of different operating parameters on the dependency of external EGR rate and indicated efficiency*. SAE Technical Papers Series, 2018-01-1736.

[61] Lauer, Thomas and Geringer, Bernhard (2013): *Bewertung der Restgastoleranz bei homogenen Brennverfahren für hohe Abgasrückführraten*. MTZ, Motortechnische Zeitschrift, 03.

[62] Laurency, P. (2013): *Funktionen wirkungsschwacher Klimaschutzabkommen*. VS Verlag für Sozialwissenschaften, 1st edition, ISBN 978-3-531-19185-0.

[63] Lewis, Andrew, Akehurst, Sam, Turner, James, Patel, Rishin and Popplewell, Andrew (2014): *Observations on the Measurement and Performance Impact of Catalyzed vs. Non Catalyzed EGR on a Heavily Downsized DISI Engine*. SAE Int. J. Engines, **7**2014-01-1196, pp. 458–467.

[64] Liu, Feilong and Pfeiffer, Jeffrey (2015): *Estimation Algorithms for Low Pressure Cooled EGR in Spark-Ignition Engines*. SAE Technical Paper Series, 2015-01-1620.

[65] Löbbert, Philipp (2006): *Möglichkeiten und Grenzen der Teillaststeuerung von Ottomotoren mit vollvariablem Ventilhub*. Ph.D. thesis.

[66] Ludwig, O. (2011): *Eine Möglichkeit zur echtzeitfähigen, physikalisch-basierten Motorprozessanalyse auf der Grundlage zeitlich fusionierter Messdaten*. Ph.D. thesis.

[67] Lumpp, C. (2011): *Echtzeitfähige Stickoxidmodellierung zur Integration im Steuergerät eines Nutzfahrzeug-Dieselmotors*. Ph.D. thesis.

[68] Maeule, Paul and Rabl, Hans-Peter (2018): *Korrelation von Oelemissionen mit Partikelemissionen bei Ottomotoren*. AVL Forum Abgas- und Partikelemissionen.

[69] MathWorks, Inc. (2019): *MATLAB Documentaion*.

[70] Merker, G. and Schwarz, C. (2009): *Grundlagen Verbrennungsmotoren: Simulation der Gemischbildung, Verbrennung und Schadstoffbildung und Aufladung*. Vieweg and Teubner,Wiesbaden, ISBN 978-3-834-89344-4.

[71] Merker, G. P. and Teichmann, R. (2014): *Grundlagen Verbrennungsmotoren*.

[72] Messing, Roman, Hahn, Joachim and Leinhos, Dirk Christian (2016): *Verfahren und Steuereinheit zur Einstellung eines Lastpunktes eines Verbrennungsmotors im dynamischen Motorbetrieb*.

[73] Metghalchi, Mohamad and Keck, James C. (1982): *Burning Velocities of Mixtures of Air with Methanol, Isooctane, and Indolene at High Pressure and Temperature*. Combustion and Flame, 48, pp. 191–210.

[74] Nitschke, H. (2014): *Erschließung von Wirkungsgradpotenzialen aufgeladener Ottomotoren mittels Ladungsverdünnung*. Ph.D. thesis.

[75] Papula, Lothar (2016): *Mathematik für Ingenieure und Naturwissenschaftler Band 3*. Vieweg + Teubner Verlag Wiesbaden, ISBN 978-3-658-11924-9.

[76] Pierburg Instruments (2002): *AMA 4000 Betriebsanleitung*.

[77] Pischinger, R., Klell, M. and Sams, Th. (2002): *Thermodynamik der Verbrennungskraftmaschine*. Springer-Verlag, Wien, 3rd edition, ISBN 978-3-7091-3826-7.

[78] Potteau, S., Lutz, P., Leroux, S. and Moroz, S. (2007): *Cooled EGR for a Turbo SI Engine to Reduce Knocking and Fuel Consumption*. SAE Technical Paper Series, 2007-01-3978.

[79] Roth, David B., Keller, Philip and Becker, Michael (2010): *Requirements of External EGR Systems for Dual Cam Phaser Turbo GDI Engines*. SAE Technical Paper Series, 2010-01-0588.

[80] Ruckstuhl, Andreas (2008): *Numerische und statistische Methoden fur Chemieingenieure: Nichtlineare Regression*, 11/17/2020.

[81] Sarikoc, Fatih (2009): *Untersuchungen zur Reduzierung der Stickoxidemissionen bei modernen Brennverfahren für Motoren mit Benzin-Direkteinspritzung*. Ph.D. thesis.

[82] Sarlashkar, Jayant, Rengarajan, Sankar and Roecker, Ryan (2016): *Transient Control of a Dedicated EGR Engine*. SAE Technical Paper Series, 2016-01-0616.

[83] Schiefer, Hartmut and Schiefer, Felix (2018): *Statisitk für Ingenieure*. Springer Fachmedien Wiesbaden, ISBN 978-3-658-20639-0.

[84] Schnittger, W., Königstein, A., Pritze, S. and Pöpperl, M. (2003): *2.2 Direct Ecotec Â– Neuer Ottomotor mit Direkteinspritzung von Opel*. MTZ, Motortechnische Zeitschrift, 64.

[85] Shelby, Michael H., Stein, Robert A. and Warren, Christopher C. (2004): *A New Analysis Method for Accurate Accounting of IC Engine Pumping Work and Indicated Work*. SAE Technical Papers Series, 2004-01-1262.

[86] Siokos, Konstantinos (2017): *Low-Pressure EGR in Spark-Ignition Engines: Combustion Effects, System Optimization, Transients & Estimation Algorithms*. Ph.D. thesis.

[87] Siokos, Konstantinos, Koli, Rohit, Prucka, Robert, Schwanke, Jason and Miersch, Julia (2015): *Assessment of Cooled Low Pressure EGR in a Turbocharged Direct Injection Gasoline Engine*. SAE Technical Paper Series, 2015-01-1253.

[88] Sjöberg, Magnus and Dec, John E. (2009): *Influence of EGR Quality and Unmixedness on the High-Load Limits of HCCI Engines*. SAE Int. J. Engines, **2**2009-01-0666, pp. 492–510.

[89] Sluder, C. Scott, Storey, John M. E., Lewis, Samuel A., Styles, Dan, Giuliano, Julia and Hoard, John W. (2008): *Hydrocarbons and Particulate Matter in EGR Cooler Deposits: Effects of Gas Flow Rate, Coolant Temperature, and Oxidation Catalyst*. SAE Int. J. Engines, **1**2008-01-2467, pp. 1196–1204.

[90] Sobh, Ahmed J. (2011): *The Aspects of Fouling on the Performance of Diesel Exhaust Gas Recirculation Coolers*. Ph.D. thesis.

[91] Stahel, Werner (2017): *Statistische Regressionsmodelle*, 11/17/2020.

[92] Styles, D.J., Hilditch, J. and Ruona, W.C. (2011): *Fixed EGR Rate System*.

[93] Surnilla, Gopichandra, Soltis, Richard, Hilditch, James, House, Christopher, Clark, Timothy and Gerhart, Mattthew (2016): *Intake Oxygen Sensor for EGR Measurement*. SAE Technical Paper Series, 2016-01-1070.

[94] Takaki, Daisuke, Tsuchida, Hirofumi, Kobara, Tetsuya, Akagi, Mitsuhiro, Tsuyuki, Takeshi and Nagamine, Morihiro (2014): *Study of an EGR System for Downsizing Turbocharged Gasoline Engine to Improve Fuel Economy*. SAE Technical Paper Series, 2014-01-1199.

