

Modellbasierte Applikation von Verbrennungsmotoren unter Berücksichtigung von Bauteiltoleranzen.

Zur Erlangung des akademischen Grades eines
DOKTORS DER INGENIEURWISSENSCHAFTEN
(Dr.-Ing.)

von der KIT-Fakultät für Maschinenbau des
Karlsruher Instituts für Technologie (KIT)
angenommene

DISSERTATION

von

M.Sc. Kadir Mourat
aus Viersen

Tag der mündlichen Prüfung:
Hauptreferent:
Korreferent:

17.12.2024
Prof. Dr. sc. techn. Thomas Koch
Prof. em. Dr. Michael Bargende

Vorwort des Herausgebers

Die Komplexität des verbrennungsmotorischen Antriebes ist seit über 100 Jahren Antrieb für kontinuierliche Aktivitäten im Bereich der Grundlagenforschung sowie der anwendungsorientierten Entwicklung. Die Kombination eines instationären, thermodynamischen Prozesses mit einem chemisch reaktiven und hochturbulenten Gemisch, welches in intensiver Wechselwirkung mit einer Mehrphasenströmung steht, stellt den technologisch anspruchsvollsten Anwendungsfall dar. Gleichzeitig ist das Produkt des Verbrennungsmotors aufgrund seiner vielseitigen Einsetzbarkeit und zahlreicher Produktvorteile für sehr viele Anwendungen annähernd konkurrenzlos. Nun steht der Verbrennungsmotor insbesondere aufgrund der Abgasemissionen im Blickpunkt des öffentlichen Interesses. Vor diesem Hintergrund ist eine weitere und kontinuierliche Verbesserung der Produkteigenschaften des Verbrennungsmotors unabdingbar. Am Institut für Kolbenmaschinen am Karlsruher Institut für Technologie wird deshalb intensiv an der Weiterentwicklung des Verbrennungsmotors geforscht. Übergeordnetes Ziel dieser Forschungsaktivitäten ist die Konzentration auf drei Entwicklungsschwerpunkte. Zum einen ist die weitere Reduzierung der Emissionen des Verbrennungsmotors, die bereits im Verlauf der letzten beiden Dekaden um circa zwei Größenordnungen reduziert werden konnten aufzuführen. Zum zweiten ist die langfristige Umstellung der Kraftstoffe auf eine nachhaltige Basis Ziel der verbrennungsmotorischen Forschungsaktivitäten. Diese Aktivitäten fokussieren gleichzeitig auf eine weitere Wirkungsgradsteigerung des Verbrennungsmotors. Der dritte Entwicklungsschwerpunkt zielt auf eine Systemverbesserung. Motivation ist beispielsweise eine Kostenreduzierung, Systemvereinfachung oder Robustheitssteigerung von technischen Lösungen. Bei den meisten Fragestellungen wird aus dem Dreiklang aus Grundlagenexperiment, Prüfstandversuch und Simulation eine technische Lösung erarbeitet. Die Arbeit an diesen Entwicklungsschwerpunkten bestimmt die Forschungs- und Entwicklungsaktivitäten des Instituts. Hierbei ist eine gesunde Mischung aus grundlagenorientierter Forschung und anwendungsorientierter Entwicklungsarbeit der Schlüssel für ein erfolgreiches Wirken. In nationalen als auch internationalen Vorhaben sind wir bestrebt, einen wissenschaftlich wertvollen Beitrag zur erfolgreichen Weiterentwicklung des Verbrennungsmotors beizusteuern. Sowohl Industriekooperationen als auch öffentlich geförderte Forschungsaktivitäten sind hierbei die Grundlage guter universitärer Forschung. Zur Diskussion der erarbeiteten Ergebnisse und Erkenntnisse dient diese Schriftenreihe, in der die Dissertationen des Instituts für Kolbenmaschinen verfasst sind. In dieser Sammlung sind somit die wesentlichen Ausarbeitungen des Instituts niedergeschrieben. Natürlich werden darüber hinaus auch Publikationen auf Konferenzen und in Fachzeitschriften veröffentlicht. Präsenz in der Fachwelt erarbeiten wir uns zudem durch die Einreichung von Erfindungsmeldungen und dem damit verknüpften Streben nach Patenten. Diese Aktivitäten sind jedoch erst das Resultat von vorgelagerter und erfolgreicher Grundlagenforschung. Jeder Doktorand am Institut beschäftigt sich mit Fragestellungen von ausgeprägter gesellschaftlicher Relevanz. Insbesondere Nachhaltigkeit und Umweltschutz als Triebfedern des ingenieurwissenschaftlichen Handelns sind die Motivation unserer Aktivität. Gleichzeitig

kann er nach Beendigung seiner Promotion mit einer sehr guten Ausbildung in der Industrie oder Forschungslandschaft wichtige Beiträge leisten.

Vorwort des Autors

Mit der vorliegenden Dissertation schlieÙe ich einen wichtigen Meilenstein meiner wissenschaftlichen Laufbahn ab. Die intensive Auseinandersetzung mit dem Thema dieser Arbeit hat mir nicht nur tiefgehende fachliche Einblicke ermöglicht, sondern auch meine wissenschaftliche Denkweise und Arbeitsweise nachhaltig geprägt.

An dieser Stelle möchte ich meinen herzlichen Dank an alle aussprechen, die mich auf diesem Weg unterstützt und begleitet haben. Mein besonderer Dank gilt meinem Betreuer, Prof. Dr. Thomas Koch, für die fachliche Expertise, die wertvollen Hinweise und die stets konstruktive Betreuung. Ebenso danke ich allen Kollegen bei der Robert Bosch GmbH, die mich im Rahmen meiner Arbeit unterstützt haben und mit einer kollegialen Atmosphäre und dem regen wissenschaftlichen Austausch wesentlich zum Gelingen dieser Arbeit beigetragen haben. Besonders hervorheben möchte ich Dr. Carola Eckstein und Dr. Hans-Peter Frisse, die mich im Rahmen der Arbeit stets unterstützt haben.

Nicht zuletzt danke ich meiner Familie und meinen Freunden für ihre Geduld, ihr Verständnis und ihre Ermutigung in den oft herausfordernden Phasen der Dissertation.

Ich hoffe, dass die Ergebnisse dieser Arbeit einen Beitrag zur Weiterentwicklung des Forschungsfeldes leisten und zukünftige wissenschaftliche Fragestellungen inspirieren können.

Inhaltsverzeichnis

1. Einleitung	1
2. Theoretische Grundlagen	3
2.1. Datenbasierte Modellbildung mit Gauß-Prozessen	3
2.2. Evaluation datenbasierter Modelle	8
2.2.1. Evaluationsmetriken zur Bewertung von Modellen	9
2.2.2. Kreuzvalidierung	12
2.2.3. Verzerrung-Varianz-Dilemma	13
2.3. Statistische Stichprobennahme bzw. Sampling	14
2.3.1. Quasi-Monte-Carlo-Sampling	16
2.3.2. Latin-Hypercube-Sampling	16
3. Versuchsaufbau	17
3.1. Versuchsträger	17
3.1.1. Reihen-Vierzylinder-Dieselmotor	17
3.1.2. V6-Dieselmotor	18
3.2. Motorprüfstand	20
3.3. Messdatenerfassung	20
3.3.1. Indizierung	20
3.3.2. Abgasanalyse	21
4. SEMO-Methode	23
4.1. Stand der Technik im Bereich der Motorsteuergeräteapplikation	23
4.2. Anforderungen an die Methodik	29
4.3. Methodenablauf	29
4.4. Kritische Bewertung und Abgrenzung der Methode gegenüber dem Stand der Technik	30
5. Versuchsmethode	33
5.1. Versuchsplanung für lokale und globale Motormodelle	33
5.2. Versuchsplanung und -durchführung	36
5.2.1. Versuchsplanung	37
5.2.2. Versuchsdurchführung mit Hilfe der ODCM-Methode	40
5.2.3. Modifizierter Online-DoE-Ansatz	42
6. Stochastisches Motormodell zur Vorhersage der Flottenemissionen	47
6.1. Anforderungen an das Motormodell	47
6.2. Wissenschaftliche Neuerungen	48
6.3. Modellbildungsprozess	49
6.3.1. Initiales nominales Motormodell	49

6.3.2. Stochastisches Motormodell	50
6.4. Implementierungsdetails	55
6.5. Untersuchung der asymptotischen Eigenschaften des stochastischen Modells	55
6.5.1. Güte des stochastischen Modells in Abhängigkeit der Stützstell- und Stichprobenanzahl	58
6.5.2. Güte des stochastischen Modells in Abhängigkeit des Stützstellen- Stichproben-Verhältnisses	61
7. Anwendungsbeispiele	65
7.1. Anwendungsbeispiel Reihen-Vierzylinder-Dieselmotor	65
7.1.1. Problemstellung	66
7.1.2. Vergleich der Methodik mit dem Stand der Technik	70
7.1.3. Vergleich der Methodik mit dem Brute-Force-Ansatz	73
7.1.4. Zusammenfassung	75
7.2. Anwendungsbeispiel V6-Dieselmotor	76
7.2.1. Validierung des stochastischen Modellierungsansatzes	76
7.2.2. Stochastische Parameteroptimierung am Beispiel des WLTC	85
8. Zusammenfassung und Diskussion	97
A. Anhang	99
A.1. Wahrscheinlichkeitstheorie	99
A.2. Algorithmen	104
Nomenklatur	105
Abbildungsverzeichnis	109
Tabellenverzeichnis	111
Literaturverzeichnis	126

1. Einleitung

Als Konsequenz der negativen Umweltauswirkungen des Verkehrssektors und eines stetig wachsenden Umweltbewusstseins in der Bevölkerung und Politik fordern Gesetzgeber und Öffentlichkeit immer strengere Richtlinien zur Regulierung von Kraftfahrzeugen [97]. Für die Fahrzeughersteller bedeutet dies die Verpflichtung, klimaschädliche Treibhausgase und gesundheitsschädliche Schadstoffe zu reduzieren. Den größten Anteil am anthropogenen Treibhauseffekt hat Kohlendioxid. Bei den gesundheitsschädlichen Emissionen stehen insbesondere Stickoxide, Kohlenmonoxid, Kohlenwasserstoffe und Partikel im Fokus. Gleichzeitig sollen natürlich Fahrleistung und Fahrkomfort weiterhin im Vordergrund stehen [69]. Um die Emissionen des Straßenverkehrs zu reduzieren, wurden die Emissionsgrenzwerte in der Vergangenheit kontinuierlich verschärft. Die Einführung des Worldwide harmonized Light vehicles Test Procedure, der Euro-6d-Norm und Maßnahmen wie In-Service-Conformity und Fuel Consumption Monitoring sind weitere Schritte in die richtige Richtung. Trotz dieser Restriktionen hat sich in der Vergangenheit gezeigt, dass die Kohlenstoffdioxid- und Schadstoffvorgaben im realen Fahrbetrieb teilweise nicht eingehalten werden. Seit der Einführung des Prüfverfahrens Real Driving Emissions (RDE) wird die Einhaltung der Grenzwerte auch in der Praxis überprüft, und zwar in Form von stichprobenartigen Tests an Flottenfahrzeugen von Fahrzeugkunden [37]. Die Einführung der RDE-Tests hat die Fahrzeughersteller dazu veranlasst, ihre Fahrzeugflotten zu überwachen und zu optimieren. Dies hat unmittelbare Auswirkungen auf die Emissionen der Fahrzeugflotte und stellt höhere Anforderungen an deren Robustheit. Je mehr sich die Grenzwerte der Nullemission annähern, desto schwieriger wird es, sie einzuhalten, da der Emissionsspielraum immer kleiner wird. Berücksichtigt man Effekte wie Fahrstil, Wetter und andere Faktoren, die Toleranzen unterliegen, wird die Einhaltung der Grenzwerte noch schwieriger. Zumal es für die Gesetzgebung nicht ausreicht, die Emissionen im Durchschnitt einzuhalten. Die Emissionen müssen unter realen Bedingungen in verschiedenen Szenarien eingehalten werden. Die Streuung der Emissionen muss so gering wie möglich sein. Um diesen Anforderungen gerecht zu werden, müssen die Fahrzeughersteller nicht nur ihre Hard- und Software kontinuierlich verbessern. Von besonderer Bedeutung ist hier auch der Einsatz neuer Methoden, die eine robuste Applikation von Verbrennungsmotoren aktiv unterstützen.

Die vorliegende Arbeit widmet sich der Entwicklung einer neuartigen Applikationsmethodik namens Stochastic Engine Modeling and Optimization (SEMO), die eine robuste Applikation von Verbrennungsmotoren unterstützt. SEMO ermöglicht es, die Auswirkungen von Toleranzeinflüssen auf den Verbrennungsmotor zu quantifizieren und mittels stochastischer Optimierungsverfahren zu minimieren. Der Schwerpunkt liegt dabei auf dem Einfluss von Bauteiltoleranzen und deren Auswirkungen auf die Emissionen.

Der Hauptbeitrag dieser Arbeit ist die Entwicklung der SEMO-Methode. Teil der Methode ist eine Versuchsmethodik zur Generieren von Daten für das Training von stochastischen Modellen. Ein weiterer Teil ist ein speziell entwickeltes stochastisches Modell zur

Quantifizierung der Toleranzeinflüsse auf einen Verbrennungsmotor. Letzter wesentlicher Bestandteil der Methode ist ein stochastisches Optimierungsverfahren, das unter Verwendung des stochastischen Modells eine Minimierung der Emissionen unter Berücksichtigung der gegebenen Toleranzen gewährleistet. Weiterhin wird untersucht, ob SEMO eine robuste Applikation von Verbrennungsmotoren ermöglicht und damit die Einhaltung von Emissionsgrenzwerten in der Praxis unterstützt.

Der Aufbau der Arbeit ist im Folgenden dargestellt. Im zweiten Kapitel werden die theoretischen Grundlagen zum Training und zur Evaluierung datengetriebener Modelle vermittelt sowie wichtige statistische Konzepte vorgestellt. Im dritten Kapitel werden der Versuchsaufbau und die Versuchsträger vorgestellt. Im vierten Kapitel wird dann die entwickelte Methode SEMO vorgestellt. Kapitel 5 behandelt die Versuchsmethodik, die Bestandteil der SEMO-Methode ist. Im Anschluss daran wird in Kapitel 6 das stochastische Motormodell vorgestellt, das einen zentralen Bestandteil der SEMO-Methode darstellt. Dabei werden insbesondere die asymptotischen Eigenschaften des Modells näher betrachtet und diskutiert. Zur Validierung der Methodik werden anschließend in Kapitel 7 zwei Anwendungsbeispiele vorgestellt. Zunächst wird die Methodik auf einen Reihenvierzylinder-Dieselmotor zur Optimierung eines einzelnen Betriebspunktes angewendet. Hierbei wird die Methodik insbesondere mit dem deterministischen Standardanwendungsszenario unter Verwendung von Sicherheitsfaktoren verglichen, um die Vorteile der Methodik aufzuzeigen. In einem weiteren Vergleich wird auf die Recheneffizienz der Methodik eingegangen. Anschließend wird die Methodik zur Optimierung eines 6-Zylinder-Dieselmotors in einem großen Teil des Motorkennfeldes eingesetzt. Zunächst werden stochastische Modelle des Motors trainiert und anhand dynamischer Prüfstandsmessungen validiert. Abschließend erfolgt eine Optimierung des Motors unter Berücksichtigung der Bauteiltoleranzen. Das Ergebnis wird mit der Serienapplikation, die die Euro 6d Norm erfüllt, verglichen.

2. Theoretische Grundlagen

In diesem Kapitel werden theoretische Grundlagen vermittelt, die für das Verständnis der Arbeit relevant sind. Da das Themenspektrum der Arbeit sehr breit gefächert ist, konzentriert sich dieses Kapitel auf die relevantesten Themen. Für den interessierten Leser werden der Vollständigkeit halber im Anhang weitere theoretische Grundlagen und Konzepte dargestellt, die im Rahmen der Arbeit erwähnt werden, aber für das Verständnis der Kernaussage nicht zwingend erforderlich sind.

Im Abschnitt 2.1 werden zunächst einige Grundlagen der datenbasierten Modellierung vermittelt, wobei der Fokus auf dem Gauß-Prozess (engl.: *Gaussian process*, GP) liegt. Anschließend wird im Abschnitt 2.2 auf die Evaluation datenbasierter Modelle eingegangen. Dabei werden in Unterabschnitt 2.2.1 wichtige Evaluationsmetriken vorgestellt, in Unterabschnitt 2.2.2 wird die Kreuzvalidierung als Validierungsverfahren vorgestellt und in Unterabschnitt 2.2.3 wird auf das Bias-Varianz-Dilemma eingegangen. Abschnitt 2.3 stellt die Stichprobenziehung bzw. das Sampling vor. Dabei werden in den Unterabschnitten 2.3.1 und 2.3.2 das Quasi-Monte-Carlo-Sampling mit Sobol-Sequenzen und das Latin-Hypercube-Sampling behandelt.

Zuvor werden einige Bezeichnungen eingeführt, die zu einer einheitlichen Schreibweise beitragen sollen. Im weiteren Verlauf der Arbeit werden alle Tensoren höherer Ordnung mit Ausnahme der Skalare fett geschrieben. Vektoren werden in Kleinbuchstaben, Matrizen und Tensoren höherer Ordnung in Großbuchstaben geschrieben. Skalare werden in kursiven Kleinbuchstaben und Zufallsvektoren in kursiven Großbuchstaben dargestellt.

2.1. Datenbasierte Modellbildung mit Gauß-Prozessen

In der Literatur werden viele verschiedene datenbasierte Modellierungsmethoden beschrieben. Diese reichen von einfacher linearer Regression über komplexere probabilistische Ansätze bis hin zu maschinellen Lernansätzen wie tiefen neuronalen Netzen. Die Wahl des Modellierungsansatzes hängt dabei von der betrachteten Problemstellung bzw. Anwendung ab. Relevant für diese Arbeit ist insbesondere der Gauß-Prozess und die Regression mit diesen.

Gauß-Prozesse sind im Bereich der Motormodellierung bereits etabliert und werden für eine Vielzahl von Modellierungsaufgaben eingesetzt. GPs basieren auf der Bayes'schen Statistik und sind nichtparametrische Modelle. Sie besitzen also keine *direkten* Parameter wie z.B. neuronale Netze. GPs besitzen sogenannte *Hyperparameter*, die auf einer übergeordneten Ebene wirken und mit der Anzahl der Neuronen oder der Tiefe eines neuronalen Netzes verglichen werden können. Aus diesem Grund spricht man bei der Optimierung von GPs

auch von Hyperparameter-Optimierung, da sich die Hyperparameter mit jedem Optimierungsschritt ändern.

Im Gegensatz zu anderen Modellierungsverfahren, die in der Regel eine Normalverteilung des Messrauschens voraussetzen, beschränken sich GPs nicht auf diese stochastische Annahme. Ein GP stellt vielmehr einen stochastischen Prozess dar. Dieser ist dadurch definiert, dass er aus einer Menge von Zufallsvariablen besteht, für die eine gemeinsame Wahrscheinlichkeitsdichtefunktion existiert. GPs können als Modellierungsansatz verstanden werden bei dem der Vorhersagefehler explizit modelliert wird [72, 117, 118]. In der Praxis bedeutet dies, dass die Punktprediktion eines GPs nicht einen einzelnen Wert, sondern eine Gauß-Verteilung darstellt. Die Notation ist dementsprechend

$$f(\mathbf{x}) \sim \mathcal{GP}(m(\mathbf{x}), k(\mathbf{x}, \mathbf{x}')), \quad (2.1)$$

wobei m den Erwartungswert und k die Kovarianzfunktion ist. Diese Verteilung ist ein besonderes Merkmal von GPs und kann als Konfidenzintervall der Vorhersage interpretiert werden und stellt einen inhärenten Vorteil gegenüber anderen Modellen dar. Ein weiterer Vorteil von GPs ist die hohe Modellgüte. Unsere Erfahrungen zeigen, dass GPs bei der Abbildung nichtlinearer Systeme, insbesondere bei kleineren Datensätzen unter 5000 Datenpunkten, anderen Modellen hinsichtlich der Modellgüte überlegen oder ebenbürtig sind. Zudem sind Gauß-Prozesse unter Maximierung der sogenannten *Marginal Likelihood* robust gegenüber optimierten Modellparametern. Nachteile beim Einsatz von GPs sind der Speicherbedarf und der Rechenaufwand beim Training und bei der Vorhersage. Der Rechenaufwand für das Training kann mit $\mathcal{O}(N^3)$ angegeben werden und skaliert kubisch mit der Anzahl der Trainingsdaten N . In der Literatur finden sich verschiedene Ansätze zur Lösung dieses Problems [23, 43, 56, 93, 80, 1, 44]. Diese reichen von der Destillation der Trainingsdaten [23] bis hin zur Nyström-Approximation der Kovarianzmatrix [56, 93, 80, 1]. Ein weiterer Nachteil von GPs sind die schlechten Extrapolationseigenschaften. Außerhalb des Trainingsdatenbereichs geht der Erwartungswert vom GP gegen dessen angenommenen Mittelwert. Dieser wird in der Regel als 0 angenommen.

Zum besseren Verständnis des GPs wird im Folgenden eine kurze Herleitung gegeben. Für ein detaillierteres Verständnis von GPs sei auf die Arbeit von Rasmussen und Williams [96] verwiesen. Darüber hinaus wird im Anhang A.1 ein Überblick über einige Konzepte der Wahrscheinlichkeitstheorie gegeben. Dies soll sowohl fachfremden als auch fachkundigen Lesern als Nachschlagewerk dienen und ihnen die Möglichkeit geben, die hier erläuterten Konzepte zu verstehen.

Für die folgenden Passagen wird angenommen, dass wir einen Trainingsdatensatz $\mathcal{D} = \{(\mathbf{x}_i, y_i)\}_{i=1, \dots, N}$ mit N Beobachtungen vorliegen haben. Hierbei stellt \mathbf{x} einen Eingangsparametervektor mit d Eingangsdimensionen dar und y stellt die skalare Ausgangsgröße bzw. Beobachtung dar. Der Trainingsdatensatz kann auch als $\mathcal{D} = (\mathbf{X}, \mathbf{y})$ zusammengefasst werden, wobei \mathbf{X} die $d \times N$ Designmatrix ist und \mathbf{y} einen Ausgangsvektor darstellt, der die Beobachtungen enthält.

Im Rahmen dieser Arbeit wird angenommen, dass der Ausgang eines stationären Systems durch eine unbekannte Funktion f und additivem Gaußsches Rauschen ϵ , d.h.

$$y(\mathbf{x}) = f(\mathbf{x}) + \epsilon \quad (2.2)$$

beschrieben werden kann, wobei

$$\epsilon \sim \mathcal{N}(0, \sigma_n^2) \quad (2.3)$$

gilt und σ_n^2 die Geräuschvarianz darstellt.

Inferenz mit Gauß-Prozessen

Die Inferenz mit GPs erfolgt über die A-Posteriori-Verteilung $P(\mathbf{w} \mid \mathcal{D})$ über den Gewichten. Die eigentliche Modellierungsaufgabe besteht darin, diese A-Posteriori-Verteilung zu bestimmen. Sie kann mit Hilfe des Bayes-Theorems als

$$P(\mathbf{w} \mid \mathcal{D}) = \frac{P(\mathbf{y} \mid \mathbf{X}, \mathbf{w}) \cdot P(\mathbf{w})}{P(\mathbf{y} \mid \mathbf{X})} \quad (2.4)$$

bestimmt werden. Hierbei ist $P(\mathbf{y} \mid \mathbf{X}, \mathbf{w})$ die Likelihood, $P(\mathbf{w})$ die A-Priori-Verteilung und $P(\mathbf{y} \mid \mathbf{X})$ die Evidenz. Die Likelihood beinhaltet hierbei Informationen über das Modellverhalten, das Geräuschverhalten und die Trainingsdaten. Die A-Priori-Verteilung trifft Annahmen bezüglich der Verteilung der Trainingsdaten. Und die Evidenz ist eine von den Gewichten unabhängige normalisierende Konstante.

A priori, d.h. noch vor der Berücksichtigung irgendwelcher Trainingsdaten, wird für jede Sammlung von Funktionswerten angenommen, dass diese Gauß-verteilt sind mit einer Mittelwertfunktion $m(\mathbf{x})$ und einer Kovarianzmatrix \mathbf{K} . Diese Verteilung über den Gewichten wird als A-priori-Verteilung bezeichnet und berücksichtigt zunächst keine Daten. Sie wird geschrieben als

$$\mathbf{w} \sim \mathcal{N}(0, \mathbf{K}). \quad (2.5)$$

Dies bedeutet, dass im Falle einer Inferenz ohne Berücksichtigung von Trainingsdaten bzw. ohne vorheriges Modelltraining die Verteilung an einem Punkt durch die A-priori-Verteilung beschrieben werden kann.

Die Likelihood kann aus der Modellannahme und der Annahme eines normalverteilten Rauschens der Beobachtungen aus Gleichung 2.3 abgeleitet werden. Sie ist die Wahrscheinlichkeitsdichte der Beobachtungen als Funktion der Parameter, faktorisiert über die Trainingsdaten. Sie ist definiert als

$$P(\mathbf{y} \mid \mathbf{X}, \mathbf{w}) = \mathcal{N}(\phi(\mathbf{X}^\top) \mathbf{w}, \sigma_n^2 \mathbf{I}_N), \quad (2.6)$$

wobei \mathbf{I}_N die $N \times N$ Einheitsmatrix darstellt. Die Funktion ϕ projiziert die Eingangsgrößen \mathbf{x} in einen höherdimensionalen Merkmalsraum und bestimmt im Wesentlichen die Kovarianzfunktion. Die Projektionen können beliebig komplex und nichtlinear sein. Solange die Projektionen feste Funktionen sind, ist das Modell bezüglich seiner Parameter linear und damit analytisch lösbar.

Die Evidenz kann durch die sogenannte Marginalisierung (engl. marginalization) der Likelihood und der A-Priori-Verteilung bestimmt werden. Sie ist definiert durch

$$\begin{aligned} P(\mathbf{y} \mid \mathbf{X}) &= \int P(\mathbf{y} \mid \mathbf{X}, \mathbf{w}) P(\mathbf{w}) d\mathbf{w} \\ &= \mathcal{N}(0, \mathbf{K} + \sigma_n^2 \mathbf{I}_N). \end{aligned} \quad (2.7)$$

Durch Einsetzen in Gleichung 2.4 und Anwendung der Regel für bedingte Normalverteilungen aus der Gleichung A.12 folgt für die A-Posteriori-Verteilung über die Gewichte

$$P(\mathbf{w} \mid \mathcal{D}) = \mathcal{N}\left(\mathbf{K}^\top (\mathbf{K} + \sigma_n^2 \mathbf{I}_N)^{-1} \mathbf{y}, \mathbf{K} - \mathbf{K}^\top (\mathbf{K} + \sigma_n^2 \mathbf{I}_N)^{-1} \mathbf{K}\right). \quad (2.8)$$

Um die a-posteriori-Verteilung für die Inferenz in einem Testpunkt \mathbf{x}_\star zu nutzen, wird das mit der A-Posteriori-Verteilung gewichtete arithmetische Mittel über alle möglichen Parameterwerte gebildet. Dieser Ansatz unterscheidet Bayesianische Verfahren von anderen Verfahren, bei denen in der Regel ein einzelner Parameter nach einem bestimmten Kriterium ausgewählt wird. Die so ermittelte Vorhersageverteilung für einen konkreten Testpunkt ergibt sich also aus dem Mittelwert der Ergebnisse aller möglichen Modelle unter Berücksichtigung der A-Posteriori-Verteilung. Die Vorhersageverteilung für $\hat{y}_\star = \hat{y}(\mathbf{x}_\star)$ ist gemäß der Definition A.6 im Anhang A.1 mittels Marginalisierung durch

$$P(\hat{y}_\star \mid \mathbf{x}_\star, \mathcal{D}) = \int P(\hat{y}_\star \mid \mathbf{x}_\star, \mathbf{w}) P(\mathbf{w} \mid \mathcal{D}) d\mathbf{w} \quad (2.9)$$

gegeben. Für die Verteilung $P(\hat{y}_\star \mid \mathbf{x}_\star, \mathbf{w})$ gilt, dass diese normalverteilt ist, so dass $P(\hat{y}_\star \mid \mathbf{x}_\star, \mathcal{D})$ ebenfalls normalverteilt ist. Mit der Definition A.7 für Randverteilungen mehrdimensionaler Normalverteilungen aus Anhang A.1 gilt

$$\begin{bmatrix} \mathbf{y} \\ \hat{y}_\star \end{bmatrix} \sim \mathcal{N}\left(0, \begin{bmatrix} \mathbf{K} + \sigma_n^2 \mathbf{I}_N & \mathbf{k}_\star \\ \mathbf{k}_\star^\top & k_{\star\star} \end{bmatrix}\right), \quad (2.10)$$

wobei $\mathbf{k}_\star = k(\mathbf{X}, \mathbf{x}_\star)$ und $k_{\star\star} = k(\mathbf{x}_\star, \mathbf{x}_\star)$. Mit Hilfe der Definition der bedingten Normalverteilung aus Anhang A.1 ergibt sich schließlich

$$P(\hat{y}_\star \mid \mathbf{x}_\star, \mathcal{D}) = \mathcal{N}\left(\mathbf{k}_\star^\top (\mathbf{K} + \sigma_n^2 \mathbf{I}_N)^{-1} \mathbf{y}, k_{\star\star} - \mathbf{k}_\star^\top (\mathbf{K} + \sigma_n^2 \mathbf{I}_N)^{-1} \mathbf{k}_\star\right). \quad (2.11)$$

Daraus folgt für den Erwartungswert, den wir durch Inferenz am Punkt \mathbf{x}_\star erhalten,

$$\mathbb{E}[\hat{y}_\star] = \mathbf{k}_\star^\top (\mathbf{K} + \sigma_n^2 \mathbf{I}_N)^{-1} \mathbf{y}. \quad (2.12)$$

Zusätzlich zum Erwartungswert erhält man die Varianz der Verteilung

$$\mathbb{V}[\hat{y}_\star] = k_{\star\star} - \mathbf{k}_\star^\top (\mathbf{K} + \sigma_n^2 \mathbf{I}_N)^{-1} \mathbf{k}_\star, \quad (2.13)$$

die als Unsicherheitsmaß dient.

Betrachtet man nun den Erwartungswert der Verteilung aus Gleichung 2.12, so stellt man fest, dass diese Gleichung auch als eine lineare Kombination von N Kernfunktionen, die jeweils auf einen Punkt zentriert sind, aufgefasst werden kann [96]. Wir können die Gleichung also umformulieren in

$$\mathbb{E}[\hat{y}_*] = \sum_{i=1}^N \alpha_i k(\mathbf{x}_i, \mathbf{x}_*), \quad (2.14)$$

wobei $\alpha = (\mathbf{K} + \sigma_n^2 \mathbf{I}_N)^{-1} \mathbf{y}$ gilt.

Kovarianzfunktion

Die Kovarianzfunktion oder der Kernel des Gauß-Prozesses bestimmt die Kovarianz zwischen zwei Punkten \mathbf{x} und \mathbf{x}' . Für den Fall, dass die Mittelwertfunktion gleich Null gesetzt wird, was in der Praxis sehr häufig der Fall ist, enthält sie praktisch die gesamte Information über das Modell. Die Kovarianzfunktion ist maßgebend für die Klasse an Kernelfunktionen, die nach Gleichung 2.14 die finale Modellvorhersage bestimmen. Auch wenn gilt, dass mit $N \rightarrow \infty$ die Wahl der Kovarianzfunktion keine Rolle spielt, sind reale Datensätze auf Werte $N < \infty$ beschränkt. Dadurch ist der Einfluss der Kovarianzfunktion auf das Modell sehr groß. Die Kovarianzfunktion ist per Definition symmetrisch positiv-definit, d.h. für jeden beliebigen Satz von Eingangsparametern ist die resultierende Kovarianzmatrix positiv-definit. Bei der Verwendung sogenannter quadratisch exponentieller (engl.: *squared exponential*, SE) Kernel ergeben sich zum Beispiel besonders glatte Modellfunktionen. Diese Eigenschaft ist auf die unendliche Differenzierbarkeit des Kernels und somit auch des GPs zurückzuführen. Der quadratisch exponentielle Kernel ist die am weitesten verbreitete Kovarianzfunktion und stellt auch in dieser Arbeit, sofern nicht anders angegeben, den verwendeten Kernel dar. Er ist definiert als

$$k(\mathbf{x}_p, \mathbf{x}_q) = \sigma_s^2 \exp \left(-\frac{1}{2} (\mathbf{x}_p - \mathbf{x}_q)^\top \mathbf{\Lambda}^{-1} (\mathbf{x}_p - \mathbf{x}_q) \right) \quad (2.15)$$

mit $\mathbf{\Lambda} = \text{diag}(l_1^2, \dots, l_d^2)$, wobei $\sigma_s^2 \in \mathbb{R}_0^+$ die Signalvarianz darstellt und l_1^2, \dots, l_d^2 die sogenannten Lengthscales sind. Die Inverse der Lengthscales bestimmt im Wesentlichen die Relevanz eines Eingangs und implementiert damit eine Art automatische Relevanzbestimmung (engl.: *automatic relevance determination*, ARD) [86]. Bei einem kleinen Wert für die entsprechende Lengthscale hat die Eingangsgröße einen großen Einfluss auf den Modellausgang und bei einem sehr großen Wert ist der Modellausgang unabhängig von dieser Eingangsgröße.

Modelltraining

Ziel des Trainingsprozesses bzw. der Optimierung Bayes'scher Modelle ist die Optimierung der Modell-Hyperparameter Θ . In unserem Fall sind dies die Rauschvarianz σ_n^2 , die Signalvarianz σ_s^2 und die Lengthscales l_1^2, \dots, l_d^2 . Das Ziel ist die Maximierung der A-Posteriori-Verteilung der Hyperparameter.

Die Optimierung ist in der Regel mit der Auswertung mehrerer Integrale verbunden. Je nach Modellannahme sind die entsprechenden Integrale analytisch lösbar oder müssen durch geeignete Verfahren wie Markov Chain Monte Carlo approximiert werden. Anstelle der A-Posteriori-Verteilung der Hyperparameter kann auch die Evidenz bezüglich der Hyperparameter $P(\mathbf{y} \mid \mathbf{X}, \boldsymbol{\Theta})$ maximiert werden, die proportional zur A-Posteriori-Verteilung ist. Diese Approximation wird auch als Maximum-Likelihood-Methode vom Typ-II bezeichnet [96] und ist im Falle einer normalverteilten Likelihood-Funktion analytisch lösbar. Wie bereits erwähnt, sind Gauß-Prozesse unter Maximierung der Evidenz robust bezüglich optimierten Hyperparameter. Dies ist eine Eigenschaft der Evidenz, die einen automatischen Trade-off zwischen Modellkomplexität und Modellanpassung gewährleistet [96]. Wir können dies auch als eine inhärente Umsetzung von *Ockhams Rasiermesser* [71, 95] interpretieren, die größtmögliche Sparsamkeit bei der Modellauswahl verlangt und komplexe Modelle sowie unnötige Annahmen meidet. Dies ist auch der Grund, warum GPs im Allgemeinen nicht zur Überanpassung neigen, zumindest wenn die Maximum-Likelihood-Methode vom Typ-II verwendet wird. Aus Optimierungsgründen wird in der Regel die negative Log-Marginal-Likelihood minimiert. Dabei wird das Optimierungsproblem durch die Log-Transformation in ein konvexes Problem umgewandelt. Die Log-Marginal-Likelihood ist gegeben durch

$$\log P(\mathbf{y} \mid \mathbf{X}, \boldsymbol{\Theta}) = -\frac{1}{2} \mathbf{y}^\top (\mathbf{K} + \sigma_n^2 \mathbf{I}_N)^{-1} \mathbf{y} - \frac{1}{2} \log |\mathbf{K} + \sigma_n^2 \mathbf{I}_N| - \frac{N}{2} \log 2\pi, \quad (2.16)$$

wobei jeder Term eine unterschiedliche Funktion inne hat. Der erste Term bezieht die Daten mit ein und ist für die Anpassung des Modells an die Beobachtungen verantwortlich. Der zweite Term bestraft eine zu hohe Komplexität. Und der dritte Term ist als Normierungskonstante zu verstehen. Das entsprechende Optimierungsproblem wird beschrieben durch

$$\begin{aligned} \boldsymbol{\Theta}_{opt} &= \arg \min_{\boldsymbol{\Theta} \in \mathbb{R}} (-\log P(\mathbf{y} \mid \mathbf{X}, \boldsymbol{\Theta})) \\ &= \arg \min_{\boldsymbol{\Theta} \in \mathbb{R}} \left(\frac{1}{2} \mathbf{y}^\top (\mathbf{K} + \sigma_n^2 \mathbf{I}_N)^{-1} \mathbf{y} + \frac{1}{2} \log |\mathbf{K} + \sigma_n^2 \mathbf{I}_N| + \frac{N}{2} \log 2\pi \right). \end{aligned} \quad (2.17)$$

2.2. Evaluation datenbasierter Modelle

In diesem Abschnitt wird der letzte Schritt in der Modellbildungskette, die Modellevaluierung, behandelt. Die Evaluierung von datenbasierten Modellen ist von entscheidender Bedeutung, da sie die Qualität der Modellvorhersagen maßgeblich bestimmt. Zu diesem Zweck sollten geeignete Evaluierungsmetriken und Evaluierungsverfahren ausgewählt und die Modelle anhand dieser Metriken und Verfahren evaluiert werden. Bei der Auswahl sollte darauf geachtet werden, dass gängige Metriken und Methoden gewählt werden. Dies ermöglicht einerseits eine bessere Vergleichbarkeit und trägt andererseits zu einer leichteren Verständlichkeit bei. Das Kapitel ist wie folgt aufgebaut. Zunächst werden in Abschnitt 2.2.1 einige gängige Metriken zur Modellevaluation vorgestellt. Anschließend wird in Abschnitt 2.2.2 auf das Kreuzvalidierungsverfahren eingegangen und eine effiziente Variante der Kreuzvalidierung vorgestellt, die im Zusammenhang mit Gauss-Prozessen verwendet werden kann.

Schließlich wird im Abschnitt 2.2.3 das Bias-Variance-Dilemma diskutiert, das eines der größten Probleme der datenbasierten Modellierung darstellt.

2.2.1. Evaluationsmetriken zur Bewertung von Modellen

Zur Bewertung der Güte empirischer Modelle werden verschiedene Bewertungsmaße bzw. Fehlermaße als Kennzahlen verwendet. Diese können je nach Aufgabenstellung als absolute Zahlen oder in normierter Form verwendet werden. Darüber hinaus gibt es auch Metriken, die nur für ganz bestimmte Problemstellungen verwendet werden können—Metriken wie die Spezifität bzw. die True-Negative-Rate eines Modells können nur für Klassifikationsprobleme ermittelt werden. Einige wichtige Bewertungsmetriken, die im Rahmen dieser Arbeit verwendet wurden, werden im Folgenden vorgestellt.

Die Wertepaare $\{\mathbf{x}_i, y_i\}_{i=1}^N$ seien eine Sammlung an Einstellparametern und zugehörigen Messwerten eines beliebigen Systems. Zudem stelle die Funktion $f: \mathcal{X} \subset \mathbb{R}^{d \in \mathbb{N}} \rightarrow \mathcal{Y} \subset \mathbb{R}$ eine Abbildung dieses Systems dar. Die Aussagen des Modells f an den Stellen $\{\mathbf{x}_i\}_{i=1}^N$ seien durch $\{\hat{y}_i\}_{i=1}^N$ gegeben.

Mittlerer absoluter Fehler

Der mittlere absolute Fehler (engl.: *mean absolute error*, MAE) stellt die mittlere absolute Abweichung zwischen den Messwerten $\mathbf{y} = (y_1, \dots, y_N)^\top$ und den entsprechenden Vorhersagen $\hat{\mathbf{y}} = (\hat{y}_1, \dots, \hat{y}_N)^\top$ dar. Der MAE ist definiert durch

$$\text{MAE}(\mathbf{y}, \hat{\mathbf{y}}) = \frac{1}{N} \sum_{i=1}^N |y_i - \hat{y}_i|. \quad (2.18)$$

Relativer und prozentualer Fehler

Der relative Fehler (engl.: *relative error*, RE) stellt die relative Abweichung zwischen den Messwerten \mathbf{y} und den entsprechenden Vorhersagen $\hat{\mathbf{y}}$ dar. Der RE ist definiert durch

$$\text{RE}(\mathbf{y}, \hat{\mathbf{y}}) = \sum_{i=1}^N \frac{y_i - \hat{y}_i}{\hat{y}_i}. \quad (2.19)$$

Der RE gibt nicht an, inwieweit zwei korrespondierende Werte y_i und \hat{y}_i voneinander abweichen. Er misst lediglich die Abweichung zwischen den kumulierten Summen von y_i und \hat{y}_i . Die Verwendung dieser Metrik ist vor allem bei Zeitreihenanalysen sinnvoll, bei denen die zu vergleichenden Vektoren einen zeitlichen Versatz und damit einen hohen mittleren Fehler aufweisen.

Der RE kann auch als prozentualer Fehler (engl.: *percentage error*, PE) ausgedrückt werden. Die Gleichung ist dann durch

$$\text{PE}(\mathbf{y}, \hat{\mathbf{y}}) = 100\% \sum_{i=1}^N \frac{y_i - \hat{y}_i}{\hat{y}_i} \quad (2.20)$$

gegeben.