[95] The Electric Vehicle World Sales Database (2018): *Global Plug-in Vehicle Sales for 2017 Â– Final Results*, 01/31/2018.

[96] Thewes, Matthias (2016): *Abgasrückführungssystem für eine Verbrennungskraftmaschine und Verfahren zum Betreiben eines solchen Abgasrückführungssystems*.

[97] Thewes, Matthias, Baumgarten, Henning, Nijs, Martin and Hoppe, Patrick (2014): *Future Fuel Consumption and Emssion Concepts for Boosted Gasoline Engines*. Motor und Umwelt.

[98] Toda, Tadashi, Sakai, Mitsuto, Hakariya, Masahi and Kato, Toshikazu (2017): *The New Inline 4 Cylinder 2.5L Gasoline Engine with Toyota New Global Architecture Concept*. Wiener Motorensymposium, 38.

[99] Urano, Y., Nakano, Y., Takada, H. and Sugita, M. (2005): *Optimization Technique for Transient Emission Reduction of Heavy Duty Diesel Engine*. SAE Technical Paper Series, 2005-01-1099.

[100] van Basshuysen, R. and Schäfer, F. (2015): *Handbuch Verbrennungsmotor*, volume 7. Springer Fachmedien Wiesbaden, ISBN 978-3-658-04677-4.

[101] van Basshuysen, R: (Hrsg.) and Spicher, U. (2013): *Ottomotor mit Direkteinspritzung*, chap. Downsizing und Downspeeding, pp. 205 – 211. Praxis. ATZ/MTZ-Fachbuch, Vieweg und Teubner Verlag / Springer Fachmedien Wiesbaden, ISBN 978-3-658-014070-0.

[102] Vitek, Oldrich, Macek, Jan, Poetsch, Christoph and Tatschl, Reinhard (2013): *Modeling Cycle-to-Cycle Variations in 0-D/1-D Simulation by Means of CombustionModel Parameter Perturbations based on Statistics of Cycle-Resolved Data*. SAE Technical Paper Series, 2013-04-08.

[103] Wagner, Marco and Fusshoeller, Hubert (2012): *Bewertung der Verfahren zur Aufteilung des indizierten Mitteldrucks in Hochdruck- und Ladungswechselanteile*.

[104] Watanabe, Eiichi and Fukutani, Itaru (1986): *Knock Reduction of Spark-Ignition Engines by EGR*. SAE Technical Paper Series, 860034.

[105] Weberbauer, Frank, Rauscher, Martin, Kulzer, Andre, Knopf, Martin and Bargende, Michael (2005): *Allgemein gültige Verlustteilung für neue Brennverfahren*. MTZ, Motortechnische Zeitschrift, 66.

[106] Wei, H., Zhu, T., Shu, G. and Tan, L. et al. (2012): *Gasoline engine exhaust gas recirculation Â– a review*. Applied Energy, volume 99, pp. 534 – 544.

[107] Wenig, Markus (2013): *Simulation der ottomotorischen Zyklenschwankungen*. Ph.D. thesis.

[108] Wiese, Ashley, Stefanopoulou, Anna, Buckland, Julia and Karnik, Amey Y. (2017): *Modelling and Control of Engine Torque for Short-Circuit Flow and EGR Evacuation*. SAE Technical Paper Series, 2017-01-0606.

[109] Wiese, A. P., Stefanopoulou, A. G., Karnik, A. Y. and Buckland, J. H. (2017): *Model Predictive Control for Low Pressure Exhaust Gas Recirculation with scavenging*. 2017 American Control Conference (ACC), pp. 3638–3643, ISSN 2378-5861.

[110] Wilhelm, Christian (2008): *Echtzeitfähige Verlustprozessmodellierung*. Ph.D. thesis.

[111] Winsel, T. (2002): *Stabile neuronale Prozessmodellierung*. Ph.D. thesis.

[112] Witt, Andreas (1999): *Analyse der thermodynamischen Verluste eines Ottomotors unter den Randbedingungen variabler Steuerzeiten*. Ph.D. thesis.

[113] Yonekawa, Akiyuki, Ueno, Masaki, Watanabe, Osamu and Ishikawa, Naohiro (2013): *Development of New Gasoline Engine for ACCORD Plug-in Hybrid*. SAE Technical Paper Series, 2013-01-1738.

[114] Yoo, Kwang Hee, Hoard, John, Boehman, Andre and Gegich, Matthew (2016): *Experimental Studies of EGR Cooler Fouling on a GDI Engine*. SAE Technical Paper Series, 2016-01-1090.

[115] Yoshioka, Mamoru, Nakamura, Takehide, Akita, Minoru and Ishii, Kazufumi (2012): *Abgasrückführungsgerät für eine Kraftmaschine*.

[116] Zellbeck, Hans, Roß, Tilo, Sens, Marc, Lautrich, Guido and Grigoriadis, Panagiotis (2015): *Aufladung von Verbrennungsmotoren*, pp. 533–559. Springer Fachmedien Wiesbaden, Wiesbaden, ISBN 978-3-658-04678-1.

[117] Zhong, Lurun, Musial, Marc, Reese, Ronald and Black, Greg (2013): *EGR Systems Evaluation in Turbocharged Engines*. SAE Technical Paper Series, 2013-01-0936.

A Appendix

A.1 Characteristic maps for roots loader

Figure A.1 below illustrates the relevant characteristic diagrams for the root loader used. They are taken from [21]. While the hydraulic power is shown on the left side, the hydraulic efficiency is depicted on the right-hand side. Both quantities are plotted in the map depending on the loader's volume flow and its pressure ratio. They are used in Section 3.3 of the thesis, where they are applied for a correction term. However, the correction term turns out to be suboptimal and is ultimately replaced by an alternative term.
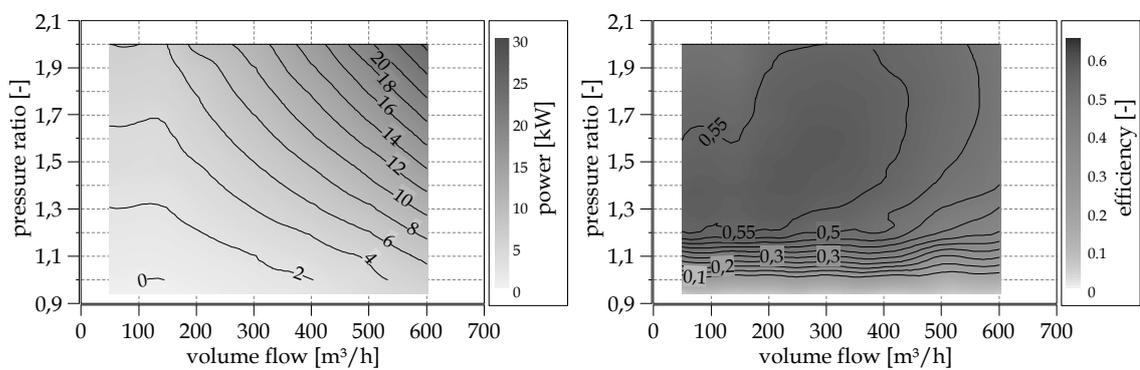
Figure A.1: Roots blower: characteristic maps of hydraulic power and isentropic efficiency [21].

A.2 Supplementary measurement results relating to the Bypass System

This chapter presents further results measured with the bypass system. At first, and in addition to the sensitivity analyses shown in the main part of the thesis, the results pertaining to the fuel-air ratio together with the target load and the throttle position follow. The results are again shown with the help of characteristic maps. Since these analyses were carried out before the optimum combination was found, the best-case values of the two actuators are not set in either of the two fields. However, this should not cause a problem, as relative statements can still easily be made.

The first diagram shows that there is virtually no cross-effect between increasing the target load and the fuel-air ratio. Only the target load mentioned in the main section shows a positive effect.
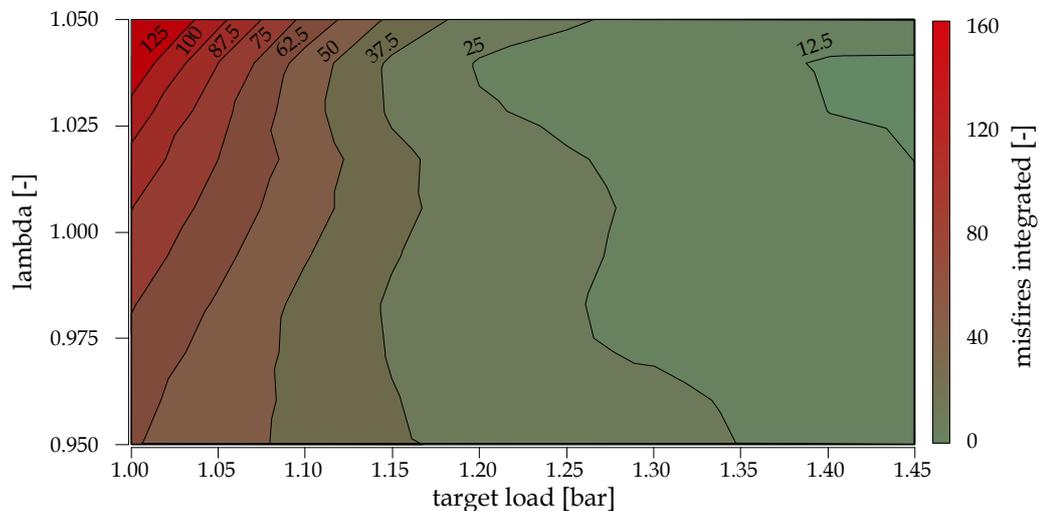
Figure A.2: Sensitivity study: fuel-air ratio and target load.

There is indeed an influence when looking at the throttle position of the main path and the fuel-air ratio. A slight enrichment results in a positive effect, while a leaner mixture leads to a further deterioration and an increase in the misfire rate. The effect is particularly relevant at throttle positions smaller than 2 %. Above that value, too little blocking of the main path and the consequently too high EGR rate in the combustion chamber predominates.

Finally, Figure A.4 illustrates the results that can be achieved using a precontrol of the bypass valve. For this purpose, the bypass valve was already opened well before the load step. Since the starting point of the worst-case load step is within the turbocharged operating range, an open bypass valve has no effect at this operating point due to the additional check valve in the bypass path. It is only when there is negative pressure in the manifold that there is an air flow through the bypass-path from the fresh air side to the manifold. However, this is the case, when the engine load is reduced. The aim of this investigation is to prevent any latency caused by the opening of the bypass valve.

The results from this diagram can be compared directly with the results from Figure 4.22 in Section 4.3. The results are better than without precontrol over almost the entire map. In addition, the non-existent latency when opening the bypass valve means that backflow via the bypass can be almost completely
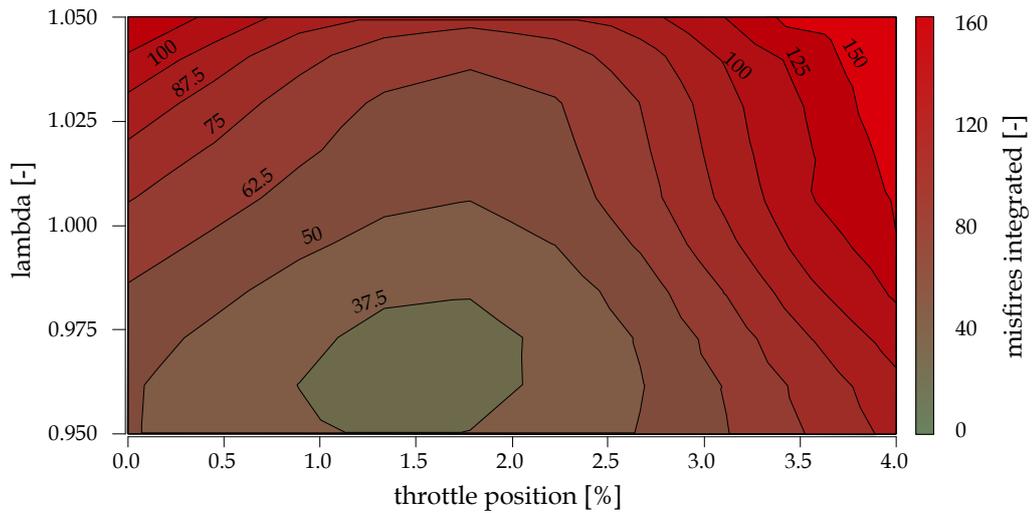
Figure A.3: Sensitivity study: fuel-air ratio and throttle position.

prevented in the range of very small throttle angles, as can be seen from the lack of misfire events at small throttle angles. This is presumably due to the significantly faster pressure compensation in the manifold area, because air can flow in more quickly via the bypass.

Thus, precontrol can generate further positive effects, but unfortunately it is not possible in all operating ranges, i.e. only where the starting load point is within the turbocharged operating range. Nevertheless, it should be considered in these operating ranges if the system is to be implemented in series production.
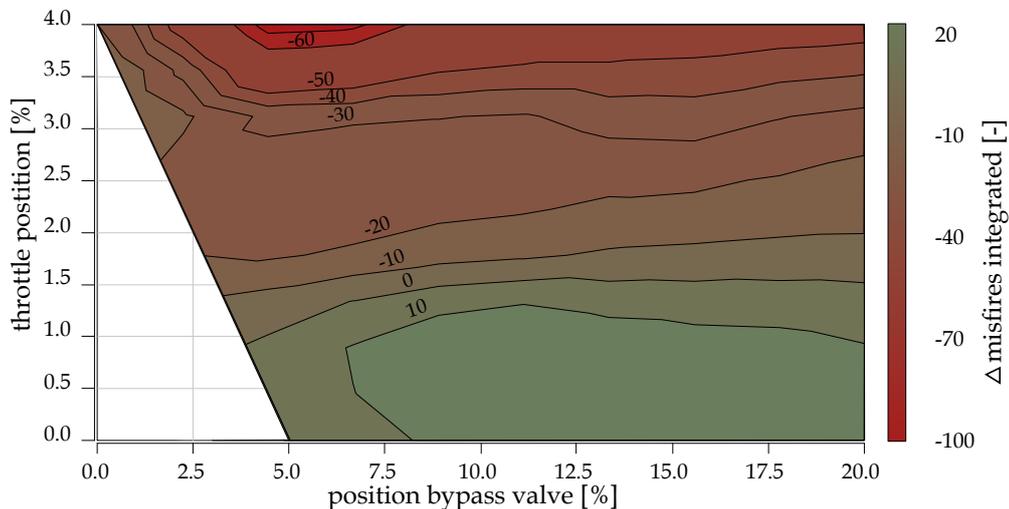
Figure A.4: Sensitivity study: precontrol of the bypass vlave depending on throttle and bypass valve position.