Mittlerer quadratischer Fehler

Eine häufig verwendete Metrik ist der mittlere quadratische Fehler (engl.: *mean squared error*, MSE). Der MSE misst die mittlere quadratische Abweichung zweier Vektoren \mathbf{y} und $\hat{\mathbf{y}}$ und ist definiert als

$$\text{MSE}(\mathbf{y}, \hat{\mathbf{y}}) = \frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2. \quad (2.21)$$

Das Quadrieren des Fehlers führt dazu, dass kleine Fehlerterme nur sehr wenig und große Fehlerterme stärker zum MSE beitragen. Dieses Fehlermaß reagiert daher stärker auf Ausreißer als Fehlermaße wie MAE. Dies kann wie folgt interpretiert werden. Ein kleines MSE bedeutet, dass sowohl die Verzerrung als auch die Varianz des Modells klein sind. Die Modellfunktion liegt also im Mittel nahe bei den zu schätzenden Funktionalen und weist eine geringe Verzerrung auf. Gleichzeitig streuen die Vorhersagen nur gering bzw. weisen eine geringe Varianz auf und liegen mit hoher Wahrscheinlichkeit in der Nähe ihres Erwartungswertes.

Der MSE ist wegen der Quadrierung schwer zu interpretieren. Stattdessen wird häufig der Root-Mean-Square-Error (RMSE) verwendet. Dieser liegt von der Größenordnung her im Bereich der anderen Fehlermaße und ist definiert als

$$\text{RMSE}(\mathbf{y}, \hat{\mathbf{y}}) = \sqrt{\frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2}. \quad (2.22)$$

Da der RMSE ein absolutes Maß ist, wird häufig auch der normalisierte Root-Mean-Square-Error (NRMSE) berechnet. Der NRMSE ist ein normierter Wert und stellt somit eine relative Größe dar, die eine bessere Vergleichbarkeit gewährleistet und zur besseren Interpretation der Ergebnisse beiträgt. Der NRMSE ist definiert als

$$\text{NRMSE}(\mathbf{y}, \hat{\mathbf{y}}) = \frac{\text{RMSE}}{\max(\mathbf{y}) - \min(\mathbf{y})}. \quad (2.23)$$

Im Rahmen dieser Arbeit wird die Min-Max-Normierung verwendet. Andere Normierungen sind möglich und sollten je nach Datengrundlage gewählt werden.

Bestimmtheitsmaß

Das Bestimmtheitsmaß (engl.: *coefficient of determination*, R^2) ist eine weitere Metrik zur Beurteilung von Regressionsmodellen. R^2 ist das Verhältnis der durch das Modell erklärten Variation zur Gesamtvariation in den Daten [121]. Es ist definiert als

$$R^2(\mathbf{y}, \hat{\mathbf{y}}) = \frac{\text{SSE}}{\text{SST}} = \frac{\sum_{i=1}^N (\hat{y}_i - \bar{\mathbf{y}})^2}{\sum_{i=1}^N (y_i - \bar{\mathbf{y}})^2} = 1 - \frac{\sum_{i=1}^N (y_i - \hat{y}_i)^2}{\sum_{i=1}^N (y_i - \bar{\mathbf{y}})^2}, \quad (2.24)$$

wobei SSE die Summe der erklärten Quadrate (*Sum of Squares Explained*), SST die Summe der zu erklärenden Quadrate (*Sum of Squares Total*) und $\bar{\mathbf{y}}$ das arithmetische Mittel der Messwerte \mathbf{y} ist. Ein R^2 von 1 bedeutet, dass die Residuen gleich 0 sind und das Modell perfekt an die Daten angepasst ist. Ein R^2 von 0 bedeutet hingegen, dass die *nicht erklärte Streuung* der gesamten *zu erklärenden Streuung* entspricht und die Eingangsparameter des Modells nichts zur Erklärung der Streuung beitragen.

Beurteilung binärer Klassifikationsmodelle

Für die Bewertung binärer Klassifikationsmodelle sind die oben beschriebenen Bewertungsmetriken in der Regel nicht ausreichend. Man stelle sich ein Klassifikationsproblem mit zwei Klassen (*positiv, negativ*) vor. Der Anteil der Daten, die zur ersten Klasse gehören, beträgt 99%. Man stelle nun eine Hypothese h auf, die als Ausgabe immer nur die erste Klasse liefert. Dies führt dazu, dass die Hypothese h mit den bisher vorgestellten Metriken als sehr gutes Modell eingestuft wird. In vielen Fällen ist diese Bewertung jedoch problematisch. Nehmen wir an, das beschriebene Problem ist die Klassifikation von Krebspatienten. In diesem Fall hätten wir ein Modell, das jeden Patienten als gesund klassifiziert und keinen Patienten als krank diagnostiziert. Mit den bisher vorgestellten Metriken würde das Modell dennoch eine Genauigkeit von nahezu 100% erreichen. Aus diesem Grund sind für Klassifikationsprobleme neben den bisher vorgestellten Metriken weitere Metriken notwendig, die die beschriebene Schwachstelle abdecken.

An dieser Stelle werden zwei Merkmale vorgestellt: Sensitivität und Spezifität. Für ein ausführlicheres Verfahren zur Bewertung von Klassifikationsmodellen wird auf [38, 99] verwiesen. Die Sensitivität ist definiert als

$$\text{Sensitivität} = \frac{\text{Anzahl richtig klassifizierter Positivexemplare}}{\text{Anzahl Positivexemplare}} \quad (2.25)$$

und gibt die Anzahl der richtig klassifizierten Positivproben im Verhältnis zur Gesamtzahl der Positivproben in den Beobachtungen an. Die Spezifität ist wiederum definiert als

$$\text{Spezifität} = \frac{\text{Anzahl richtig klassifizierter Negativexemplare}}{\text{Anzahl Negativexemplare}} \quad (2.26)$$

und gibt die Anzahl der richtig klassifizierten Negativproben im Verhältnis zur Gesamtzahl der Negativproben in den Beobachtungen an. Für das obige Beispiel ergäbe sich somit

eine Sensitivität von 1 und eine Spezifität von 0, was die Einseitigkeit der aufgestellten Hypothese verdeutlicht.

2.2.2. Kreuzvalidierung

Kreuzvalidierung (engl.: *cross-validation*, CV) ist ein statistischer Test, der zur Modellauswahl verwendet wird. Die Grundidee des Verfahrens besteht darin, den Trainingsdatensatz in zwei disjunkte Datensätze aufzuteilen, von denen einer tatsächlich für das Training verwendet wird, während der andere als Validierungsdatensatz dient. Die Leistung des auf der Trainingsmenge trainierten Modells im Vergleich zur Validierungsmenge wird als Indikator für den Generalisierungsfehler verwendet, und die Modellauswahl erfolgt auf der Grundlage dieses Maßes. Der Generalisierungsfehler bezieht sich hier auf den Fehler, den das Modell bei noch nicht gesehenen Daten macht.

Ein Nachteil dieses Verfahrens in der Praxis besteht darin, dass nur ein Bruchteil des gesamten Datensatzes für das Training verwendet werden kann und dass bei einem kleinen Validierungsdatensatz der berechnete Generalisierungsfehler eine große Varianz aufweisen kann. Um dieses Problem zu vermeiden, wird fast immer die k -fache Kreuzvalidierung angewandt. Bei der k -fachen Kreuzvalidierung werden die Daten in k disjunkte Mengen gleicher Größe aufgeteilt. Die Validierung erfolgt an einer einzigen Teilmenge, und das Training wird durch Vereinigung der verbleibenden Teilmengen durchgeführt. Diese Prozedur wird k -mal wiederholt, wobei die Validierungsmenge jedes Mal eine andere ist. Auf diese Weise kann ein großer Teil der Daten für das Training und alle Daten für die Validierung verwendet werden. Der Nachteil ist, dass k Modelle trainiert werden müssen. Typische Werte für k liegen zwischen 3 und 10 [96].

Ein Sonderfall k -facher Kreuzvalidierung ist die sogenannte Leave-One-Out-Kreuzvalidierung (engl.: *leave-one-out cross-validation*, LOOCV) [82, 109]. Hierbei wird die Trainingsmenge \mathcal{X} in $k = N_{obs}$ Mengen $\{\mathcal{X}_1, \dots, \mathcal{X}_{N_{obs}}\}$ mit jeweils einem Datenpunkt aufgeteilt und eine k -fache Kreuzvalidierung durchgeführt. Der Rechenaufwand für LOOCV ist aufgrund der hohen Anzahl der zu trainierenden Modelle unverhältnismäßig hoch. In bestimmten Fällen, wie z.B. bei der Verwendung von GPs, gibt es geeignete Approximationen für die Berechnung einer Bewertungsmetrik mittels LOOCV. Eine Möglichkeit zur Berechnung einer solchen Approximation wird in [113] beschrieben. Danach kann für ein GP-Modell f , das aus der Menge der Trainingsdaten trainiert wurde, eine Approximation für den Term \hat{y}_{-i} berechnet werden. Der Term \hat{y}_{-i} stellt dabei die Vorhersage eines Modells f_{-i} , das ohne den i -ten Datenpunkt trainiert wurde, an der Stelle x_i dar. Die Approximation ist gegeben durch

$$\hat{y}_{-i} = y_i - w_{ii}^{-1} v_i, \quad (2.27)$$

wobei w_{ii} der i te Diagonaleintrag der Matrix \mathbf{W} und v_i der i te Eintrag des Vektors \mathbf{v} ist. \mathbf{W} und v_i können als

$$\mathbf{W} = (\mathbf{K} + \sigma_n^2 \mathbf{I})^{-1} \quad (2.28)$$

und

$$\mathbf{v} = \mathbf{W} \mathbf{y} \quad (2.29)$$

notiert werden, siehe Gleichung 2.12. Das Residuum an einer Stelle x_i ist gegeben durch

$$e_{-i} = \hat{y}_{-i} - y_i = -w_{ii}^{-1} v_i. \quad (2.30)$$

2.2.3. Verzerrung-Varianz-Dilemma

Beim Training empirischer Modelle aus Daten stellt sich immer die Frage nach der *Modellkomplexität* und der *Modellgenauigkeit*. Es muss also festgelegt werden, wie komplex das zu trainierende Modell sein darf, um das betrachtete System hinreichend genau abzubilden. Hierbei können beliebige Annahmen bezüglich der Modelltopologie, der einzelnen Modellparameter oder auch bezüglich des Beobachtungsfehlers getroffen werden. Ein direkt damit zusammenhängendes Thema ist das so genannte Verzerrung-Varianz-Dilemma (engl. *bias-variance dilemma*).

Das Verzerrung-Varianz-Dilemma beschreibt den Konflikt zwischen der Verzerrung und der Varianz eines Modells. Der inhärente Trade-off zwischen diesen beiden Fehlerquellen führt dazu, dass die Minimierung der Verzerrung zu einer Erhöhung der Varianz führt und umgekehrt. Dies macht es für Lernalgorithmen schwer, zu generalisieren. Eine hohe Verzerrung resultiert dabei aus einer zu geringen Modellkomplexität. Bei zu hoher Verzerrung ist das Modell nicht in der Lage, die Zusammenhänge zwischen Modelleingängen und Modellausgabe zu erkennen. Man spricht hier auch von Unteranpassung (engl. *underfitting*). Dies ist insbesondere dann der Fall, wenn die reale Problemstellung einen komplexen Kontext aufweist, die Modellstruktur aber sehr einfach gewählt ist und z.B. nur lineare Zusammenhänge abbilden kann. Die Verzerrung nimmt mit steigender Komplexität bzw. Flexibilität des Modells deutlich ab. Die Varianz hingegen ist der Fehler, der sich aus der Streuung der Beobachtungen ergibt. Die Beobachtungen streuen also um den wahren Funktionsverlauf. Eine hohe Varianz in Kombination mit einer hohen Modellkomplexität führt in der Regel dazu, dass sich das Modell an die Daten anpasst und die Trainingsdaten perfekt, aber nicht den wahren Funktionsverlauf nachbildet. Dieses Phänomen wird auch als Überanpassung (engl. *overfitting*) bezeichnet und ist in Abbildung 2.1 dargestellt. Die Varianz steigt linear mit der Komplexität des Modells [59].

Der Modellfehler kann nach [59] in drei Komponenten unterteilt werden. Dies sind der Verzerrungsfehler, der Varianzfehler und ein irreduzibler Fehleranteil, der vom probleminternen Rauschen herrührt. Der erwartete quadratische Modellfehler ist definiert durch

$$e(\mathbf{x})^2 = (\mathbb{E} [\hat{y}(\mathbf{x}) - y(\mathbf{x})])^2 + \mathbb{E} [(\hat{y}(\mathbf{x}) - \mathbb{E} [\hat{y}(\mathbf{x})])^2] + \sigma_{irr}^2, \quad (2.31)$$

wobei y die Beobachtung und \hat{y} die Vorhersage ist. Der erste Term steht für den Verzerrungsfehler, der zweite für den Varianzfehler und der dritte für den irreduziblen Anteil. Daraus wird ersichtlich, dass der Bias die Abweichung zwischen Vorhersage und Beobachtung angibt. Die Varianz gibt an, wie stark das Messrauschen das Lernen beeinflusst. Der irreduzible Fehleranteil wiederum stellt eine untere Schranke für die zu erwartende Abweichung bei bisher unbekannten Testdaten dar.

In der Literatur finden sich je nach verwendetem Algorithmus verschiedene Methoden zur Reduzierung des Modellfehlers. Der Verzerrungsfehler kann insbesondere durch eine geeig-

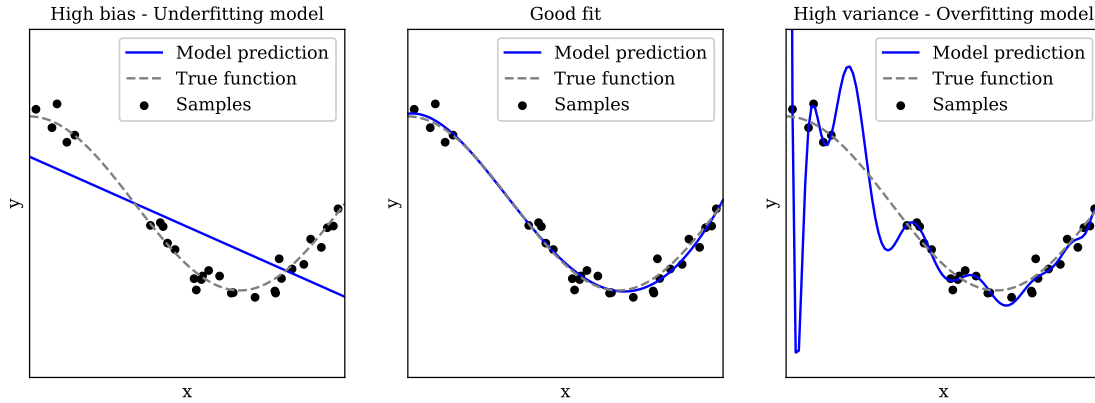


Abbildung 2.1.: Von links nach rechts: Modell mit Unteranpassung. Gut angepasstes Modell. Modell mit Überanpassung.

nete Wahl der Modellkomplexität reduziert werden. Die Varianz hingegen kann durch den Einsatz von Regularisierungsverfahren, wie dem Early-Stopping bei neuronalen Netzen und diversen anderen Methoden reduziert werden [94, 8]. Die Idee hinter der Regularisierung ist, dass die Modellparameter aufgrund der Regularisierungsfunktion nicht beliebig stark ansteigen können. Dies führt dazu, dass das Modell nicht so stark an die Trainingsdaten angepasst wird, was wiederum die Varianz reduziert. Die Regularisierungsfunktion kann beliebig gewählt werden. In der Literatur werden häufig die L1- und L2-Regularisierung verwendet. Die L1-Regularisierung führt dazu, dass die Modellparameter auf eine Summe der Absolutwerte der Parameter beschränkt werden. Die L2-Regularisierung führt hingegen zu einer Beschränkung der Modellparameter auf eine Summe der quadratischen Werte der Parameter.

Wie in Abschnitt 2.1 bereits erwähnt, sind Gauß-Prozesse unter Maximierung der Evidenz mit Hilfe der Maximum-Likelihood-Methode vom Typ-II robust bezüglich der optimierten Hyperparameter und weisen einen automatischen Trade-off zwischen Modellkomplexität und Modellanpassung auf. Sie weisen eine möglichst geringe Modellkomplexität unter möglichst hoher Modellanpassung auf und neigen in der Regel nicht zur Überanpassung. Dies schließt jedoch eine Über- oder Unteranpassung nicht aus. In der Praxis treten diese Phänomene auf und müssen entsprechend berücksichtigt werden.

2.3. Statistische Stichprobennahme bzw. Sampling

Unter Stichprobenziehung (engl.: *Sampling*) versteht man in der Statistik die Ziehung einer statistischen Stichprobe von Individuen aus einer statistischen Grundgesamtheit zur Schätzung von Merkmalen der gesamten Grundgesamtheit. Ziel ist es, die Grundgesamtheit repräsentativ abzubilden.

Definition 2.1 X sei eine multivariate Zufallsvariable auf einem Wahrscheinlichkeitsraum (Ω, Σ, P) mit Ergebnisraum Ω , Sigma-Algebra Σ und Wahrscheinlichkeitsmaß P . Dann ist

$Y = g(X)$ eine Transformation von X unter g und wiederum eine Zufallsvariable auf Ω . Ihre kumulative Verteilungsfunktion ist gegeben durch

$$F_Y(y) = P(Y \leq y). \quad (2.32)$$

Ein in industriellen Anwendungen häufig anzutreffendes Szenario ist das folgende: Ein physikalisches System kann durch eine Funktion f mit Eingangsparametern \mathbf{x} beschrieben werden. Die Eingangsparameter des Systems sind wiederum zufällig und diese Zufallsereignisse können durch eine stetige multivariate Verteilung beschrieben werden. Die Aufgabe des Anwenders besteht darin, die Verteilungsfunktion $F_Y(y)$ herzuleiten. In solchen Szenarien ist die Lösung dieser Aufgabe mit analytischen Methoden, abgesehen von einigen Sonderfällen, oft nicht möglich. Die Bestimmung der Verteilung ist dann häufig nur mit Hilfe sogenannter Stichprobenverfahren realisierbar.

Eine Möglichkeit zur Approximation der Verteilung bietet zum Beispiel das klassische Monte-Carlo (MC)-Verfahren [79], das ein reines Zufallsexperiment darstellt. Für das hier beschriebene Beispiel kann das Verfahren in drei Schritten beschrieben werden:

- **Stichprobennahme / Sampling:** Entnehme N zufällige Stichproben aus der multivariaten Eingangsparameterverteilung. Die resultierende Stichprobenpopulation kann durch die Menge $\mathcal{D} = \{\mathbf{x}_j\}_{j=1}^N$ zusammengefasst werden.
- **Fortpflanzung / Propagation:** Propagiere jede Stichprobe $\mathbf{x}_j \in \mathcal{D}$ mit dem Modell f . Die resultierenden Ausgangswerte können durch die Menge $\mathcal{F} = \{f_j\}_{j=1}^N$ zusammengefasst werden, wobei $f_j = f(\mathbf{x}_j)$ gilt.
- **Zusammenfassen / Aggregieren:** Fasse die Menge \mathcal{F} zu einer Wahrscheinlichkeitsverteilung zusammen.

Die Genauigkeit der Verteilung hängt von zwei Faktoren ab. Zum einen spielt die Genauigkeit des Modells f eine Rolle, da diese die Abweichung der Realität von der vorhergesagten Größe bestimmt. Zum anderen spielt die Anzahl der Stichproben eine Rolle, da die aus der Menge \mathcal{F} ermittelte Wahrscheinlichkeitsverteilung eine Schätzung darstellt und diese Schätzung mit zunehmender Anzahl der Stichproben an Genauigkeit gewinnt. Nach dem Gesetz der großen Zahlen [36] konvergiert die geschätzte Verteilung gegen die wahre Verteilung.

Definition 2.2 Gesetz der großen Zahlen: Die relative Häufigkeit eines Zufallsergebnisses nähert sich stetig an die tatsächliche Wahrscheinlichkeit des Experiments an, je öfter das Zufallsexperiment durchgeführt wird [106].

Das oben beschriebene MC-Verfahren verwendet Pseudozufallsfolgen zur Stichprobenziehung. Der große Nachteil dieses Verfahrens ist das langsame Konvergenzverhalten. Die Konvergenzordnung bzw. der Stichprobenfehler des Monte-Carlo-Verfahrens kann mit $\mathcal{O}(\frac{1}{\sqrt{N}})$ angegeben werden [3], wobei N die Anzahl der Stichproben angibt. Alternativ zur Monte-Carlo-Stichprobenziehung mit Pseudozufallszahlen können auch Verfahren mit geschichteten Zufallsstichproben oder Quasi-Monte-Carlo-Verfahren eingesetzt werden. Diese Verfahren weisen eine bessere Konvergenzordnung auf und werden daher häufig bevorzugt.

2.3.1. Quasi-Monte-Carlo-Sampling

Im Gegensatz zu den beim MC-Verfahren verwendeten Pseudozufallsfolgen werden beim Quasi-Monte-Carlo (QMC)-Verfahren Niederdiskrepanz-Folgen verwendet. Eine der bekanntesten hier verwendeten Folgen ist die Sobol-Folge [107], die erstmals 1967 von dem russischen Mathematiker Ilya Meyerovich Sobol eingeführt wurde. Wenn im Folgenden von QMC die Rede ist, ist damit die Verwendung von Sobol-Sequenzen impliziert. Diese Art der Verteilung bzw. des Versuchsdesigns füllt den gesamten Versuchsraum aus und wird daher als raumfüllendes Versuchsdesign bezeichnet [61]. Durch die gleichmäßige Verteilung der Stichproben wird, wie der Name schon sagt, die Varianz der Stichprobe reduziert. QMC-Verfahren gehören damit zu den Varianzreduktionsverfahren [106]. Sie konvergieren in der Regel deutlich schneller als die Standard-Monte-Carlo-Methode [3]. Die Konvergenzordnung für QMC-Verfahren beträgt etwa $\mathcal{O}(\frac{1}{N})$ [106].

In der Literatur finden sich verschiedene Algorithmen, mit denen sich Sobol-Folgen erzeugen lassen [107, 2, 41, 92, 108]. Darüber hinaus existieren bereits implementierte Algorithmen in verschiedenen Programmiersprachen [27], die zur Generierung von Sobol-Folgen verwendet werden können. Eine Möglichkeit zum Erzeugen einer Sobol-Folge ist in Anhang A.2 beschrieben.

2.3.2. Latin-Hypercube-Sampling

Eine weitere Möglichkeit der Stichprobengenerierung stellt das Latin-Hypercube-Sampling dar [77]. Ein Latin-Hypercube-Design ist wie ein Sobol-Design raumfüllend und gehört ebenfalls zu den Varianzreduktionsverfahren. Bei dieser Art der Stichprobenziehung wird die Grundgesamtheit in Schichten bzw. Gruppen unterteilt und die Stichprobe aus den einzelnen Schichten gezogen. Die Schichten unterscheiden sich dabei in einem oder mehreren Merkmalen. Außerdem wird die Anzahl der zu ziehenden Stichproben pro Schicht festgelegt. Dadurch wird die reine Zufallsauswahl der Stichprobe eingeschränkt. Ein Vorteil dieser Methode ist die Vermeidung extrem heterogener Punktverteilungen. So kann beispielsweise vermieden werden, dass alle Stichproben in der gleichen Schicht liegen. Die Konvergenzordnung des Latin-Hypercube-Sampling ist in der Regel identisch mit der Konvergenzordnung des Standard-Monte-Carlo-Ansatzes [78]. In der Literatur finden sich verschiedene Ansätze zur Generierung von Stichproben, die dem Latin-Hypercube-Design folgen, siehe [106].

3. Versuchsaufbau

In diesem Kapitel wird der Versuchsaufbau für die Experimente beschrieben. In Abschnitt 3.1 wird auf die beiden Versuchsträger eingegangen. Anschließend wird in Abschnitt 3.2 der Aufbau des Prüfstandes dargestellt. In Abschnitt 3.3 wird die Messdatenerfassung inklusive Indizierung und Abgasanalyse beschrieben.

3.1. Versuchsträger

Für die Versuche im Rahmen dieser Arbeit wurden zwei unterschiedliche Versuchsträger verwendet. Ein Reihen-Vierzylinder-Dieselmotor, der zu Forschungszwecken bei der Robert Bosch GmbH eingesetzt wird und ein V6-Dieselmotor, der voll appliziert und bereits mit Euro 6d-Temp Norm in Serie ist.

Zur besseren Abbildung der realen Fahrsituation im Prüfstandsbetrieb sind beide Motoren mit einer vollständigen Abgasnachbehandlung und einer variablen Abgasklappe ausgestattet. Darüber hinaus wurde bei beiden Motoren darauf geachtet, dass sie möglichst mit serienmäßigen Komponenten ausgestattet sind. Dies ergibt sich unter anderem aus den Modellannahmen.

3.1.1. Reihen-Vierzylinder-Dieselmotor

Der erste Versuchsträger ist ein aufgeladener Reihen-Vierzylinder-Dieselmotor mit Direkteinspritzung, der bei der Robert Bosch GmbH zu Forschungszwecken eingesetzt wird. Der Motor ist nahezu baugleich mit dem in [66] vorgestellten Versuchsträger. Die wichtigsten Spezifikationen des Versuchsträgers sind in Tabelle 3.1 dargestellt.

Tabelle 3.1.: Spezifikationen des Vierzylinder-Forschungsmotors.

Motortyp	PKW-Dieselmotor
Aufbau	Reihenmotor
Zylinder / Ventile pro Zylinder	4 / 4
Hubraum	1667 cm ³
Verdichtungsverhältnis	16.2

Der Motor ist mit einem Bosch CRS2-22 Common-Rail-System mit Magnetventil-Injektoren ausgestattet. Dies ermöglicht eine Variation des Einspritzmusters mit bis zu 9 Einspritzungen pro Verbrennungszyklus und einem Mindestabstand zwischen zwei Ventilöffnungen von

180 μ s. Er ist mit einem Turbolader mit variabler Turbinengeometrie (VTG) ausgestattet, um unterschiedliche Ladedrücke zu erreichen. Die Abgasrückführung (AGR) wird über ein gekühltes Hochdruck-AGR-System mit Bypass und ein gekühltes Niederdruck-AGR-System mit Bypass realisiert. Dies ermöglicht den Einsatz hoher AGR-Raten unter verschiedenen Bedingungen. Zusätzlich können die einzelnen Zylinder über eine Drallklappe gedrosselt werden. Dies dient als Maßnahme zur Steuerung des Drallniveaus des Motors. Hinter dem Turbolader regelt ein flüssigkeitsgekühlter Ladeluftkühler (LLK) die Ansauglufttemperatur. Zusätzlich regelt eine Abgasklappe den Endrohrgegendruck am Prüfstand.

Auf der Abgasseite ist der Motor mit einer Reihe von Katalysatoren ausgestattet. Hinter dem Turbolader befindet sich zunächst ein Dieseloxydationskatalysator (DOC). Dem DOC folgt ein eng an ihn gekoppelter Katalysator, der eine Kombination aus selektiver katalytischer Reduktion (SCR) und Dieselpartikelfilter (DPF) darstellt und als SCR beschichteter DPF (SDPF) bezeichnet wird. Hinter dem SDPF befindet sich ein Unterflur-SCR-Brick mit zonaler Beschichtung, der einen Ammoniumschlupfkatalysator darstellt. Standardmäßig ist der Motor mit einem NO_x -Speicherkatalysator (engl.: *Lean NOx Trap*, LNT) anstelle des DOC ausgestattet. Um thermische Schäden am LNT zu vermeiden, wird stattdessen ein identischer DOC eingebaut. Die Versuche werden durch diesen Umbau nicht wesentlich beeinflusst, da für die Untersuchungen im Rahmen dieser Arbeit nur die Abgaszusammensetzung des Rohgases von Bedeutung ist. Eine schematische Darstellung des Motors ist in Abbildung 3.1 zu sehen.

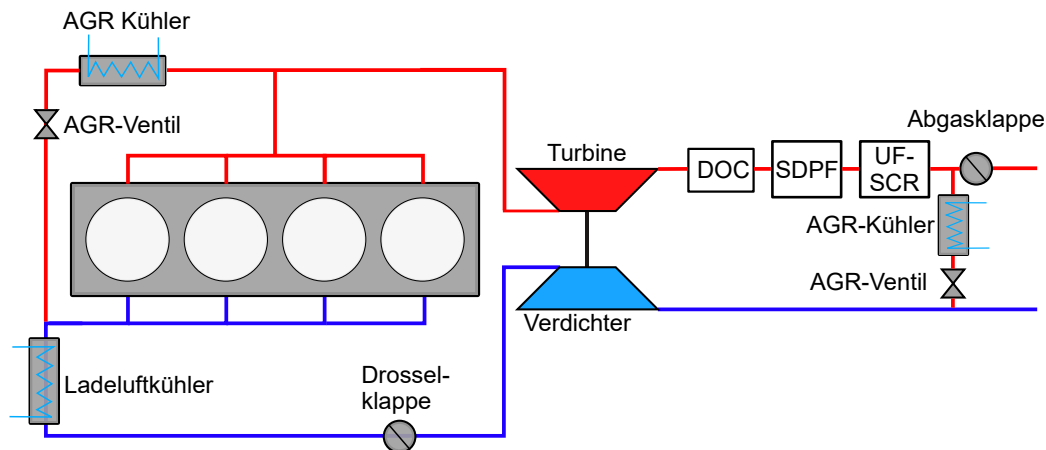


Abbildung 3.1.: Schematischer Aufbau des Vierzylindermotors.

Der Motor ist mit einem MDG1-Steuergerät von Bosch ausgerüstet. Das Steuergerät ist zusätzlich mit einem elektronischen Tastkopf von Bosch ausgestattet und wird mit der Software INCA über ein Schnittstellenmodul vom Typ ES590 von ETAS angesteuert.

3.1.2. V6-Dieselmotor

Bei dem zweiten Versuchsträger handelt es sich um einen aufgeladenen Sechszylinder-Dieselmotor in V-Bauweise mit Direkteinspritzung. Der Versuchsträger ist ein vollständig

applizierter Serienmotor der Euro 6d-Temp Norm. Der Datenstand des V6-Motors unterscheidet sich von der Serie insofern, als weitere Anpassungen an der Applikation vorgenommen wurden, um die Emissionen weiter zu reduzieren. Die wichtigsten Spezifikationen des Versuchsfahrzeugs sind in Tabelle 3.2 dargestellt.

Tabelle 3.2.: Spezifikationen des Sechszylinder-V-Motors.

Motortyp	PKW-Dieselmotor
Aufbau	V-Motor
Zylinder / Ventile pro Zylinder	6 / 4
Hubraum	2967 cm ³
Verdichtungsverhältnis	16.5

Der Motor ist mit einem Common-Rail-System CRS2-25 von Bosch mit Magnetventil-Injektoren und Needle-Closing-Control ausgestattet [119]. Wie beim CRS2-22-System ist damit eine sehr präzise Steuerung des Einspritzverlaufs möglich. Das Needle-Closing-Control-System ermöglicht zudem durch den Closed-Loop-Ansatz eine sehr genaue Dosierung der einzelnen Einspritzmengen. Der Motor ist mit einem VTG-Turbolader ausgerüstet. Wie beim Vierzylindermotor sind ein gekühltes Hochdruck-AGR-System mit Bypass und ein gekühltes Niederdruck-AGR-System mit Bypass integriert. Zusätzlich können die Zylinder über eine Drallklappe gedrosselt werden. Damit wird der Drall des Motors gesteuert. Ein flüssigkeitsgekühlter Ladeluftkühler regelt die Temperatur der Ansaugluft hinter dem Turbolader. Zusätzlich regelt eine Abgasklappe den Endrohrgegendruck am Prüfstand.

Abgasseitig ist auch bei diesem Versuchsträger eine Reihe von Katalysatoren eingebaut. Hinter dem Turbolader ist zunächst ein LNT verbaut. Dem LNT folgt ein SDPF, der eng mit dem LNT gekoppelt ist. Dem SDPF ist ein Unterflur-SCR-Katalysator nachgeschaltet, dem ein Ammonium-Shift-Katalysator folgt. Eine Abgasklappe regelt den Abgasgegendruck. Eine schematische Darstellung des Motors zeigt Abbildung 3.2.

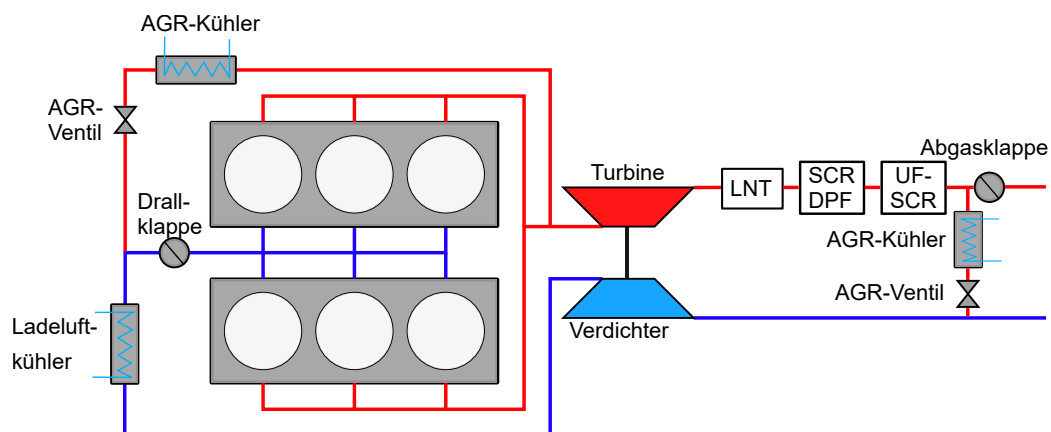


Abbildung 3.2.: Schematischer Aufbau des V6-Motors.

Auch dieser Motor ist mit einem MDG1-Steuergerät von Bosch ausgerüstet. Das Steuergerät ist zusätzlich mit einem elektronischen Tastkopf von Bosch ausgestattet und wird von INCA über ein Schnittstellenmodul vom Typ ES590 von ETAS angesteuert.

3.2. Motorprüfstand

Der Motorenprüfstand ist ein von der Firma AVL List GmbH konfiguriertes System. Die Steuerung des Prüfstandes erfolgt über das System PUMA OPEN von AVL. Die Motordrehzahl wird durch einen Asynchron-Elektromotor vom Typ APA 302 der Firma AVL geregelt. Der Elektromotor kann sowohl ein negatives als auch ein positives Drehmoment erzeugen. Das Drehmoment wird mit einer Drehmomentmesswelle gemessen. Der Kraftstoffverbrauch wird mit einem Kraftstoffverbrauchsmesssystem AVL 735S gemessen und mit dem System AVL 753S aufbereitet. Die Ansaugluft wird mit einem ABB Sensyflow P-100 gemessen. Die Temperatur der Ansaugluft wird bei allen Prüfungen auf $T=23^{\circ}\text{C}$ gehalten. Zusätzlich wird die Luftfeuchtigkeit überwacht und geregelt.

3.3. Messdatenerfassung

In diesem Abschnitt wird die Messdatenerfassung und die zugehörige Messsensorik veranschaulicht. In Unterabschnitt 3.3.1 wird die Indiziermesstechnik beschrieben. Und in Unterabschnitt 3.3.2 ist die Abgasanalyse dargestellt.

3.3.1. Indizierung

Die Überwachung und Aufzeichnung der indizierten Größen erfolgt mit der Software Indicom von AVL über ein Indiset Advanced Plus 641 der gleichen Firma. Die überwachten Größen sind in Tabelle 3.3 dargestellt. Der Commonraildruck wird mit piezoelektrischen Hochdrucksensoren vom Typ 4067A3000 der Firma Kistler gemessen. Diese werden mit einem Kistler Klemmadapter vom Typ 6533A12 am Commonrail befestigt. Jeder einzelne Zylinder des Vier- und Sechszylindermotors ist mit einem piezoelektrischen Niederdrucksensor vom Typ 4075A10 von Kistler ausgestattet. Diese werden über einen Glühkerzenadapter in den Zylinder integriert. Die Ladung an den Sensoren wird mit Ladungsverstärkern vom Typ 5064 von Kistler aufbereitet. Die Erfassung des Injektorsignals erfolgt über eine Stromzange der Firma Chauvin Arnoux vom Typ E3N. Zur Überwachung der Gastemperaturen sind die Motoren ausgangsseitig mit Thermoelementen Typ K und eingangsseitig mit PT100-Temperaturfühlern ausgerüstet. Die Motoren sind außerdem mit Drucksensoren vom Typ C-FEM-P der Firma AVL ausgestattet, um die Gasdrücke an verschiedenen Positionen zu überwachen. Das Verbrennungsgeräusch wird aus den erfassten Größen mit einer zylinderdruckbasierten Methode berechnet [76].

Tabelle 3.3.: Messdatenerfassung und -sensorik.

Messgröße	Sensorik
Commonraildruck	Kistler 4067A3000
Zylinderdruck	Kistler 4075A10
Ladung	Kistler 5064
Injektorsignal	Chauvin Arnoux E3N
Gastemperatur	Thermoelemente Typ K, PT100
Gasdruck	AVL C-FEM-P
Verbrennungsgeräusch	zylinderdruckbasiert

3.3.2. Abgasanalyse

Die Abgasemissionen beider Motoren werden mit einer Vielzahl von Analysegeräten gemessen. Dabei werden sowohl die festen als auch die gasförmigen Abgasbestandteile erfasst.

Die Partikelkonzentration wird im stationären Betrieb mit einem AVL 415S Smoke Meter und im transienten Betrieb mit einem AVL 483 Micro Soot Sensor bestimmt. Die Probenahmestelle befindet sich bei beiden Motoren nach dem Turbolader und vor dem Eintritt in den motornahen Katalysator. Im stationären Betrieb wird zusätzlich ein AVL 439 Opazimeter zur Überwachung der Partikelkonzentration eingesetzt.

Die gasförmigen Abgasbestandteile werden mit zwei Abgasanalysatoren vom Typ MEXA-ONE der Firma Horiba bestimmt. Die erste Probenahmestelle befindet sich auch hier nach dem Turbolader und vor dem Eintritt in den motornahen Katalysator. Die zweite Probenahmestelle befindet sich nach der Drosselklappe und dient zur Bestimmung der physikalischen AGR-Rate. Tabelle 3.4 gibt einen Überblick über die Abgasanalysegeräte.

Tabelle 3.4.: Übersicht der Analyseeinheiten des Horiba MEXA-ONE.

Abgaskomponente	Messprinzip	Messbereiche
CO	Nichtdispersiver Infrarotsensor	0-5000ppm; 5000-50000ppm
CO ₂	Nichtdispersiver Infrarotsensor	0-20vol%
HC	Flammenionisationsdetektor	0-1000ppm; 1000-60000ppm
NO, NO ₂	Chemilumineszenzdetektor	0-1000ppm; 1000-10000ppm
O ₂	Magnetopneumatischer Detektor	0-25vol%

4. SEMO-Methode

Das primäre Ziel dieser Arbeit ist die Entwicklung einer Applikationsmethodik, die einen Paradigmenwechsel im Umgang mit Toleranzeinflüssen einleitet. Im Rahmen des Applikationsprozesses soll anstelle eines einzelnen Nominalfahrzeugs die gesamte Fahrzeugflotte betrachtet werden. Die vorgestellte Methodik setzt hierbei bereits bei der Basisapplikation an. Es wird ein modellbasierter Applikationsansatz verfolgt, der eine Applikation des Motors unter Berücksichtigung der gegebenen Bauteiltoleranzen ermöglicht. Die Methodik lässt sich in drei Themenbereiche unterteilen. Diese sind die in Kapitel 5 vorgestellte Versuchsplanung, das in Kapitel 6 vorgestellte stochastische Motormodell und der in Kapitel 7 vorgestellte stochastische Optimierungsansatz. Die Methodik wurde bereits in [85] anhand eines Anwendungsbeispiels vorgestellt. Ein wesentlich komplexeres Anwendungsbeispiel wurde in [84] dargestellt. Teile dieser Ergebnisse werden auch in dieser Arbeit verwendet. Die Methodik wurde bereits in einem Anwendungsbeispiel in [85] angewendet. Ein wesentlich komplexeres Anwendungsbeispiel wurde in [84] dargestellt. Teile dieser Ergebnisse werden auch in dieser Arbeit verwendet.

Dieses Kapitel ist folgendermaßen aufgebaut. In Abschnitt 4.1 wird zunächst der Stand der Technik im Bereich der Motorsteuergeräte-Applikation aufgearbeitet. Abschnitt 4.2 beschreibt die Anforderungen an die Methodik. In Abschnitt 4.3 wird der Ablauf der Methode beschrieben. Der Abschnitt 4.4 schließt das Kapitel mit einer kritischen Bewertung und Abgrenzung der Methodik gegenüber dem Stand der Technik ab.