Another possible means of optimization in series implementation is shown below in Figure A.5. The idea here is to introduce the fresh air via the bypass directly into the intake ports of the engine, such that stratification occurs in the combustion chamber between the fresh air, which is shown in the figure as a yellow line, and the EGR-air mixture, which is shown as a blue line.
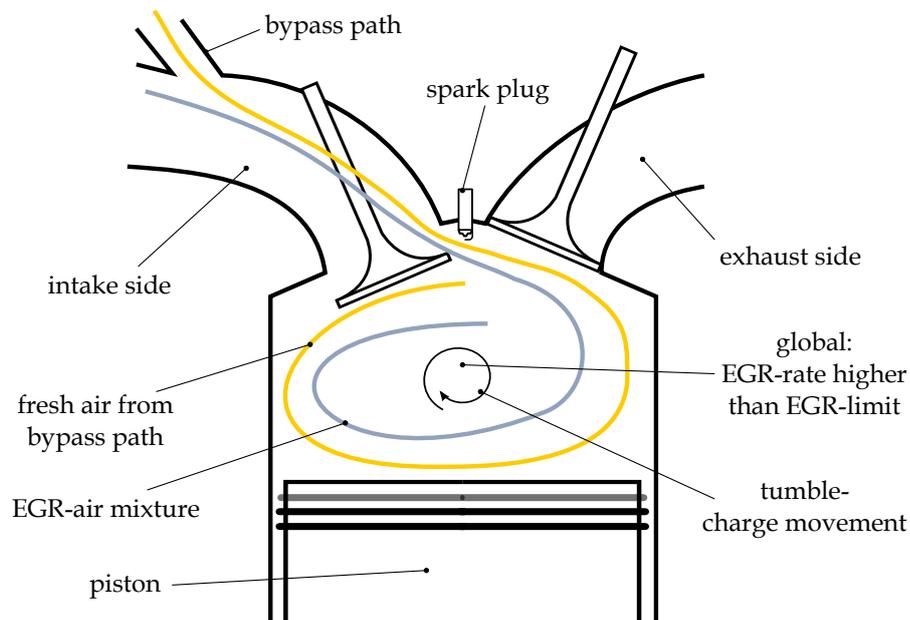
Figure A.5: Stratified charge using the Bypass-System.

By intentionally stratifying the two, a mixture of fresh air and fuel and a particularly low EGR rate can theoretically be provided locally around the spark plug. It is already known that the initial phase of the combustion has a particularly decisive influence on the cyclic fluctuations. If a virtually EGR-free mixture is present around the spark plug, the initial phase of combustion is particularly favored and cyclic fluctuations are minimized. For a detailed analysis of this type of system, 3D CFD simulations are necessary. Based on these, a specific design of the intake system and the bypass system can be made. Since this work was based on an existing engine into which the bypass path had to be integrated, the optimization option could not be implemented within the existing system. Nevertheless, it is shown as a further theoretical potential of the system.

A.3 Additional weighting factors β from Engine 1

This section and Figures A.6 to A.10 present the outstanding results of the single parameters β_{k,a_j} of Matrix *B* for Engine 1. The diagrams correspond to those shown in Chapter 5.4 with respect to sequence and layout. The results are listed here for completion, but do not include anything that requires additional explanation.
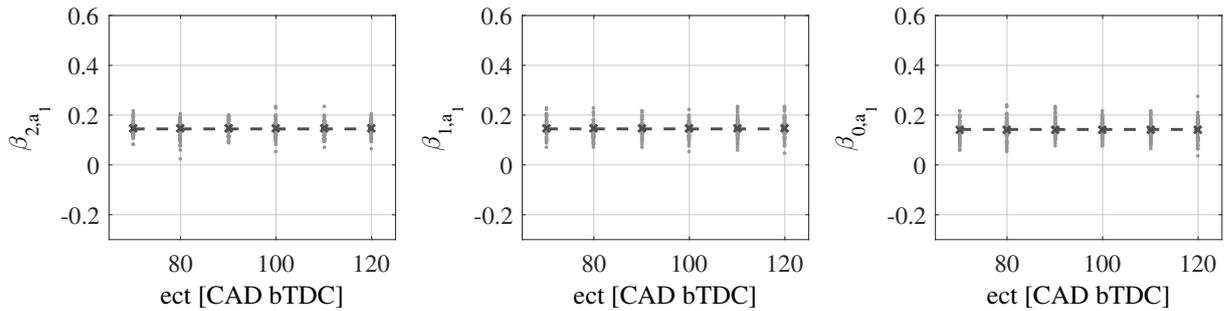
Figure A.6: Parameters β_{2,a_1}, β_{1,a_1} and β_{0,a_1} depending on exhaust cam timing (*ect*).
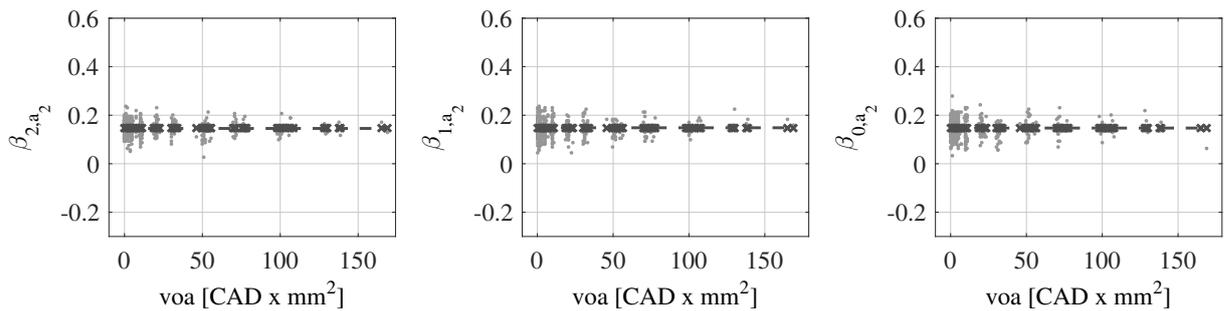
Figure A.7: Parameters β_{2,a_2}, β_{1,a_2} and β_{0,a_2} depending on valve overlapping area (*voa*).
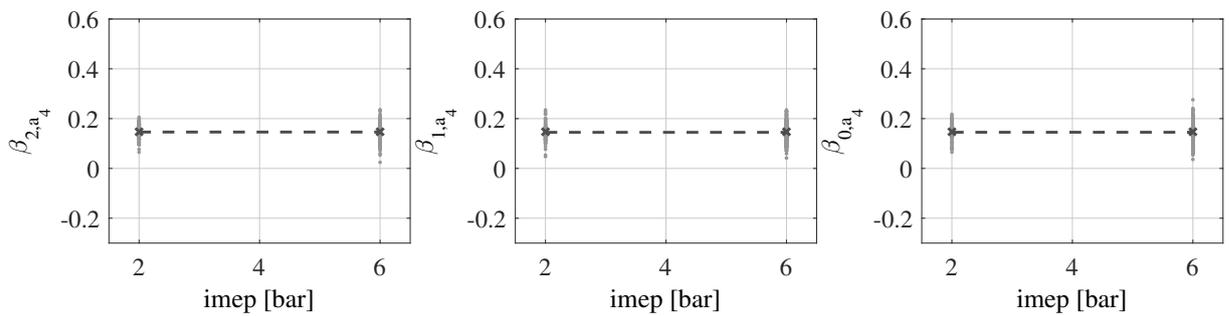
Figure A.8: Parameters β_{2,a_4}, β_{1,a_4} and β_{0,a_4} depending on engine load (*imep*).
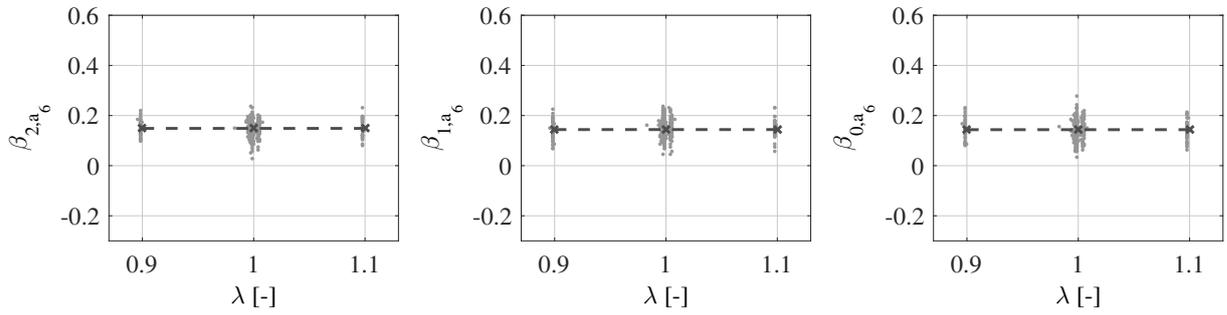
Figure A.9: Parameters β_{2,a_6}, β_{1,a_6} and β_{0,a_6} depending on air fuel ratio (λ).
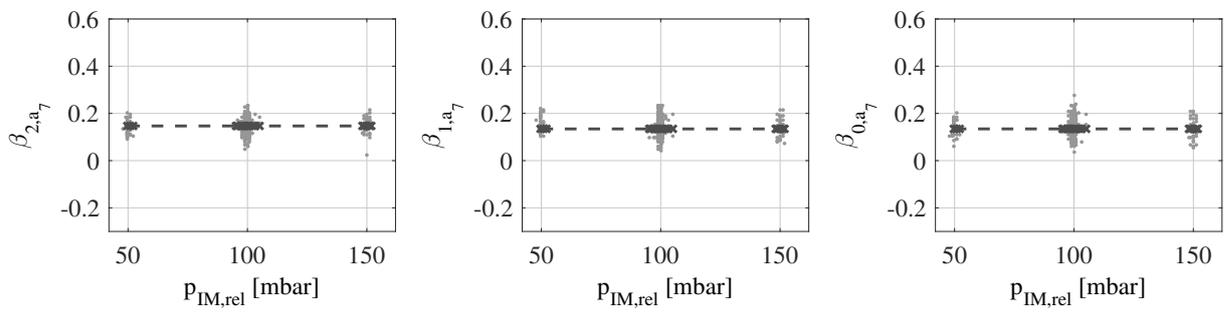
Figure A.10: Parameters β_{2,a_7}, β_{1,a_7} and β_{0,a_7} depending on intake manifold pressure ($p_{IM,rel}$).

A.4 Additional weighting factors β from Engine 2

As in Section A.3, the weighting factors of Engine 2 are now presented at this point. The summary of the values from Figures A.11 to A.17 can be found in Matrix B_{Eng2} or Equation 5.20. Just as in the previous section, none of the diagrams requires additional explanation; they round off the documentation.
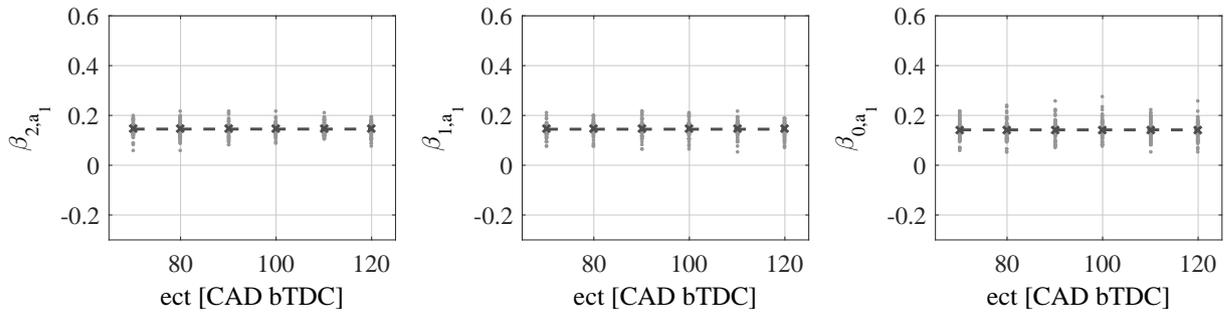
Figure A.11: Parameters β_{2,a_1}, β_{1,a_1} and β_{0,a_1} depending on exhaust cam timing (*ect*).
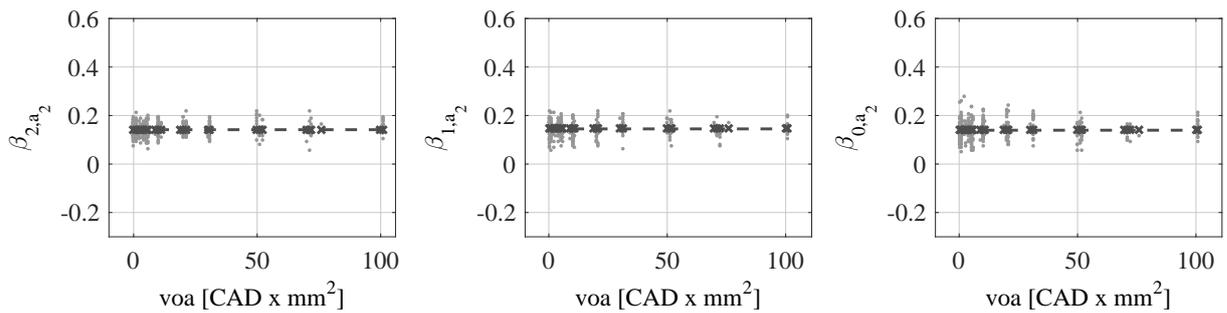
Figure A.12: Parameters β_{2,a_2}, β_{1,a_2} and β_{0,a_2} depending on valve overlapping area (*voa*).
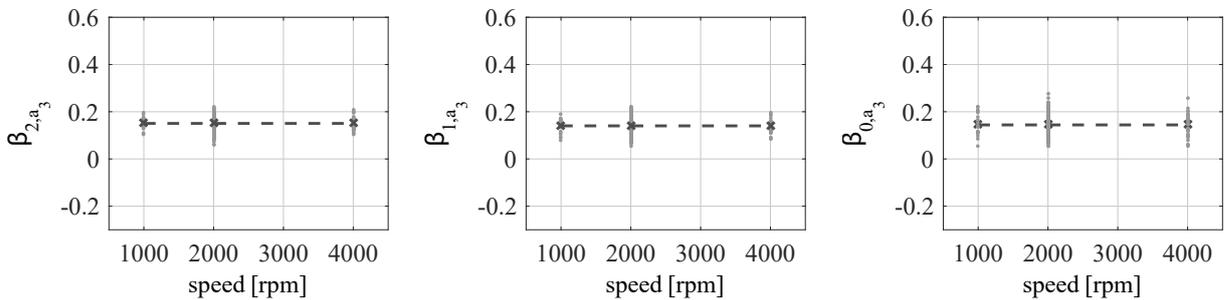
Figure A.13: Parameters β_{2,a_3}, β_{1,a_3} and β_{0,a_3} depending on engine speed (n_{mot}).
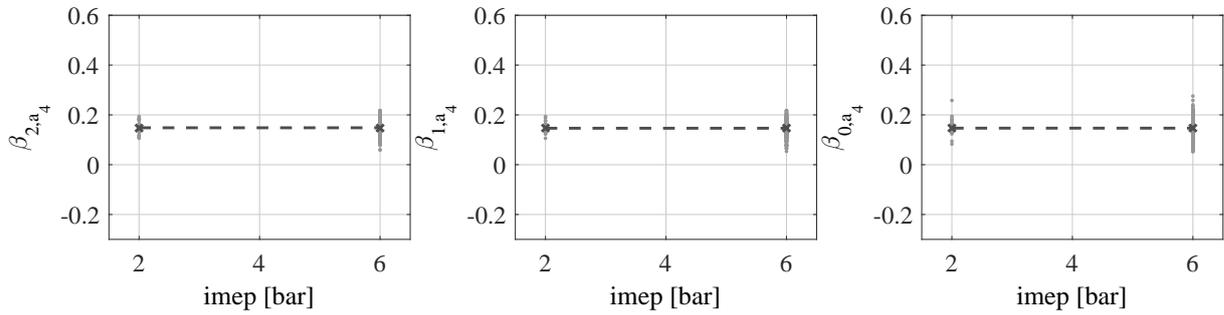
Figure A.14: Parameters β_{2,a_4}, β_{1,a_4} and β_{0,a_4} depending on engine load (*imep*).
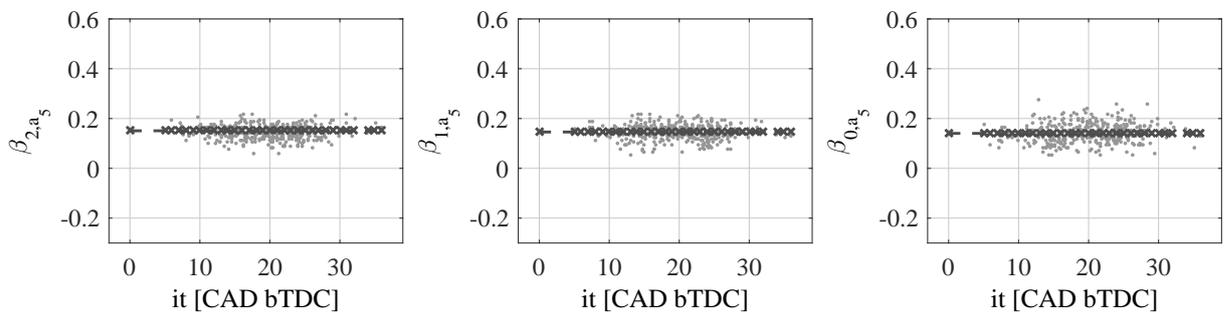
Figure A.15: Parameters β_{2,a_5}, β_{1,a_5} and β_{0,a_5} depending on ignition timing (*it*).
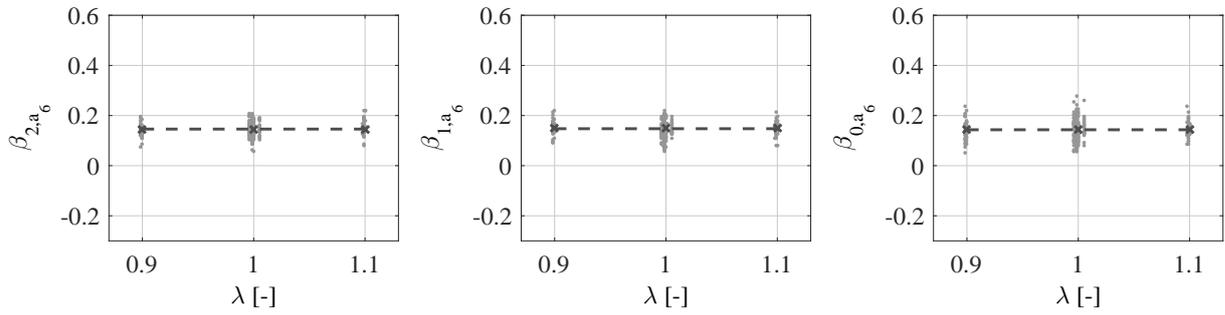
Figure A.16: Parameters β_{2,a_6}, β_{1,a_6} and β_{0,a_6} depending on air fuel ratio (λ).
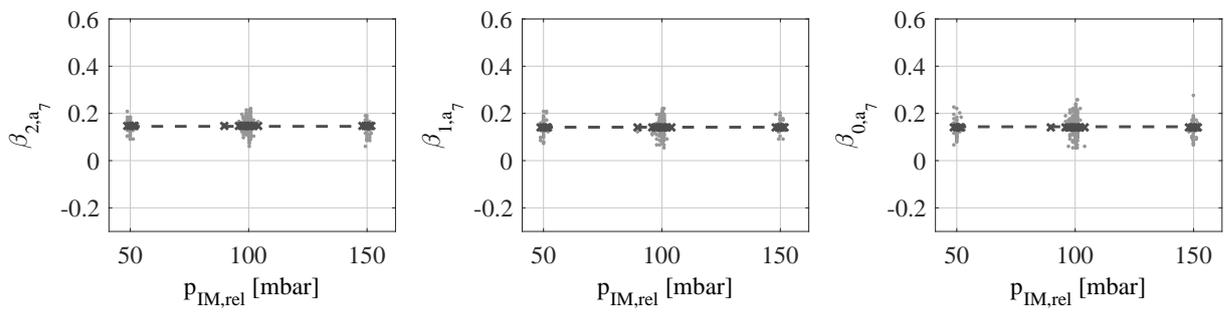
Figure A.17: Parameters β_{2,a_7}, β_{1,a_7} and β_{0,a_7} depending on intake manifold pressure ($p_{IM,rel}$).