4.1. Stand der Technik im Bereich der Motorsteuergeräteapplikation

Seit der Einführung der Bosch D-Jetronic im Jahr 1967 tragen Motorsteuergeräte entscheidend zur Optimierung von Motoren bei [19]. Seitdem ist auch die Zahl der im Motor und im Steuergerät eingesetzten Technologien stark angestiegen [19]. Und mit jeder neuen Motortechnologie steigt die Flexibilität, die physikalischen und chemischen Prozesse im Motor zu beeinflussen. Um das Potenzial dieser Technologien auszuschöpfen, ist es jedoch notwendig, die einzelnen Steller bzw. Aktuatoren des Motors präzise zu steuern. Diese Aufgabe übernimmt das Motorsteuergerät, ein Mikrocontroller, der aus Hardwarekomponenten und der Software besteht [103]. Die Software kann in Programm- und Datenstand unterteilt werden. Dies ist in Abbildung 4.1 dargestellt. Der Programmstand beinhaltet den im Steuergerät abgebildeten Funktionsrahmen mit den einzelnen Funktionen und deren Abhängigkeiten. Diese berechnen aus den veränderlichen und unveränderlichen Parametern sowie den gemessenen Sensorgrößen die Ausgabewerte des Steuergeräts. Der Datenzustand umfasst die Gesamtheit der veränderlichen Parameter des Steuergeräts. Dies können unterschiedliche Strukturen wie Kennlinien und Kennfelder, aber auch skalare Werte sein. Die Kalibrierung

des Datenstandes bzw. der Steuergeräteparameter zur Beeinflussung des Motorverhaltens wird allgemein als Motorapplikation bzw. *Applikation* bezeichnet. Bei der Applikation wird häufig auch von der Applikation von Funktionen gesprochen. Damit ist die Veränderung der mit diesen Funktionen verbundenen Steuergeräteparameter gemeint.

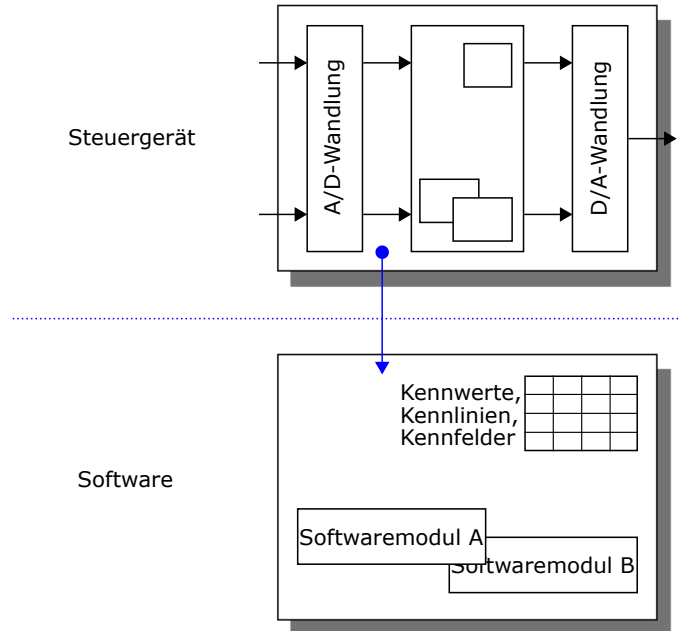


Abbildung 4.1.: Aufbau eines Motorsteuergeräts.

Übersicht der Applikationsthemen

Die Anzahl der Steuergeräteparameter für das Dieselmotorsteuergerät liegt im hohen fünfstelligen Bereich und steigt immer weiter an [70]. Aufgrund der hohen Anzahl an Parametern und der daraus resultierenden Komplexität wird die Applikation in der Regel in verschiedene Themenbereiche unterteilt. Betrachtet man den Gesamtablauf der Applikation, so wird zunächst die *Basisapplikation* realisiert, gefolgt von der *stationären Applikation*, der *dynamischen Applikation* und der *Fahrzeugapplikation*. Diese einzelnen Themenbereiche setzen sich wiederum aus kleineren Bausteinen wie der Modellierung von Zustandsgrößen, der Emissionsmodellierung, der Leerlaufregelung, dem Motor Start-Stopp, der On-Board-Diagnose und der Realisierung von Betriebszuständen wie dem dynamischen Motorverhalten zusammen. Aufgrund der zahlreichen Wechselwirkungen zwischen den einzelnen Funktionen hat eine Änderung von Steuergeräteparametern fast immer einen Einfluss auf verschiedenste Funktionen. Dies macht die Applikation zu einem iterativen Prozess, der ein hohes Maß an Abstimmung zwischen den Applikateuren erfordert und mit hohem Aufwand und hohen Kosten verbunden ist.

Zu Beginn eines Projektes ist eine grundlegende Parametrierung des zu applizierenden Motors erforderlich. Diese wird in der Regel als *Basisapplikation* bezeichnet. Die Basisapplika-

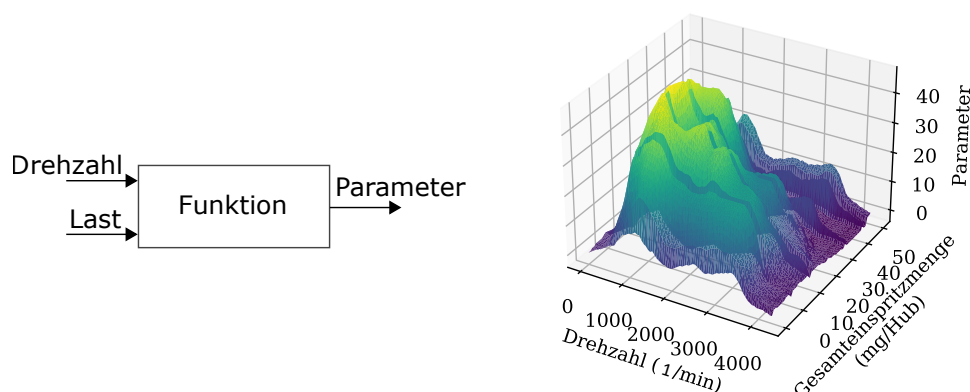


Abbildung 4.2.: Schematische Darstellung der Motorparameter über der Drehzahl und der eingespritzten Kraftstoffmasse.

tion enthält unter anderem *Basiskennfelder*, die systemrelevant und für den sicheren Betrieb des Motors unabdingbar sind, und stellt damit eine Voraussetzung für alle weiteren Applikationsschritte dar. Die Basiskennfelder enthalten für jeden Betriebspunkt Vorsteuerwerte für betriebsrelevante Parameter wie Commonraildruck, Abgasrückführrate, Einspritzmenge und Einspritzzeitpunkt. Die Parametrierung der Basiskennfelder erfolgt am *Motorprüfstand* anhand von stationären Messungen. Die Basiskennfelder sind, wie in Abbildung 4.2 dargestellt, eine Funktion der Drehzahl und der gesamten eingespritzten Kraftstoffmasse pro Hub. Die Abhängigkeit von der Gesamteinspritzmenge ergibt sich aus der beim Dieselmotor prinzipbedingten Gemischmengenregelung, bei der das Drehmoment bzw. die Motorlast über die Menge des zugeführten Gemisches geregelt wird. Bei Nachfolgemotoren werden häufig die Applikationsdaten des Vorgängermotors als Basis verwendet. Darauf aufbauend werden die Parameter in verschiedenen Betriebspunkten nach unterschiedlichen Kriterien optimiert. Im Leerlaufbereich spielt vor allem die Laufruhe eine große Rolle. Im Teillastbereich stehen Verbrauch, Emissionen und Verbrennungsgeräusch im Vordergrund und im Volllastbereich ist die Leistung das dominierende Kriterium. Die einzelnen Motorbetriebsbereiche sind in Abbildung 4.3 dargestellt. Neben der Betriebspunktabhängigkeit sind auch Motor- und Katalysatortemperaturen zu berücksichtigen. Daher sind Kennfelder sowohl für den Warmlauf als auch für den Kaltlauf zu parametrieren. Hat der Katalysator seine Betriebstemperatur noch nicht erreicht, steuern zusätzliche Katalysatorheizkennfelder den Motor.

Im Anschluss an die Basisapplikation folgen weitere Applikationsmaßnahmen. Zu diesen gehören zum einen die Fahrbarkeitsapplikation. Hierbei werden Anpassungen der Komfortfunktionen und Funktionen, welche für mehrere Komponenten relevant sind (z.B. Schaltvorgänge für Motor und Getriebe) vorgenommen. Zum anderen gehören hierzu aber auch die Emissions- und Verbrauchsoptimierung sowie die Diagnoseapplikation, bei der die Diagnosefunktionen appliziert werden. Neben dieser Klassifizierung kann die Applikation wie oben bereits geschildert in stationäre Applikation, dynamische Applikation und Fahrzeugapplikation unterteilt werden. Die stationäre sowie die dynamische Applikation unterscheiden sich hauptsächlich durch die Art der Vermessung, die für die Parametrierung der Funktionen benötigt werden. Bei stationärer Applikation werden für die Applikation der einzelnen

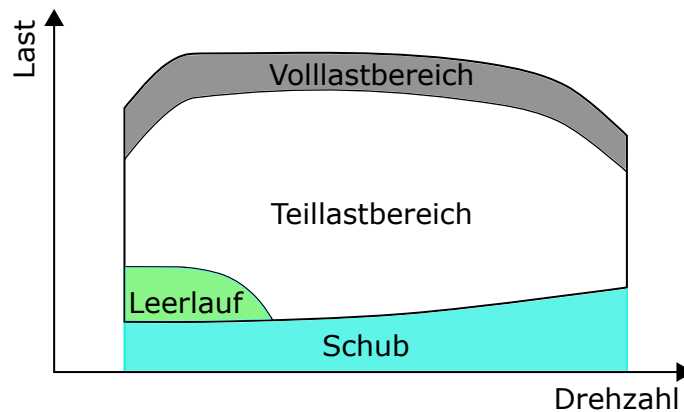


Abbildung 4.3.: Motorbetriebsraum über Drehzahl und Last.

Funktionen zeitinvariante Messdaten verwendet. Hierbei wird am Motorprüfstand der Betriebspunkt vor Beginn der Messung konstant gehalten und *stabilisiert*. Nach der Stabilisierung beginnt die eigentliche Messung, welche einen arithmetischen Mittelwert über eine vordefinierte *Messzeit* darstellt. Im Gegensatz hierzu, kommen in der dynamischen Applikation instationäre bzw. zeitabhängige Daten zum Einsatz. Die Erhebung der Daten erfolgt sowohl am Motorprüfstand als auch am *Fahrzeugrollenprüfstand*. Hierbei werden neben einfachen Beschleunigungsmanövern und Lastwechseln auch Zyklen wie der *Worldwide Harmonized Light-Duty Vehicles Test Cycle* (WLTC) gefahren. Aus den gewonnenen Messdaten werden entweder Sofortmaßnahmen vor Ort getroffen oder Verbesserungen im Rahmen des Post-Processing abgeleitet. Die Fahrzeugapplikation unterscheidet sich von der stationären und dynamischen Applikation darin, dass die Messungen Live im Fahrzeug erfolgen und Applikationsanpassungen oftmals bereits im Fahrzeug vorgenommen werden. Das Ziel der Fahrzeugapplikation ist die Anpassung des Motors an reale Fahrsituationen. Hierzu gehört vor allem die Überprüfung von Kälte- und Höhenkorrekturen, die den ordnungsgemäßen Betrieb des Fahrzeugs bei kalter Luft und niedrigem Umgebungsdruck gewährleisten. Zur Realisierung dieser Umgebungsrandbedingungen werden i.d.R. Fahrzeugerprobungen in dafür geeigneten Regionen durchgeführt. Alternativ können die Umgebungsrandbedingungen auch auf einem Höhen- und Kälterollenprüfstand simuliert werden. Eine Erprobung ersetzen diese jedoch nicht.

Modellbasierte Applikation von Verbrennungsmotoren

Immer strengere Abgas- und CO₂-Emissionsvorschriften erfordern neben neuen Motortechnologien auch immer komplexere Motormanagementsysteme. Dies führt zu einem stetigen Anstieg der im Steuergerät zu implementierenden Funktionen. In Verbindung mit den immer höheren Anforderungen an die Qualität der einzelnen Funktionen ist es heute kaum noch möglich, die Applikation mit konventionellen Verfahren, *manuell*, durchzuführen. Bei der manuellen Applikation werden die Funktionen direkt am Motorprüfstand appliziert. Dabei werden die Funktionsparameter nacheinander so lange variiert, bis sich ein vom Applikateur als optimal erachtetes Funktionsverhalten einstellt. Dieses Verfahren wird auch als *one-factor-at-a-time* Methode bezeichnet. Dieses Optimierungsproblem ist in

niedrigdimensionalen Parameterräumen gut lösbar. Bei einer großen Anzahl von Funktionsparametern wird das Problem jedoch für den Menschen sehr unübersichtlich und das Finden des Optimums praktisch unmöglich. Die Folge ist, dass der Anwender viel Zeit mit der Optimierung eines Betriebspunktes verbringt und trotzdem suboptimale Parameter appliziert. Eine Möglichkeit, diesen Prozess effizienter zu gestalten, sind automatisierte Messkampagnen bzw. Testläufe unter Verwendung von *vollfaktoriellen Versuchsplänen*. Der Hauptvorteil automatisierter Testläufe besteht darin, dass die Prüfstandsautomatisierung in einem vordefinierten Zeitfenster deutlich mehr Messungen durchführen kann als der Bediener. Außerdem kann der Prüfstand nachts und am Wochenende betrieben werden. Die Applikation kann dann am Arbeitsplatz erfolgen. Und die Arbeit am Prüfstand dient der Validierung. Allerdings unterliegen vollfaktorielle Versuchspläne dem so genannten *Fluch der Dimensionalität*. D.h. die Anzahl der zu messenden Betriebspunkte steigt exponentiell mit der Anzahl der Dimensionen. Dieser Zusammenhang wird durch

$$N_{\text{Messungen}} = N_{\text{Stützstellen}}^d \quad (4.1)$$

beschrieben, wobei $N_{\text{Messungen}}$ die Anzahl an benötigten Messungen, d die Anzahl der zu variierenden Parameter und $N_{\text{Stützstellen}}$ die Anzahl der Stützstellen pro Parameter darstellt. Bei einem dreistufigen Versuchsplan, d.h. drei Stützstellen mit acht zu variierenden Parametern, beträgt die Anzahl der zu messenden Punkte 6561. Automatisierte Messkampagnen unter Verwendung vollfaktorieller Versuchspläne sind jedoch nur eine von vielen Möglichkeiten, die Komplexität des Anwendungsprozesses zu reduzieren.

Eine Alternative zu der bereits erwähnten manuellen Anwendung und den automatisierten Messkampagnen ist die Applikation mit Hilfe von Motormodellen. Dieses Verfahren wird auch als modellbasierte Applikation bezeichnet. Von einer modellbasierten Applikation spricht man, wenn zunächst auf Basis von Messungen Motormodelle erstellt werden, anhand derer das physikalische System stellvertretend optimiert wird. Das Verfahren der modellbasierten Applikation ist keine neue Erfindung und hat sich sowohl in der Forschung [49, 69, 81] als auch in der Praxis [57, 101, 120] etabliert. Sie wird sowohl im Dieselmotorenbereich [51, 101, 120] als auch im Ottomotorenbereich [49, 67] eingesetzt. Hierzu stehen verschiedene Softwaretools zur Verfügung, die den Anwender bei diesem Prozess unterstützen [45, 46, 74]. Weiterhin wird zwischen modellbasierter Offline-Optimierung [50, 57, 104] und modellbasierter Online-Optimierung [12, 110] unterschieden. Der Unterschied zwischen modellbasierter Offline- und Online-Optimierung liegt im Zusammenspiel von Prüfstandsautomatisierung und Auswertumgebung. Bei der modellbasierten Offline-Optimierung sind Prüfstandsautomatisierung und Auswertumgebung voneinander entkoppelt und es findet keine Interaktion zwischen Messsystem, Modellbildung und Optimierung statt. Im Gegensatz dazu stehen bei der Online-Optimierung beide Ebenen in ständiger Wechselwirkung.

Die Online-Optimierung erfolgt in der Regel in drei Schritten. Im ersten Schritt wird ein einfacher Versuchsplan erstellt und vermessen. Mit den gewonnenen Messdaten werden initial Modelle trainiert. Im zweiten Schritt erfolgt eine adaptive Versuchsplanung, bei der die Modelle zur Bestimmung neuer Messpunkte verwendet werden. Die Modelle werden nach jeder Messung angereichert und neu trainiert, wodurch der Versuchsplan ständig adaptiert wird. Im dritten und letzten Schritt wird das zuvor formulierte Optimierungsproblem

gelöst. Die so ermittelten Parameter können sofort verifiziert und die Modelle in diesen Bereichen verfeinert werden. Dadurch können Messungen mit geringem Informationsgehalt vermieden und Messungen mit hohem Informationsgehalt angestrebt werden. Die modellbasierte Online-Optimierung ermöglicht somit eine deutliche Reduktion des experimentellen Aufwandes [30, 50, 110]. Diese Art von Online-Optimierungsalgorithmen gehört in den Bereich des *sequential hypothesis testing*. Zu dieser Problemgruppe gehören auch das *mehr-armige Banditenproblem* [4, 62] und die im letzten Jahrzehnt populär gewordenen *Bayesian Optimization*-Verfahren [18, 39, 73]. Diese Probleme unterliegen alle dem sogenannten *Exploration-Exploitation-Dilemma* [14]. Dieses beschreibt den Kompromiss zwischen der Notwendigkeit, neues Wissen zu erlangen, und der Notwendigkeit, dieses Wissen zu nutzen, um die Leistung zu verbessern. Eine optimale Leistung erfordert in der Regel ein gewisses Gleichgewicht zwischen explorativem und exploitativem Verhalten. Im Falle der adaptiven Versuchsplanung bedeutet dies, dass einerseits die Messpunkte in Bereichen platziert werden sollten, in denen die Optimierungskriterien bestmöglich erfüllt werden. Andererseits ist es aber auch notwendig, wenig oder nicht explorierte Bereiche zu vermessen, da diese möglicherweise das globale Optimum enthalten. Einen guten Kompromiss zwischen diesen beiden Gegensätzen zu finden, ist die größte Schwierigkeit bei der Online-Optimierung und bestimmt maßgeblich die Effizienz des Verfahrens sowie die Güte der Optima.

Im Gegensatz zur Online-Optimierung ist bei der modellbasierten Offline-Applikation die Messdatenerfassung vom Modelltraining und der Optimierung entkoppelt. Die Messdatenerfassung ist dabei den weiteren Applikationsschritten vorgeschaltet und läuft im Wesentlichen wie folgt ab. Der Applikateur legt zunächst die zu variierenden Steuergeräteparameter und die zu messenden Zielgrößen fest. Dieser Schritt ist im Grunde genommen die *Feature selection* oder die Parametersелеktion. Die Grenzen des Variationsraumes für die zu variierenden Parameter werden vom Applikateur festgelegt. Dies erfolgt zum einen am Motorprüfstand durch Messungen und zum anderen anhand empirischer Daten. Dabei werden insbesondere kritische Parameterkombinationen überprüft, die zu einer Schädigung des Motors führen können. Auf Basis dieser Informationen wird ein Versuchsplan nach den Prinzipien der statistischen Versuchsplanung (engl.: *design of experiments*, DoE) erstellt, um die Anzahl der Messpunkte zu reduzieren [106]. Der Versuchsplan wird am Motorprüfstand vermessen. Diese Art der Versuchsplanung, bei der der gesamte Versuchsablauf vor Beginn der Messungen feststeht, wird auch als Offline-Versuchsplanung bezeichnet. Hierbei wird der Versuchsplan vor Beginn der Messungen vollständig festgelegt. Anschließend werden anhand der erhobenen Daten für jede Zielgröße datengetriebene Modelle trainiert. Unter Vorgabe geeigneter Optimierungskriterien wird anhand der Modelle eine Parameteroptimierung durchgeführt. Dabei wird z.B. für einen festen Betriebspunkt der Verbrauch unter Einhaltung eines maximalen auszustoßenden Rußes minimiert. Die so ermittelten optimalen Parameter werden anschließend durch Messungen am Motorprüfstand validiert. Damit wird die Vorhersagekraft des Modells und die Wirksamkeit der Applikationsänderungen überprüft. Anschließend werden die optimierten Parameter in den Datenstand übernommen. Diese Art der modellbasierten Applikation ermöglicht die Identifikation und Modellierung des Motors und damit die mathematische Formulierung und Lösung des zugrundeliegenden Optimierungsproblems. Dies ermöglicht es dem Anwender, ohne großen Aufwand das tatsächliche globale Optimum des Problems zu finden. Damit ist das Optimierungsergebnis nicht mehr von der Expertise des Bedieners abhängig und beruht vor allem nicht mehr auf

dessen subjektiver Wahrnehmung. Zum anderen reduziert sich sowohl der Messaufwand als auch die Gesamtzeit, die für die Applikation der Funktion benötigt wird [29].

4.2. Anforderungen an die Methodik

Die im Rahmen dieser Arbeit entwickelte Applikationsmethodik muss eine Reihe vorgegebener Anforderungen erfüllen. Diese Anforderungen werden nachfolgend erläutert.

Ein aus heutiger Sicht wichtiges Applikationsthema ist der Umgang mit Bauteiltoleranzen. Diese werden in der Regel während des Applikationsprozesses nur indirekt berücksichtigt und gegen Ende eines Projektvorhabens mit Hilfe von Grenzmustern überprüft. Die Berücksichtigung von Toleranzeffekten erfolgt in der Regel über empirische Faktoren. Dabei werden Sicherheitskoeffizienten zur Absicherung auf die einzuhaltenden Grenzwerte angewendet. Die Koeffizienten sind empirisch ermittelte Werte und stellen häufig Schätzungen oder Erfahrungswerte aus früheren Projekten dar. Eine Überprüfung der Toleranzlagen erfolgt prinzipbedingt erst nach Abschluss der übrigen Anwendungsthemen im Rahmen der Absicherung. Im Falle einer nicht konformen Anwendung aufgrund extremer Toleranzlagen sind daher eine Reihe von Iterationsschritten erforderlich. Daraus ergibt sich die wesentliche Anforderung an die Methodik, dass Toleranzeinflüsse bereits im Anwendungsprozess berücksichtigt werden. Eine weitere Anforderung an die Methodik ist die globale Modellierung des Motors. Es soll nicht nur eine Vorhersage für einen einzelnen Betriebspunkt gemacht werden. Vielmehr sollen Vorhersagen für den gesamten Betriebsbereich getroffen werden, um eine Bilanzierung durchführen zu können. Daraus ergibt sich, dass ein effizientes Verfahren zur Vermessung globaler Versuchspläne eingesetzt und ggf. Anpassungen vorgenommen werden müssen. Weiterhin muss die Methodik die Optimierung des Motors unter Berücksichtigung der ermittelten Toleranzeinflüsse ermöglichen. Hierfür müssen die Toleranzeinflüsse quantifiziert werden. Und es muss ein Motormodell definiert werden, welches den Einfluss dieser Toleranzen auf den Modellausgang beschreibt bzw. quantifiziert. Als Ergebnis soll die Methodik eine Applikation liefern, die die Toleranzlage der modellierten Ausgangsgrößen berücksichtigt und somit robust gegenüber diesen Ausgangsgrößen ist. Neben der Funktionalität der Methodik soll auch die Benutzerfreundlichkeit in Bezug auf Datenaufbereitung, Visualisierung, Modellbildung, Optimierung und Datenexport im Vordergrund stehen.

4.3. Methodenablauf

Die entwickelte Applikationsmethodik stellt einen mehrstufigen Prozess dar. Sie kann grob in Versuch, Modellierung und Optimierung unterteilt werden. Im Folgenden wird ein Überblick über den methodischen Ablauf gegeben. Eine detailliertere Beschreibung erfolgt in den nachfolgenden Kapiteln.

Im ersten Schritt werden der Versuchsträger und der Prüfstand vorbereitet. Dabei ist darauf zu achten, dass es sich bei den einzelnen Komponenten des Versuchsträgers um nominale oder annähernd nominale Komponenten handelt. Das heißt, die einzelnen Komponenten

weisen keine oder nur minimale Toleranzen auf. Anschließend wird der Versuch aufgesetzt und durchgeführt. Konkret bedeutet dies, dass ein Versuchsplan erstellt und vermessen wird. Die gemessenen Daten bilden die Grundlage für die datenbasierten Modelle. Dazu müssen zunächst die zu verstellenden Motorstellgrößen identifiziert werden. Anschließend wird ein erster Versuchsplan erstellt, der mit Hilfe eines dafür vorgesehenen Algorithmus vermessen wird. Die genauen Details des Algorithmus werden in Kapitel 5 beschrieben.

Der zweite Schritt ist die Modellierung. Diese erfolgt in mehreren Schritten. Zunächst werden mit den zuvor gesammelten Daten sogenannte nominale Modelle des Motors abgebildet. Der Begriff des nominalen Modells bezieht sich darauf, dass wir bei unseren Betrachtungen den untersuchten Versuchsträger als ein Exemplar mit nominalen Komponenten betrachten. Anschließend werden die Parametertoleranzen des Motors bestimmt. Hierzu werden Erfahrungswerte und Datenblätter herangezogen. Aus den nominalen Modellen und den ermittelten Parametertoleranzen werden nun stochastische Modelle des Versuchsträgers trainiert. Das stochastische Modell liefert im Gegensatz zum nominalen Modell keine deterministische Schätzung. Vielmehr zielt es unter Hinzunahme der Toleranzbeiträge auf eine ganzheitliche Betrachtung der gesamten Fahrzeugflotte ab. Anstelle einer Vorhersage für ein einzelnes Versuchsfahrzeug modelliert es eine Verteilungsstatistik der betrachteten Ausgangsgröße. Diese Verteilungsstatistik kann z.B. das 95%-Quantil der Rußemissionsverteilung darstellen. Das ist der Wert, unter dem die Rußemissionen der Motorenflotte in 95% aller Fälle liegen.

Im dritten und letzten Schritt wird der Versuchsträger mit Hilfe der trainierten stochastischen Modelle optimiert. Dazu wird eine stochastische Formulierung des Optimierungsproblems aufgestellt. Dies kann für einen einzelnen Betriebspunkt, aber auch für einen oder mehrere zuvor diskretisierte Zyklen erfolgen. Anschließend wird das Problem mit geeigneten Optimierungsalgorithmen gelöst. Die fertige Applikation kann als Datensatz exportiert werden und muss noch validiert und abgesichert werden. Anschließend kann die endgültige Emissionsverteilung der Flotte simuliert werden.

4.4. Kritische Bewertung und Abgrenzung der Methode gegenüber dem Stand der Technik

In diesem Kapitel wurde zunächst der Stand der Technik im Bereich der Verbrennungsmotorenapplikation dargestellt. Anschließend wurden die Anforderungen an die entwickelte Methodik aufgezeigt und der methodische Ablauf kurz dargestellt. Im Folgenden wird der Ansatz kritisch bewertet und zum Stand der Technik abgegrenzt.

Nach dem aktuellem Stand der Technik findet die Toleranzabsicherung eines Motors gegen Ende des Projekts statt. Dies ist ein manuelles Verfahren und dient der Überprüfung der Toleranzlagen. Werden im Rahmen der Toleranzabsicherung Emissionsgrenzwerte überschritten, folgen Iterationsschritte zur Korrektur der Parametrierung. Die entwickelte Methodik zielt hingegen darauf ab, die sich aus den Toleranzen ergebende Emissionsverteilung bereits während der Grundapplikation zu berücksichtigen. Dadurch wird ein intuitiver Applikationsprozess möglich und die Abhängigkeit gegenüber empirischen Werten entfällt. Die Applikation kann somit hinsichtlich der Flottentoleranz optimiert werden. Zudem können

Iterationen bei der Toleranzabsicherung auf ein Minimum reduziert werden, da Extremlagen bereits durch das Modell erkannt und bei der Optimierung vermieden werden. Der aus dem datengetriebenen Ansatz resultierende Messaufwand stellt einen zusätzlichen Aufwand dar. Darüber hinaus ist es notwendig, die individuellen Bauteiltoleranzen zu bestimmen.

Die im Rahmen dieser Arbeit erarbeitete Methodik wurde bereits in [85] und [84] vorgestellt und berücksichtigt erstmals den Einfluss von Bauteiltoleranzen im Applikationsprozess. In [116] stellen die Autoren ebenfalls einen Ansatz zur Behandlung von Toleranzeinflüssen vor. Dabei gehen sie auf Einflüsse wie unterschiedliche Straßen- und Umweltbedingungen sowie sich ändernde Verkehrssituationen ein. Die Flottenverteilung wird dann unter Berücksichtigung dieser Einflüsse optimiert. Dieser Ansatz unterscheidet sich von unserem insofern, als dass er keine Bauteiltoleranzen berücksichtigt, deren Einfluss also weder abgebildet wird noch in die Optimierung einfließt. In [115] wird ein Ansatz vorgestellt, der sowohl Verkehrs- und Umweltbedingungen als auch den Einfluss von Bauteiltoleranzen berücksichtigen kann. Die Autoren setzen dabei auf sogenannte Gaussian-Mixture-Models [15, 55]. Im Gegensatz zu unserer Arbeit wird jedoch kein Ansatz verfolgt, bei dem eine Verteilungsschätzung approximiert wird. Vielmehr wird versucht, die einzelnen Zufallsvariablen durch Normalverteilungen abzubilden und den Erwartungswert der Ausgangsverteilung durch eine Taylor-Reihenentwicklung zweiter Ordnung zu approximieren. Dies hat den Vorteil, dass auf eine rechenintensive Monte-Carlo-Simulation verzichtet werden kann. Der Nachteil besteht aus unserer Sicht darin, dass multimodale Verteilungen sowie Verteilungen, die von der Normalverteilung abweichen, nicht hinreichend genau abgebildet werden können. Dies kann zu Ungenauigkeiten in der Optimierung führen. Unser Ansatz verlagert hier das Monte-Carlo-Sampling in den Modellierungsprozess, dessen Rechenaufwand im Vergleich zum Optimierungsschritt vernachlässigbar ist. Dadurch ist der Rechenaufwand im Optimierungsschritt vergleichsweise gering—sogar geringer als bei einer Taylor-Reihen-Entwicklung.

5. Versuchsmethode

In diesem Kapitel wird das Versuchsplanungsverfahren beschrieben, die der SEMO-Methode zugrunde liegt. Zunächst wird in Abschnitt 5.1 eine Einführung in die Versuchsplanung zur Modellierung und Optimierung lokaler und globaler Motormodelle gegeben und auf die damit verbundene Komplexität hingewiesen. In Abschnitt 5.2 wird die Versuchsplanung und -durchführung mit Hilfe der ODCM-Methode beschrieben, die einen integralen Bestandteil der Versuchsplanung darstellt. Anschließend wird in Abschnitt 5.2.2 ein modifizierter Online-DoE-Ansatz auf Basis des ODCM-Verfahrens vorgestellt, der im Rahmen dieser Arbeit entwickelt wurde.

5.1. Versuchsplanung für lokale und globale Motormodelle

An dieser Stelle soll ein Überblick über den Einsatz stationärer Versuchsplanungsverfahren zur Unterstützung der Modellbildung und des anschließenden Optimierungsprozesses gegeben werden. Dabei wird sowohl auf die Versuchsplanung für lokale Motormodelle als auch auf die Versuchsplanung für globale Motormodelle eingegangen. Nach unserem Verständnis ist ein lokales Motormodell ein Modell, das für einen festen Betriebspunkt gültig ist. Eine Vorhersage für einen beliebigen anderen Betriebspunkt ist mit einem lokalen Modell daher ohne Extrapolation nicht möglich. Im Gegensatz dazu bezieht sich der Begriff des globalen Motormodells auf die Variation von Drehzahl und Last. Das Motormodell gilt also nicht für einen einzelnen Betriebspunkt, sondern für einen gesamten Betriebsbereich.

Nach der in [21] vorgeschlagene Kategorisierung kann die modellbasierte Applikation von Verbrennungsmotoren auf der Basis von stationären Motormodellen in vier Kategorien eingeteilt werden. Die erste Kategorie stellen lokale Modelle in Verbindung mit lokalen Optimierungsalgorithmen, siehe [57]. Hierbei werden Motormodelle für einzelne Betriebspunkte trainiert und unabhängig voneinander optimiert. Der Vorteil dieser Methode ist, dass bereits mit wenigen Messpunkten ein gültiges Motormodell für einen bestimmten Betriebspunkt trainiert werden kann. Nachteilig ist vor allem, dass beim Trainieren mehrerer Modelle für weitere Betriebspunkte die Messpunkteanzahl stark erhöht werden muss. Insgesamt benötigt man somit deutlich mehr Trainingsdaten als wenn man ein einziges global gültiges Modell trainieren würde. Außerdem müssen für jeden Betriebspunkt die Grenzen des Versuchsraums festgelegt und ein DoE geplant und durchgeführt werden. Der größte Nachteil dieses Ansatzes ist jedoch, dass Optima die mit verschiedenen lokalen Modellen ermittelt wurden, vollkommen unabhängig voneinander sind. Zwei benachbarte Punkte können bei dieser Art der Optimierung also völlig unterschiedliche Parameterkombinationen aufweisen, das wiederum die Fahrbarkeit im Betrieb komprimieren kann.

Eine weitere Kategorie stellen lokale Modelle in Verbindung mit globalen Optimierungsalgorithmen dar [90, 33]. Auch hier werden Motormodelle für einzelne Betriebspunkte trainiert. Das Optimierungsproblem wird jedoch global formuliert, so dass die einzelnen Betriebspunkte in Abhängigkeit voneinander optimiert werden. Der Vorteil dieses Verfahrens gegenüber einer rein lokalen Optimierung ist die Berücksichtigung der Abhängigkeiten zwischen den einzelnen Betriebspunkten. Dadurch ist eine Minimierung der integralen Ausgangsgrößen und eine Gewichtung der einzelnen Betriebspunkte möglich. Häufig angefahrte Betriebspunkte können somit stärker gewichtet und bei der Optimierung gesondert berücksichtigt werden. Weiterhin kann durch geeignete Nebenbedingungen eine Glättung der optimierten Kennfeldbereiche erreicht werden, wodurch die Fahrbarkeit der optimierten Anwendung gewährleistet wird. Der Nachteil dieser Methode ist wiederum die Verwendung lokaler Modelle und der damit verbundene Informationsverlust bzw. höhere Messaufwand. Zwar gibt es Ansätze wie Local Linear Model Tree (LoLiMoT) [88, 87] und Hierarchical Local Model Tree (HiLoMoT) [54], die darauf abzielen, die Anzahl der Messpunkte durch angepasste Versuchspläne und einfachere Modellstrukturen zu reduzieren. Globale Modelle mit entsprechend komplexen Modellstrukturen erreichen jedoch im Vergleich zu diesen bei gleichem Messaufwand eine höhere Modellgüte [54] und besitzen vor allem bessere Generalisierungseigenschaften.

Ein weiterer Ansatz, der in der Literatur häufig verwendet wird, ist die Fusion mehrerer lokaler Modelle zu einem globalen Modell. Hierbei werden lokale DoE's an mehreren Betriebspunkten gemessen und mit den Daten ein globales Modell trainiert. Alternativ kann auch ein einziges DoE mit diskreten Betriebspunktstützstellen erstellt werden [12]. Auch in diesem Fall kann ein globaler Optimierungsansatz gewählt werden. Der Vorteil dieses Ansatzes ist die Gültigkeit des Modells über einen größeren Betriebsbereich. Zudem findet bei der Modellbildung kein Informationsverlust statt, da alle Datenpunkte in die Modellbildung einfließen. Der Nachteil des Ansatzes besteht vor allem darin, dass mehrere DoE erstellt und vermessen werden müssen. Außerdem müssen jedes Mal die Grenzen des Versuchsraums festgelegt werden. Weiterhin ist durch die Diskretisierung der Betriebspunkte die Verteilung der Messpunkte suboptimal. Je nach Modelltyp sind z.B. raumfüllende Designs zu bevorzugen [106].

Die vierte Kategorie stellen globale Modelle in Verbindung mit globalen Optimierungsalgorithmen dar [84]. Hier steht die Messdatenerfassung unter kontinuierlicher Variation von Drehzahl und Last im Vordergrund. Durch die bessere Verteilung der Messpunkte ist die Informationsdichte der Daten im Vergleich zu den bisherigen Ansätzen deutlich höher. Bei gleicher Anzahl von Datenpunkten kann somit eine höhere Modellgüte erreicht werden [54]. Ein weiterer Vorteil ist die Gültigkeit des Modells über einen größeren Betriebsbereich. Außerdem gehen bei der Modellierung keine Informationen verloren. Der große Nachteil dieses Ansatzes ist die hohe Komplexität und die damit verbundenen Probleme. Diese werden im Folgenden diskutiert.

Die vierte Kategorie stellen globale Modelle im Zusammenhang mit globalen Optimierungsalgorithmen dar [84]. Hierbei steht vor allem die Messdatenerfassung unter kontinuierlicher Variation der Drehzahl und Last im Vordergrund. Durch die bessere Verteilung der Messpunkte ist die Informationsdichte der Daten im Vergleich zu den vorherigen Ansätzen deutlich höher. Bei der gleichen Anzahl an Datenpunkten kann somit eine höhere Modellgüte

erzielt werden [54]. Ein weiterer Vorteil ist die Gültigkeit des Modells über einen größeren Betriebsbereich. Zudem geht bei der Modellbildung keine Information verloren. Der große Nachteil dieses Ansatzes ist ihre hohe Komplexität und die damit einhergehenden Probleme. Eine Diskussion hierüber findet im Folgenden statt.

Komplexität globaler Versuchsplanung und Versuchsdurchführung

Die Modellierung von datenbasierten globalen Motormodellen ist an sich nicht komplexer als die Modellierung von lokalen Modellen. Der Unterschied besteht darin, dass nun zwei weitere Eingangsgrößen zur Verfügung stehen. Diese Art der Regression ist in der Literatur bereits ausführlich beschrieben [15]. Die eigentliche Herausforderung liegt vielmehr in der Versuchsplanung und -durchführung. Die wichtigsten damit verbundenen Probleme werden im Folgenden kurz erläutert.

Ein Problem, das sowohl die Versuchsplanung als auch die Versuchsdurchführung betrifft, ist die Festlegung der Versuchsraumgrenzen. Diese ist für die automatisierte Vermessung von vordefinierten Versuchsplänen unumgänglich, da eine ungünstige Wahl der zu variierenden Parameter zu einer Schädigung des Versuchsträgers führen kann. Als Beispiel sei hier ein zu hoher Spitzendruck im Zylinder genannt. Die Festlegung der Versuchsraumgrenzen erfolgt in der Regel manuell und ist in Abbildung 5.1 schematisch dargestellt. Zunächst wird ein zentraler Startpunkt gewählt und jeder Parameter einzeln variiert. Dies geschieht schrittweise bis an die Grenzen des vorgegebenen Variationsbereichs oder bis ein Grenzwert überschritten wird. Ein Grenzwert ist hier eine Größe wie z.B. der Spitzendruck, der zum Schutz des Motors überwacht wird. Im Anschluss an die Messung wird ein rudimentärer Versuchsplan erstellt und vermessen. Abbildung 5.1 zeigt jedoch, dass die manuelle Bestimmung der Versuchsraumgrenzen sehr schnell an ihre Grenzen stößt. Der Aufwand steigt mit der Anzahl der Dimensionen exponentiell an und die Erfassung kritischer Parameterkombinationen ist nicht immer gewährleistet. In der Literatur existieren verschiedene Verfahren und Algorithmen zur Lösung dieses Problems. Diese reichen von der einfachen schrittweisen Parametervariation bis hin zu sogenannten *Boundary-Estimation-Algorithmen* [7, 32, 53], bei denen die Grenzen des fahrbaren Versuchsraums modelliert werden. Unter dem fahrbaren Versuchsraum wird hier der Betriebsbereich verstanden, in dem der Motor ohne Schäden oder Folgeschäden betrieben werden kann. Das in dieser Arbeit angewandte ODCM-Verfahren kann ebenfalls als Boundary-Estimation-Verfahren klassifiziert werden.

Ein weiteres wesentliches Problem der globalen Versuchsplanung, das besonders bei Dieselmotoren im Vordergrund steht, ist die Messpunktereihenfolge. In vielen Computereperimenten spielt diese oftmals keine Rolle. In einem physikalischen System wie dem Verbrennungsmotor beeinflusst die Reihenfolge der Messpunkte jedoch sowohl die gewonnenen Messdaten als auch die Versuchsdauer. Zum einen sind zeitliche Veränderungen des Systems bzw. Drifts, zu berücksichtigen, die zu unerwünschten Trends in der modellierten Funktion führen können. Andererseits erschweren physikalische Phänomene die Versuchsdurchführung und erhöhen die Versuchsdauer, die aus Kostengründen möglichst kurz gehalten werden muss, um ein Vielfaches. Abbildung 5.2 verdeutlicht diese Problematik. Die Abbildung zeigt schematisch den Ablauf einer stationären Messung am Motorprüfstand. Zum

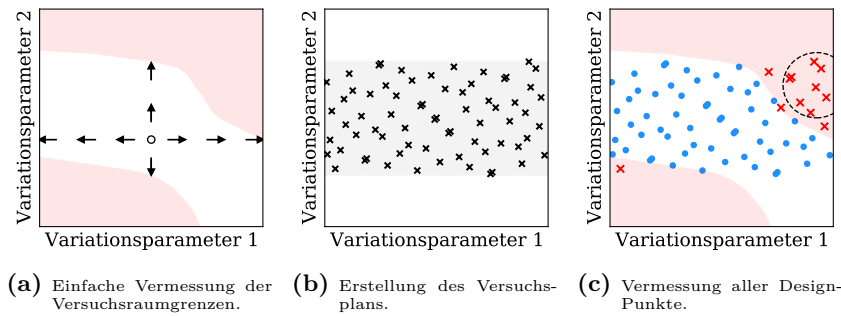


Abbildung 5.1.: Ablauf einer klassischen DoE. Schwarzer Kreis: Ausgangspunkt für die Festlegung der Versuchsraumgrenzen, schwarze Kreuze: geplante Messungen, hellblaue Punkte: (vermessene) fahrbare Messpunkte, rote Kreuze: (abgebrochene) nicht fahrbare Messpunkte, hellrote Fläche: (unbekannter) nicht fahrbarer Versuchsraum, schwarz gestrichelt eingekreiste Fläche: nicht fahrbare Messpunkte, die zu einer möglichen Beschädigung des Motors führen können.