A.5 Additional model validation

Figures A.18 to A.21 below show how the maximum indicated efficiency of the EGR variation $\eta_{ind,EGR,max}$ and the corresponding EGR rate $x_{EGR,max,tol}$ depend on the model input variables that are omitted in Chapter 5.4. All results are again taken from Engine 2 and provide the model input variables that are not shown in the diagrams in the main part. The results confirm the validation results from the main part, where the relative behavior of the individual variables, in particular, can be reproduced very well by the model.
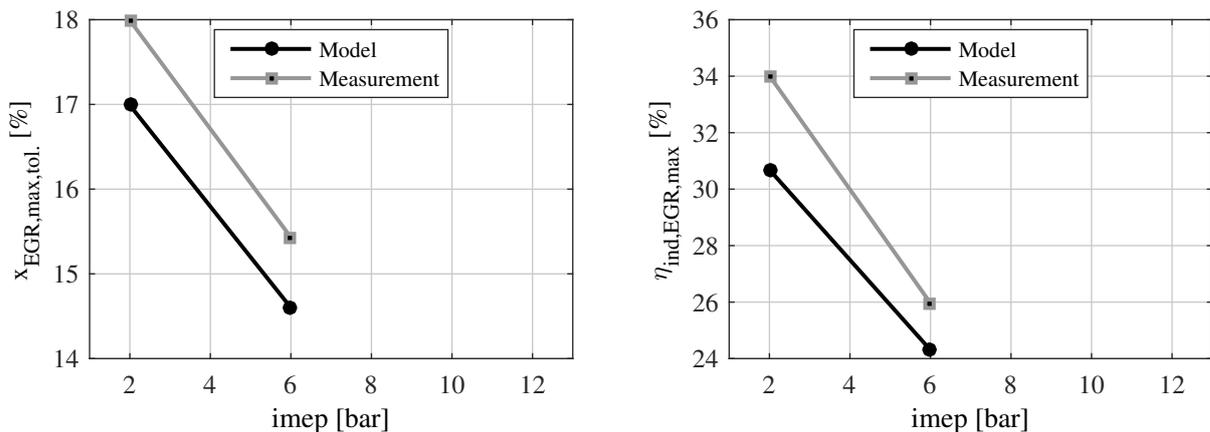
Figure A.18: $x_{EGR,max,tol}$ and $\eta_{ind,EGR,max}$ depending on the engine load @1000 *rpm* engine speed, all other model input parameters constant.
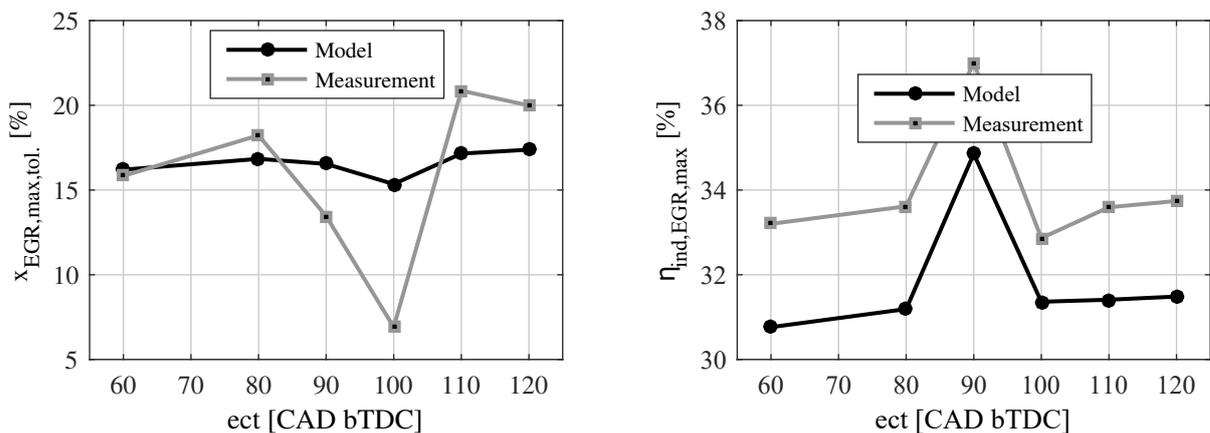
Figure A.19: $x_{EGR,max,tol}$ and $\eta_{ind,EGR,max}$ depending on the exhaust cam timing @6 *bar* engine load and @2000 *rpm* engine speed, all other model input parameters constant.

Just as the maximum values are considered depending on the model input variables in the previous figures, Figures A.22 to A.24 now show the detailed EGR variations for a fixed set of model input variables. Two sets of model input variables are illustrated per diagram; only one of them is modified. In each case, the modeled variation (dashed) and the measured variation (solid) are shown.
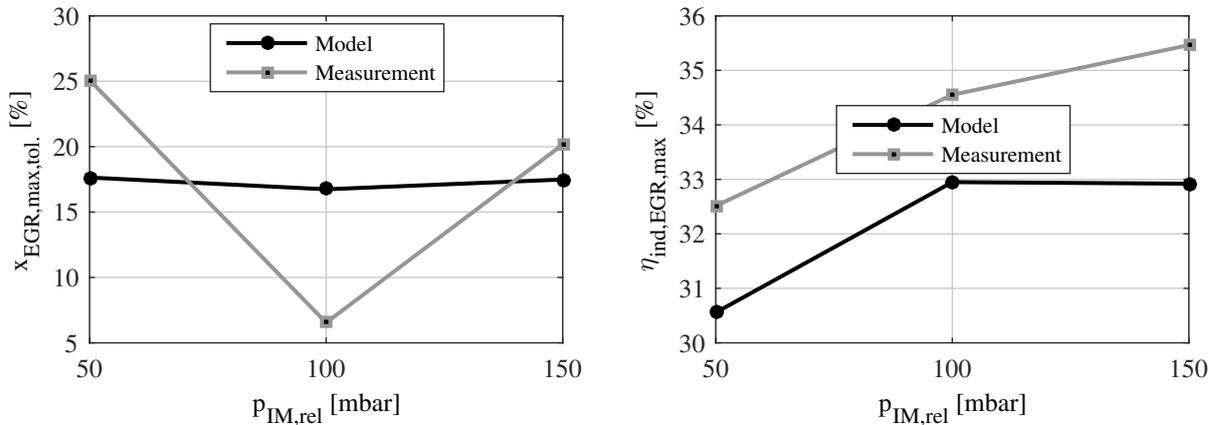
Figure A.20: $x_{EGR,max,tol}$ and $\eta_{ind,EGR,max}$ depending on the intake maniflod pressure @6 *bar* engine load and @2000 *rpm* engine speed, all other model input parameters constant.
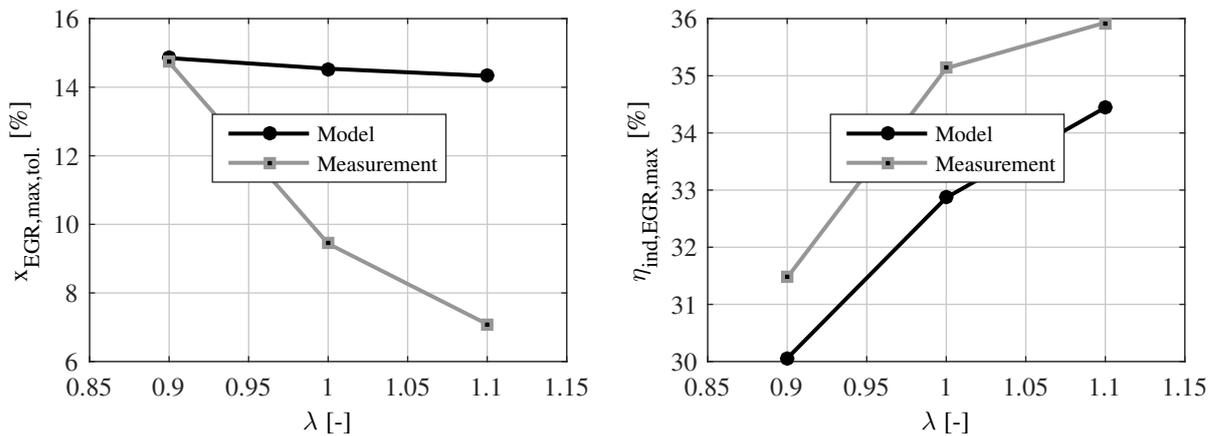
Figure A.21: $x_{EGR,max,tol}$ and $\eta_{ind,EGR,max}$ depending on the fuel-air ratio @6 *bar* engine load and @4000 *rpm* engine speed, all other model input parameters constant.
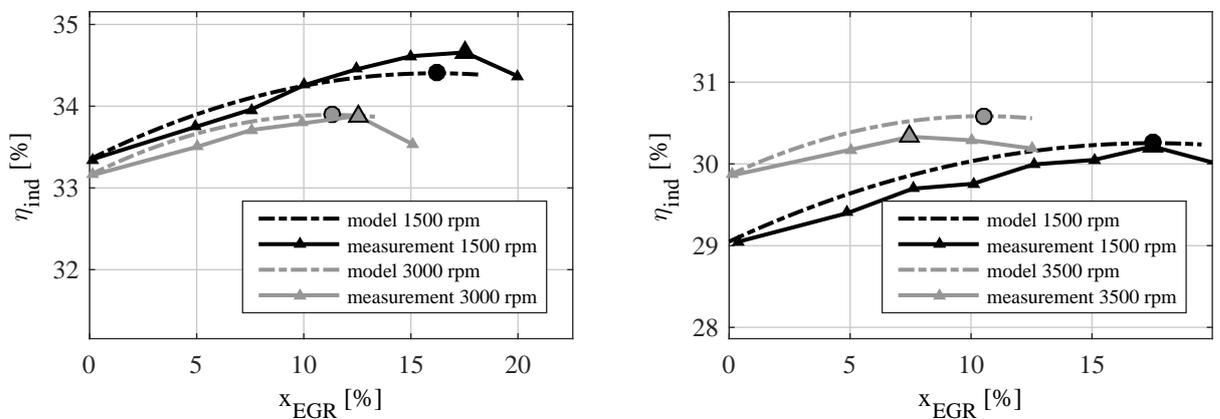
Figure A.22: Comparison of measured and modeled data - ω variation for Engine 1 @4 *bar* (left-hand side) and for Engine 2 @3 *bar* (right-hand side, all other model input parameters constant).
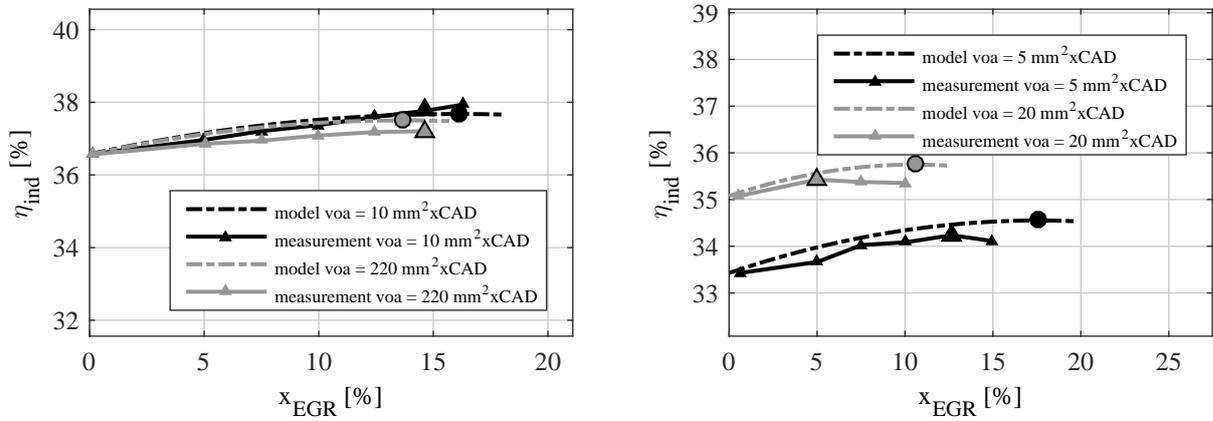
Figure A.23: Comparison of measured and modeled data - *voa* variation for Engine 1 @3000 *rpm*, 10 *bar* (left-hand side) and for Engine 2 @1500 *rpm*, 4 *bar* (right-hand side, all other model input parameters constant
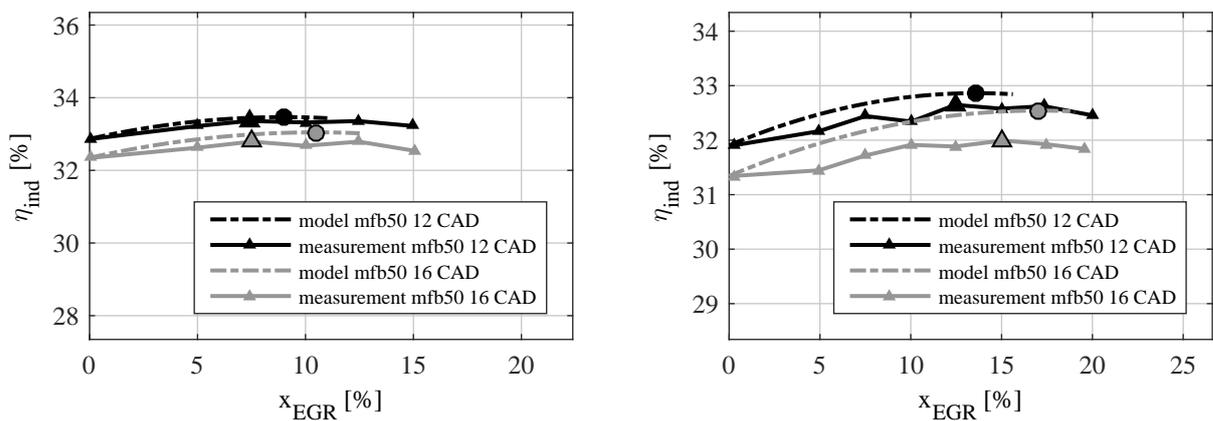