Zeitpunkt $t = 0$ werden die Variationsparameter des Systems eingestellt und es erfolgt eine Systemantwort. Nach der Stabilisierung der Ausgangsgrößen erfolgt die eigentliche Messung. Einige Ausgangsgrößen reagieren relativ schnell auf die Änderung und erreichen nach kurzer Zeit einen stationären Einschwingzustand, während andere Ausgangsgrößen eher träge sind. Für das Experiment bedeutet dies, dass die Einschwingzeit vor der Messung ausreichend groß gewählt werden muss, um eine Verfälschung der Messergebnisse zu vermeiden. Bei Größen wie Temperaturen ist die Einschwingzeit nicht konstant und hängt stark von der Temperaturdifferenz zwischen Anfangs- und Endpunkt ab. Insbesondere Abkühlvorgänge von Hochlastpunkten zu Niedriglastpunkten nehmen entsprechend viel Zeit in Anspruch und führen somit zu einer deutlichen Verlängerung der Versuchsdauer. Um die Versuchsdauer zu minimieren, sollte also ein ständiger Wechsel zwischen stark unterschiedlichen Betriebspunkten vermieden und in der Versuchsplanung berücksichtigt werden. Diese Art der Sortierung ist in der Literatur auch als temperaturoptimale Sortierung bekannt [102].

5.2. Versuchsplanung und -durchführung

Dieser Abschnitt beschreibt die Versuchsplanung und -durchführung. Zusammen bilden sie den ersten Block der datenbasierten Modellierung. Ziel ist die Beschaffung der für die Modellierung erforderlichen Messdaten. Im Detail werden die Auswahl der Modellparameter sowie die Erstellung des initialen Versuchsplans dargestellt. Weiterhin wird die Versuchsdurchführung mit Hilfe der ODCM-Methode dargestellt. Abschließend werden Anpassungen des ODCM-Algorithmus beschrieben, die zu einer effizienten Versuchsdurchführung beitragen.

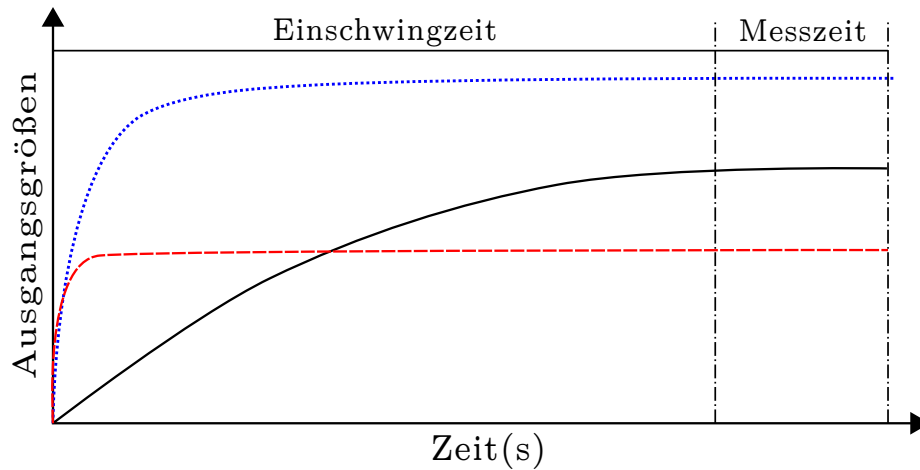


Abbildung 5.2.: Schematischer Ablauf einer stationären Messung am Motorprüfstand [68]. Dargestellt sind drei beispielhafte Messgrößen mit unterschiedlichen Einschwingzeiten. Aufgetragen sind auch die für die Messung angenommene Einschwingzeit und die Messdauer.

5.2.1. Versuchsplanung

Die Grundlage jeder datenbasierten Modellierungsaufgabe sind die für die Modellbildung notwendigen Trainingsdaten. Je nach Aufgabenstellung liegen diese bereits vor oder müssen noch gemessen werden. Bei der modellbasierten Applikation von Verbrennungsmotoren sind die benötigten Daten sehr spezifisch und müssen in der Regel am Motorprüfstand erhoben werden. Dies geschieht in Form einer Messkampagne, deren Grundlage die Versuchsplanung ist. Wie der Begriff schon sagt, zielt die Versuchsplanung auf einen geplanten Versuchsablauf ab. Dabei werden zum einen die zu variierenden Motorstellgrößen festgelegt. Zum anderen werden der zu vermessende Versuchsraum und die zu messenden Parameterkombinationen bzw. Messpunkte festgelegt. Die Versuchsplanung berücksichtigt dabei alles, was zur Identifikation der Motormodelle notwendig ist. Um einen Überblick über die Versuchsplanung zu geben, werden im Folgenden die einzelnen Schritte der Versuchsplanung kurz erläutert.

Feature Selection

Der Vorgang der Auswahl der Modellparameter bzw. der Festlegung der zu variierenden Motorstellgrößen ist in der Literatur als *Feature Selection* bekannt [16, 47, 63]. Die für die Modellierung ausgewählten einzelnen Modellparameter werden auch als sogenannte *Features* bezeichnet. Für die Bestimmung der einzelnen Modellparameter sind verschiedene Lösungsansätze denkbar. Dabei spielt vor allem das a priori Wissen über das zu modellierende System eine Rolle. Nach [47] sollte man sich bei der Feature Selection einige Fragen stellen. Im Kontext dieser Arbeit lauten diese Fragen:

- Sind die Motorparameter, die die Verbrennung beeinflussen, bekannt? Liegen empirische Erkenntnisse vor?

- Haben die einzelnen Modellparameter die gleiche Größenordnung?
- Gibt es relevante Abhängigkeiten zwischen den einzelnen Modellparametern und müssen diese einzeln abgebildet werden?
- Können Modellparameter vernachlässigt werden, ohne dass die Modellgüte darunter leidet?¹ - Auch bekannt als *Feature Pruning* [6].
- Wie viel Zeit, Ressourcen und Trainingsdaten stehen zur Verfügung? Soll eine detaillierte Analyse durchgeführt werden?

Die Modellierung von Dieselmotoren mit Hilfe von Black-Box-Modellen im Kontext der Parameteroptimierung ist in der Literatur ausführlich behandelt worden [12, 60, 69, 84]. Obwohl die im Motorsteuergerät abgebildeten Funktionen tausende von Stellgrößen umfassen, beschränkt man sich bei der Auswahl der Modellparameter in der Regel auf wenige zu applizierende Funktionen, die die Verbrennung im Motor maßgeblich beeinflussen. Der Grund hierfür liegt vor allem in der Aufgabenstellung der Parameteroptimierung. Die Optimierungsergebnisse müssen im Fahrbetrieb darstellbar sein und die optimierten Parameter entsprechend im Steuergerät hinterlegt werden. Wechselseitige Abhängigkeiten zwischen den einzelnen Modellparametern sind in der Regel vernachlässigbar. Insbesondere bei der Wahl komplexer nichtlinearer Modelle, die diese Abhängigkeiten abbilden können. Daher wird bei der Auswahl der Modellparameter auf bereits vorhandene Expertise in der Modellierung von Dieselmotoren zurückgegriffen. In Tabelle 5.1 sind die Modellparameter für den Reihen-Vierzylinder-Dieselmotor aus Abschnitt 3.1.1 dargestellt. Insgesamt werden 13 Modellparameter variiert.

Initialer Versuchsplan

Im Anschluss an die Feature Selection wird ein initialer Versuchsplan erstellt. Der Versuchsplan dient als Grundlage für den ODCM-Algorithmus. Der durch den Versuchsplan abgedeckte Betriebsbereich ist in Abbildung 5.3 dargestellt. Der Variationsbereich der übrigen Modellparameter ist in Tabelle 5.2 dargestellt. Dieser ist bewusst groß gewählt, da unerwünschte, nicht fahrbare Parameterkombinationen vom ODCM-Algorithmus erkannt und vermieden werden. Um die Anzahl der Messpunkte gering und den Informationsgehalt möglichst hoch zu halten, wird ein raumfüllender Versuchsplan erstellt [106]. Wir verwenden hier ein Quasi-Monte-Carlo-Verfahren und verteilen die Messpunkte nach einer Sobol-Sequenz. Eine genauere Beschreibung des Quasi-Monte-Carlo-Verfahrens [20] und der Sobol-Folgen [107] findet sich in den Abschnitten 2.3.1 und 2.3.2.

¹Eine Reihe von Studien [11, 31, 52] zeigt, dass überdimensionierte Machine-Learning-Modelle typischerweise zu überbestimmten Systemen führen. Die resultierenden Modelle sind aufgrund der höheren Trainingszeit und der höheren Inferenzkosten naturgemäß weniger recheneffizient. Zudem erhöhen redundante Modellparameter das Risiko des Overfitting [122].

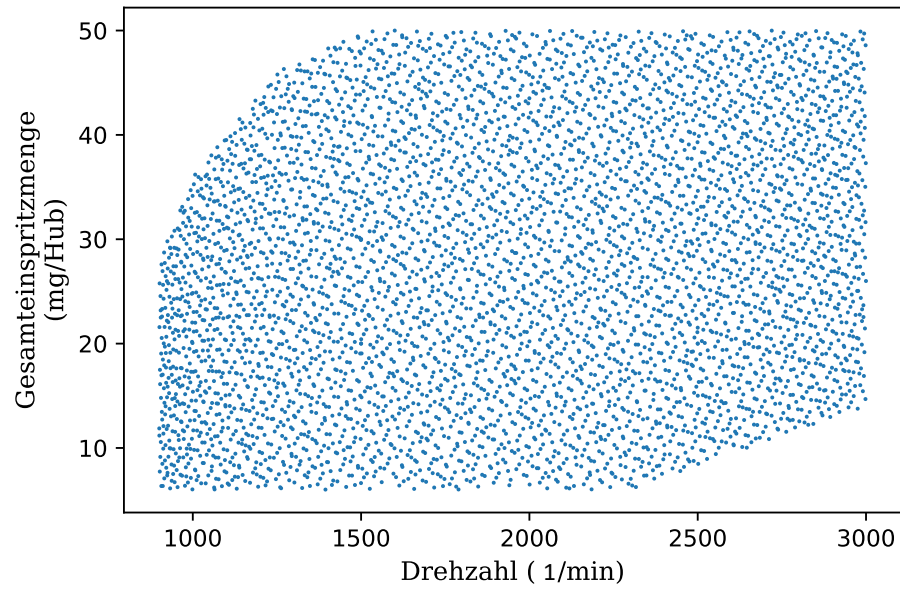


Abbildung 5.3.: Initialer Versuchsplan aufgespannt über der Motordrehzahl und der Gesamtein-spritzmenge.

Tabelle 5.1.: Modellparameterübersicht.

Modellparameter	Abkürzung
Motordrehzahl	n_{Mot}
Gesamtkraftstoffmenge	m_{Tot}
Saugrohrdruck	PSR
AGR-Rate	X_{AGR}
Massenanteil der Niederdruck-AGR	$\psi_{\text{AGR,ND}}$
Commonraildruck	PCR
Ansteuerbeginn der Haupteinspritzung	ϕ_{HE}
Menge der ersten Voreinspritzung	m_{PE1}
Menge der zweiten Voreinspritzung	m_{PE2}
Menge der Nacheinspritzung	m_{NE}
Zeitabstand zwischen der ersten und zweiten Voreinspritzung	t_{PE1}
Zeitabstand zwischen der zweiten Voreinspritzung und der Haupteinspritzung	t_{PE2}
Zeitabstand zwischen der Haupt- und Nacheinspritzung	t_{NE}

5.2.2. Versuchsdurchführung mit Hilfe der ODCM-Methode

Ein wesentliches Ziel dieser Arbeit ist die Entwicklung einer praxistauglichen Applikationsmethodik. Daher steht neben der Funktionalität der Methodik auch die Umsetzbarkeit im Fokus der Arbeit. Für die Versuchsplanung und -durchführung bedeutet dies vor allem einen geringen Zeit- und Ressourcenaufwand. Um den Personalaufwand und die Prüfstandsbelastung gering zu halten, sind daher ein hoher Automatisierungsgrad und effiziente Algorithmen erforderlich. Dem stehen zum einen die hohen Genauigkeitsanforderungen an die Motormodelle entgegen, die mit der modellbasierten Applikation und insbesondere der vorgestellten Methodik einhergehen. Zum anderen muss eine effiziente Identifikation globaler Motormodelle gewährleistet sein.

Tabelle 5.2.: Variationsbereich der Modellparameter.

Modellparameter	Variationsbereich	Einheit
PSR	1000 – 2900	mbar
X_{AGR}	0 – 60	%
$\psi_{AGR,ND}$	0 – 100	%
PCR	300 – 2400	bar
ϕ_{MI}	-7 – 12	°KW vor Zünd-OT
m_{PE1}	0 – 4.5	mg/Hub
m_{PE2}	0 – 3	mg/Hub
m_{NE}	0 – 6	mg/Hub
t_{PE1}	0 – 1500	μs
t_{PE2}	0 – 1500	μs
t_{NE}	0 – 1600	μs

Um diesen Anforderungen gerecht zu werden, wurde im Rahmen dieser Arbeit eine klare Vorgehensweise für die Versuchsplanung und -durchführung definiert. Den Kern bildet die ODCM-Methode, die erstmals in [53] vorgestellt wurde. Abbildung 5.4 zeigt den schematischen Ablauf des ODCM-Verfahrens. In einem ersten Schritt wird der vom Applikateur definierte DoE-Plan geladen. Es ist wichtig zu beachten, dass dieser Versuchsplan rudimentärer Natur ist. Der Applikateur legt lediglich die Minimal- und Maximalwerte der Variationsparameter und die Anzahl der Messpunkte fest. Die Verteilung der Punkte ist raumfüllend zu wählen. Ausgehend von einem Startpunkt misst die Prüfstandsautomatisierung die vom ODCM-Algorithmus vorgegebenen Messpunkte. Nach jeder Messung findet ein Austausch zwischen der Prüfstandsautomatisierung und der ODCM-Software statt. Die Prüfstandsautomatisierung liefert für jeden gemessenen Punkt die Information, ob der Punkt fahrbar oder nicht fahrbar war. Ein Punkt ist nicht fahrbar, wenn während der Messung ein Grenzwert überschritten wurde. Diese Information wird verwendet, um ein Klassifikationsmodell zu trainieren, das eine Wahrscheinlichkeitsvorhersage über die Fahrbarkeit für anstehen-

de Messungen liefert. Das Modell wird nach jeder Messung neu trainiert und aktualisiert. Das Klassifikationsmodell ermittelt vor jeder anstehenden Messung im Versuchsplan eine Wahrscheinlichkeit für die Fahrbarkeit des Punktes. Wenn das Modell den Punkt als wahrscheinlich nicht fahrbar einstuft, wird der Punkt übersprungen. Im Falle einer fahrbaren Klassifikation wird der Punkt angefahren. Das ODCM-Verfahren verwendet einen binären Klassifikator, der zwischen den beiden Klassen *fahrbar* bzw. Eins und *nicht fahrbar* bzw. Null unterscheidet. Die zugrundeliegende Modellstruktur ist ein Gauß-Prozess, der aus dem binären Datensatz lernt. Dementsprechend macht das Modell Vorhersagen, die zwischen Null und Eins liegen können. Dies wird als Wahrscheinlichkeit interpretiert, dass ein Punkt fahrbar ist. Für die GP wird ein Prior mit einem Mittelwert von Null und einem isotropen quadratisch exponentiellen Kernel verwendet. Die Hyperparameter werden durch Maximierung der Log-Likelihood trainiert. Um das Training robust zu gestalten, sind zahlreiche Vorkehrungen auf der Optimierungsseite notwendig. Parallel zum Modelltraining wird die Messpunktereihenfolge ständig aktualisiert. Dazu wird zunächst aus der Menge aller fahrbaren Punkte ein örtlich sicherer Punkt bestimmt, der im Falle einer Grenzwertüberschreitung als Rücksprungpunkt dient. Anschließend wird der Abstand aller im Versuchsplan enthaltenen Messpunkte zum örtlich sicheren Punkt berechnet. Als Abstandsmaß wird der euklidische Abstand zwischen zwei Punkten gewählt. Durch diese Strategie wird der Versuchsplan ständig angepasst, so dass der nächste zu vermessende Punkt immer den geringsten Abstand zum lokal sicheren Punkt hat. Diese Strategie gewährleistet eine sichere und systematische Exploration des fahrbaren Parameterraums. Der Algorithmus stoppt, nachdem eine vorgegebene Anzahl an fahrbaren Punkten gemessen wurde oder nachdem alle Punkte des Versuchsplans gemessen wurden.

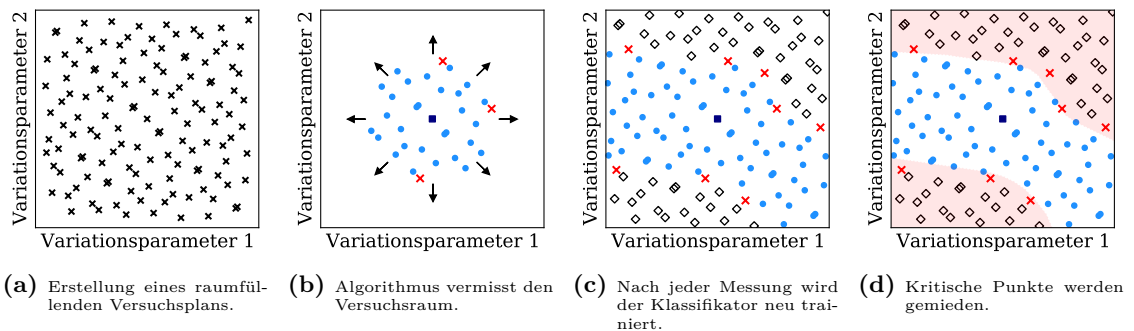


Abbildung 5.4.: ODCM-Verfahren. Schwarze Kreuze: geplante Messungen, dunkelblaues Quadrat: sicherer Punkt, hellblaue Punkte: fahrbare Messpunkte, rote Kreuze: nicht fahrbare Messpunkte, schwarze Raute: übersprungener Punkt, hellroter Bereich: nicht fahrbarer Versuchsraum.

Die gemessenen Punkte werden dann für die Modellierung und Optimierung der Applikationsparameter verwendet. Die Verwendung der Informationen des ODCM-Klassifikationsmodells kann die Optimierung deutlich verbessern, da es den fahrbaren Bereich als Randbedingung liefert. In Algorithmus 1 ist der Pseudo-Code des ODCM-Algorithmus gegeben.

Algorithmus 1 Pseudo-Code des ODCM-Algorithmus.

```

1: Lade initialen Versuchsplan mit  $n$  Messpunkten
2:  $i \leftarrow 1$ 
3: while  $i \leq n$  do
4:   Lese nächsten Punkt aus dem Versuchsplan
5:   Bewerte, ob der Punkt fahrbar ist
6:   if Punkt ist als fahrbar bewertet then
7:     Starte Messung
8:     Kennzeichne Punkt als fahrbar oder nicht-fahrbar
9:     Trainiere das Klassifikationsmodell neu
10:    Berechne neuen lokal sicheren Punkt und adaptiere den Versuchsplan
11:  else
12:    Überspringe den Punkt
13:  end if
14: end while

```

5.2.3. Modifizierter Online-DoE-Ansatz

Der oben vorgestellte ODCM-Ansatz wurde für die Zwecke dieser Arbeit angepasst. Das primäre Ziel der Anpassungen war die effiziente Vermessung globaler DoEs unter Variation von Drehzahl und Last. Das Hauptproblem, das sich aus der Drehzahl-Last-Variation ergibt, ist die starke Zunahme der Stabilisierungszeiten zwischen zwei Messungen, siehe Abschnitt 5.1. Um lange Einschwingzeiten zu vermeiden, wurde eine Anpassung des ODCM-Algorithmus vorgenommen, die eine temperaturoptimale Verteilung der Messpunkte bevorzugt. Die Änderungen bzw. Erweiterungen werden im Folgenden beschrieben.

Die erste Erweiterung betrifft die Versuchsplanerstellung. Hierbei wird der initiale erstellte Versuchsplan in zusammenhängende Blöcke von Betriebspunkten unterteilt. Dies ist beispielhaft in Abbildung 5.5a dargestellt. Weiterhin wird für jeden Betriebspunktblock ein sicherer Punkt als initialer Rücksprungpunkt definiert. Ausgehend hiervon wird zunächst ein erster Block vermessen und das Ganze für die restlichen Blöcke wiederholt. Die Messreihenfolge der Blöcke spielt keine wesentliche Rolle und kann beliebig gewählt werden. Durch die Maßnahme wird die Reihenfolge der Messpunkte so verändert, dass große Drehzahl- und Lastsprünge zwischen den Messungen und damit lange Abkühlzeiten vermieden werden. Lediglich beim Wechsel zwischen zwei Blöcken können größere Sprünge auftreten. Diese sind jedoch selten und haben nur einen marginalen Einfluss auf die Versuchsdauer.

Die zweite Änderung betrifft die Bestimmung des sicheren Rückzugspunktes. Wie oben beschrieben, wird jedem Betriebspunktblock ein sicherer Punkt zugeordnet. Wie beim Standard-ODCM-Verfahren wird parallel zu den Messungen eine Anpassung der Reihenfolge der Messpunkte vorgenommen. Im Unterschied zum Standardverfahren erfolgt die Bestimmung des örtlich sicheren Punktes jedoch auf Basis der fahrbaren Punkte in einem Block. Somit wird bei einer Grenzwertverletzung je nach Betriebspunktblock auf einen anderen sicheren Punkt zurückgegriffen. In jedem Betriebspunktblock wird also die Reihenfolge der Messpunkte ständig angepasst und ein neuer lokaler sicherer Punkt berechnet. Die Wahl dieser Strategie hat einen Grund. Der lokal sichere Punkt dient als Ausgangspunkt

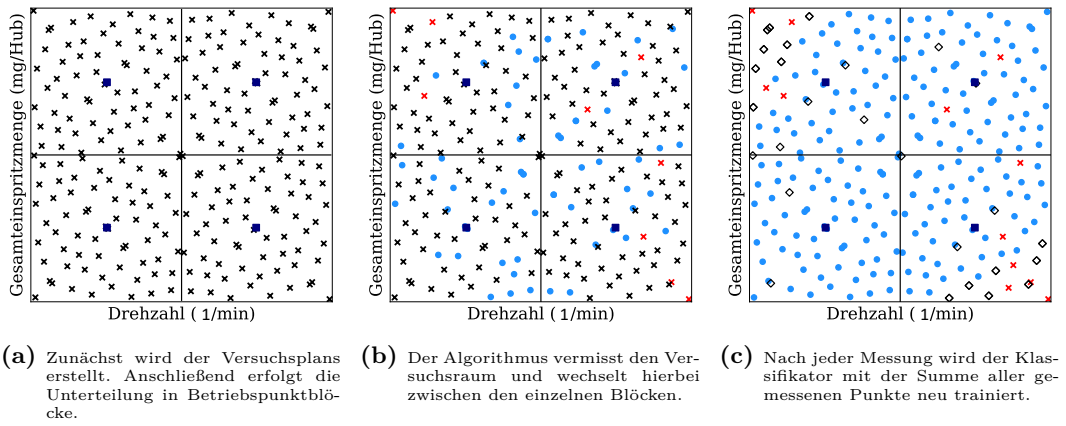


Abbildung 5.5.: Modifiziertes Online-DoE-Verfahren. Schwarze Kreuze: geplante Messungen, dunkelblaues Quadrat: sicherer Punkt, hellblaue Punkte: fahrbare Messpunkte, rote Kreuze: nicht fahrbare Messpunkte, schwarze Raute: übersprungener Punkt.

für die Erkundungsstrategie des Algorithmus. Durch die räumlich zentrale Positionierung dieses Punktes wird der Geschwindigkeits-Last-Sprung zwischen diesem und dem zu vermessenden Punkt minimiert. Im Normalbetrieb ist der Vorteil dieser Strategie nicht sehr groß, da die Anzahl der nicht fahrbaren Punkte vergleichsweise geringer ist. Insbesondere bei der Vermessung von Sonderbetriebsarten wie NSC-Regenerationsbetrieb, Heizbetrieb oder DPF-Regenerationsbetrieb ist die Anzahl der nicht fahrbaren Punkte jedoch deutlich höher und der Vorteil dieser Strategie signifikant.

Die letzte Änderung betrifft das Klassifikationsmodell. Der Grund hierfür ist die unzureichende Flexibilität des Modells bei einer steigenden Anzahl an Eingangsparametern. Hier wird Anstelle des isotropen SE-Kernels ein ARD SE-Kernel verwendet. Das bedeutet, dass jedem Eingangsparameter eine eigene Lengthscale zugewiesen wird. Hierdurch kann das Modell die Bedeutung der einzelnen Eingangsparameter gewichten und unbedeutende Parameter ausblenden [118]. Abbildung 5.6 zeigt einen Vergleich der beiden Kernel für ein globales Fahrbarkeitsmodell eines Dieselmotors im Heizbetrieb gegeben. Insgesamt werden 19 Eingangsparameter variiert, darunter Drehzahl und Gesamteinspritzmenge. Die Modelle werden jeweils für eine variierende Anzahl an Trainingsdaten trainiert. Die Trainingsdaten werden zufällig aus einer Grundgesamtheit von 770 fahrbaren und 159 nicht-fahrbaren ausgewählt. Für jedes Modell werden NRMSE, Spezifität und Selektivität auf der Grundlage separater Testdaten berechnet. Die Testdaten bestehen aus 324 fahrbaren und 75 nicht-fahrbaren Punkten. Aus der Datenbasis geht hervor, dass es sich hierbei um einen unausgebalancierten (engl. *imbalanced*) Datensatz handelt. Das Modell mit isotropem SE-Kernel weist bei einer geringen Anzahl von Trainingsdaten eine Spezifität nahe Null und eine Selektivität nahe Eins auf. Dies bedeutet, dass das Modell jeden Punkt als fahrbar klassifiziert und keine Zusammenhänge lernt. Es hat also keinen Vorteil gegenüber dem Szenario ohne Modell. Erst ab ca. 400 Messpunkten zeigt sich, dass die Spezifität zunimmt, die Sensitivität leicht abnimmt und das Modell lernt. Aber auch hier ist die Sensitivität mit ca. 0.18 relativ gering. Das Modell mit ARD SE-Kernel verhält sich deutlich anders. Die Sensitivität ist auch bei diesem Modell mit einem Wert von über 0.9 sehr hoch. Im Gegensatz zum Modell

mit isotropem Kernel ist die Spezifität jedoch mit ca. 0.31 deutlich höher. Weiterhin ist zu beobachten, dass das Modell bereits mit relativ wenigen Daten lernt.

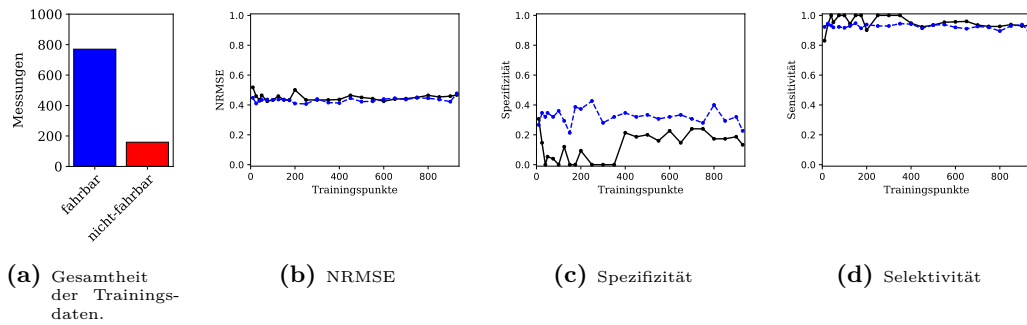


Abbildung 5.6.: NRMSE, Spezifität und Selektivität zweier GP mit unterschiedlichem Kernel über der Anzahl an Trainingsdaten, die für die Modellbildung verwendet wurden. Die Trainingsdaten für das jeweilige Modell werden zufällig aus der Gesamtheit der Trainingsdaten initialisiert. Schwarz durchgezogene Linie: isotroper SE Kernel, Blau gestrichelte Linie: ARD SE Kernel.

Mögliche Maßnahmen zur weiteren Steigerung der Effizienz der Datenerfassung

An dieser Stelle werden kurz mögliche Maßnahmen zur weiteren Effizienzsteigerung der Datenerhebung diskutiert. Diese Maßnahmen sollen nur als Anregung für zukünftige Arbeiten dienen und sind nicht Bestandteil dieser Arbeit, da die Versuchsplanung zwar eine Notwendigkeit, aber nicht der Schwerpunkt dieser Arbeit ist.

In der Literatur gibt es verschiedene Ansätze, die im Zusammenhang mit dem SEMO-Ansatz zur Effizienzsteigerung bei der Messdatenerfassung beitragen können. Schillinger schlägt beispielsweise in [100] ein diskriminatives Modell in Kombination mit einer Active-Learning-Strategie zur Exploration des Versuchsraums vor. Das Modell wird hierbei nur aus fahrbaren Messungen trainiert und schätzt die Fahrbarkeitsgrenze aus weiteren Größen wie Abgastemperatur und Zylinderdruck ab. Ein Vorteil dieser Methode gegenüber dem ODCM-Verfahren ist, dass für das Training des Modells keine gefährlichen nicht-fahrbaren Messpunkte benötigt werden. Im Idealfall muss also keine einzige nicht fahrbare Parameterkombination realisiert werden. Zusätzlich wird die Auswahl der nächsten Messpunkte durch ein Varianzkriterium bestimmt. Die nächsten Messpunkte werden somit in Bereichen mit geringem Informationsgehalt platziert, um eine möglichst effektive Identifikation zu ermöglichen. Auch relativ einfache Maßnahmen, wie das sogenannte Resampling (engl. *Resampling*) [25, 26], können zu einer effektiveren Exploration des Versuchsraumes beitragen. Beim Resampling werden unausgeglichene Mehrklassendatensätze künstlich angepasst, um eine homogenere Verteilung zwischen den einzelnen Klassen zu erreichen. Abbildung 5.7 zeigt einen Fall von *Oversampling* [22] am Beispiel des in Abschnitt 5.2.3 beschriebenen Klassifikationsproblems. Oversampling bedeutet in diesem Fall die künstliche Erhöhung der nicht fahrbaren Punkte durch Vervielfachung dieser.

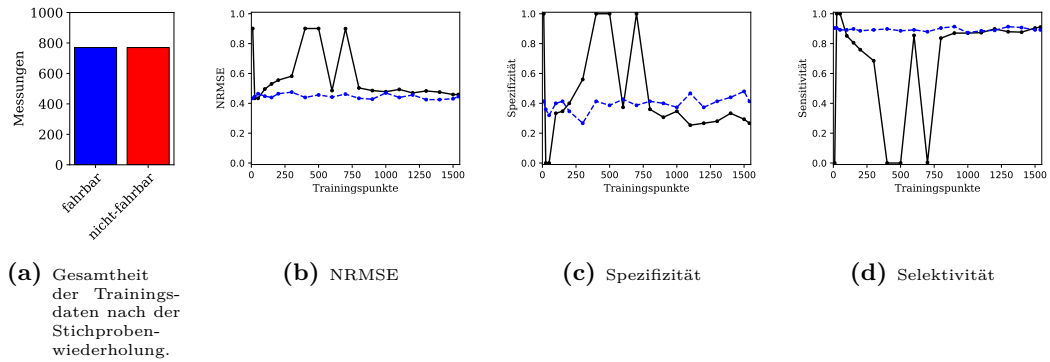


Abbildung 5.7.: NRMSE, Spezifität und Selektivität zweier GP mit unterschiedlichem Kernel über der Anzahl an Trainingsdaten, die für die Modellbildung verwendet wurden. Die Trainingsdaten für das jeweilige Modell werden zufällig aus der Gesamtheit der Trainingsdaten (mit Stichprobenwiederholung) initialisiert. Schwarze durchgezogene Linie: isotroper SE Kernel, Blau gestrichelte Linie: ARD SE Kernel.

Zusammenfassung der Maßnahmen

In diesem Abschnitt wurde eine Erweiterung bzw. Anpassung des ODCM-Verfahrens vorgestellt, welches eine noch effiziente und sichere Identifikation des Versuchsträgers ermöglicht. Die Erweiterung des ODCM-Verfahrens umfasst die folgenden Maßnahmen:

- Unterteilung des Versuchsplans in mehrere Betriebspunktblöcke, die zeitlich versetzt durchgeführt werden, um längere Abkühlzeiten zu meiden und damit eine geringere zu Messdauer zu ermöglichen.
- Einführung eines sicheren Rücksprungpunktes für jeden Betriebspunktblock, um bei einem nicht fahrbaren Versuchspunkt den Versuchsträger in einen naheliegenden sicheren Rücksprungpunkt zu bringen.
- Anpassung des Klassifikationsmodells, um ein flexibleres und robusteres Modell zu erhalten, welches auch bei geringer Anzahl an Trainingsdaten eine gute Klassifikation ermöglicht.

Zudem wurden mögliche Maßnahmen zur weiteren Effizienzsteigerung der Datenerfassung diskutiert. Diese Maßnahmen sollen nur als Anregung für zukünftige Arbeiten dienen und sind nicht Bestandteil dieser Arbeit.

6. Stochastisches Motormodell zur Vorhersage der Flottenemissionen

Kern der in dieser Arbeit vorgestellten Applikationsmethodik sind datenbasierte Motormodelle, die auf stationären Messungen am Motorprüfstand basieren. Im Gegensatz zu anderen Arbeiten, in denen ein nominelles Modell des Versuchsträgers erstellt wird [10, 12], zielen die in dieser Arbeit vorgestellten Modelle auf die Vorhersage der gesamten Flottenverteilung bzw. von aus dieser Verteilung abgeleiteten statistischen Größen ab. In Abschnitt 6.5 werden z.B. Quantile der Rohemissionsverteilung von Ruß dargestellt.

Dieses Kapitel ist folgendermaßen aufgebaut. In Abschnitt 6.1 werden die Anforderungen an das entwickelte Modell erläutert. Abschnitt 6.2 beschreibt die wissenschaftlichen Neuerungen, die dem Modell zugrunde liegen. In Abschnitt 6.3 wird eine Formulierung für den Modellbildungsprozess gegeben. In Abschnitt 6.4 wird die Implementierung des Modells beschrieben. Und in Abschnitt 6.5 werden die asymptotischen Eigenschaften des vorgestellten Modells untersucht und diskutiert.

6.1. Anforderungen an das Motormodell

Das in diesem Kapitel vorgestellte Motormodell muss eine Reihe vorgegebener Anforderungen erfüllen. Diese Anforderungen ergeben sich aus der Aufgabenstellung und werden im Folgenden erläutert.

Ziel der Motorapplikation ist die optimale Parametrierung der betrachteten Softwarefunktionen. Dieser Parametrierungsprozess stellt ein klassisches nichtlineares Optimierungsproblem dar. Der modellbasierte Applikationsansatz stellt hier ein Werkzeug zur Verfügung, das einerseits den Parametrierungsprozess automatisiert und andererseits den Prozess der Parametersuche unterstützt. Die manuelle Bestimmung des globalen Optimums ist nämlich aufgrund der Komplexität des Optimierungsproblems praktisch unmöglich. Für einen Computer hingegen ist die Suche nach dem globalen Optimum hingegen ein einfaches Problem. Die Vorgehensweise lässt sich wie folgt beschreiben. Zunächst werden Modelle der zu optimierenden Ausgangsgrößen erstellt, deren Eingangsgrößen die zu optimierenden Parameter sind. Anschließend wird aus den erstellten Modellen eine zu optimierende Kostenfunktion gebildet, die unter Hinzunahme geeigneter Ansätze optimiert wird. Daraus ergibt sich die Anforderung an das Modell, eine Sensitivität gegenüber den zu optimierenden Steuergeräteparametern aufzuweisen. Eine weitere Anforderung ist die Genauigkeitsanforderung an das Modell, da für die Anwendung der betrachteten Motorsteuergerätefunktionen eine geringe Diskrepanz zwischen modellierter Größe und Realität bestehen muss.

Der Beitrag dieser Arbeit liegt in der Entwicklung eines neuen modellbasierten Applikationsansatzes. Unser Ansatz zielt im Gegensatz zu den bisherigen Verfahren auf eine ganzheitliche Betrachtung der gesamten Flotte durch die Berücksichtigung von Toleranzbeiträgen ab. Das zu entwickelnde Modell soll die Auswirkungen der einzelnen Toleranzbeiträge auf das Flottenergebnis abbilden—konkret soll z.B. die Emissionsverteilung oder eine statistische Größe, die die Verteilung beschreibt, durch das Modell abgebildet werden. Der Schwerpunkt der Arbeit liegt dabei auf den Bauteiltoleranzen. Entscheidend ist auch die ordnungsgemäße Fortpflanzung der randomisierten Eingangsparameter des Modells. Die betrachteten Parametertoleranzen sind nicht notwendigerweise einheitlich oder normalverteilt und können Korrelationen aufweisen. Für das Modell bedeutet dies, dass weniger rechenintensive Approximationen, wie z.B. die Most-Probable-Points-Methode [34], keine valide Option zur Schätzung der Ausgangsverteilung darstellen. Eine Vernachlässigung oder unsachgemäße Darstellung von Modellintensitäten würde nämlich zu einer ungenauen bzw. verzerrten Verteilung für die betrachtete Ausgangsgröße führen und unsere Genauigkeitsanforderungen nicht erfüllen. Da eine analytische Lösung dieses Problems ebenfalls nicht in Frage kommt, stellt die numerische Lösung des Problems mit Hilfe von Zufalls- bzw. Pseudozufallsexperimenten die einzig gültige Möglichkeit dar.

Aus Anwendersicht sind zudem kurze Rechenzeiten für den Optimierungsprozess wünschenswert. Dies vereinfacht vor allem die Bedienung des fertigen Softwareprodukts. Bei sehr langen Rechenzeiten ist ein Einsatz ohne zusätzliche Rechenkapazitäten häufig nicht möglich und mindert aus Anwendersicht die Benutzerfreundlichkeit. Diese Anforderung steht im Widerspruch zu der zuvor abgeleiteten Notwendigkeit einer numerischen Lösung der Problemstellung.

Für das in dieser Arbeit vorgestellte Modell lassen sich die Anforderungen wie folgt zusammenfassen:

- Die Eingangsgrößen des Modells entsprechen den zu optimierenden Steuergeräteparametern.
- Die Modellausgangsgröße stellt eine Verteilung oder eine statistische Größe einer Verteilung dar.
- Die Modellausgangsgröße muss mit hinreichender Genauigkeit abgebildet werden.
- Das Modell muss geringe Vorhersagekosten aufweisen, um eine kurze Optimierungszeit zu gewährleisten.
- Die Toleranzbeiträge der einzelnen Komponenten müssen für die Modellbildung bekannt sein.

6.2. Wissenschaftliche Neuerungen

Aufgrund des stochastischen Charakters der in dieser Arbeit behandelten Problemstellung stellt die Verwendung üblicher datenbasierter Modelle, wie sie aus der Literatur bekannt sind, keine sinnvolle Option zur Modellierung der gewünschten Zielgrößen dar [12, 40].

Vielmehr ist die Entwicklung von speziell für diese Aufgabenstellung konzipierten Motormodellen erforderlich. Im Vergleich zu deterministischen, datenbasierten Modellen, die den Stand der Technik repräsentieren, weist unser Modell einige wissenschaftliche Neuerungen auf. Diese werden im Folgenden beschrieben.

Die Neuerung des vorgestellten Modells besteht darin, dass a priori Informationen über die Toleranzbeiträge in Form von Wahrscheinlichkeitsverteilungen und einer Korrelationsmatrix in die Modellbildung einfließen. Damit ist das Modell in der Lage, die aus den Toleranzbeiträgen resultierende Ausgangsverteilung bzw. aus der Ausgangsverteilung ableitbare statistische Größen abzubilden. Mit Hilfe dieses stochastischen Modells kann somit die gesamte Flottenverteilung optimiert werden, anstatt wie bisher ein einzelnes nominales Versuchsträgers. Dazu wird eine Modellformulierung hergeleitet, die aus einem Satz von Trainingsdaten, den zugehörigen Toleranzbeiträgen und der zugehörigen Korrelationsmatrix eine statistische Größe der resultierenden Ausgangsverteilung vorhersagt. Unser Ansatz basiert nicht auf einer einfachen Faltung der Eingangsverteilungen über die modellierte Zielfunktion. Vielmehr wird eine Abbildung zwischen einer Menge von Eingangsgrößen des Modells und den zugehörigen Werten der zu prädizierenden statistischen Größe der Ausgangsverteilung erstellt.

Die wissenschaftlichen Neuerungen des Modells lassen sich wie folgt zusammenfassen

- Informationen über Toleranzbeiträge werden a priori in Form von Wahrscheinlichkeitsverteilungen und einer Korrelationsmatrix in das Modell eingebracht.
- Das Modell ist in der Lage, die aus den Toleranzbeiträgen resultierende Verteilung der modellierten Größe bzw. verschiedene statistische Größen dieser Verteilung (arithmetisches Mittel, Varianz, Quantile) abzubilden.
- Das Modell kann in einem nachgeschalteten Prozess zur Optimierung der gesamten Flottenverteilung verwendet werden.

6.3. Modellbildungsprozess

Die Erstellung des stochastischen Modells ist ein mehrstufiger Prozess. Zunächst wird in Abschnitt 6.3.1 ein nominales Modell der zu modellierenden Größe erstellt. Anschließend findet in Abschnitt 6.3.2 die eigentliche Modellbildung des stochastischen Modells statt. Der gesamte Modellierungsprozess ist in Abbildung 6.1 schematisch dargestellt und wird im Folgenden näher beschrieben.

6.3.1. Initiales nominales Motormodell

Grundlage des stochastischen Modells ist ein zuvor trainiertes Gauß-Prozessmodell der zu modellierenden Größe, das als

$$f(\mathbf{x}) \sim \mathcal{GP}(m, k) \quad (6.1)$$

notiert werden kann. Dieses wird mit stationären Messdaten aus einer DoE-Messkampagne trainiert und stellt ein sogenanntes nominales Modell dar. Der Begriff nominales Modell

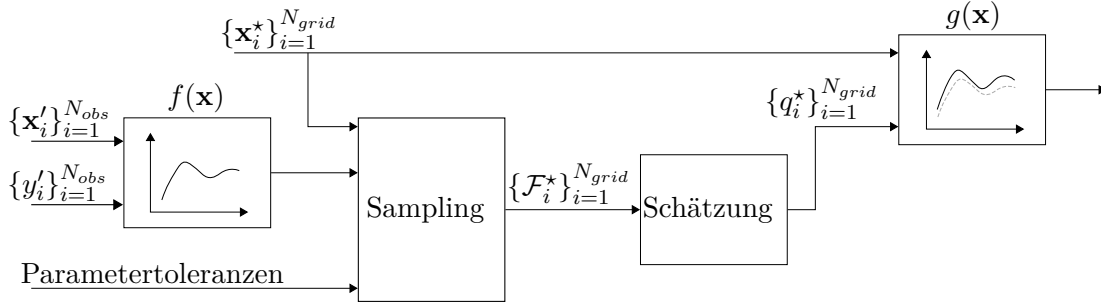


Abbildung 6.1.: Schematische Darstellung des Modellbildungsprozesses.

bezieht sich darauf, dass wir in unseren Betrachtungen den untersuchten Versuchsträger als ein Exemplar mit nominalen Komponenten betrachten. Die einzelnen Komponenten sind also nominal und weisen keine Streuungen auf. Dazu werden die relevanten Komponenten, wie z.B. die Injektoren oder der Heißfilm-Luftmassenmesser, gesondert ausgewählt. Der Ansatz ist modell-agnostisch und daher auch mit anderen Modellierungsansätzen wie z.B. künstlichen neuronalen Netzen oder auch mit einem Multilayer-Perzeptron [58] kompatibel. Unsere Untersuchungen zeigen jedoch, dass verschiedene Modelltopologien wie mehrschichtige Neuronale Netze oder auch Deep-Believe-Netze wie der Deep Gaussian Process [28] keine nennenswerten Genauigkeitsvorteile gegenüber dem GP aufweisen—zumindest nicht für den in dieser Arbeit betrachteten Fall. Zudem sind Gauß-Prozesse in der Verbrennungsmodellierung weit verbreitet und haben sich über die Jahre als robust und zuverlässig erwiesen [13, 112].

Die Datenbasis für das nominale Motormodell sind stationäre Prüfstandsmessungen aus einem DoE. Das resultierende Gauß-Prozessmodell stellt einen stochastischen Prozess dar und liefert für jede Vorhersage neben einem Erwartungswert auch eine Verteilung. Diese ist für einen Gauß-Prozess Gauß- oder normalverteilt und kann durch den Erwartungswert und die Standardabweichung beschrieben werden. Dieser stochastische Prozess ist nicht zu verwechseln mit dem stochastischen Modell, das in Abschnitt 6.3.2 vorgestellt wird. An dieser Stelle ist nur der Erwartungswert des GP von Interesse, der eine deterministische Vorhersage darstellt. Die Verteilung des GP wird hier zunächst nicht betrachtet.

6.3.2. Stochastisches Motormodell

Nachdem in Abschnitt 6.3.1 das nominale Motormodell vorgestellt wurde, wird im Folgenden das stochastische Motormodell vorgestellt. Im Gegensatz zum nominalen Verbrennungsmodell, das eine deterministische Schätzung liefert, zielt das stochastische Motormodell durch die Hinzunahme von Toleranzbeiträgen auf eine ganzheitliche Betrachtung der gesamten Flotte ab. Anstelle einer Mittelwertvorhersage, die eine Vorhersage für ein einzelnes Versuchsträger darstellt, werden Verteilungsstatistiken der betrachteten Ausgangsgröße modelliert.

Modellformulierung

Nehmen wir an, dass die Menge $\mathcal{X}' = \{\mathbf{x}'_i\}_{i=1}^{N_{obs}}$ eine Sammlung von Parametervektoren \mathbf{x}'_i mit d Eingangsdimensionen ist und $\mathcal{Y}' = \{y'_i\}_{i=1}^{N_{obs}}$ die Sammlung der zugehörigen Messwerte y'_i —in diesem Fall, Prüfstandsmessungen aus einem stationären DoE. Nehmen wir weiter an, dass $f: \mathcal{X}' \subset \mathbb{R}^{d \in \mathbb{N}} \rightarrow \mathcal{Y}' \subset \mathbb{R}$ ein nominales Modell der hier betrachteten Zielgröße darstellt—z.B. ein für Applikationszwecke trainiertes Ruß-Rohemissionsmodell. Weiterhin sei die multivariate Verteilung der Modellparameter bekannt, die aus den Parametertoleranzen des betrachteten Systems abgeleitet wurde. Die Menge $\mathcal{D} = \{\mathbf{x}_{i,j}\}_{j=1}^{N_{samp}}$ sei eine aus der multivariaten Eingangsparameterverteilung durch Stichprobenziehung an der Stelle \mathbf{x}_i gewonnene Stichprobenpopulation, wobei N_{samp} die Anzahl der Stichproben ist. Für einen Punkt \mathbf{x}_i , dessen Eingangsstreuung durch die multivariate Zufallsvariable X_i gegeben ist, ist die Verteilungsfunktion durch P_{X_i} gegeben. Zudem umfasse die Menge $\mathcal{F}_i = \{f_{i,j}\}_{j=1}^{N_{samp}}$, die durch Propagation ermittelten einzelnen Ergebnisse $f_{i,j} = f(\mathbf{x}_{i,j})$ aus denen die Ausgangsverteilung hervorgeht. Für eine Menge $\mathcal{X}^* = \{\mathbf{x}^*_i\}_{i=1}^{N_{grid}}$ an Punkten des Eingangsraumes kann dann durch Propagation eine Sammlung von Mengen $\{\mathcal{F}^*_i\}_{i=1}^{N_{grid}}$ ermittelt werden. Für die Wahl der Stützstellen $\{\mathbf{x}^*_i\}_{i=1}^{N_{grid}}$ können verschiedene Strategien angewendet werden. Eine valide Strategie ist es, die Stichproben \mathbf{x}^*_i zufällig aus einer Menge $\mathcal{S} = [\mathbf{x}^l, \mathbf{x}^u] \in \mathbb{R}^d$ zu ziehen, wobei $\mathbf{x}^l = (x^l_1, \dots, x^l_d)^\top$ die untere Grenze und $\mathbf{x}^u = (x^u_1, \dots, x^u_d)^\top$ die obere Grenze des durch die Menge \mathcal{X}' aufgespannten Eingangsparameterraumes darstellen. Eine Diskussion zu den unterschiedlichen Strategien findet weiter unten statt.

Definition 6.1 *Gegeben sei eine Zufallsvariable X . Dann ist das p -Quantil von X mit $p \in (0, 1)$ gegeben durch*

$$Q_X(p) = F_X^{-1} = \inf\{x \in \mathbb{R} : P(X \leq x) \geq p\}, \quad (6.2)$$

wobei F_X^{-1} die inverse Verteilungsfunktion von X ist.

Die empirischen p -Quantile der Ausgangsverteilungen, die durch $\{\mathcal{F}^*_i\}_{i=1}^{N_{grid}}$ gegeben sind, lassen sich zu $\mathcal{Q}^* = \{q^*_i\}_{i=1}^{N_{grid}}$ zusammenfassen. Das p -Quantil einer Verteilung ist ein Merkmal, unterhalb dessen ein vorgegebener Anteil p aller Ereignisse der Verteilung liegt, wobei p die Unterschreitungswahrscheinlichkeit ist.

Das Wertepaar $\{\mathbf{x}^*_i, q^*_i\}_{i=1}^{N_{grid}}$ stellt den Trainingsdatensatz des stochastischen Modells $g: \mathcal{X}^* \subset \mathbb{R}^{d \in \mathbb{N}} \rightarrow \mathcal{Q}^* \subset \mathbb{R}$ dar. Der Begriff *stochastisches* Modell wird hier verwendet, um die stochastische Natur der zu modellierenden statistischen Größe bzw. des zu modellierenden Merkmals zu verdeutlichen—anstelle des Quantils der Verteilung können auch andere Merkmale wie z.B. die Varianz als zu modellierende Größe gewählt werden. Das Modell wird unter der Annahme eines additiv unabhängigen und identisch verteilten Gaußschen Fehler als Gauß-Prozess trainiert und kann wie folgt dargestellt werden

$$g(\mathbf{x}) \sim \mathcal{GP}(m(\mathbf{x}), k(\mathbf{x}, \mathbf{X}^*)). \quad (6.3)$$

Dabei ist $\mathbf{X}^* = (\mathbf{x}_1^*, \dots, \mathbf{x}_{N_{grid}}^*)$ die Designmatrix der Form $d \times N_{grid}$. Das resultierende Gauß-Prozessmodell stellt einen stochastischen Prozess dar und liefert als Vorhersage eine Verteilung der Form

$$g(\mathbf{x}) \sim \mathcal{N}(\mathbb{E}[g(\mathbf{x})], \mathbb{V}[g(\mathbf{x})]), \quad (6.4)$$

wobei $\mathbb{E}[g(\mathbf{x})]$ der Erwartungswert der Verteilung und $\mathbb{V}[g(\mathbf{x})]$ die Kovarianz der Verteilung ist. Diese können wie folgt geschrieben werden

$$\mathbb{E}[g(\mathbf{x})] = k(\mathbf{x}, \mathbf{X}^*)^\top (k(\mathbf{X}^*, \mathbf{X}^*) + \sigma_n^2 \mathbf{I})^{-1} \mathbf{q}^*, \quad (6.5a)$$

$$\mathbb{V}[g(\mathbf{x})] = k(\mathbf{x}, \mathbf{x}) - k(\mathbf{x}, \mathbf{X}^*)^\top (k(\mathbf{X}^*, \mathbf{X}^*) + \sigma_n^2 \mathbf{I})^{-1} k(\mathbf{x}, \mathbf{X}^*) \quad (6.5b)$$

mit $\mathbf{q}^* = (q_1^*, \dots, q_{N_{grid}}^*)^\top$, wobei k die Kovarianzfunktion des Modells, σ_n die Fehlervarianz und \mathbf{I} die $d \times d$ Einheitsmatrix ist.

Algorithmus 2 gibt den Pseudocode des oben beschriebenen Modellierungsprozeds an.

Algorithmus 2 Pseudo-Code zur Modellbildung des stochastischen Modells.

- 1: **Daten:** $\mathcal{X}' = \{\mathbf{x}_i'\}_{i=1}^{N_{obs}}$, $\mathcal{Y}' = \{\mathbf{y}_i'\}_{i=1}^{N_{obs}}$
 - 2: Trainiere ein Modell $f: \mathcal{X}' \subset \mathbb{R}^{d \in \mathbb{N}} \rightarrow \mathcal{Y}' \subset \mathbb{R}$
 - 3: Wähle die Stützstellenanzahl N_{grid} und die Stichprobenanzahl N_{samp} aus.
 - 4: Wähle die Unterschreitungswahrscheinlichkeit p aus.
 - 5: $i \leftarrow 1$
 - 6: **while** $i \leq N_{grid}$ **do**
 - 7: Ziehe eine zufällige Stichprobe \mathbf{x}_i^* aus der geschlossenen Menge \mathcal{S} .
 - 8: $j \leftarrow 1$
 - 9: **while** $j \leq N_{samp}$ **do**
 - 10: Ziehe eine zufällige Stichprobe $\mathbf{x}_{i,j}^*$ aus der multivariaten Verteilung $P_{\mathbf{x}_i^*}$.
 - 11: Berechne $f(\mathbf{x}_{i,j}^*)$.
 - 12: **end while**
 - 13: Fasse die Ergebnisse zusammen zu $\mathcal{F}_i^* = \{f(\mathbf{x}_{i,j}^*)\}_{j=1}^{N_{samp}}$.
 - 14: Berechne das empirische p -Quantil q_i^* aus \mathcal{F}_i^* .
 - 15: **end while**
 - 16: Fasse die Ergebnisse zum Trainingsdatensatz $\{\mathbf{x}_i^*, q_i^*\}_{i=1}^{N_{grid}}$ zusammen.
 - 17: Trainiere das stochastische Modell $g: \mathcal{X}^* \subset \mathbb{R}^{d \in \mathbb{N}} \rightarrow \mathcal{Q}^* \subset \mathbb{R}$.
-

Einflussfaktoren

Die Güte des stochastischen Modells hängt von verschiedenen Faktoren ab. Das stochastische Modell kann hierbei unabhängig vom nominalen Modell betrachtet werden. Die Modellgüte hängt in keiner Weise vom nominalen Modell ab. Die Güte des stochastischen Modells bestimmt lediglich, wie gut es die Vorhersage des nominalen Modells reproduziert. Die Modellgüte des nominalen Modells muss gesondert betrachtet werden und beeinflusst maßgeblich die Abweichung zwischen realer Messung und Modellvorhersage. Die beiden Fehlerterme des nominalen und des stochastischen Modells sind nicht additiv. Es sollte jedoch angestrebt werden, beide Fehlerterme zu minimieren.

Definition 6.2 *Die Parameterunsicherheit beschreibt den Fehler zwischen den Parametern eines Modells und den tatsächlichen Werten der Parameter, die während des Experiments eingestellt wurden, jedoch den Experimentatoren unbekannt sind.*

Die oben genannten Einflussfaktoren, die sich auf die Güte des stochastischen Modells auswirken, sind zum einen die Wahl des Modells. Zum anderen hängt die Güte des Modells von den Trainingsdaten ab—hier spielen vor allem die Anzahl und Verteilung sowie die Güte der Daten eine Rolle. Die Anzahl und Verteilung der Daten wird durch die Wahl der Stützstellen bestimmt und ist frei wählbar. Die Güte der Daten wird hier ausschließlich durch den Approximationsfehler bestimmt, der bei der Schätzung der Merkmale entsteht. Andere Fehlerquellen wie Parameterunsicherheiten können vernachlässigt werden, da die Basis des Modells ein anderes Modell und kein physikalisches Experiment ist.

Stützstellenauswahl

Eine wichtige Rolle bei der Modellbildung spielt die Wahl der \mathbf{x}_i^* , da die Anzahl und die Lage der Stützstellen wesentlich zur Güte des endgültigen Modells beiträgt. Es sind verschiedene Strategien denkbar. Für alle Strategien wird vorausgesetzt, dass eine Menge \mathcal{S} existiert, aus der die Stützstellen entnommen werden können.

Die wohl einfachste Strategie ist die Ziehung der einzelnen Stützstellen durch Monte-Carlo-Sampling mit Pseudozufallszahlen. Der Vorteil dieser Methode liegt in der einfachen Durchführbarkeit. Der Nachteil liegt in der ungleichen Verteilung der Stichproben. Dies kann dazu führen, dass trotz einer hohen Anzahl von Stützstellen einige Bereiche des Eingangsparameterraumes nicht ausreichend abgedeckt werden. Im Kontext des überwachten Lernens ist die Verteilung der Punkte entscheidend für die Güte des Modells, da viele Annahmen auf die Ähnlichkeit zwischen Datenpunkten abzielen. Eine grundlegende Ähnlichkeitsannahme ist, dass zwei Punkte mit ähnlichen Eingangsparametern, die nahe beieinander liegen, wahrscheinlich ähnliche Zielwerte haben. Daher sollten Trainingspunkte, die nahe an einem Testpunkt liegen, aufschlussreich für die Vorhersage an diesem Punkt sein. Daraus folgt auch, dass die Vorhersagekraft des Modells in dünn besiedelten Gebieten abnimmt oder schlecht ist.

Definition 6.3 *Sei X eine multivariate Zufallsvariable auf einem Wahrscheinlichkeitsraum (Ω, Σ, P) mit Ergebnisraum Ω , Sigma-Algebra Σ und Wahrscheinlichkeitsmaß P . Dann ist die Varianz*

$$\text{Var}(X) := \mathbb{E}((X - \mathbb{E}(X))^2) = \int_{\Omega} (X - \mathbb{E}(X))^2 dP \quad (6.6)$$

definiert als die erwartete quadratische Abweichung dieser Zufallsvariablen von ihrem Erwartungswert \mathbb{E} , sofern dieser existiert.

Um dem oben beschriebenen Problem der ungleichen Verteilung zu begegnen, eignen sich u.a. das Quasi-Monte-Carlo-Sampling oder die Ziehung einer geschichteten Zufallsstichprobe, siehe Abschnitt 2.3. Beide Verfahren zielen auf eine Verringerung der Stichprobenvarianz

bzw. eine gleichmäßigere Verteilung der gezogenen Stichproben ab und zählen zu den Varianzreduktionsverfahren. Bei dem QMC-Sampling ergibt sich die Gleichverteilung aus den Eigenschaften der gewählten Niederdiskrepanz-Folge. Bei der geschichteten Zufallsstichprobe wird die Grundgesamtheit der Stichprobe in sinnvolle Gruppen unterteilt und aus jeder Gruppe eine Stichprobe gezogen, wodurch eine gleichmäßigere Verteilung erreicht wird.

Neben den oben beschriebenen Verfahren sind auch andere Strategien zur Auswahl der Stützstellen denkbar. Eine Stärke von GPs ist die resultierende Verteilung. Diese kann im Allgemeinen durch den Erwartungswert und die Standardabweichung oder Varianz beschrieben werden. Die Varianz des GP an einem Punkt \mathbf{x} kann als eine Art Unsicherheit über die Modellvorhersage interpretiert werden. Eine Strategie zur Reduzierung der Varianz ist daher die Platzierung von Punkten in Bereichen hoher Varianz. Ein prominenter Vertreter dieser Platzierungsstrategien ist das sogenannte Bayesian Optimization, bei der die Stichproben nach einem Optimalitätskriterium ausgewählt und platziert werden [39, 42, 105]. Da die Trainingszeit des Gesamtmodells jedoch in der Regel überschaubar ist, erübrigt sich die Betrachtung weiterer Ansätze.

Stichprobenauswahl für die Merkmalsschätzung

Eine ebenso wichtige Rolle bei der Modellbildung spielt die Stichprobenauswahl für die Merkmalsschätzung, da auch hier die Anzahl und Lage der Stützstellen wesentlich zur endgültigen Güte des Modells beiträgt. Auch hier sind verschiedenste Strategien der Stichprobenauswahl denkbar und es wird vorausgesetzt, dass eine Verteilung $P_{X_i^*}$ existiert, aus der die Stichproben gezogen werden können.

In der Literatur finden sich neben der Stichprobenziehung weitere Ansätze zur Approximation von Wahrscheinlichkeitsverteilungen. Diese Ansätze sind jedoch immer mit bestimmten Annahmen verbunden. Zum Beispiel, dass zwei Eingangsverteilungen unkorreliert sind—im Bereich der Verbrennungsmodellierung ist diese Annahme nicht zulässig, da z.B. die einzelnen Einspritzungen miteinander korrelieren und dieser Einfluss nicht vernachlässigbar ist. Außerdem wurde mit der Stichprobenziehung ein Verfahren gewählt, das unter allen Umständen eine gültige Lösung liefert.

Das Monte-Carlo-Sampling mit Pseudozufallszahlen ist auch bei der Stichprobenwahl für die Merkmalsschätzung die einfachste Strategie zur Bestimmung der einzelnen Stichproben dar. Der Vorteil der Methode liegt, wie bereits erläutert, in der einfachen Durchführbarkeit. Der Nachteil liegt in der hohen Varianz der gezogenen Stichproben, wodurch das Verfahren sehr langsam konvergiert. Daher ist es wichtig, die Varianz so gering wie möglich zu halten. Die Konvergenzordnung der Standardabweichung der Stichproben kann mit $\mathcal{O}(\frac{1}{\sqrt{N_{smp}}})$ angegeben werden. Wie bereits erwähnt stellen Varianzreduktionsverfahren eine gängige Alternative zum Monte-Carlo-Sampling mit Pseudozufallszahlen dar. Die Konvergenzordnungen von Quasi-Monte-Carlo-Verfahren liegen beispielsweise im Bereich von $\mathcal{O}(\frac{1}{N})$. Damit ist es möglich mit bereits wenigen Stichproben eine gute Approximation der geschätzten Verteilung zu erhalten.

Der hier vorgestellte Ansatz ist nicht nur auf die Modellierung von Quantilen beschränkt. Es können auch andere Merkmale wie z.B. die Momente der geschätzten Verteilung modelliert

werden. Die Wahl des Merkmals hängt dabei von der Formulierung des Optimierungsproblems ab. So ist es z.B. denkbar, dass das Optimierungsziel die Minimierung der Streuung der Ausgangsverteilung ist. Ein geeignetes Merkmal für diesen Anwendungsfall wäre die Varianz der Ausgangsverteilung.

6.4. Implementierungsdetails

Die vorgestellte Methodik wurde für den praktischen Einsatz als Software implementiert. Dabei wurde zum einen eine produktive Version implementiert, die für den Applikationsprozess vorgesehen ist. Zum anderen wurde eine weitere Python-Implementierung erstellt, die für Untersuchungen eingesetzt wird, bei denen eine hohe Transparenz und Justierbarkeit der einzelnen Methoden erforderlich ist. Für ein funktionierendes Gesamtkonstrukt werden neben dem Kernalgorithmus in der Regel noch eine Reihe von Hilfsfunktionalitäten und Schnittstellen benötigt. Die Menge an generiertem Code ist daher nicht trivial und stellt einen wesentlichen Beitrag dieser Arbeit dar.

Die produktive Version der Software ist in Matlab implementiert und ist als Software-Plugin für die kommerzielle Software ASCMO der ETAS GmbH konzipiert. Dies hat den Vorteil, dass viele nützliche ASCMO-Werkzeuge wie die Modellierungs- und Optimierungsalgorithmen, Analysewerkzeuge, Visualisierungsmethoden und Methoden zur abschließenden Generierung von Software-Datenständen bereits implementiert und verfügbar sind.

Zusätzlich zur Produktivversion wurde die Software vollständig in Python implementiert. Dies hat den Hintergrund, dass zur Durchführung der in Abschnitt 6.5 vorgestellten Untersuchungen eine hohe Transparenz und Justierbarkeit der einzelnen Methoden erforderlich ist. Die von ASCMO bereitgestellten Methoden sind in dieser Hinsicht nicht flexibel genug bzw. nicht dafür vorgesehen. Neben dem selbst geschriebenen Code verwenden wir auch einige externe Python-Bibliotheken: Zur Erzeugung pseudo-zufälliger Sobol-Sequenzen verwenden wir eine Implementierung von Chisari C. [27]. Dieser ist als offener Quellcode unter https://people.sc.fsu.edu/~jburkardt/py_src/sobol/sobol.html verfügbar. Zur Erstellung und zum Training von GP-Modellen nutzen wir die GPy-Bibliothek [111], dessen Quellcode ebenfalls frei verfügbar ist. Für die Maximum-Likelihood-Schätzung verwenden wir die im Paket enthaltene Implementierung des *Stochastic-Conjugate-Gradient*-Verfahrens. Zur Lösung des nichtlinearen Optimierungsproblems mit Nebenbedingungen wird das Sequential-Least-Squares-Programming (SLSQP)-Verfahren [65] aus der Scientific Computing Library [114] verwendet. Korrelierte Zufallsvariablen können durch Cholesky-Zerlegung erzeugt werden. Dies wird in [85] näher beschrieben.

6.5. Untersuchung der asymptotischen Eigenschaften des stochastischen Modells

In diesem Abschnitt werden die asymptotischen Eigenschaften der vorgestellten stochastischen Modelle diskutiert. Am Beispiel eines Reihen-Vierzylinder-Dieselmotors wird die Güte

eines stochastischen Ruß-Rohemissionsmodells in Abhängigkeit von der Anzahl der Stützstellen N_{grid} und der Anzahl der Stichproben N_{samp} , die zur Merkmalsschätzung verwendet werden, untersucht. Weiterhin wird untersucht, welches Verhältnis von N_{grid} zu N_{samp} die höchste Modellgüte bei einer konstanten Anzahl von Rechenoperationen liefert.

Erklärung

Teile der vorgestellten Ergebnisse wurden bereits in *Automotive and Engine Technology* durch Springer Nature veröffentlicht:

- **Mourat, Kadir**; Eckstein, Carola; Koch, Thomas: A stochastic design optimization methodology to reduce emission spread in combustion engines. In: *Automotive and Engine Technology* 6 (2021), Nr. 1, S. 15–29.

Beiträge der Autoren:

K.M., C.E. und T.K. haben den Inhalt des Papers konzipiert. **K.M.** und C.E. konzipierten die Experimente. **K.M.** führte die Experimente durch. **K.M.** führte die Nachbearbeitung der Daten durch. **K.M.** führte die Simulationen und die Optimierungen durch. **K.M.** schrieb den Computercode. **K.M.** erstellte die Grafiken. **K.M.** schrieb den Artikel mit Beiträgen von C.E. und T.K.

Zu den veröffentlichten Ergebnissen gehören insbesondere die Abbildungen 6.3 und 6.4 sowie die Tabellen 6.1, 6.2 und 6.3. Diese wurden mit Genehmigung von Springer Nature wiederverwertet.

Nominales Ruß-Rohemissionsmodell

Die für den Vergleich trainierten stochastischen Modelle basieren auf dem selben nominalen Modell. Datenbasis für das nominale Modell sind stationäre Prüfstandsmessungen aus einer automatisierten Messkampagne mit dem im Abschnitt 3.1.1 vorgestellten Reihen-Vierzylinder-Dieselmotor. Der zugehörige Versuchsplan wurde mittels DoE erstellt.

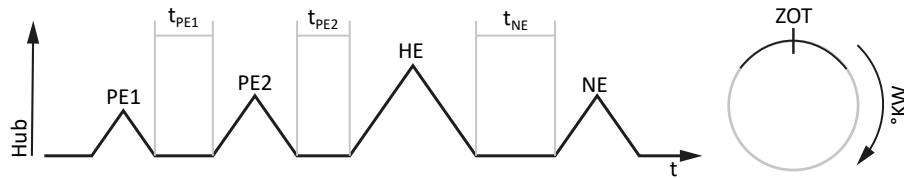


Abbildung 6.2.: Qualitative Darstellung des Einspritzmusters mit zwei Voreinspritzungen und einer Nacheinspritzung um den oberen Totpunkt [84], mit Genehmigung von Springer Nature.

Der Motor wird für einen fixen Betriebspunkt von 2000 min^{-1} und 400 kPa effektivem Mitteldruck vermessen. Das Modell ist daher nur für diesen Betriebspunkt gültig. Der DoE-Versuchsplan wird mit dem Softwarepaket ETAS ASCMO [57] erstellt. Die Grenzen des Versuchsplans sind in Tabelle 6.1 angegeben. Insgesamt werden $N_{obs} = 675$ Messungen aufgezeichnet. Die Verteilung der von ASCMO generierten Messpunkte folgt einer

Sobol-Sequenz. Die verstellten Steuergeräteparameter können in Luftsystemparameter und Kraftstoffversorgungsparameter unterteilt werden. Auf der Seite des Luftsystems werden die Frischluftmasse (m_L), der Ladedruck (p_2) und die Drallklappenstellung (ψ_{SW}) variiert. Auf der Seite des Kraftstoffsystems werden der Commonraildruck (p_{CR}), die Menge der ersten Voreinspritzung (m_{PE1}), die Menge der zweiten Voreinspritzung (m_{PE2}), die Menge der Nacheinspritzung (m_{NE}), der Ansteuerbeginn der Haupteinspritzung (ϕ_{HE})¹ und die Abstände zwischen den Einspritzungen variiert. Das Einspritzmuster ist in Abbildung 6.2 dargestellt. Die AGR-Rate wird in den Versuchen nicht aktiv variiert. Dies ermöglicht eine maximale Abgasrückführaten (X_{AGR}) und vermeidet Parameterkombinationen, die zu hohen Stickoxidkonzentrationen im Abgas führen. Das Niederdruck-AGR-Ventil und die Drosselklappe sind während der Versuche vollständig offen.

Tabelle 6.1.: Betriebsraumgrenzen des DoE-Plans [85].

Variationsparameter	Untere Grenze	Obere Grenze
m_L (mg/Hub)	250	375
p_2 (kPa)	105	125
ψ_{SW} (-)	0.10	0.45
p_{CR} (MPa)	45	90
m_{PE1} (mg/Hub)	0.8	2
m_{PE2} (mg/Hub)	0.8	3
m_{NE} (mg/Hub)	1	2.5
ϕ_{HE} (°KW n. OT)	1	9
t_{VE1} (µs)	500	1000
t_{VE2} (µs)	500	1000
t_{NE} (µs)	550	900

Für das Training des nominalen Modells werden folgende Annahmen getroffen. Zur Modellierung wird ein GP verwendet. Der Fehlerterm des GP wird als additiver Gaußscher Fehler angenommen. Als a-priori-Kovarianzfunktion wird ein quadratisch exponentieller-Kernel mit automatischer Relevanzbestimmung gewählt. Die Varianz der modellierten Ausgangsgröße wird mit Hilfe einer Box-Cox-Transformation stabilisiert [17]. Die Hyperparameter des GP werden durch Maximierung der Log-Likelihood bestimmt.

Zur Evaluierung des Modells werden NRMSE und R^2 als Metrik verwendet. Da die Anzahl der Trainingsdaten gering ist, wird der Datensatz nicht in Test- und Trainingsdatensatz aufgeteilt. Die Evaluationsmetriken werden zum einen für den Trainingsdatensatz und zum anderen auf Basis der LOOCV-Methode berechnet. Letztere dient zur Bestimmung des Generalisierungsfehlers bzw. als Ersatz für die Evaluation des Testdatensatzes. Die Ergebnisse der Evaluation sind in Tabelle 6.2 dargestellt. Für die Trainingsdaten erreicht das Modell

¹Wird in Grad Kurbelwinkel nach dem oberen Totpunkt (°KW n. OT) angegeben.

ein NRMSE von 0.0011 und ein R^2 von 0.98. Dies zeigt eine nahezu perfekte Vorhersage für die Trainingsdaten. Der Generalisierungsfehler ergibt sich aus den mit der LOOCV-Methode berechneten NRMSE und R^2 . Er beträgt 0.047 für NRMSE und 0.948 für R^2 . Der Unterschied zwischen dem Trainings- und Generalisierungsfehler ist gering, was auf eine minimale Überanpassung des Modells an die Trainingsdaten hindeutet.

Tabelle 6.2.: NRMSE und R^2 der Trainingsdaten sowie NRMSE und R^2 mit der LOOCV-Methode für das nominale Ruß-Rohemissionsmodell [85].

	Training		LOOCV	
	NRMSE	R^2	NRMSE	R^2
Schwärzungszahl (-)	0.0011	0.980	0.047	0.948

Parametertoleranzen des Reihen-Vierzylinder-Dieselmotors

Die Parametertoleranzen des Reihen-Vierzylinder-Dieselmotors sind in Tabelle 6.3 in Form von Wahrscheinlichkeitsverteilungen und einer Korrelationsmatrix aufgelistet. Die Werte wurden aus Datenblättern und empirischen Erfahrungswerten ermittelt. Die Ladedrucktoleranz und die Commonraildruck-Toleranz entsprechen den jeweiligen Sensortoleranzen. Die Luftmassentoleranz ist gleich der HFM-Toleranz gewählt. Für die Toleranz der Drallklappe wurde eine Normalverteilung mit einer Standardabweichung von einem Prozent der Gesamtventilöffnung angenommen. Die Mengentoleranzen der Injektoren wurden aus den Datenblättern der Injektoren abgeleitet. Der Toleranzbeitrag der Vor- und Nacheinspritzungen kann durch eine Normalverteilung mit einem Mittelwert von -0.05 und einer Standardabweichung von 0.135 angenähert werden. Die Einspritzmengen werden als korreliert angenommen. Die Abweichung des Ansteuerbeginns der Haupteinspritzung wird durch eine Normalverteilung mit einer Standardabweichung von 0.09 angenommen. Die Toleranzen für die Einspritzabstände werden dem Datenblättern der Injektoren entnommen. Für die Einspritzabstände wird eine Korrelation untereinander angenommen.

6.5.1. Güte des stochastischen Modells in Abhängigkeit der Stützstell- und Stichprobenanzahl

Zunächst wird anhand des Ruß-Rohemissionsmodells die Güte der stochastischen Modelle in Abhängigkeit von der Anzahl der Stützstellen N_{grid} und der Anzahl der Stichproben N_{samp} , die zur Merkmalsschätzung verwendet werden, untersucht. Für die Untersuchung werden verschiedene Quantilmodelle für eine variierende Anzahl von Stützstellen und Stichproben trainiert. Die Anzahl der Stützstellen variiert zwischen 256, 512 und 1024 und die Anzahl der Stichproben variiert zwischen 128, 256, 512, 1024 und 2048. Die Modelle werden dann anhand des NRMSE verglichen. Das NRMSE wird für einen Testdatensatz berechnet, der sich vom Trainingsdatensatz unterscheidet. Der Testdatensatz umfasst 100 Datenpunkte.

Tabelle 6.3.: Parametertoleranzen des Reihen-Vierzylinder-Dieselmotors [85].

Parameter	Verteilung	Correlation matrix
m_L (mg/Hub)	$\mathcal{N}(0, 12.25)$	1 0 0 0 0 0 0 0 0 0 0 0
p_2 (kPa)	$\mathcal{N}(0, 0.49)$	0 1 0 0 0 0 0 0 0 0 0 0
ψ_{SW} (-)	$\mathcal{N}(0, 0.0001)$	0 0 1 0 0 0 0 0 0 0 0 0
p_{CR} (MPa)	$\mathcal{N}(0, 1)$	0 0 0 1 0 0 0 0 0 0 0 0
m_{PE1} (mg/Hub)	$\mathcal{N}(-0.05, 0.135)$	0 0 0 0 1 1 1 0 0 0 0 0
m_{PE2} (mg/Hub)	$\mathcal{N}(-0.05, 0.135)$	0 0 0 0 1 1 1 0 0 0 0 0
m_{NE}	$\mathcal{N}(-0.05, 0.135)$	0 0 0 0 1 1 1 0 0 0 0 0
ϕ_{HE} (°KW n. OT)	$\mathcal{N}(0, 0.09)$	0 0 0 0 0 0 0 0 1 0 0 0
t_{PI1} (µs)	$\mathcal{N}(0, 12.5)$	0 0 0 0 0 0 0 0 0 1 1 1
t_{PI2} (µs)	$\mathcal{N}(0, 12.5)$	0 0 0 0 0 0 0 0 0 1 1 1
t_{PoI} (µs)	$\mathcal{N}(0, 12.5)$	0 0 0 0 0 0 0 0 0 1 1 1

Zusätzlich wird die Güte der Modelle mit einer Monte-Carlo- und einer Quasi-Monte-Carlo-Approximation verglichen.

Für das Training der einzelnen Gaußprozessmodelle wurden folgende Annahmen getroffen. Für die Unterschreitungswahrscheinlichkeit der Quantilmodelle wird $p = 0.95$ gewählt. Der Fehlerterm der GPs wird als additiver Gaußscher Fehler angenommen. Als a-priori-Kovarianzfunktion wird wie beim nominalen Modell ein SE-Kernel mit automatischer Relevanzbestimmung gewählt. Die Varianz der modellierten Ausgangsgröße wird durch eine Box-Cox-Transformation stabilisiert. Die Hyperparameter des GPs werden durch Maximierung der Log-Likelihood bestimmt. Zusätzlich zu den Modellen berechnen wir für den vorliegenden Testdatensatz die MC- und QMC-Approximation der Quantile. Diese werden für 128, 256, 512, 1024 und 2048 Stichproben approximiert. Bei der Berechnung des NRMSE der verschiedenen Approximatoren ist das wahre Quantil q_i für einen beliebigen Punkt nicht bekannt. Daher wird es durch MC-Sampling approximiert und als \tilde{q}_i notiert. Um eine möglichst gute Annäherung an den wahren Wert q_i zu erhalten, werden 10^6 Ziehungen durchgeführt. Das Residuum der verschiedenen Approximatoren für einen beliebigen Punkt ist somit gegeben durch

$$e_i = \tilde{q}_i - \hat{q}_i. \quad (6.7)$$

Da es sich bei den beschriebenen Experimenten um Zufallsexperimente handelt, variiert auch die resultierende NRMSE bei Wiederholung des jeweiligen Experiments. Um eine aussagekräftige Statistik zu erhalten, wird jedes Experiment insgesamt 50 Mal wiederholt. Zur Reproduktion der Ergebnisse wird zunächst jeweils eine Liste von *Seed-Keys* erzeugt. Anschließend wird das Experiment für jeden Seed-Key wiederholt.

Die Untersuchungsergebnisse sind in Abbildung 6.3 in Form von *Box-Plots* dargestellt. Dabei ist der NRMSE für die MC- und QMC-Approximationen sowie für die Modelle mit 256,

512 und 1024 Stützstellen über der Stichprobenanzahl $N_{samp} = 128, 256, 512, 1024, 2048$ aufgetragen. Die X-Koordinatenwerte sind nicht äquidistant. Ausreißer sind in den Box-Plots nicht dargestellt. Der Median, das untere und obere Quantil sowie die unteren und oberen *Whisker* sind aufgetragen. Die Whisker stellen hier das 2.5%-Quantil und das 97.5%-Quantil dar.

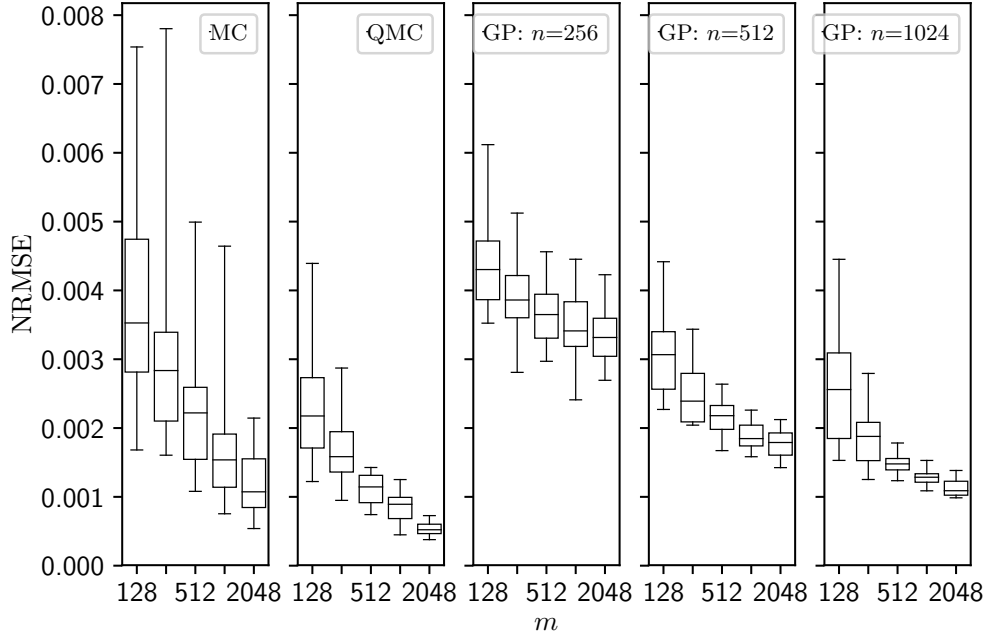


Abbildung 6.3.: NRMSE der Quantilmodelle mit 256, 512 und 1024 Stützstellen sowie der MC und QMC-Approximationen über der Stichprobenanzahl N_{samp} . Die Schwärzungszahl, gemessen an der ersten Probenahmestelle vor dem DOC, wird modelliert. Die Unterschreitungswahrscheinlichkeit des Quantils beträgt $p = 0.95$. Der NRMSE wird für 100 Testpunkte berechnet, die sich von den Trainingspunkten unterscheiden. Zur Bestimmung der Box-Plots werden die Experimente 50 Mal wiederholt [85].