Figure A.24: Comparison of measured and modeled data - *mfb*50 variation for Engine 1 @3000 *rpm*, 4 *bar* (left-hand side) and for Engine 2 @3000 *rpm*, 6 *bar* (right-hand side, all other model input parameters constant).

Ehrenerklärung

„Ich versichere hiermit, dass ich die vorliegende Arbeit ohne unzulässige Hilfe Dritter und ohne Benutzung anderer als der angegebenen Hilfsmittel angefertigt habe. Die Hilfe eines kommerziellen Promotionsberaters habe ich nicht in Anspruch genommen. Dritte haben von mir weder unmittelbar noch mittelbar geldwerte Leistungen für Arbeiten erhalten, die im Zusammenhang mit dem Inhalt der vorgelegten Dissertation stehen. Verwendete fremde und eigene Quellen sind als solche kenntlich gemacht. Ich habe insbesondere nicht wissentlich:

- Ergebnisse erfunden oder widersprüchliche Ergebnisse verschwiegen,
- statistische Verfahren absichtlich missbraucht, um Daten in ungerechtfertigter Weise zu interpretieren,
- fremde Ergebnisse oder Veröffentlichungen plagiiert,
- fremde Forschungsergebnisse verzerrt wiedergegeben

Mir ist bekannt, dass Verstöße gegen das Urheberrecht Unterlassungs- und Schadensersatzansprüche des Urhebers sowie eine strafrechtliche Ahndung durch die Strafverfolgungsbehörden begründen kann.

Ich erkläre mich damit einverstanden, dass die Dissertation ggf. mit Mitteln der elektronischen Datenverarbeitung auf Plagiate überprüft werden kann.

Die Arbeit wurde bisher weder im Inland noch im Ausland in gleicher oder ähnlicher Form als Dissertation eingereicht und ist als Ganzes auch noch nicht veröffentlicht.“

München, 15.03.2023