Die MC-Approximation zeigt eine sehr hohe Varianz bzw. Streuung in der Vorhersagekraft. Die Varianz ist deutlich höher als die der QMC-Approximation und der Quantilmodelle. Der Unterschied zwischen der MC- und der QMC-Approximation erklärt sich durch die höhere Konvergenzrate des QMC-Verfahrens, siehe Abschnitt 2.3. Weiterhin ist zu erkennen, dass die Varianz aller Approximatoren für 128 und 256 Stichproben recht hoch ist. Erst ab 512 Stichproben wird eine statistisch signifikante Stichprobenanzahl und damit eine signifikante Varianzreduktion erreicht. Insgesamt kann festgehalten werden, dass bei zu geringer Stichprobenanzahl die Approximation der Quantile sowohl für MC als auch für QMC sehr ungenau ist und streut. Darunter leidet auch die Modellgüte. Die Varianz des Quantilmodells mit $N_{grid} = 256$ Stützstellen ist auch bei steigender Stichprobenanzahl hoch. Dies ist darauf zurückzuführen, dass der Eingangsparameterraum nicht ausreichend abgedeckt ist und das Modell auch mit sehr genauen Quantilwerten die komplexen Zusammenhänge

bzw. Nichtlinearitäten des modellierten Systems nicht abbilden kann. Für die Modelle mit 512 und 1024 Stützstellen ist ab 512 Stichproben eine deutliche Reduktion der Varianz zu erkennen. Die Varianz des Modells mit $N_{grid} = 1024$ Stützstellen ist geringer als die der QMC-Approximation—die Ausnahme bildet die Approximation mit 2048 Stichproben. Eine geringere Varianz des Modells ist ab einer hohen Punktdichte im Eingangsparameterraum zu erwarten, da durch die Maximierung der Log-Likelihood der Erwartungswert des GPs die wahrscheinlichste Funktion annimmt. Im Falle eines hinreichend fein benetzten Eingangsparameterraumes sollte der Erwartungswert des GPs bei kleinen Perturbationen in den Mess- bzw. Quantilwerten nahezu unverändert bleiben, da die Perturbationen durch den Fehlerterm des GPs revidiert werden. Im Vergleich zur QMC-Approximation ist der NRMSE der Modelle höher. Mit steigender Anzahl der Stützstellen ist jedoch eine Annäherung an deren Performanz zu erkennen. Insgesamt ist sichtbar, dass der NRMSE der Quantilmodelle mit steigender Stützstellen- und Stichprobenanzahl abnimmt.

6.5.2. Güte des stochastischen Modells in Abhängigkeit des Stützstellen-Stichproben-Verhältnisses

Anschließend wird anhand des Ruß-Rohemissionsmodells die Güte der stochastischen Modelle in Abhängigkeit vom Verhältnis $\frac{N_{grid}}{N_{samp}}$ der Stützstellen zu den Stichproben untersucht. Für die Untersuchung werden Quantilmodelle für verschiedene Verhältnisse $\frac{N_{grid}}{N_{samp}}$ trainiert. Dieses Vorgehen wird in Abhängigkeit von der Anzahl der Funktionsaufrufe, die zur Berechnung der Quantilwerte $q_1^*, \dots, q_{N_{grid}}^*$ notwendig sind, für $N_{grid} \cdot N_{samp} = 2^{17}, 2^{18}, 2^{19}, 2^{20}$ wiederholt. Anstelle des Begriffs Funktionsaufrufe kann auch der Begriff Rechenkapazität verwendet werden, der in diesem Fall analog ist.

Die Ergebnisse der Untersuchung sind in Abbildung 6.4 dargestellt. Der NRMSE ist über dem Verhältnis $\frac{N_{grid}}{N_{samp}}$ für $N_{grid} \cdot N_{samp} = 2^{17}, 2^{18}, 2^{19}, 2^{20}$ Funktionsaufrufe aufgetragen. Die x-Achse ist logarithmisch aufgetragen. Da die Experimente randomisiert sind, stellt der NRMSE das arithmetische Mittel von 50 Experimenten dar. Die Abbildung zeigt, dass die Modellgenauigkeit für ein festes Verhältnis $\frac{N_{grid}}{N_{samp}}$ mit steigender Rechenkapazität zunimmt. Es ist jedoch zu erkennen, dass je nach Verhältnis $\frac{N_{grid}}{N_{samp}}$ eine höhere Modellgüte teilweise auch mit weniger Rechenkapazität erreicht werden kann. Wir hatten zuvor geschlussfolgert, dass bei unzureichender Abdeckung des Eingangsparameterraumes das Modell auch bei hoher Stichprobenzahl die komplexen Zusammenhänge bzw. Nichtlinearitäten des modellierten Systems nicht abbilden kann. Abbildung 6.4 bestätigt diese Aussage und zeigt, dass Modelle mit wenigen Stützstellen auch mit steigender Stichprobenzahl keine signifikante Verbesserung aufweisen.

Stone hat bereits in [24] bewiesen, dass für differenzierbare nicht-parametrische Regressionsfunktionen asymptotische Optimalität gilt, d.h. es existiert eine optimale Konvergenzrate für das Problem. Darauf aufbauend haben Plumlee et al. in [91] den Beweis auf Probleme erweitert, bei denen statt einer Ein-Punkt-Vorhersage eine Wahrscheinlichkeitsverteilung geschätzt wird. Sie zeigen, dass unter bestimmten Randbedingungen ein solcher nicht-parametrischer Prädiktor mit $N_{grid}, N_{samp} \rightarrow \infty$ asymptotisch konsistent und effizient

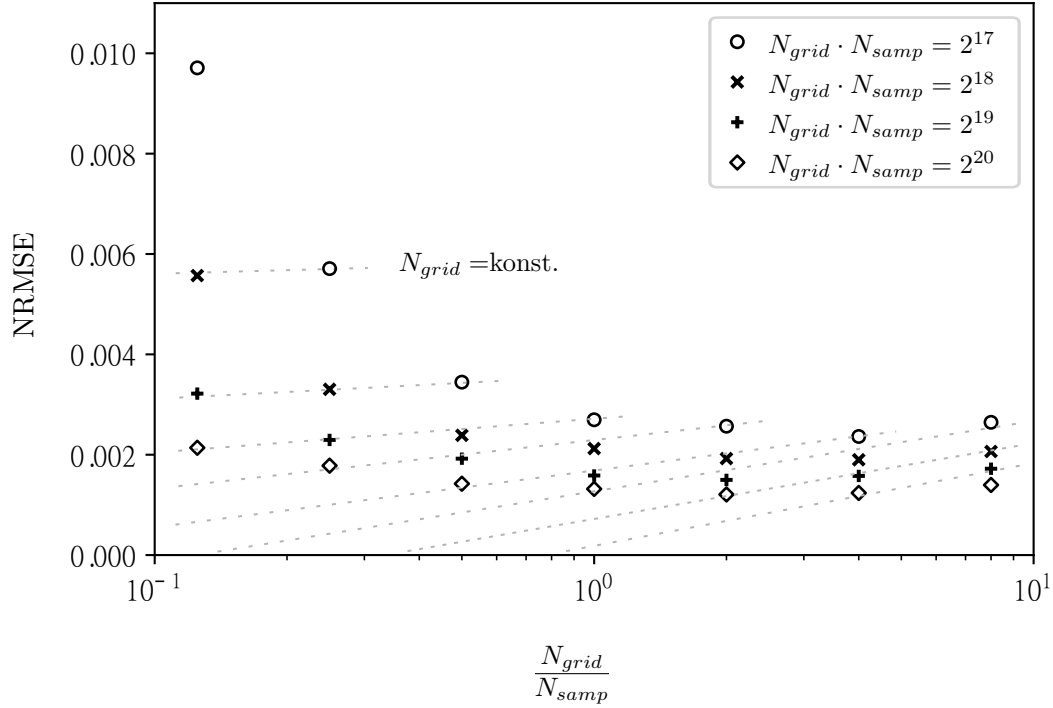


Abbildung 6.4.: NRMSE der Quantilmodelle über dem Verhältnis der Anzahl der Stützstellen zur Anzahl der Stichproben $\frac{N_{grid}}{N_{samp}}$ für $N_{grid} \cdot N_{samp} = 2^{17}$ (Kreis), $N_{grid} \cdot N_{samp} = 2^{18}$ (Kreuz), $N_{grid} \cdot N_{samp} = 2^{19}$ (Plus) und $N_{grid} \cdot N_{samp} = 2^{20}$ (Raute) Funktionsaufrufe des Nominalmodells, die zur Berechnung der Quantilwerte notwendig sind. Es wird die Schwärzungszahl modelliert, die an der ersten Probenahmestelle vor dem DOC gemessen wurde. Die Unterschreitungswahrscheinlichkeit am Quantil beträgt $p = 0.95$. Der NRMSE wird für 100 Testpunkte berechnet, die sich von den Trainingspunkten unterscheiden. Die gestrichelten Linien stellen Geraden konstanter Stützstellenanzahl dar [85].

ist—d.h. mit $N_{grid}, N_{samp} \rightarrow \infty$ konvergiert die modellierte Verteilung gegen die tatsächliche Verteilung. Die Randbedingungen lauten:

- Der modellierte Eingangsparameterraum wird mit zunehmender Anzahl von Punkten N_{grid} gefüllt.
- Die Annahme des stetigen Bereichs für den Gauß-Prozess ist gültig, d.h. die Funktion ist stetig und differenzierbar.
- Es gibt ein Verhältnis zwischen N_{grid} und N_{samp} , das immer die bestmögliche Modelleistung erreicht.

Plumlee et al. geben dieses Verhältnis als

$$N_{grid} \asymp N_{samp}^{\gamma} \quad (6.8)$$

an, wobei

$$\gamma > 2\frac{\tau}{d}. \quad (6.9)$$

τ ist ein Maß für die Glätte der geschätzten Funktion. Ist τ klein, so weist die Funktion eine eher raue Struktur auf und es werden mehr Stichproben benötigt, um sie zu Approximieren. Dies ist darauf zurückzuführen, dass bei rauen Funktionen der Informationsgewinn durch eine Erhöhung der Anzahl der Stützstellen geringer ist als durch eine Erhöhung der Anzahl der Stichproben. Die Ergebnisse von Plumlee et al. [91] gelten auch für die Schätzung eines einzelnen Quantils. Insofern können wir sie auf die von uns entwickelten Quantilmodelle übertragen. Aus Abbildung 6.4 ist ersichtlich, dass für den untersuchten Bereich das optimale Verhältnis $\frac{N_{grid}}{N_{samp}} \leq 5$ beträgt. Damit ergibt sich für den Parameter τ mit $d = 11$ nach Einsetzen in die Gleichungen 6.8 und 6.9: $\tau < 7.24$. Dieser Wert für τ deutet auf relativ glatte Verteilungsfunktionen bzw. Quantile des untersuchten Modells hin. Aus den Untersuchungen lässt sich für das Verhältnis $\frac{N_{grid}}{N_{samp}}$ als geeigneter Wertebereich $\frac{N_{grid}}{N_{samp}} \in [2, 5] \subset \mathbb{R}$ angeben.

7. Anwendungsbeispiele

In diesem Kapitel wird die in den vorhergehenden Kapiteln vorgestellte Methodik zur Bewertung und Validierung auf verschiedene Problemstellungen angewendet. Der erste Anwendungsfall bezieht sich auf die Optimierung eines Betriebspunktes des Vierzylinder-Reihendieselmotors. Im zweiten Anwendungsfall wird die gesamte Grundapplikation des V6-Dieselmotors hinsichtlich Stickoxid (NO_x) und Ruß am Beispiel des Worldwide harmonized Light-duty vehicles Test Cycle (WLTC) optimiert.

Das Kapitel beginnt mit dem Abschnitt 7.1, in dem die entwickelte Methodik an einem Reihen-Vierzylinder-Dieselmotor erprobt wird. Hier wird in Unterabschnitt 7.1.1 zunächst die Problemstellung formuliert. Anschließend wird in Unterabschnitt 7.1.2 die Methodik mit einer deterministischen Optimierung, die den Stand der Technik darstellt, verglichen. Anschließend wird in Unterabschnitt 7.1.3 die Methodik mit einem Brute-Force-Ansatz verglichen, bei dem das stochastische Optimierungsproblem rein numerisch gelöst wird. Abschließend werden die Ergebnisse in Abschnitt 7.1.4 zusammengefasst. In Abschnitt 7.2 wird das Verfahren auf die Optimierung eines V6-Dieselmotors angewendet. Zunächst wird in Unterabschnitt 7.2.1 eine Validierung des stochastischen Modellierungsansatzes durchgeführt. Anschließend wird in Unterabschnitt 7.2.2 der Motor mit Hilfe des SEMO-Ansatzes hinsichtlich des WLTC optimiert.

7.1. Anwendungsbeispiel Reihen-Vierzylinder-Dieselmotor

In diesem Abschnitt wird die entwickelte Methodik auf die Optimierung eines einzelnen Betriebspunktes eines Reihen-Vierzylinder-Dieselmotor angewendet. Das Optimierungsziel für den betrachteten Betriebspunkt ist die Minimierung der NO_x -Emissionen unter Einhaltung von Zielwerten für den Ruß und das Verbrennungsgeräusch. Die Aufgabenstellung stellt ein Optimierungsproblem mit Nebenbedingungen dar und wird zum Vergleich sowohl als deterministisches als auch als stochastisches Problem formuliert. Die Ergebnisse beider Optimierungsszenarien werden miteinander verglichen.

Erklärung

Teile der vorgestellten Ergebnisse wurden bereits in *Automotive and Engine Technology* durch Springer Nature veröffentlicht:

- **Mourat, Kadir;** Eckstein, Carola; Koch, Thomas: A stochastic design optimization methodology to reduce emission spread in combustion engines. In: *Automotive and Engine Technology* 6 (2021), Nr. 1, S. 15–29.

Beiträge der Autoren:

K.M., C.E. und T.K. haben den Inhalt des Artikels konzipiert. **K.M.** und C.E. konzipierten die Experimente **K.M.** führte die Experimente durch. **K.M.** führte die Nachbearbeitung der Daten durch. **K.M.** führte die Simulationen und die Optimierungen durch. **K.M.** schrieb den Computercode. **K.M.** erstellte die Grafiken. **K.M.** schrieb den Artikel mit Beiträgen von C.E. und T.K.

Die veröffentlichten Ergebnisse umfassen die Abbildungen 7.1 und 7.2 sowie die Tabellen 7.1–7.5. Diese wurden mit Genehmigung von Springer Nature wiederverwertet.

7.1.1. Problemstellung

Um den Nutzen der vorgestellten Applikationsmethodik zu beurteilen, vergleichen wir sie mit der standardmäßigen modellbasierten Applikation, die den Stand der Technik repräsentiert. Da es sich bei den betrachteten Zielgrößen in diesem Fall um deterministische Größen handelt, bezeichnen wir dieses Vorgehen im Folgenden als *deterministisches Optimierungsszenario*. Analog dazu wird der Optimierungsansatz mit der vorgestellten Methodik als *stochastisches Optimierungsszenario* bezeichnet, da die zu optimierenden Zielgrößen stochastische Größen darstellen. Wie bereits erwähnt, handelt es sich bei der Aufgabenstellung um ein Optimierungsproblem mit Nebenbedingungen. Ziel solcher Optimierungsprobleme ist es, eine entsprechende Zielfunktion zu minimieren und gleichzeitig sicherzustellen, dass eine Reihe von Nebenbedingungen eingehalten werden.

Die den Untersuchungen zugrunde liegende Messkampagne wurde bereits in Abschnitt 6.5 beschrieben. Der Motor wurde für einen festen Betriebspunkt von 2000 min^{-1} und einem effektiven Mitteldruck 400 kPa vermessen. Insgesamt wurden elf Parameter variiert und $N_{\text{obs}} = 675$ Messungen aufgezeichnet. Die Grenzen des Versuchsplans sind in Tabelle 6.1 angegeben. Neben der Schwärzungszahl (engl.: *Filter Smoke Number*, FSN) wurden während der Versuche auch das Verbrennungsgeräusch und die Rohemissionen NO_x gemessen. Die Entnahmestelle des Abgasanalysators befindet sich zwischen dem DOC und der Abgas-turbine. Das Verbrennungsgeräusch ist eine aus dem indizierten Zylinderdruck berechnete Größe.

Deterministische Formulierung des Problems

Bei der Formulierung des deterministischen Optimierungsproblems orientieren wir uns an der gängigen Praxis. Stand der Technik sind hier empirisch ermittelte Werte und die Verwendung von Sicherheitsabständen bzw. Sicherheitsfaktoren dar. So werden für einen Betriebspunkt empirische Grenzwerte der zu optimierenden Größen ermittelt und mit einem Sicherheitsfaktor versehen, der u.a. Toleranzeinflüsse berücksichtigen soll. So werden für das vorliegende Optimierungsszenario als obere Grenzwerte eine ein Schwärzungszahl von 0.5 und ein maximales Verbrennungsgeräusch von 86dBA gewählt. Für das Verbrennungsgeräusch wird ein Sicherheitsabstand 1dBA gewählt. Für die Schwärzungszahl wird ein Sicherheitsabstand von 30% angesetzt, was einem in der Praxis üblichen Wert entspricht.

Tabelle 7.1.: Sicherheitsfaktoren für das deterministische Optimierungsszenario [85].

	j	$c_{j,lim}$	$c_{j,fac}$
Ruß (FSN)	1	0.5	0.15
Verbrennungsgeräusch (dBA)	2	86	1

In Tabelle 7.1 sind die genauen Werte angegeben. Die jeweiligen Grenzwerte und Sicherheitsabstände werden durch $c_{j,lim}$ und $c_{j,fac}$ dargestellt.

Für das Training der nominalen bzw. deterministischen Motormodelle werden folgende Annahmen getroffen. Es werden Gauß-Prozessmodelle mit additivem Gauß-Fehler trainiert. Für die Kovarianzfunktion der einzelnen Modelle wird der quadratisch exponentieller-Kernel mit automatischer Relevanzbestimmung gewählt. Die Varianz der modellierten Ausgangsgrößen wird mittels Box-Cox-Transformation stabilisiert. Die Hyperparameter der GPs werden durch Maximierung der Log-Likelihood-Funktion bestimmt.

Tabelle 7.2 gibt einen Überblick über die relevanten Evaluierungsmetriken für die trainierten Modelle. Das NO_x-Rohemissionsmodell zeigt eine nahezu perfekte Übereinstimmung mit den Trainingsdaten. Darüber hinaus ist auch der LOOCV-Fehler äußerst gering. Auch das Verbrennungsgeräuschmodell und das Ruß-Rohemissionsmodell weisen eine nahezu ähnlich hohe Modellgüte auf. Der untersuchte Eingangsparameterraum, in dem das Modell gültig ist, kann durch die Menge $\mathcal{S} = [\mathbf{x}^l, \mathbf{x}^u] \in \mathbb{R}^d$ beschrieben werden, wobei $\mathbf{x}^l = (x_1^l, \dots, x_d^l)^\top$ die Minimalwerte und $\mathbf{x}^u = (x_1^u, \dots, x_d^u)^\top$ die Maximalwerte der jeweiligen Eingangsdimension darstellen. Die genauen Werte sind in der Tabelle 6.1 in Abschnitt 6.5 angegeben.

Tabelle 7.2.: NRMSE und R^2 der Trainingsdaten sowie NRMSE und R^2 mit der LOOCV-Methode für die nominalen Motormodelle (Reihen-Vierzylinder-Dieselmotor) [85].

	Training		LOOCV	
	NRMSE	R^2	NRMSE	R^2
NO _x (g/kWh)	0.0003	0.997	0.012	0.995
Ruß (FSN)	0.0011	0.980	0.047	0.948
Verbrennungsgeräusch (dBA)	0.0015	0.954	0.033	0.968

Das deterministische Optimierungsproblem kann als

$$\min_{\mathbf{x}} f_0(\mathbf{x}) \quad (7.1a)$$

$$\text{s.t.} \quad c_{j,lim} - c_{j,fac} - f_j(\mathbf{x}) \leq 0 \quad \forall \quad j \in \{1, 2\}, \quad (7.1b)$$

$$x_i^l \leq x_i \leq x_i^u \quad \forall \quad i \in \{1, \dots, d\} \quad (7.1c)$$

formuliert werden, wobei f_0 die zu minimierende deterministische Zielfunktion bzw. das NO_x-Modell und f_j die einzugrenzenden Funktionen darstellen.

Stochastische Formulierung des Problems

Im Gegensatz zum deterministischen Standardansatz basiert der im Rahmen dieser Arbeit entwickelte stochastische Ansatz nicht auf empirisch ermittelten Werten und Faktoren. Vielmehr bietet er einen systematischen Ansatz zur Lösung des Optimierungsproblems, indem die aus den Bauteiltoleranzen resultierende Systemstreuung quantifiziert und bei der Applikation bzw. Optimierung berücksichtigt wird. Für die Formulierung des Optimierungsproblems bedeutet dies, dass keine zusätzlichen Faktoren benötigt werden. Weiterhin kann das Problem so formuliert werden, dass der vorgegebene Grenzwert mit einer vorgegebenen Unterschreitungswahrscheinlichkeit eingehalten wird.

Für die vorliegende Problemstellung werden drei Quantilmodelle trainiert—ein Modell für jede Ausgangsgröße. Die Parametertoleranzen für die Modelle sind in Tabelle 6.3 in Abschnitt 6.5 angegeben. Die Unterschreitungswahrscheinlichkeit p wird für alle Modelle mit 0.95 angegeben. Zur Bestimmung der Modellstützstellen bzw. Trainingspunkte werden mittels Quasi-Monte-Carlo-Sampling 500 Stichproben aus \mathcal{S} gezogen. Zur Bestimmung der zugehörigen empirischen Quantile werden mittels Quasi-Monte-Carlo-Sampling 1000 Stichproben aus der Eingangsparameterverteilung gezogen. Dies entspricht einem Stützstellen-Stichprobenverhältnis von $\frac{N_{grid}}{N_{samp}} = 0.5$ und liegt außerhalb des in Abschnitt 6.5 als optimal angesehenen Bereichs $[2, 5]$. Die Modellgüte ist für das gewählte Verhältnis jedoch nur geringfügig schlechter als für $\frac{N_{grid}}{N_{samp}} = 0.5$. Unter Berücksichtigung der Vorhersagekosten eines GP mit $\mathcal{O}(n)$ reduziert sich der Rechenaufwand für eine Vorhersage deutlich, wodurch sich der Zeitaufwand für den Optimierungsschritt verringert. Der Fehlerterm der GPs wird als additiver Gaußscher Fehler angenommen. Als A-Priori-Kovarianzfunktion wird wie beim nominalen Modell ein quadratisch exponentieller-Kernel mit automatische Relevanzbestimmung gewählt. Die Varianz der modellierten Ausgangsgröße wird durch eine Box-Cox-Transformation stabilisiert. Die Hyperparameter des GP werden durch Maximierung der Log-Likelihood bestimmt.

Die Modellgüten sind in Tabelle 7.3 zusammengefasst. Sie wurden sowohl für den Trainingsdatensatz als auch für einen Testdatensatz mit 100 Datenpunkten berechnet. Der NRMSE ist für alle drei Modelle verschwindend gering—dies gilt sowohl für die Trainings- als auch für die Testdaten. Dies zeigt, dass die Quantilmodelle die tatsächlichen Quantile nahezu perfekt abbilden und im Allgemeinen auch sehr gut generalisieren.

Tabelle 7.3.: NRMSE und R^2 der Trainings- und Testdaten für die Quantilmodelle [85].

	Training		Test	
	NRMSE	R^2	NRMSE	R^2
NO _x (g/kWh)	0	1	0.001	1
Ruß (FSN)	0	1	0.002	1
Verbrennungsgeräusch (dBA)	0.001	1	0.011	0.9975

Das entsprechende stochastische Optimierungsproblem kann somit als

$$\min_{\mathbf{x}} q_0(\mathbf{x}) \quad (7.2a)$$

$$\text{s.t. } c_{j,lim} - q_j(\mathbf{x}) \leq 0 \quad \forall j \in \{1, 2\}, \quad (7.2b)$$

$$x_i^l \leq x_i \leq x_i^u \quad \forall i \in \{1, \dots, d\} \quad (7.2c)$$

formuliert werden, wobei die Funktionen q die jeweiligen Quantilmodelle darstellen. Für die Unterschreitungswahrscheinlichkeiten gilt $p = 0.95$.

Brute-Force-Ansatz

Zu Vergleichs- und Validierungszwecken wird das durch die Gleichungen 7.2 beschriebene Optimierungsproblem zusätzlich zu der hier entwickelten Methodik mit einem *Brute-Force*-Ansatz gelöst—Brute-Force bedeutet in diesem Zusammenhang, dass die Funktionswerte $q(\mathbf{x})$ durch die entsprechenden empirischen Quantile ersetzt werden, die mit MC bzw. in diesem Fall QMC ermittelt wurden.

Zur Berechnung der empirischen Quantile wird die Ausgangsverteilung der Funktion an einem Punkt \mathbf{x} mittels QMC geschätzt. Dazu werden mittels QMC $N_{smp,BF} = 500$ quasi zufällige Stichproben gezogen. Im Gegensatz zu den oben beschriebenen modellbasierten Ansätzen liefert die numerische Approximation keinen Gradienten. Daher muss der Gradient numerisch mit Hilfe der partiellen Ableitungen berechnet werden. Die partielle Ableitung einer Funktion $f(\mathbf{x})$ in Richtung x_i ist gegeben durch

$$\frac{\partial f}{\partial x_i}(\mathbf{x}^*) = \lim_{h \rightarrow 0} \frac{f(x_1^*, \dots, x_i^* + h, \dots, x_d^*) - f(x_1^*, \dots, x_i^*, \dots, x_d^*)}{h}, \quad (7.3)$$

wobei h möglichst klein zu wählen ist. Der Gradient von f ist dann gegeben durch

$$\nabla f(\mathbf{x}) = \left(\frac{\partial f}{\partial x_1}(\mathbf{x}^*), \dots, \frac{\partial f}{\partial x_d}(\mathbf{x}^*) \right)^\top. \quad (7.4)$$

Aus den obigen Gleichungen ist ersichtlich, dass diese Art der numerischen Approximation des Gradienten im Vergleich zu einer analytischen Lösung zufallsbehaftet ist, so dass am

gleichen Punkt \mathbf{x} bei mehrfacher Berechnung unterschiedliche Gradienten auftreten können. Die Genauigkeit des Gradienten nimmt jedoch mit der Anzahl der Stichproben zu.

7.1.2. Vergleich der Methodik mit dem Stand der Technik

In diesem Abschnitt werden die Ergebnisse der deterministischen Optimierung mit den Ergebnissen des im Rahmen dieser Arbeit entwickelten SEMO-Ansatzes verglichen. Der deterministische Optimierungsansatz stellt hierbei den Stand der Technik dar. Dazu werden die in den Gleichungen 7.1 und den Gleichungen 7.2 formulierten Optimierungsprobleme mit Hilfe des SLSQP-Verfahrens gelöst.

Die Optimierungsparameter werden für beide Szenarien identisch gewählt. Die Funktionstoleranz wird auf 10^{-5} gesetzt. Um das globale Optimum zu finden und lokale Optima zu vermeiden, wird das Problem jeweils $k = 20$ mal für verschiedene Initialisierungspunkte gelöst. Die Initialisierungspunkte $\mathbf{x}_{init,k}$ werden quasi zufällig aus dem Eingangsparameterraum \mathcal{S} gewählt. Außerdem werden dem Optimierer die analytischen Gradienten der GPs zur Verfügung gestellt. Für die restlichen Einstellungen werden die Voreinstellungen des Optimierers verwendet.

Tabelle 7.4.: Optimierungsergebnisse für das deterministische und das stochastische Optimierungsszenario. Dargestellt sind die Ergebnisse für die NO_x - und Ruß-Rohemissionen sowie für das Verbrennungsgeräusch. Aus den mit jeweils 10^6 QMC-Stichproben approximierten Ausgangsverteilungen wurden die verschiedenen Statistiken der jeweiligen Optima berechnet [85].

	NO_x (g/kWh)			Ruß (FSN)			Geräusch (dBA)		
	Nom.	Mean	$Q_{0.95}$	Nom.	Mean	$Q_{0.95}$	Nom.	Mean	$Q_{0.95}$
Deterministisch	1.122	1.129	1.279	0.351	0.367	0.450	85.000	85.015	85.236
SEMO	0.227	0.231	0.281	0.425	0.430	0.504	85.350	85.383	85.841

Die Ergebnisse der Optimierung sind in der Tabelle 7.4 und den Abbildungen 7.1 und 7.2 dargestellt. Die Werte in der Tabelle 7.4 stellen die Verteilungsstatistiken am jeweiligen Optimum dar. Die Statistiken der jeweiligen Optima wurden aus den Ausgangsverteilungen berechnet, die jeweils mit 10^6 QMC-Stichproben approximiert wurden. Für beide Optima sind jeweils der Nominalwert, der arithmetische Mittelwert der Verteilung und das 95%-Quantil der Verteilung angegeben. Im Gegensatz dazu zeigen die Abbildungen 7.1 und 7.2 jeweils die nominalen Ergebnisse der beiden Optima und 200 quasi zufällige Realisationen. Zusätzlich sind auf den Sekundärachsen der Abbildungen die zugehörigen Wahrscheinlichkeitsdichteverteilungen dargestellt.

Obwohl beide Optimierungsszenarien dasselbe Ziel verfolgen, sind deutliche Unterschiede in den Ergebnissen bzw. zwei distinkte Ergebnisse zu beobachten. Im Falle des deterministischen Szenarios sind die NO_x -Emissionen mit einem nominalen Wert von 1.122g/kWh, einem Mittelwert von 1.129g/kWh und einem 95%-Quantil von 1.279g/kWh relativ hoch. Der SEMO-Ansatz liefert hier einen nominalen Wert von 0.227g/kWh und einen Mittelwert

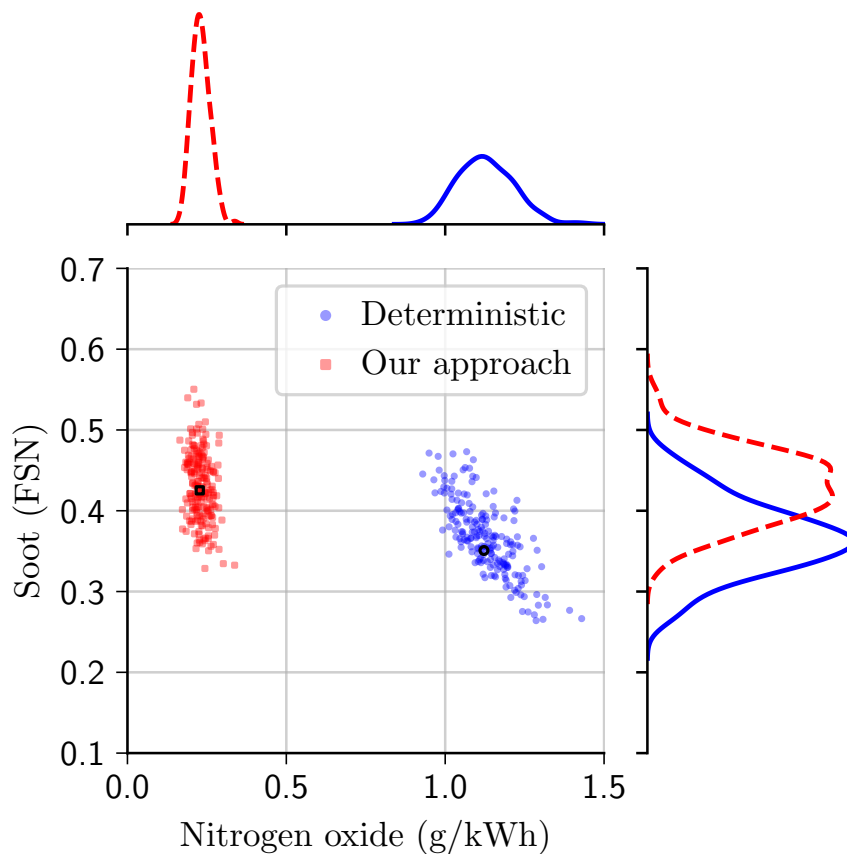


Abbildung 7.1.: Ergebnisse der NO_x -Rohemissionen und der Schwärzungszahl für das deterministische und das stochastische Optimierungsszenario. Deterministisches Szenario: nominales Ergebnis (schwarzer Kreis), quasi zufällige Realisierungen (blaue Kreise), Wahrscheinlichkeitsdichteverteilung (blaue durchgezogene Linie). Stochastisches Szenario: nominales Ergebnis (schwarzes Quadrat), quasi zufällige Realisierungen (rote Quadrate), Wahrscheinlichkeitsdichteverteilung (rot gestrichelte Linie) [85].

von 0.231g/kWh. Das 95%-Quantil liegt bei 0.281g/kWh. Es ist also insgesamt eine deutlich stärkere Minimierung der NO_x -Emissionen zu erkennen.

Für die deterministische Optimierung zeigt sich, dass der Zielwert von 0.35FSN mit 0.351FSN nahezu perfekt eingehalten wird. Die nominalen und mittleren Schwärzungszahl-Werte liegen mit 0.351FSN bzw. 0.367FSN deutlich unter den stochastischen Ergebnissen mit 0.425FSN bzw. 0.43FSN. Die $Q_{0.95}$ -Werte liegen mit 0.45FSN für das deterministische und 0.504FSN für das stochastische Szenario wiederum näher beieinander, da die Verteilung der Schwärzungszahl im deterministischen Fall eine höhere Varianz aufweist. Insgesamt ist für die Schwärzungszahl zu beobachten, dass im deterministischen Fall die a priori angenommenen Sicherheitsabstände zu streng gewählt wurden—dies ist eine der größten Schwächen des Standardansatzes. Darüber hinaus wählt der von uns entwickelte stochastische Ansatz

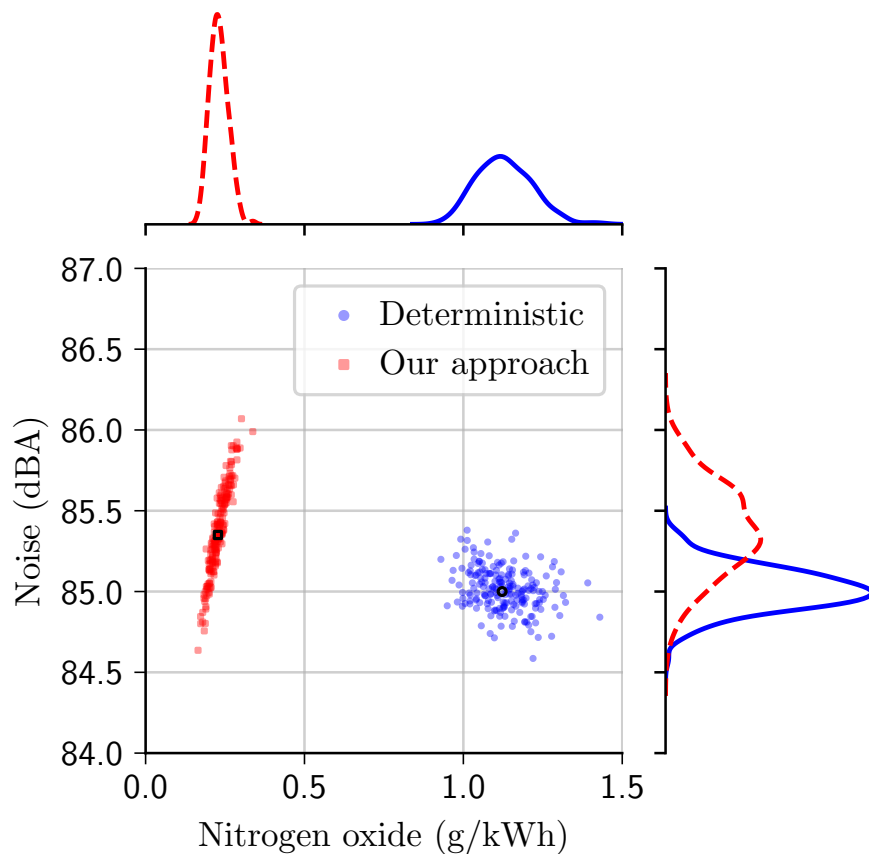


Abbildung 7.2.: Ergebnisse der NO_x -Rohemissionen und des Verbrennungsgeräuschs für das deterministische und stochastische Optimierungsszenario. Deterministisches Szenario: nominales Ergebnis (schwarzer Kreis), quasi zufällige Realisierungen (blaue Kreise), Wahrscheinlichkeitsdichteverteilung (blaue durchgezogene Linie). Stochastisches Szenario: nominales Ergebnis (schwarzes Quadrat), quasi zufällige Realisierungen (rote Quadrate), Wahrscheinlichkeitsdichteverteilung (rote gestrichelte Linie) [85].

ein Ergebnis, das in Bezug auf die Schwärzungszahl robuster ist und die Vorgaben im Allgemeinen einhält. 95% aller Realisierungen erreichen eine Schwärzungszahl von weniger als 0.504FSN, was in etwa der Vorgabe von 0.5FSN entspricht. Interessant ist auch, dass die Verteilung im stochastischen Fall multimodal ist. Dies bestätigt unsere anfänglichen Annahmen bezüglich der Komplexität des Problems. Aufgrund der Nichtlinearität des Systems und der Korrelationen zwischen den einzelnen Eingangsparametern weichen die resultierenden Ausgangsverteilungen von der Normalverteilung ab und können durch entsprechende Verfahren, die eine Normalverteilung voraussetzen, nicht wahrheitsgetreu approximiert werden.

Die deterministische Optimierung hält auch für das Verbrennungsgeräusch die Zielvorgabe von 85dBA für den Nominalwert ein. Der Mittelwert der Verteilung und das 95%-Quantil liegen mit 85.015dBA und 85.236dBA sehr nahe am nominalen Wert. Die Varianz des Ver-

brennungsgeräusches ist für den deterministischen Fall eher gering. Insgesamt ist auch hier eine zu strenge Wahl des Sicherheitsabstandes zu beobachten. Für das stochastische Optimierungsszenario weist die Verteilung eine deutlich höhere Varianz auf. Der Nominalwert und das arithmetische Mittel betragen 85.35dBA bzw. 85.383dBA. Der $Q_{0.95}$ -Wert liegt mit 85.841dBA unter dem vorgegebenen Grenzwert von 86dBA.

Insgesamt ist festzustellen, dass die Grenzwerte für FSN und Verbrennungsgeräusch im deterministischen Fall zu streng gewählt sind. Durch die Wahl kleinerer Sicherheitsabstände wäre eine Annäherung der Verteilungen an die tatsächlichen Grenzwerte theoretisch möglich. Die Wahl eines kleineren Sicherheitsabstandes stellt jedoch keine sinnvolle Option dar, da dies auch zu einer Überschreitung des Grenzwertes führen kann, da in der Praxis die Emissionsverteilung für den deterministischen Fall nicht bekannt ist. Für den stochastischen Fall beobachten wir sowohl für die FSN als auch für das Verbrennungsgeräusch eine sehr starke Annäherung der Ausgangsverteilungen an die tatsächlichen Grenzwerte. Dies ist darauf zurückzuführen, dass der von uns entwickelte stochastische Ansatz im Gegensatz zum deterministischen Standardansatz auf empirisch ermittelte Faktoren verzichtet und das zugrundeliegende Optimierungsproblem deutlich zielgerichteter formuliert. Durch die gezieltere Optimierung ist eine deutlich höhere Reduktion der NO_x -Emissionen möglich. Während im deterministischen Fall ein Ergebnis von 1.279g/kWh für das 95%-Quantil erreicht wird, kann dieses im stochastischen Fall auf 0.281g/kWh reduziert werden, was einer Reduktion von ca. 78% entspricht.

7.1.3. Vergleich der Methodik mit dem Brute-Force-Ansatz

In diesem Abschnitt vergleichen wir die Ergebnisse der deterministischen und stochastischen Optimierung mit einer Brute-Force-Lösung des stochastischen Optimierungsproblems. Brute-Force bedeutet in diesem Zusammenhang, dass die Quantilwerte in den Gleichungen 7.2 durch die entsprechenden empirischen Quantile approximiert werden. Das Optimierungsproblem wird wie im deterministischen und stochastischen Setting mit Hilfe der SLSQP-Methode gelöst. Anschließend werden die Ergebnisse insbesondere hinsichtlich der Rechenzeit verglichen. Es soll gezeigt werden, dass der Rechenaufwand des entwickelten Ansatzes vergleichbar mit der deterministischen Variante und deutlich geringer als bei der Brute-Force-Methode ist.

Die Optimierungsparameter sind identisch zu denen in Abschnitt 7.1.2. Die einzige Ausnahme ist der Gradient der Brute-Force-Methode. Da hier kein analytischer Gradient zur Verfügung steht, wird der numerische Gradient verwendet. Alle Berechnungen werden auf einem Computer mit einem Intel Xeon E3-1505M Prozessor, 32GB Arbeitsspeicher und einer Nvidia Quadro M2000M Grafikkarte durchgeführt.

Die Ergebnisse der Optimierung sind in der Tabelle 7.5 dargestellt. Die Werte in der Tabelle stellen die Verteilungsstatistiken der jeweiligen Optima dar. Die Statistiken der jeweiligen Optima wurden aus den Ausgangsverteilungen berechnet, die jeweils mit 10^6 QMC-Stichproben approximiert wurden. Für beide Optima sind jeweils der Nominalwert, der arithmetische Mittelwert der Verteilung und das 95%-Quantil der Verteilung angegeben.

Tabelle 7.5.: Optimierungsergebnisse für das stochastische Optimierungsszenario sowie für den Brute-Force-Ansatz. Dargestellt sind die Ergebnisse für die NO_x - und Ruß-Rohemissionen sowie für das Verbrennungsgeräusch. Aus den mit jeweils 10^6 QMC-Stichproben approximierten Ausgangsverteilungen wurden die verschiedenen Statistiken der jeweiligen Optima berechnet. [85]

	NO_x (g/kWh)			Ruß (FSN)			Geräusch (dBA)		
	Nom.	Mean	$Q_{0.95}$	Nom.	Mean	$Q_{0.95}$	Nom.	Mean	$Q_{0.95}$
Deterministisch	1.122	1.129	1.279	0.351	0.367	0.450	85.000	85.015	85.236
Unser Ansatz	0.227	0.231	0.281	0.425	0.430	0.504	85.350	85.383	85.841
Brute-Force	0.218	0.220	0.268	0.432	0.436	0.505	85.483	85.520	85.977

Das Ergebnis der Brute-Force-Optimierung ist nahezu identisch mit dem Ergebnis des stochastischen Szenarios. Es ist zu erkennen, dass sowohl die NO_x - als auch die Ruß- und Verbrennungsgeräuschwerte auf einem sehr ähnlichen Niveau liegen. Dies ist zu erwarten, da das Optimierungsproblem identisch formuliert ist. Durch die hohe Anzahl an Zufallsinitialisierungen ist zudem sichergestellt, dass in beiden Fällen das globale Minimum gefunden wird. Unterschiede zwischen den beiden Optimierungsergebnissen sind auf das Abbruchkriterium des Optimierers zurückzuführen. Die Rechenzeiten der einzelnen Optimierungsansätze sind in Tabelle 7.6 aufgelistet. Der deterministische und der stochastische Optimierungsansatz benötigen 15.87s bzw. 9.26s zur Lösung des Optimierungsproblems, während die Brute-Force-Variante mit 1691.76s deutlich rechenintensiver ist. Der Rechenaufwand für eine einzelne Vorhersage eines klassischen GP-Prädiktors kann für den Mittelwert des Prädiktors mit $\mathcal{O}(N)$ und für die Varianz des Prädiktors mit $\mathcal{O}(N^2)$ angegeben werden [35]. Dabei steht N für die Anzahl der Trainingsdaten. Dementsprechend können die Vorhersagekosten im deterministischen Fall mit $\mathcal{O}(N_{obs})$, im stochastischen Fall mit $\mathcal{O}(N_{grid})$ und für den Brute-Force-Ansatz mit $\mathcal{O}(N_{obs}N_{samp,BF})$ angegeben werden, wobei $N_{obs} = 675$, $N_{grid} = 500$ und $N_{samp,BF} = 500$ gelten. Unter der Annahme, dass alle drei Optimierungsprobleme die gleiche Anzahl von Iterationen zur Lösung des Optimierungsproblems benötigen, kann der theoretische Rechenaufwand als Produkt aus der Anzahl der Prädiktionen und den Kosten für eine einzelne Prädiktion approximiert werden. Der so berechnete theoretische Rechenaufwand für den deterministischen und den stochastischen Fall stimmt mit den in Tabelle 7.6 gemessenen Werten überein. Wir vergleichen auch den Aufwand des stochastischen Ansatzes mit dem der Brute-Force-Methode, die das gleiche Problem numerisch löst. Der theoretische Aufwand des Brute-Force-Ansatzes ist um den Faktor 675 höher als der des stochastischen Ansatzes. Die Messung zeigt jedoch einen Mehraufwand, der nur um den Faktor 183 größer ist. Der Unterschied zwischen theoretischem und gemessenem Rechenaufwand ist auf den Einfluss der Modellladezeiten zurückzuführen. Das Modell wird aufgrund der Implementierung bei jeder Prognose neu aus dem Speicher geladen. Beim stochastischen Ansatz ist die Rechenzeit für das vorliegende Optimierungsproblem sehr gering, so dass die Ladezeit einen großen Anteil am Gesamtaufwand hat. Bei der Brute-Force-Optimierung ist die Ladezeit dagegen kaum von Bedeutung, da sie nur einen kleinen Bruchteil der Gesamtdauer ausmacht. Bei komplexeren Optimierungsproblemen dürfte der Einfluss deutlich geringer sein und der Rechenzeitgewinn näher am theoretischen Zeitge-

winn liegen. Der Gesamtzeitgewinn für das vorliegende Problem mag gering erscheinen. Es ist jedoch zu beachten, dass hier nur ein Betriebspunkt optimiert wurde. In der Praxis werden in der Regel komplexere Problemstellungen gelöst, bei denen mehrere Betriebspunkte gleichzeitig optimiert werden. Zudem sind die Trainingsdatensätze in der Regel größer und es müssen deutlich mehr Randbedingungen erfüllt werden. Gleichzeitig muss die Kostenfunktion für eine große Anzahl von Fahrttrajektorien minimiert werden, was zu einer sehr großen Anzahl von Funktionsauswertungen führt [64]. Die Lösung solcher Optimierungsprobleme kann mehrere Stunden dauern. Vor diesem Hintergrund wird deutlich, dass der Brute-Force-Ansatz für den praktischen Einsatz keine sinnvolle Option darstellt. Zumindest nicht ohne die Notwendigkeit des sogenannten Parallel Computing, also der Parallelisierung der einzelnen Rechenoperationen auf mehreren CPUs oder GPUs. Darüber hinaus kommen wir zu dem Schluss, dass die entwickelte Methodik, ebenso wie der Brute-Force-Ansatz, das globale Optimum für das formulierte Optimierungsproblem findet und gleichzeitig den geringsten Rechenaufwand der verglichenen Verfahren aufweist.

7.1.4. Zusammenfassung

In diesem Abschnitt wurde die entwickelte SEMO-Methode zur Optimierung des in Abschnitt 3.1.1 vorgestellten Reihen-Vierzylinder-Dieselmotors verwendet. Zunächst wurde ein DoE erstellt und nominale Modelle sowie Quantilmodelle des Motors für die Rohemissionen von NO_x und Ruß sowie für das Verbrennungsgeräusch trainiert. Unter Hinzunahme der jeweiligen Modelle wurden ein deterministisches und ein stochastisches Optimierungsszenario definiert—im Falle der deterministischen Optimierung wurden Sicherheitsfaktoren angewendet. Anschließend wurden die Ergebnisse beider Szenarien miteinander verglichen. Für den deterministischen Fall wurde deutlich, dass die empirisch ermittelten Grenzwerte für den Ruß und das Verbrennungsgeräusch zu streng gewählt waren. Für den stochastischen Fall wurden die Grenzwerte nahezu perfekt eingehalten. Außerdem war hier durch die gezieltere Optimierung eine deutlich höhere Reduktion der NO_x -Rohemissionen möglich. Im nächsten Schritt wurden die beiden Optimierungsszenarien mit einem dritten Szenario verglichen, in dem das stochastische Problem numerisch gelöst wurde. Dabei wurde neben dem Emissionsergebnis auch der Zeitaufwand für den Optimierungsprozess betrachtet. Das Emissionsergebnis dieser sogenannten Brute-Force-Variante war vergleichbar mit dem Ergebnis des stochastischen Optimierungsszenarios. Der Rechenaufwand für die stochastische Optimierung war mit 9.26s am geringsten, während die Brute-Force-Variante mit 1691.76s deutlich rechenintensiver war.

Tabelle 7.6.: Rechenaufwand für die deterministische und stochastische Optimierung sowie für den Brute-Force-Ansatz. Angegeben ist die Rechenzeit in Sekunden.

	Rechendauer
Deterministisch	15.87s
Unser Ansatz	9.26s
Brute-Force	1691.76s

7.2. Anwendungsbeispiel V6-Dieselmotor

In Abschnitt 7.1 wurde die in dieser Arbeit entwickelte Methodik an einem Anwendungsbeispiel demonstriert. Dabei wurde die Methode erfolgreich zur Optimierung eines Reihen-Vierzylinder-Dieselmotors in einem festen Betriebspunkt eingesetzt und mit einer analogen deterministischen Optimierung verglichen. Weiterhin wurde der Rechenaufwand des Verfahrens mit dem Aufwand der deterministischen Lösung und dem Rechenaufwand einer rein numerischen Lösung des stochastischen Problems verglichen. Dieser Abschnitt unterscheidet sich inhaltlich insofern vom vorhergehenden, als die Anwendung des Verfahrens auf die Optimierung mehrerer Betriebspunkte bzw. Kennfelder erweitert wird. Hierzu wird ein Modell trainiert, das über einen sehr weiten Betriebspunktbereich des Motors gültig ist und daher als globales Motormodell bezeichnet wird. Darüber hinaus wird in Abschnitt 7.2.1 ein aufwendiger Validierungsansatz verfolgt, bei dem die trainierten Nominal- und Quantilmodelle anhand von transienten Prüfstandsmessungen validiert werden. Dazu werden Änderungen am Motoraufbau vorgenommen, um unterschiedliche Toleranzniveaus abzubilden, und im Rahmen von transienten Messzyklen vermessen. Der Vergleich zwischen Modell und Messung soll ebenfalls zur Validierung des entwickelten Ansatzes dienen. Die Untersuchungen werden an einem V6-Dieselmotor durchgeführt. Dieser stellt ein appliziertes Serienfahrzeug dar und wurde bereits in Kapitel 3 vorgestellt. Anschließend wird in Abschnitt 7.2.2 die Parametrierung des Motors ausgehend von der Serienapplikation des Motors optimiert. Ziel der Optimierung ist die Minimierung der Ruß-Rohemissionen zur Reduzierung der Partikelfilterbeladung am Beispiel des WLTC. Dies geschieht unter Einhaltung eines vorgegebenen Stickoxidgrenzwertes für die gesamte Fahrzeugflotte. Die Ergebnisse der Optimierung werden anschließend mit der Serienapplikation verglichen, die auf die Einhaltung der Euro-6d-Norm ausgelegt ist.

Teile der vorgestellten Ergebnisse wurden bereits in *Automotive and Engine Technology* durch Springer Nature veröffentlicht:

- **Mourat, Kadir**; Eckstein, Carola; Koch, Thomas: Application of stochastic design optimization to a passenger car diesel engine to reduce emission spread in a vehicle fleet. In: *Automotive and Engine Technology* 6 (2021), Nr. 1, S. 99–112.

Beiträge der Autoren:

K.M., C.E. und T.K. haben den Inhalt des Artikels konzipiert. **K.M.** konzipierte und führte die Experimente durch. **K.M.** führte die Nachbearbeitung der Daten durch. **K.M.** führte die Simulationen und die Optimierungen durch. **K.M.** schrieb den Computercode. **K.M.** erstellte die Grafiken. **K.M.** schrieb den Artikel mit Beiträgen von C.E. und T.K.

Zu den veröffentlichten Ergebnissen gehören die Abbildungen 7.3–7.11 sowie die Tabellen 7.7–7.9. Diese wurden mit Genehmigung von Springer Nature wiederverwertet.

7.2.1. Validierung des stochastischen Modellierungsansatzes

Dieser Abschnitt befasst sich mit der Validierung des stochastischen Modells am Versuchsträger. In Kapitel 2 und in Abschnitt 7.1.1 dieses Kapitels wurde die Validierung von Modellen bereits behandelt. Das deterministische Modell wurde durch Kreuzvalidierung mit

den Trainingsdaten validiert und das stochastische Modell wurde mit dem deterministischen Modell validiert. Dieser Ansatz ist in sich konsistent, sofern die Annahmen bezüglich der Toleranzbeiträge korrekt sind. In diesem Abschnitt wollen wir an dieser Stelle anknüpfen und diese Annahmen überprüfen. Dazu wird ein Ansatz verfolgt, bei dem drei Motor-konfigurationen mit unterschiedlichen Toleranzniveaus im transienten Betrieb vermessen werden. Für jede Konfiguration wird ein eigenes Modell trainiert und die Messung mit der Modellvorhersage verglichen. Die Trainingsdaten für das Modell stammen aus einem zuvor durchgeführten stationären DoE. Der Grund für die Wahl von transienten Tests anstelle von stationären Einzelpunkten liegt darin, dass zum einen deutlich mehr Messpunkte mit weniger Aufwand gemessen werden können. Zum anderen kann gleichzeitig der Einfluss transients Phänomene untersucht werden. Bei dem Versuchsfahrzeug handelt es sich um ein Serienfahrzeug mit einem V6-Dieselmotor, dessen Versuchsaufbau bereits im Kapitel 3 vorgestellt wurde.

DoE-Kampagne

In diesem Abschnitt werden mehrere Modelle betrachtet, die in drei Kategorien unterteilt werden können: Nominalmodell, Worst-Case-Modell (WC-Modell) und Best-Case-Modell (BC-Modell). Ein Nominalmodell modelliert dabei die Emissionen eines nominellen Systems. Ein WC-Modell modelliert ein System, dessen Emissionen gerade besser sind als die schlechtesten 5% der Fahrzeugflotte. Und ein BC-Modell modelliert wiederum ein System, dessen Emissionen gerade noch schlechter sind als die besten 5% der Fahrzeugflotte.

Tabelle 7.7.: Liste der, durch das ODCM-Verfahren, überwachten Grenzwerte [84].

Überwachte Größe	oberer Grenzwert
Temperatur v. Turbine (°C)	840
Zylinderdruck (kPa)	21
Opazität (%)	60
Unverbrannte Kohlenwasserstoffe (UHC) (ppm)	3000
LNT Temperatur (°C)	700

Grundlage für die in diesem Abschnitt vorgestellten Modelle ist ein DoE, der mit Hilfe einer automatisierten Messkampagne auf einem stationären Motorprüfstand vermessen wird. Zur Vermessung des DoE und zur Vermeidung kritischer Grenzwertüberschreitungen wird der vorgestellte modifizierte Online DoE with Constraint Modelling (ODCM)-Ansatz gewählt. In Tabelle 7.7 sind die wichtigsten Grenzwerte aufgeführt. Die Betriebsraumgrenzen des DoE-Versuchsplans sind in Tabelle 7.8 dargestellt. Neben der Motordrehzahl und dem Motordrehmoment bzw. der eingespritzte Gesamtkraftstoffmasse werden zwölf weitere Steuergeräteparameter variiert. Diese sind der Ladedruck, die Abgasrückführrate, die Niederdruckanteil der Abgasrückführrate ($\psi_{AGR,ND}$), die Drallklappenstellung, die Commonraildruck, die Menge der ersten Voreinspritzung, die Menge der zweiten Voreinspritzung, die Menge der Nacheinspritzung, die Ansteuerbeginn der Haupteinspritzung und die

Einspritzintervalle. Es ist zu beachten, dass die Motordrehzahl und der eingespritzte Gesamtkraftstoffmasse (m_{Ges}) kontinuierlich variiert wurden. Das Einspritzmuster entspricht im Allgemeinen dem in Abbildung 6.2. Eine Ausnahme bilden Parameterkombinationen, bei denen eine der Einspritzmengen Null ist. Es wurde darauf geachtet, dass die Messpunkte im Versuchsplan einer Sobol-Sequenz folgen und somit raumfüllend sind. Die minimale Einspritzmenge wurde auf einen Wert von 0.8mg/Hub festgelegt. Werte, die aufgrund der raumfüllenden Verteilung unterhalb dieser Grenze liegen, werden auf Null gesetzt. In gleicher Weise wurde mit den Einspritzintervallen verfahren. Der Mindestabstand zwischen zwei Injektionen beträgt hier 200 μ s. Diese Untergrenze entspricht der mit den Injektoren realisierbaren Minimalmenge. Der Motor wird mit einer AGR-Ratenregelung betrieben. Die Luftmasse und das Kraftstoff-Luft-Verhältnis werden über die vorgegebene Soll-AGR-Rate und den Soll-Ladedruck eingestellt. Zur Bestimmung des initialen Variationsbereichs wurden Erfahrungswerte herangezogen. Der Variationsbereich ist so gewählt, dass das Motorverhalten in einem möglichst großen Parameterraum möglichst genau abgebildet wird. Lediglich die Einspritzabstände wurden begrenzt, um eine Ölverdünnung zu vermeiden. Untersuchungen zeigen, dass sehr frühe Einspritzzeitpunkte in Kombination mit niedrigen Gastemperaturen unter 800K zu Kraftstoffwandauftrag und Ölverdünnung führen können [48]. Aufgezeichnet wurden die FSN und die NO_x-Rohemissionen.

Tabelle 7.8.: Betriebsraumgrenzen des DoE-Versuchsplans für die Untersuchungen am Sechszylinder-Motor [84].

Variationsparameter	Untere Grenze	Obere Grenze
Motordrehzahl (n_{Mot}) (1/min)	750	3000
m_{Ges} (mg/Hub)	0	50
p_2 (kPa)	100	290
X_{AGR} (%)	0	60
$\psi_{\text{AGR,ND}}$ (%)	0	100
ψ_{SW} (-)	0.40	1.00
p_{CR} (MPa)	30	240
m_{PE1} (mg/Hub)	0	4.5
m_{PE2} (mg/Hub)	0	3
m_{NE} (mg/Hub)	0	6
ϕ_{HE} (°KW n. OT)	-7	12
t_{VE1} (μ s)	0	1500
t_{VE2} (μ s)	0	1500
t_{NE} (μ s)	0	1600

Nominales Motormodell

In diesem Abschnitt werden zunächst zwei globale Motormodelle, ein NO_x- und ein Ruß-Rohemissionsmodell, trainiert. Anschließend wird gezeigt, dass die Vorhersagen der trainier-

ten Modelle unter bestimmten Bedingungen auf dynamische Messzyklen übertragbar sind. Sowohl die gemessenen Emissionen als auch die über einen bestimmten Zeitraum kumulierten Emissionen können sehr gut vorhergesagt werden. Darüber hinaus dienen die Modelle als Grundlage für die im folgenden Abschnitt vorgestellten stochastischen Modelle.

Der Trainingsdatensatz stammt aus dem in Abschnitt 7.2.1 beschriebenen DoE und umfasst ca. 1500 Trainingsdatenpunkte. Als Modell werden GPs verwendet. Der Fehler am Modellausgang wird als additiver Gaußscher Fehler angenommen. Als Kernelfunktion wird für beide Modelle ein SE-Kernel mit ARD verwendet. Zusätzlich wird eine Box-Cox-Transformation zur Stabilisierung der Varianz der Modellausgabe verwendet. Die Optimierung der Hyperparameter des GPs erfolgt durch Maximierung der Likelihood-Funktion. Auch bei diesem Motor wird davon ausgegangen, dass es sich bei dem untersuchten Versuchsträger um ein Exemplar mit nominellen Komponenten handelt. Die relevanten Motorkomponenten wurden beim Motoraufbau entsprechend ausgewählt.

In Tabelle 7.9 sind die Evaluationsergebnisse der beiden Modelle dargestellt. Die Evaluierung von NRMSE und R^2 erfolgt sowohl für die Trainingsdaten als auch mit der LOOCV-Methode. Der Modellfehler für das NO_x -Rohemissionsmodell ist sowohl für die Trainingsdaten als auch für die LOOCV-Methode extrem klein. Das Modell liefert somit sehr genaue Vorhersagen sowohl für bereits beobachtete als auch für noch unbekannte Parameterkombinationen. Eine vergleichsweise geringere, aber immer noch sehr hohe Modellgüte weist das Ruß-Rohemissionsmodell auf. Das R^2 über den Trainingsdaten liegt bei ca. 0.9 und das R^2 für das LOOCV-Verfahren bei ca. 0.8. Dies ist darauf zurückzuführen, dass die aleatorische Unsicherheit¹ des Messapparates [5] vergleichsweise hoch ist.

Tabelle 7.9.: NRMSE und R^2 der Trainingsdaten sowie NRMSE und R^2 mit der LOOCV-Methode für das nominale Ruß-Rohemissionsmodells (V6-Dieselmotor) [84].

	Training		LOOCV	
	NRMSE	R^2	NRMSE	R^2
NO_x (g/h)	-	0.996	-	0.991
Partikelmasse (g/h)	-	0.896	-	0.796

Zusätzlich zu den stationären Daten werden die Modelle im Folgenden auch mit transienten Daten evaluiert. Dazu werden transiente Messungen am Motorprüfstand durchgeführt. Es wird ein WLTC-Fahrprofil gemessen. Zur Generierung der Drehzahl-Lastspuren am Motorprüfstand wird angenommen, dass es sich bei dem Fahrzeug um einen Mittelklasse-Sport-Utility-Vehicle (SUV) mit einem Gewicht von ca. 2000kg handelt. Zur Messung der Rohemissionen wird das Smoke-Meter durch einen Micro-Soot-Sensor der Firma AVL ersetzt, der eine kontinuierliche Emissionsmessung ermöglicht. Die Probenahmestelle bleibt unverändert.

¹In der Literatur wird Unsicherheit üblicherweise in aleatorische und epistemische Unsicherheit unterteilt [75, 83, 89, 98]. Aleatorische Unsicherheit ist bis zu einem gewissen Grad nicht reduzierbar und hängt mit der zufälligen Variabilität innerhalb eines Systems oder seiner Umgebung zusammen. Epistemische Unsicherheit ist reduzierbar und systemimmanent. Sie ist mit einem gewissen Mangel an Wissen über das System verbunden.

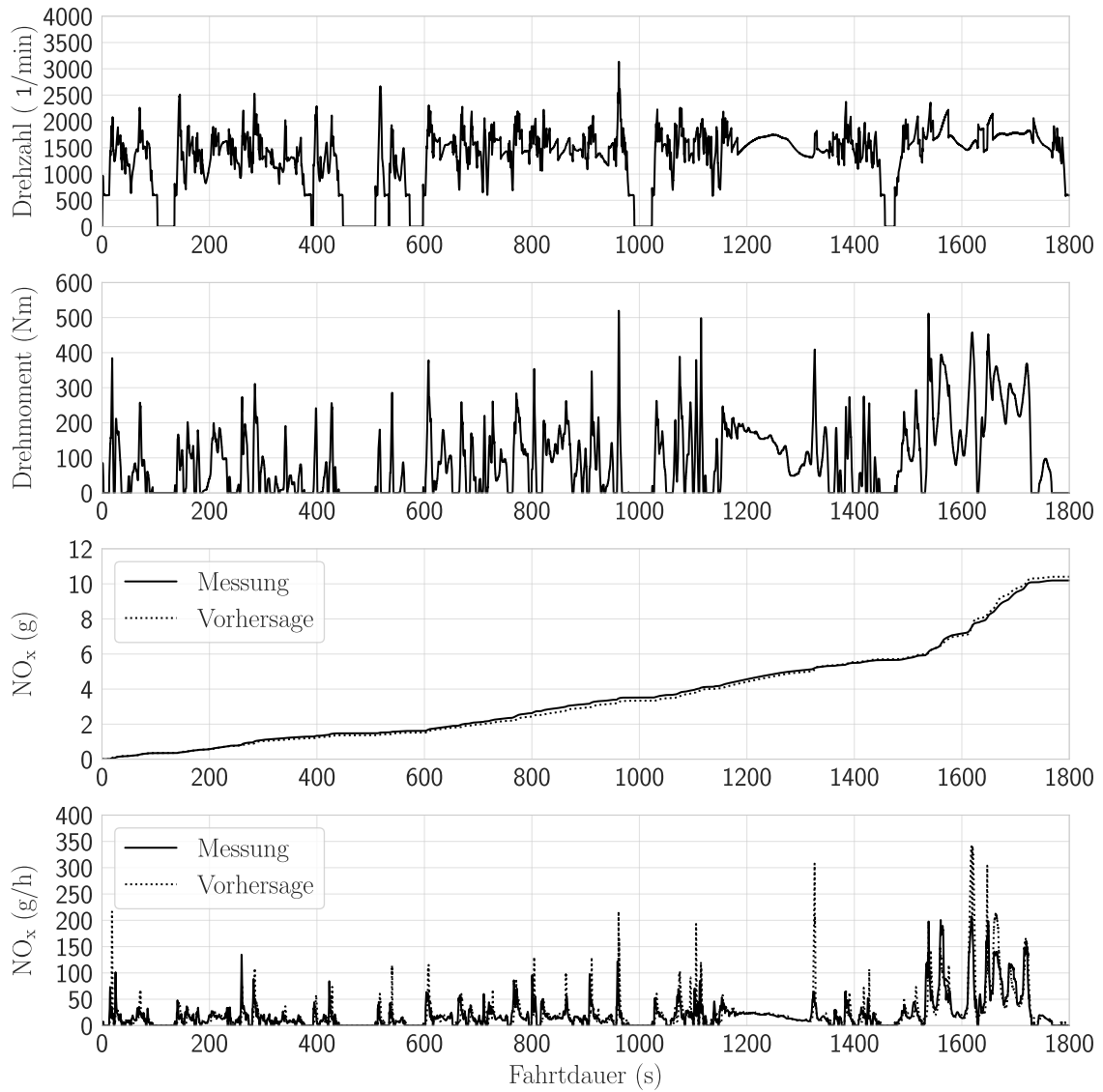


Abbildung 7.3.: Motordrehzahl, Motordrehmoment sowie kumulierte und absolute NO_x -Rohemissionen im WLTC (Messung und Vorhersage) [84].

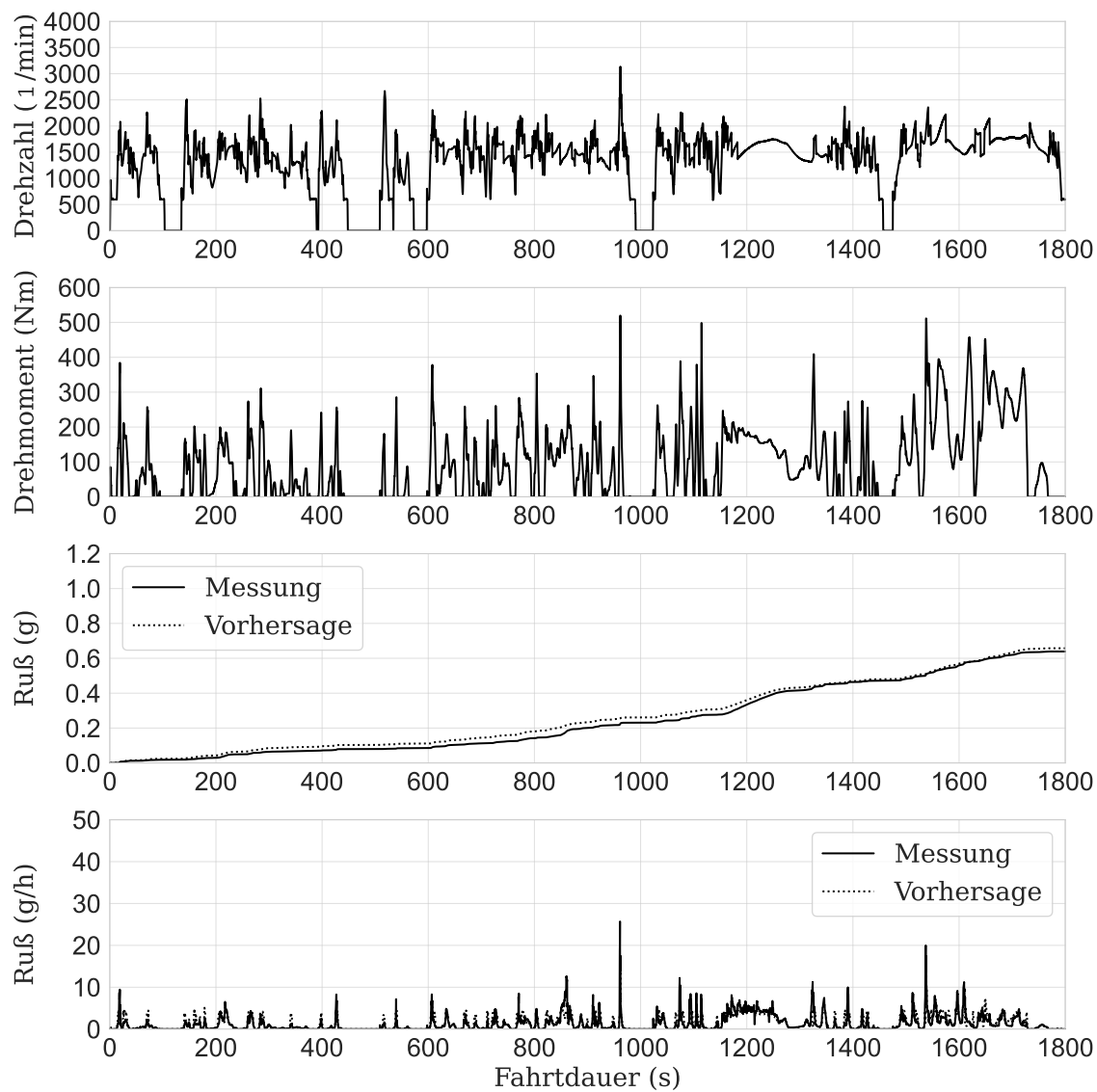


Abbildung 7.4.: Motordrehzahl, Motordrehmoment, sowie kumulierte und absolute Ruß-Rohemissionen im WLTC (Messung und Vorhersage) [84].

Abbildung 7.3 zeigt das Drehzahlprofil, das Drehmomentprofil, die gemessenen und vorhergesagten NO_x -Rohemissionen sowie die gemessenen und vorhergesagten kumulierten NO_x -Rohemissionen für den WLTC. Die Abbildung zeigt, dass die gemessenen kumulierten NO_x -Rohemissionen durch das Modell sehr genau abgebildet werden. Der prozentuale Fehler zwischen Messung und Vorhersage beträgt etwa -2.1% . Es zeigt sich auch, dass die Diskrepanz zwischen gemessenen und vorhergesagten Emissionen bis auf wenige Ausnahmen sehr gering ist. Größere Diskrepanzen sind nur in Gebieten mit hoher Beschleunigung zu beobachten. Dies ist jedoch zu erwarten, da es sich um ein stationäres Modell handelt, das die Dynamik im Luftpfad nicht abbildet. Es ist jedoch anzumerken, dass die Diskrepanz zwischen Messung und Modellvorhersage sehr gering ist und für die Parameteroptimierung ausreichend gut ist. Darüber hinaus ist auch von einem Modell, das die Dynamik im Luftpfad abbilden kann, keine wesentliche Verbesserung der Genauigkeit zu erwarten. Die geringe Abweichung ist auch darauf zurückzuführen, dass der Motor über einen schnell ansprechenden Turbolader verfügt und die Steuergerätefunktionen für den Luftpfad sinnvoll parametrisiert sind. Wir gehen davon aus, dass die Simulationsergebnisse schlechter ausfallen können, wenn diese Voraussetzungen nicht erfüllt sind, da die Diskrepanz zwischen Ist- und Sollzustand im Luftpfad größer ist. Insgesamt spielt der Modellfehler für das NO_x -Rohemissionsmodell keine große Rolle, da der allgemeine Trend der Emissionsentwicklung sehr gut wiedergegeben wird. Die Rußemissionen sind in Abbildung 7.4 dargestellt. Hier sind wieder die absoluten und kumulierten Rohemissionen der Messung und der Vorhersagekosten für den WLTC dargestellt. Der prozentuale Fehler für die Ruß-Rohemissionen ist mit ca. -2.9% ebenfalls sehr gering. Allerdings ist die Abweichung zwischen Messung und Vorhersage im kumulierten Verlauf im Zeitintervall zwischen 130s und 1320s nach Start leicht erhöht. Der Grund für die erhöhten Emissionen liegt darin, dass das Modell in diesem Bereich mehrere Ruß-Peaks leicht überschätzt, die sich in Summe auf das kumulierte Ergebnis auswirken. Eine Verringerung der Abweichung erfolgt insbesondere durch eine Unterschätzung der Emissionen im Bereich um 850s nach dem Start und im Bereich zwischen 1180s und 1250s. Darüber hinaus ist eine leichte Zunahme der Diskrepanz zwischen Messung und Vorhersage im letzten Abschnitt ab ca. 1550s zu beobachten. Die im Vergleich zum NO_x -Modell erhöhte Diskrepanz kann zum einen auf die geringere Modellgüte des Ruß-Modells zurückgeführt werden, siehe Tabelle 7.9. Zum anderen spielen beim Ruß auch An- und Ablagerungseffekte eine Rolle, die das Ergebnis beeinflussen können. Dabei können sich Rußpartikel bei niedrigen Temperaturen im Bereich des Abgassystems anlagern und bei höheren Temperaturen wieder ablagern. Daher muss die Modellvorhersage nicht unbedingt falsch sein, solange die kumulierten Emissionen in etwa übereinstimmen. Bei näherer Betrachtung der Abbildung 7.4 scheint dies zutreffend zu sein, da das Modell insbesondere zu Beginn bis zu einer Fahrzeit von ca. 850s zu hohe Emissionen vorhersagt. Dies deutet auf eine Ablagerung von Rußpartikeln hin. Auch die Abgastemperatur vor der Turbine ist in diesem Zeitintervall vergleichsweise niedrig, was die Annahme bekräftigt, siehe Abbildung 7.5. Luftsystemseitige dynamische Effekte scheinen im Ruß-Modell eine eher untergeordnete Rolle zu spielen. Zumindest sind keine signifikanten Indikatoren erkennbar, die dies bestätigen würden. Insgesamt ist jedoch festzustellen, dass das Ruß-Modell trotz der genannten Einflüsse gute Generalisierungseigenschaften aufweist. Der generelle Trend der produzierten Ruß-Emissionen wird gut erfasst und die Vorhersage für die kumulierten Emissionen stimmt sehr gut mit den Messungen überein.

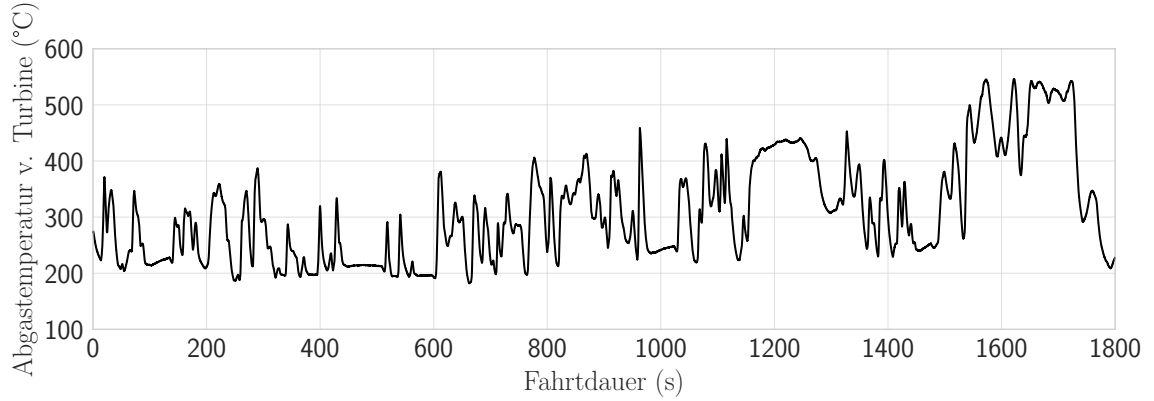


Abbildung 7.5.: Temperatur vor Turbolader im WLTC [84].

Stochastisches Motormodell

In diesem Abschnitt wird der im Rahmen der Arbeit entwickelte stochastische Modellierungsansatz validiert. Dazu werden die Rohemissionen von NO_x und Ruß im WLTC am Motorprüfstand für zwei unterschiedliche Motorkonfigurationen gemessen. Die Konfigurationen bestehen aus zwei Motoren mit grenzwertigen HFMs, die jeweils eine Toleranz von $\pm 2\sigma$ aufweisen. Die übrigen Motorkomponenten entsprechen der im vorherigen Abschnitt beschriebenen nominalen Motorkonfiguration. Zusätzlich werden beide Konfigurationen durch ein NO_x - und ein Ruß-Modell repräsentiert. Dazu werden auf Basis der in Abschnitt 7.2.1 trainierten nominalen Modelle jeweils zwei Quantilmodelle trainiert, die die HFM-Toleranz berücksichtigen. Die Modellvorhersage wird anschließend anhand der Messergebnisse validiert.

Im Folgenden wird angenommen, dass der betrachtete Motor nur eine Bauteiltoleranz für den Heißfilm-Luftmassenmesser (HFM) aufweist. Alle anderen Bauteile sind nominal. Diese Annahme wird getroffen, um die HFM-Toleranz separat zu betrachten. Die Sensortoleranz des HFM kann nach Herstellerangaben durch eine Normalverteilung mit einem Erwartungswert von Null und einer relativen Standardabweichung von 1.2% beschrieben werden. Die relative Standardabweichung bezieht sich auf den Luftmassenstrom des HFM. Wegen der Regelung der AGR-Rate des Motors im Normalbetrieb ist die Ansaugluftmasse keine Modelleingangsgröße. Sie ist vielmehr ein Produkt der anderen Eingangsparameter und beeinflusst diese. Die HFM-Toleranz beeinflusst hauptsächlich den Ladedruck, die AGR-Rate und den Massenanteil der Niederdruck-AGR-Rate. Im Vergleich zur AGR-Rate ist die Empfindlichkeit der anderen Parameter in Bezug auf PM und NO_x eher gering. Die Eingangsparameter-toleranz für diese beiden Parameter, die sich aus der HFM-Toleranz ergibt, ist ebenfalls gering. Daher ist die dominierende Parameterunsicherheit die AGR-Rate und die beiden anderen Parameter sind vernachlässigbar. Im Folgenden wird die Abweichung der AGR-Rate hergeleitet. Sie ist gegeben durch

$$X_{\text{EGR}} = 1 - \frac{m_{\text{IAM}}}{m_{\text{CC}}}, \quad (7.5)$$

wobei m_{IAM} die Ansaugluftmasse und m_{CC} die Zylinderfüllung ist. Die Abweichung der AGR-Rate kann als eine Taylorreihe erster Ordnung notiert werden

$$\Delta X_{EGR} = \Delta m_{IAM} \cdot \frac{\delta X_{EGR}}{\delta m_{IAM}} + \Delta m_{CC} \cdot \frac{\delta X_{EGR}}{\delta m_{CC}}, \quad (7.6)$$

wobei $\frac{\delta f}{\delta x}$ die partielle Ableitung von f nach x ist. Durch Lösen dieser Gleichung ergibt sich

$$\Delta X_{EGR} = (1 - X_{EGR}) \cdot \left(\frac{\Delta m_{CC}}{m_{CC}} - \frac{\Delta m_{IAM}}{m_{IAM}} \right). \quad (7.7)$$

Unter der Prämisse eines überwiegend ladedruckgeregelten Motorbetriebs kann für einen Betriebspunkt mit vorgegebenem Ladedruck davon ausgegangen werden, dass die Zylinderfüllung nahezu konstant ist und somit $\Delta m_{CC} = 0$ gilt. Daher kann die absolute Toleranz der AGR-Rate wie folgt definiert werden

$$\Delta X_{EGR} = (X_{EGR} - 1) \cdot \frac{\Delta m_{IAM}}{m_{IAM}}, \quad (7.8)$$

wobei $\frac{\Delta m_{IAM}}{m_{IAM}}$ die relative HFM-Toleranz ist. Wie aus der Gleichung hervorgeht, hängt die Toleranz der AGR-Rate nicht nur von der HFM-Toleranz, sondern auch von der Soll-AGR-Rate ab. Unter Hinzunahme der so ermittelten Toleranz der AGR-Rate, können Quantilmodelle des Motors trainiert werden. Die Stützstellen der stochastischen Modelle werden gleich den Stützstellen der nominalen Modelle gesetzt. Die Unterschreitungswahrscheinlichkeit wird für die WC-Modelle auf $p = 0.977$ und für die BC-Modelle auf $p = 0.023$ gesetzt. Die beiden Wahrscheinlichkeiten entsprechen in etwa $\mu \pm 2\sigma$ in Bezug auf die Normalverteilung. Die Anzahl der Stichproben N_{samp} wird auf 500 gesetzt und die Stichproben werden mittels Latin-Hypercube-Sampling gezogen.

Zur Validierung der Quantilmodelle werden erneut transiente Messungen am Motorprüfstand durchgeführt. Das zuvor generierte WLTC-Fahrprofil wird erneut vermessen. Im Gegensatz zur vorherigen Messkampagne definieren wir nun drei verschiedene Fahrscenarien mit unterschiedlicher Motorkonfiguration:

- **Konfiguration A:** Nominaler Motor - keine Soft- oder Hardwareänderungen,
- **Konfiguration B:** Motor mit $+2\sigma$ -HFM - Der physikalische Luftmassenstrom ist kleiner als der gemessene Luftmassenstrom,
- **Konfiguration C:** Motor mit -2σ -HFM - Der physikalische Luftmassenstrom ist größer als der gemessene Luftmassenstrom,

wobei 1σ einer relativen Luftmassenabweichung von 1.2% entspricht. Die Abweichung des HFM wird softwareseitig realisiert, was zwei große Vorteile hat. Zum einen müssen keine Hardwareänderungen vorgenommen werden, die immer die Gefahr unerwünschter Leckagen in sich bergen. Zum anderen kann die gewünschte HFM-Toleranz über die Software sehr genau eingestellt werden. Zwei exakte $\pm 2\sigma$ -HFMs ausfindig zu machen ist hingegen eher schwer zu realisieren. Unter der Annahme, dass das zugrundeliegende Motorverhalten in unmittelbarer Nähe der Stützstellen monoton ist, bilden das BC- NO_x -Modell und das WC-Ruß-Modell das Motorverhalten mit Konfiguration B ab. Und das WC- NO_x -Modell und das

BC-Ruß-Modell bilden das Verhalten des Motors mit Konfiguration C ab. Konfiguration A entspricht den zuvor trainierten nominellen Modellen.

In Abbildung 7.6 sind für die drei Konfigurationen die Drehzahl, das Drehmoment und die gemessenen und vorhergesagten kumulierten NO_x -Rohemissionen im WLTC dargestellt. Zusätzlich sind in Abbildung 7.7 die gemessenen und vorhergesagten absoluten NO_x -Rohemissionen gegeben. Der prozentuale Fehler zwischen der Messung mit Konfiguration B und der Vorhersage mit dem BC- NO_x -Modell beträgt in etwa 3.5%. Und der prozentuale Fehler zwischen der Messung mit Konfiguration C und der Vorhersage mit dem WC- NO_x -Modell beträgt in etwa -1.1%. Die kumulierten Verläufe verhalten sich ähnlich wie beim nominalen Modell. Es sind lediglich sehr geringe Abweichungen zu erkennen. Diskrepanzen zwischen Messung und Vorhersage sind für alle drei Konfigurationen im gleichen Zeitabschnitt zu beobachten. Dieses Ergebnis ist jedoch zu erwarten, da die stochastischen Modelle ebenso wie das nominale Modell die Dynamik im Luftpfad nicht abbilden und somit die gleichen Probleme bei der Vorhersage aufweisen sollten. Es ist deutlich zu erkennen, dass alle drei Messungen ein unterschiedliches Emissionsniveau aufweisen und eine deutliche Verschiebung zwischen den gemessenen Signalen erkennbar ist. Die unterschiedlichen Emissionsniveaus resultieren aus den unterschiedlichen Toleranzlagen. Die beiden Modelle erfassen dieses Verhalten sehr gut und geben den Trend der Emissionsverläufe sehr gut wieder. Die Ruß-Emissionen für die drei verschiedenen Konfigurationen sind in Abbildung 7.8 und Abbildung 7.9 dargestellt. Der prozentuale Fehler zwischen der Messung mit Konfiguration B und der Vorhersage mit dem WC-Ruß-Modell beträgt in etwa -2.2%. Und der prozentuale Fehler zwischen der Messung mit Konfiguration C und der Vorhersage mit dem BC-Ruß-Modell beträgt in etwa -3.3%. Für alle drei Modelle ist eine Diskrepanz zwischen Messung und Vorhersage im kumulierten Verlauf in dem Zeitintervall von 130s bis 1320s zu erkennen. Die Diskrepanz ist für das WC-Modell am größten und für das BC-Modell am geringsten. Dies kann wiederum durch die Korrelation mit den absoluten Emissionsniveaus der einzelnen Messungen erklärt werden.

Im vorhergehenden Abschnitt wurde die leicht erhöhte Diskrepanz in diesem Zeitintervall auf An- und Ablagerungseffekte von Rußpartikeln zurückgeführt. Diese Annahme wird auch hier getroffen, da in beiden Toleranzlagen das gleiche Verhalten beobachtet werden kann. Lediglich das Ausmaß der An- und Ablagerungen unterscheidet sich von der Messung mit der nominalen Konfiguration, was jedoch auf die Summe der emittierten Emissionen zurückzuführen ist. Zusammenfassend kann festgestellt werden, dass alle drei Modelle sowohl den Emissionsverlauf als auch den allgemeinen Emissionstrend im WLTC sehr gut wiedergeben.

7.2.2. Stochastische Parameteroptimierung am Beispiel des WLTC

In diesem Abschnitt wird die Parametrierung des Motors am Beispiel des WLTC hinsichtlich NO_x und Ruß optimiert. Die Ausgangslage für die Optimierung stellt die Serienapplikation des Motors dar, der zur Einhaltung der Euro-6-Norm ausgelegt ist.

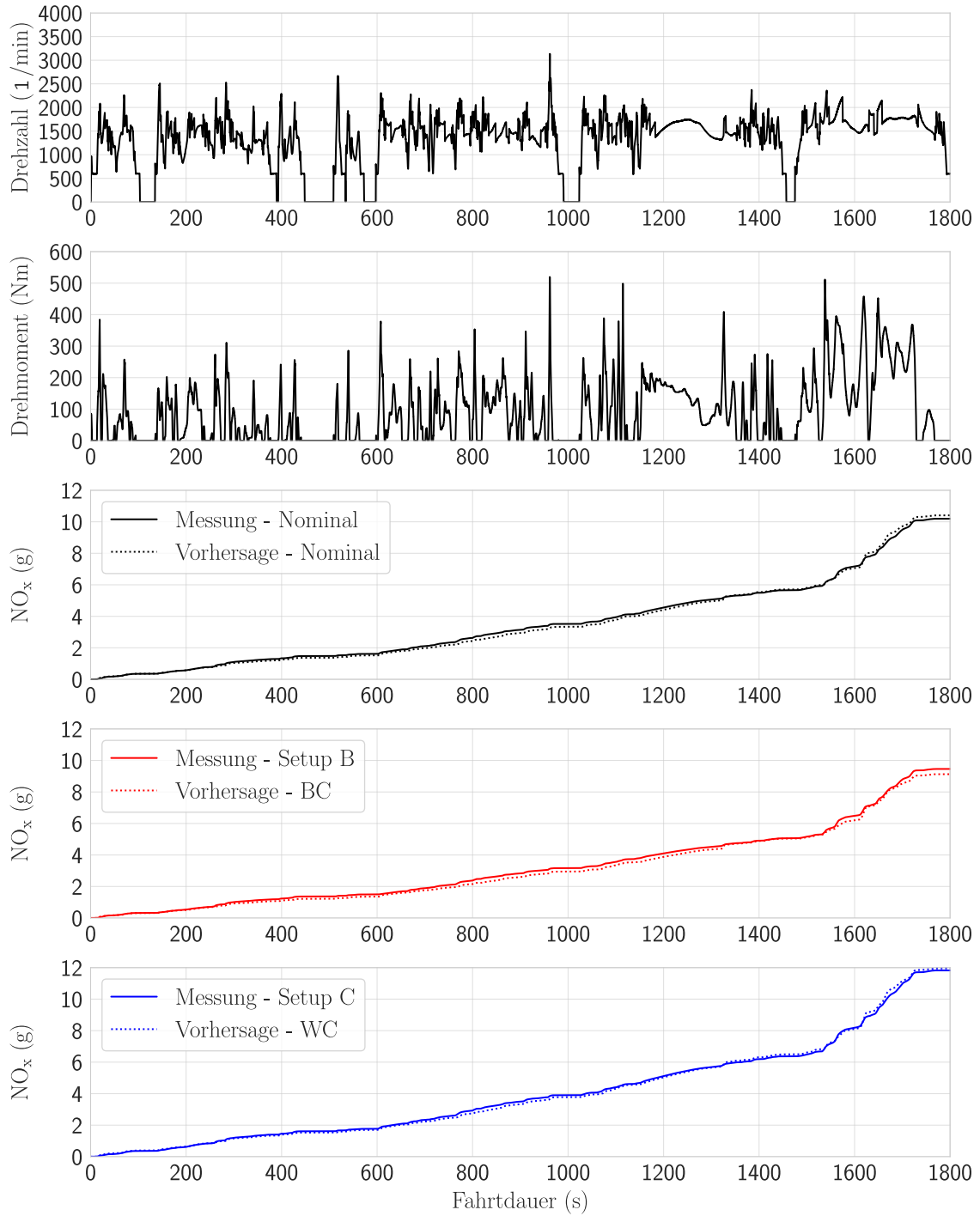


Abbildung 7.6.: Motordrehzahl, Motordrehmoment und kumulierte NO_x-Rohemissionen im WLTC (Messung und Vorhersage) für die verschiedenen Konfigurationen [84].

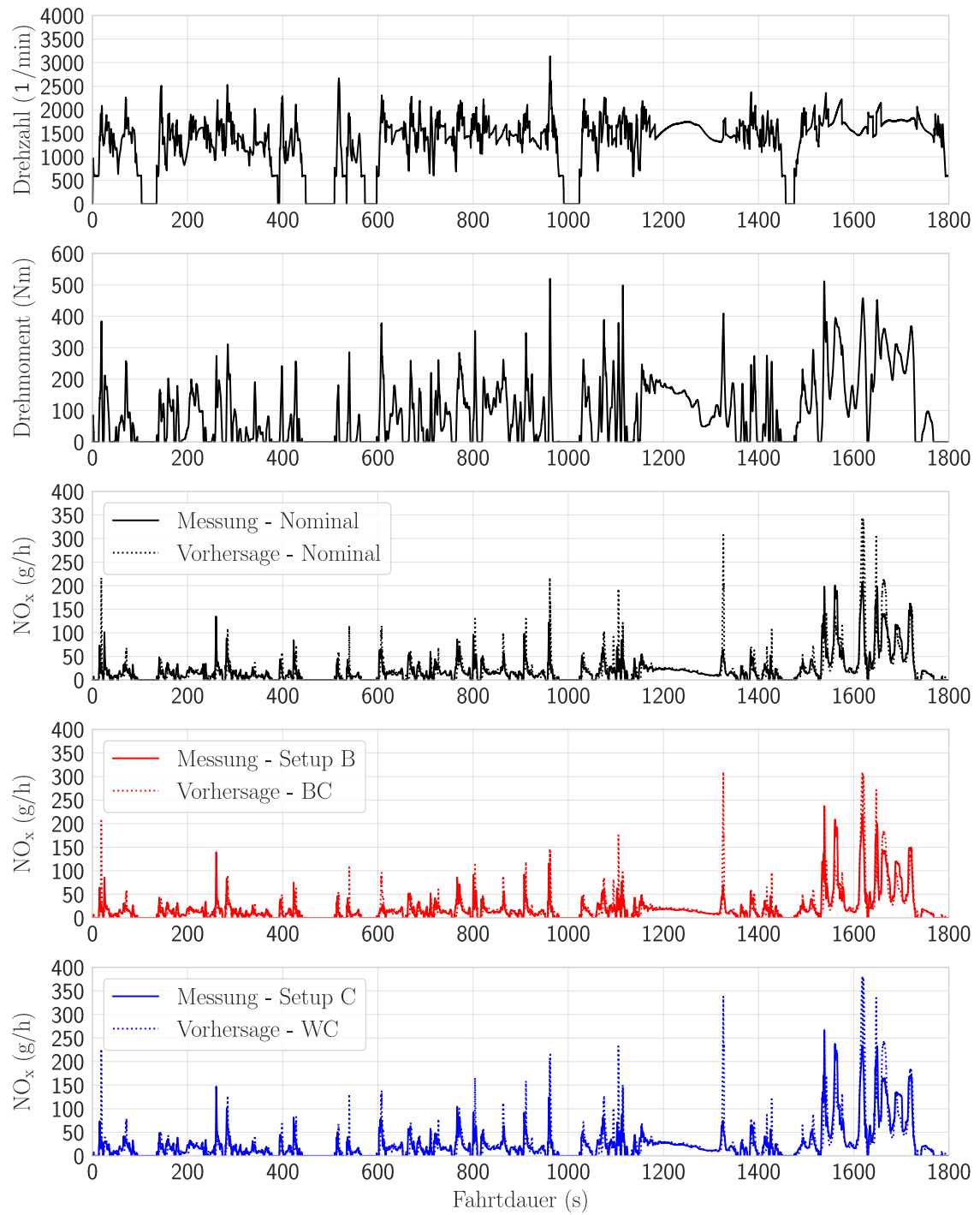


Abbildung 7.7.: Motordrehzahl, Motordrehmoment und absolute NO_x-Rohemissionen im WLTC (Messung und Vorhersage) für die verschiedenen Konfigurationen [84].

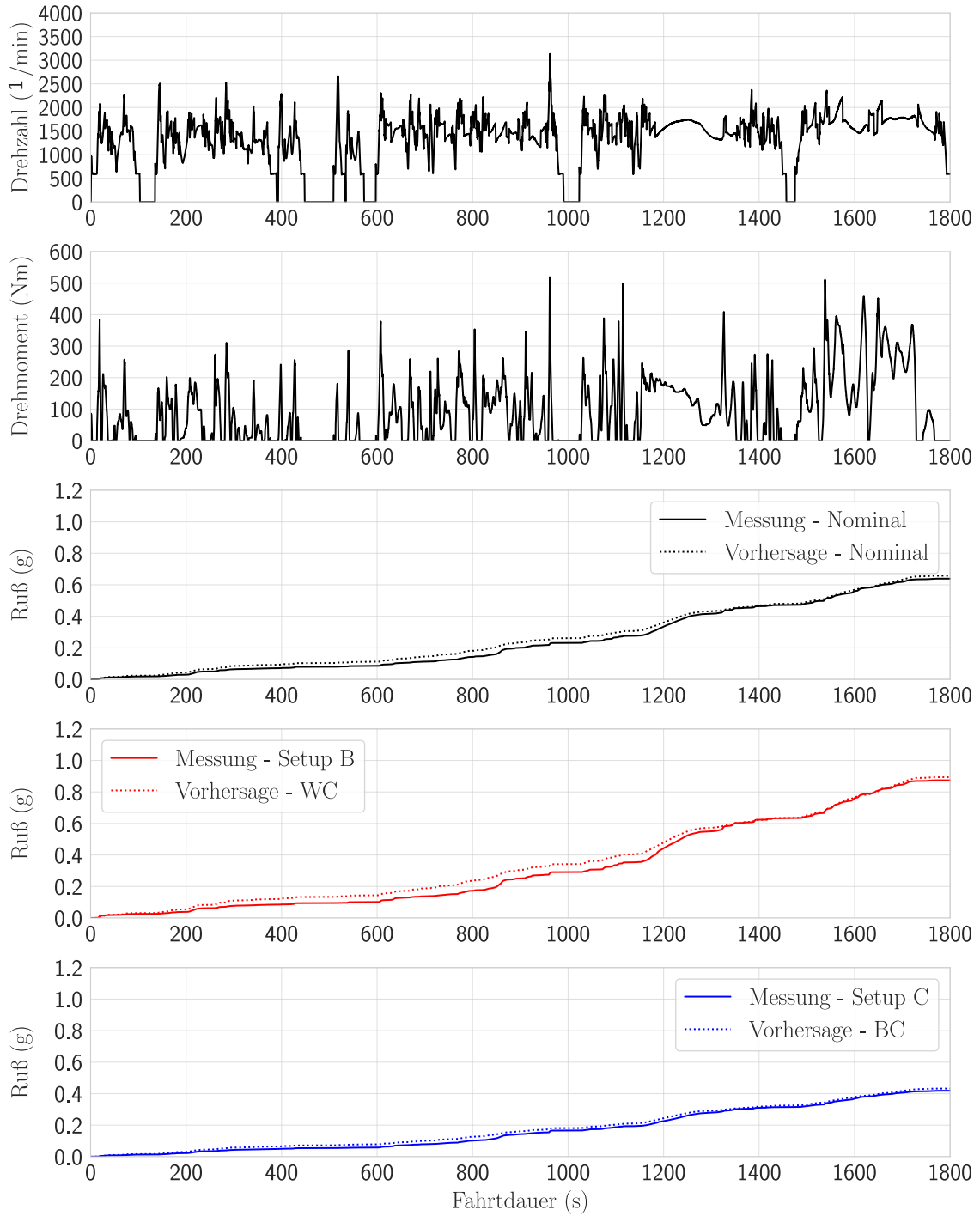


Abbildung 7.8.: Motordrehzahl, Motordrehmoment und kumulierte Ruß-Rohemissionen im WLTC (Messung und Vorhersage) für die verschiedenen Konfigurationen [84].

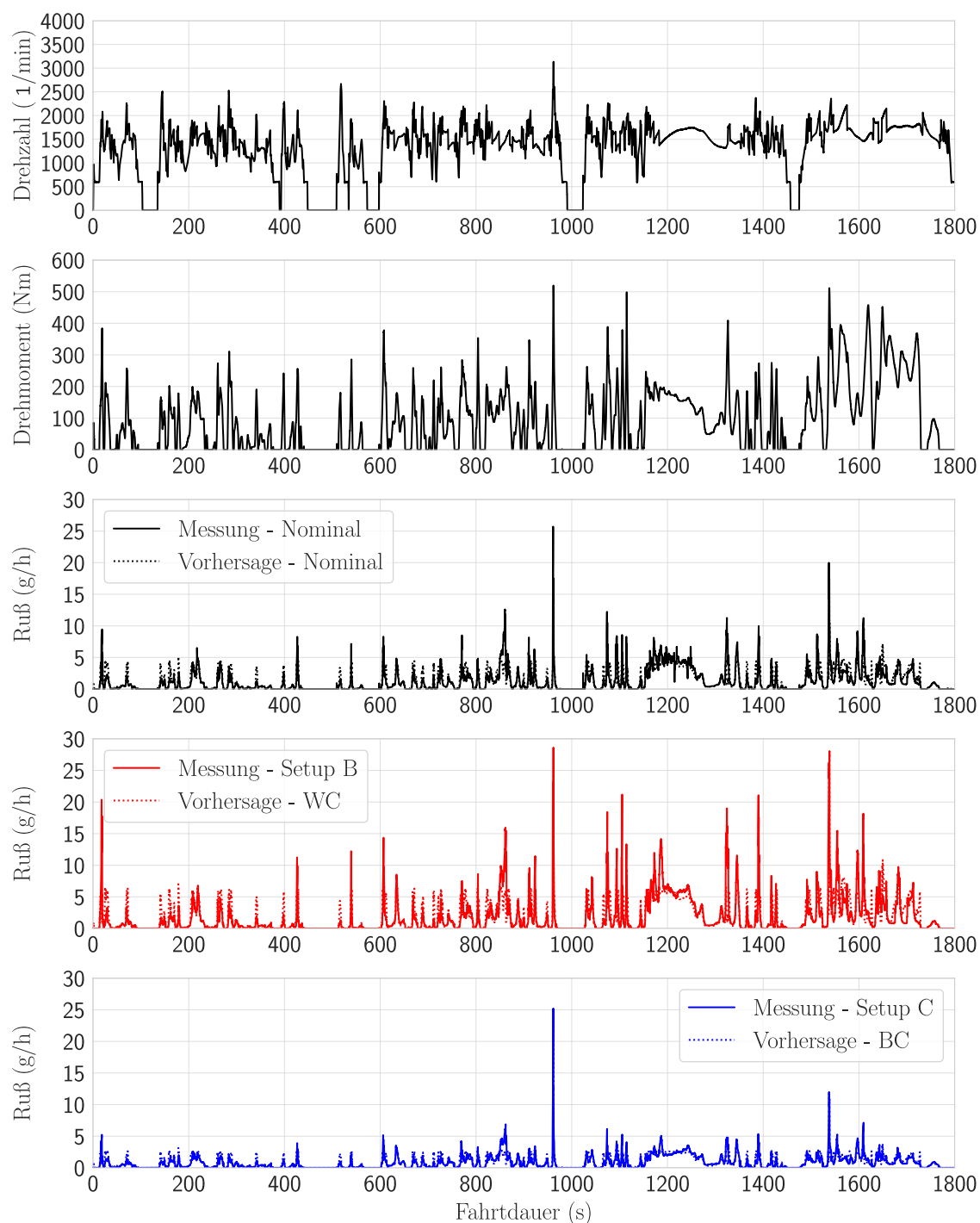


Abbildung 7.9.: Motordrehzahl, Motordrehmoment und absolute Ruß-Rohemissionen im WLTC (Messung und Vorhersage) für die verschiedenen Konfigurationen [84].

Modell

Zur Formulierung des Optimierungsproblems werden, ausgehend von den in Abschnitt 7.2.1 trainierten nominalen Modellen, zunächst stochastische NO_x - und Ruß-Rohemissionsmodelle trainiert. Für die Unterschreitungswahrscheinlichkeit wird in beiden Fällen $p = 0.95$ gewählt. Die Stützstellen der stochastischen Modelle werden erneut den Stützstellen der nominalen Modelle gleichgesetzt. Die Stichprobenanzahl N_{samp} wird auf 1000 gesetzt und die Stichproben werden mittels Latin-Hypercube-Sampling gezogen. Die Verteilung der Eingangsparametertoleranzen des Motors ist in Tabelle 7.10 dargestellt.

Tabelle 7.10.: Parametertoleranzen des V6-Dieselmotors [84].

Parameter	Verteilung	Correlation matrix
m_{Ges}	$f(m_{\text{Ges}}, m_{\text{PE1}}, m_{\text{PE2}}, m_{\text{NE}})$	1 0 0 0 0 0 0 0 1 1 1 0 0 0
p_2	$\mathcal{N}(0, 12\text{kPa})$	0 1 0 0 0 0 0 0 0 0 0 0 0 0
X_{AGR}	$\mathcal{N}(0, 0.9\%)$	0 0 1 0 0 0 0 0 0 0 0 0 0 0
$\psi_{\text{AGR,ND}}$	$\mathcal{N}(0, 1.5\%)$	0 0 0 1 0 0 0 0 0 0 0 0 0 0
ψ_{SW}	$\mathcal{N}(0, 1\%)$	0 0 0 0 1 0 0 0 0 0 0 0 0 0
p_{CR}	$\mathcal{N}(0, 10\text{MPa})$	0 0 0 0 0 1 0 0 0 0 0 0 0 0
ϕ_{HE}	$\mathcal{N}(0, 0.33^\circ\text{KW})$	0 0 0 0 0 0 1 0 0 0 0 0 0 0
m_{PE1}	$\mathcal{N}(0, 0.1\text{mg/Hub})$	1 0 0 0 0 0 0 0 1 1 1 0 0 0
m_{PE2}	$\mathcal{N}(0, 0.1\text{mg/Hub})$	1 0 0 0 0 0 0 0 1 1 1 0 0 0
m_{NE}	$\mathcal{N}(0, 0.1\text{mg/Hub})$	1 0 0 0 0 0 0 0 1 1 1 0 0 0
t_{VE1}	$\mathcal{N}(0, 4\mu\text{s})$	0 0 0 0 0 0 0 0 0 0 0 1 0 0
t_{VE2}	$\mathcal{N}(0, 4\mu\text{s})$	0 0 0 0 0 0 0 0 0 0 0 0 1 0
t_{NE}	$\mathcal{N}(0, 4\mu\text{s})$	0 0 0 0 0 0 0 0 0 0 0 0 0 1

Die Werte für die Eingangsparametertoleranzen wurden aus Datenblättern und Erfahrungswerten ermittelt. Die Toleranzen für den Ladedruck und den Commonraildruck entsprechen den jeweiligen Sensortoleranzen. Die Toleranz der AGR-Rate wird aus den Toleranzbeiträgen von HFM, AGR-Ventil, Ladedrucksensor sowie Verdichtungsverhältnis ermittelt. Dabei wird angenommen, dass die einzelnen Toleranzen unabhängige und identisch verteilte Zufallsvariablen darstellen. Der absolute Toleranzbeitrag des AGR-Ventils auf die AGR-Rate wird unter Hinzunahme von Erfahrungswerten mit $\mathcal{N}(0, 0.33\%)$ angenommen. Die verbleibenden Beiträge werden nach Gleichung 7.8 ermittelt, wobei eine konstante AGR-Rate von 50% angenommen wird. Die relative Toleranz des HFM wird mit $\mathcal{N}(0, 1.2)$ angenommen. Der relative Toleranzbeitrag des Ladedrucksensors wird mit $\mathcal{N}(0, 0.8)$ angenommen und der relative Beitrag des Verdichtungsverhältnisses mit $\mathcal{N}(0, 1.0)$. Nach Definition A.5 ergibt sich aus den Einzelbeiträgen eine absolute Toleranz der AGR-Rate von etwa $\mathcal{N}(0, 0.9\%)$. Die Toleranz für die $\psi_{\text{AGR,ND}}$ wird aus Erfahrungswerten abgeleitet. Die Mengentoleranzen der Injektoren werden aus den entsprechenden Datenblättern abgeleitet. Der Toleranzbeitrag der Vor- und Nacheinspritzungen wird mit einem absoluten Wert von $\mathcal{N}(0, 0.1\text{mg/Hub})$ angenähert. Der Toleranzbeitrag der Gesamteinspritzmenge setzt sich aus den Beiträgen der Einzeleinspritzungen zusammen, wobei der Toleranzbeitrag der Haupteinspritzung

mit $\mathcal{N}(0, 0.008 \cdot m_{HE})$ angenähert wird. Die Abweichung des Ansteuerbeginns der Haupteinspritzung wird gleich der Toleranz des Kurbelwellensensors gesetzt. Die Einspritzabstände werden dem Injektordatenblatt entnommen. Korrelationseinflüsse untereinander und mit den Einspritzmengen sind aufgrund ihrer Größenordnung vernachlässigbar.

Formulierung des Optimierungsproblems

Angenommen, $\Omega \subset \mathbb{R}^2$ ist der Betriebsraum eines Verbrennungsmotors, der aus allen physikalisch realisierbaren Betriebspunktkombinationen besteht. Dabei stellt ein Betriebspunkt $\mathbf{z} \in \Omega$ einen Vektor dar, der sich aus n_{Mot} und m_{Ges} zusammensetzt. Die Aufgabe der Basisapplikation besteht nun darin, die Applikationsparameter der Motorsteuererätefunktionen für eine Reihe von Betriebspunkten so zu verändern, dass eine zuvor definierte Kostenfunktion optimal erfüllt wird. Das Ziel der Optimierung kann z.B. die Minimierung des Kraftstoffverbrauchs bei gleichzeitiger Einhaltung der Emissionsgrenzwerte sein. Dieses Ziel soll für jeden denkbaren zeitkontinuierlichen Fahrzyklus, wie z.B. den WLTC, eingehalten werden. Die Applikationsparameter für einen Betriebspunkt können durch den Vektor $\mathbf{x}(\mathbf{z}) \in \mathbb{R}^d$ beschrieben werden. Im Steuergerät werden sie in Form von Kennfeldern gespeichert, die für jeden Parameter je nach Betriebspunkt eine eigene Funktion darstellen. Bei der Optimierung ganzer Kennfelder sind neben der Minimierung von Verbrauch und Emissionen auch weitere Kriterien wie Kennfeldglättung und Realisierbarkeit von Bedeutung. Steile Gradienten in der Parametrierung des Ladedruckkennfeldes können beispielsweise dazu führen, dass der Ladedruck nicht schnell genug nachgeregelt werden kann. Dies wiederum stellt ein schlechtes dynamisches Verhalten des Motors dar und ist für den Fahrer spürbar, was sich auf die Fahrbarkeit und den Fahrkomfort auswirkt. Andererseits gibt es auch Parameterkombinationen, die von vornherein nicht realisierbar sind. Beispiele hierfür sind, Parameterkombinationen die eine physikalische Grenze darstellen oder zu einer Schädigung des Motors führen. Eine physikalische Grenze wäre beispielsweise ein zu hoch gewählter Ladedruck, der aufgrund der physikalischen Spezifikation des Turboladers nicht realisierbar ist. Eine Schädigung wiederum könnte durch Parameterkombinationen hervorgerufen werden, die zu extrem hohen Zylinderdrücken oder Abgastemperaturen führen, welche die Festigkeitsgrenzen einzelner Motorkomponenten überschreiten. In der Praxis wird das beschriebene Optimierungsproblem für eine endliche Anzahl von Fahrzyklen gelöst, da die Lösung des Problems für eine unendliche Anzahl von Zyklen ein unlösbares Problem darstellt. Die Problemstellung wird weiter von einem zeitkontinuierlichen in ein zeitdiskretes Problem umformuliert, wodurch ein Zyklus als ein Tupel $\mathcal{A} = (\mathcal{Z}, \mathcal{W})$ von Betriebspunkten und Gewichten notiert werden kann. Dabei ist $\mathcal{Z} = \{\mathbf{z}_1, \dots, \mathbf{z}_N\}$ die Menge der Betriebspunkte und $\mathcal{W} = \{w_1, \dots, w_N\}$ eine diskrete Menge von Gewichten, wobei jedes Gewicht den normierten Zeitanteil des jeweiligen Betriebspunktes darstellt.

f_{NOx} und $f_{\text{Ruß}}$ seien die Quantilmodelle der jeweiligen Emissionen. Beide hängen vom Betriebspunkt z_i und den Stellgrößen $x(z_i) = x^i$ ab. $f_{\text{Fahrbarkeit}}$ sei die Varianz eines der nominalen GP-Modelle und stellt ein Maß für die Fahrbarkeit eines Betriebspunktes mit der Stellgröße x^i dar. Zusätzlich ist $f_{\text{Kennfeldglättung}}$ eine Straffunktion, die hohe Gradienten

zwischen benachbarten Kennfeldstützstellen bestraft und somit glatte Kennfelder gewährleistet. Die Straffunktion sei gegeben durch

$$f_{\text{Kennfeldglättung}}(\bar{\mathbf{x}}) = \sum_{i=1}^K (a_k \bar{x})^2, \quad (7.9)$$

wobei $\bar{\mathbf{x}} = (x^1, \dots, x^N)^\top$ ist und a_k für $k = 1, \dots, K$ ein Vektor der Länge N ist. Dabei ist K die Anzahl der Kantenverbindungen zwischen den Stützstellen im Kennfeld. Die Kostenfunktion für das multikriterielle Optimierungsproblem ist dann gegeben durch

$$\begin{aligned} \sum_{i=1}^N \mathbf{w}_i (f_{\text{NO}_x}(\mathbf{z}_i, \mathbf{x}^i) \cdot p_{\text{NO}_x} + f_{\text{Ruß}}(\mathbf{z}_i, \mathbf{x}^i) \cdot p_{\text{Ruß}}) + f_{\text{Fahrbarkeit}}(\bar{\mathbf{x}}) \cdot p_{\text{Fahrbarkeit}} \\ + f_{\text{Kennfeldglättung}}(\bar{\mathbf{x}}) \cdot p_{\text{Kennfeldglättung}}, \end{aligned} \quad (7.10)$$

wobei \mathbf{w}_i die Gewichte der Betriebspunkte darstellen. Die Gewichte werden durch die Anzahl der Fahrzyklen und die Dauer eines Zyklus bestimmt. Die Parameter p_{NO_x} , $p_{\text{Ruß}}$, $p_{\text{Fahrbarkeit}}$ und p_{KFG} stellen die relative Gewichtung der einzelnen Kriterien dar.

Ein Problem, das bei der Verwendung der obigen Gleichung mit Quantilmodellen auftritt, ist die daraus resultierende Änderung der effektiven Leistung. Zur Veranschaulichung soll folgendes Beispiel dienen: Das vorhergesagte Quantil repräsentiert einen Motor mit unterschiedlichen Stellgrößen. Dadurch ändert sich nicht nur die Abgaszusammensetzung, sondern auch die mittlere effektive Leistung. Dies führt dazu, dass die effektive Leistung über den optimierten Zyklus nicht der geforderten Leistung über den Originalzyklus entspricht. Dies verfälscht das Optimierungsergebnis und muss durch eine Korrektur berücksichtigt werden. Zu diesem Zweck modellieren wir in unseren Betrachtungen die leistungsbezogenen Größen immer in g/kWh und normieren die kumulierten spezifischen Emissionen auf die Referenzleistung P_{Ref} , die der für den Zyklus mit dem Nennmotor benötigten Leistung entspricht. Aus der zuvorigen Kostenfunktion wird somit

$$\begin{aligned} K(x^1, \dots, x^N) = \sum_{i=1}^N \mathbf{w}_i P_{\text{Ref}} (f_{\text{NO}_x}(\mathbf{z}_i, \mathbf{x}^i) \cdot p_{\text{NO}_x} + f_{\text{Ruß}}(\mathbf{z}_i, \mathbf{x}^i) \cdot p_{\text{Ruß}}) \\ + f_{\text{Fahrbarkeit}}(\bar{\mathbf{x}}) \cdot p_{\text{Fahrbarkeit}} + f_{\text{Kennfeldglättung}}(\bar{\mathbf{x}}) \cdot p_{\text{KFG}}, \end{aligned} \quad (7.11)$$

wobei f_{NO_x} und $f_{\text{Ruß}}$ die spezifischen Emissionen in g/kWh darstellen.

Da nun eine Formulierung für die Kostenfunktion vorliegt, kann das endgültige Optimierungsproblem wie folgt formuliert werden: Ziel ist die Minimierung des 95%-Quantils der Ruß-Rohemissionen. Zusätzlich wird das 95%-Quantil der kumulierten NO_x -Rohemissionen auf 12.5 g begrenzt—dieser Wert entspricht dem 95%-Quantilwert der Serienapplikation. Für die Kennfeldgradienten wird in ASCMO ein Wert von 0.1 zugelassen, der eine moderate Glättung erlaubt. Der Grenzwert für die Fahrbarkeit wird gleich der maximalen Varianz des Trainingsdatensatzes gewählt. Alle 12 Stellgrößen werden in der Optimierung berücksichtigt.

Ergebnisse und Diskussion

Im Folgenden werden die Ergebnisse der Optimierung dargestellt und diskutiert. Die Optimierungsergebnisse werden mit der Serienapplikation verglichen. Der Vergleich zeigt nicht nur die nominalen Ergebnisse, sondern auch die Ausgangsverteilung unter Berücksichtigung der Toleranzeinflüsse. Zur Visualisierung der Ergebnisse werden 200 quasi zufällige Stichproben gezogen, so dass die Verteilung der Flotte repräsentativ dargestellt wird—jede Stichprobe stellt einen Motor mit Toleranzen dar. Neben den Stichproben ist auch die Verteilungsfunktion dargestellt.

Wie bereits erwähnt, muss bei der Darstellung von nicht-nominalen Stichproben auch die Änderung der effektiven Leistung berücksichtigt werden. Die Vorgehensweise ist in Abbildung 7.10 dargestellt. Bei der Optimierung wird die Änderung der effektiven Leistung durch eine Normalisierung berücksichtigt. Zur Visualisierung simulieren wir die Situation im geschlossenen Regelkreis durch ein iteratives Schema und passen den Betriebspunkt entsprechend der gewünschten effektiven Leistung an. Dabei führt der aus der Unsicherheit resultierende Fehler zu einer Änderung des Verbrennungsprozesses, was wiederum zu Unterschieden in den Emissionen und im Drehmoment führt. Das Soll-Drehmoment und das Ist-Drehmoment werden verglichen und der Betriebspunkt sowie die übrigen Stellgrößen entsprechend angepasst. Dies geschieht iterativ, bis die Drehmomentdifferenz unter einem Grenzwert von 0.1Nm liegt. Dieser Vorgang wird für jede Stichprobe durchgeführt.

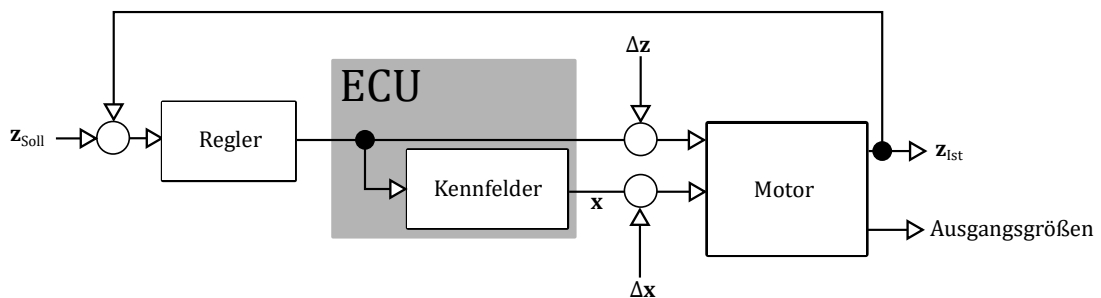


Abbildung 7.10.: Schematische Darstellung des Regelkreises. z_{Soll} und z_{Ist} sind der Soll- und der Ist-Betriebspunkt zu einem bestimmten Zeitschritt. Δz ist die zugehörige Betriebspunkttoleranz (Haupteinspritzmengenabweichung) und Δx die Parametertoleranz [84].

In Abbildung 7.11 sind die Simulationsergebnisse der Serienapplikation und der optimierten Applikation dargestellt. Ein Punkt innerhalb des Diagramms stellt jeweils die kumulierten NO_x - und Ruß-Rohemissionen für einen WLTC dar. Der orangefarbene Punkt stellt das WLTC-Ergebnis für eine nominale Motorkonfiguration dar. Die blau gepunkteten Ergebnisse entsprechen dagegen Motoren mit toleranzbehafteten Komponenten. Auf den Achsen sind weiterhin die Kerndichteschätzungen (engl. kernel density estimates) der NO_x - und Ruß-Rohemissionen aufgetragen. Zusätzlich sind die 95%-Quantile der zugrundeliegenden Verteilungen hervorgehoben. Es ist deutlich zu erkennen, dass die Verteilung der Ruß-Rohemissionen unter Einhaltung des NO_x -Grenzwertes deutlich reduziert wurde. Der nominale Ruß wurde um 0.36g von 0.66g auf 0.3g reduziert. Das 95%-Quantil der Rußverteilung wurde um 0.66g von 1.11g auf 0.45g verringert. Die Kerndichteschätzungen der

Ruß-Rohemissionen zeigen, dass der rechte Schweif und die Varianz der Verteilung sehr stark reduziert wurden. Die nominalen NO_x -Rohemissionen sind mit 10.4g und 10.3g nahezu identisch. Das 95%-Quantil der NO_x -Rohemissionsverteilung wurde um 0.47g von 12.77g auf 12.3g reduziert. Damit liegt das 95%-Quantil unter dem angegebenen Grenzwert von 12.5g. Weiterhin werden sowohl das Varianzkriterium als auch das Glättungskriterium in allen optimierten Betriebspunkten eingehalten.

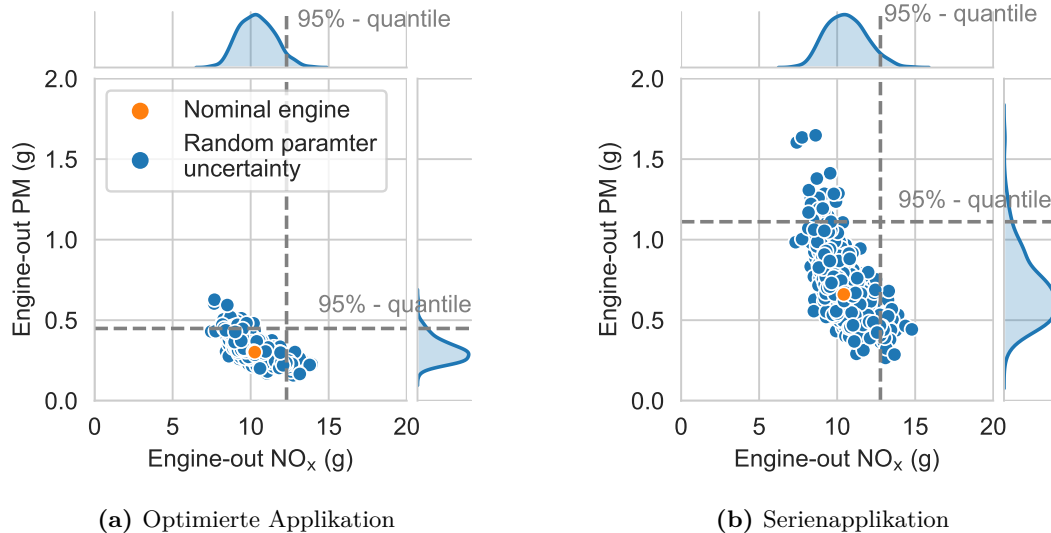


Abbildung 7.11.: WLTC-Simulationsergebnisse mit Kerndichteschätzungen für einen nominalen Motor (orange) und 200 Motoren mit zufälligem Toleranzverhalten (blau). Dargestellt sind die Ergebnisse für die NO_x - und Ruß-Rohemissionen [84].

Insgesamt kann festgestellt werden, dass alle auferlegten Grenzwerte bzw. Beschränkungen eingehalten wurden. Gleichzeitig wurden der Nominalwert, das 95%-Quantil und die Verteilung der Ruß-Rohemissionen in allen Bereichen deutlich reduziert. Ein vergleichbares nominelles Ergebnis ist mit einem konventionellen Applikationsansatz unter Verwendung von Sicherheitskoeffizienten sicherlich erreichbar, aber eher unwahrscheinlich. Eine bewusste Beeinflussung der Emissionsverteilung setzt die Kenntnis der Verteilung oder eine Abschätzung der Verteilung voraus. Wie Abbildung 7.11 zeigt, ändert sich die Verteilung jedoch in Abhängigkeit von den Applikationsparametern. Daher wird die Verwendung eines empirischen Koeffizienten zur Bestimmung der Emissionsverteilung nicht empfohlen.

Zusammenfassung

In diesem Abschnitt wurde die entwickelte SEMO-Methode zur Optimierung eines V6-Dieselmotors verwendet. Zunächst wurde ein DoE erstellt und nominale Ruß- und NO_x -Modelle des Motors trainiert. Anschließend wurden stochastische Motormodelle trainiert, welche eine HFM-Toleranz abbilden. Diese wurden mit dynamischen Prüfstandsmessungen validiert. Anschließend wurden erneut stochastische Motormodelle trainiert, die verschiedene Bauteiltoleranzen berücksichtigten. Mit Hilfe dieser Modelle wurde der Motor unter

Berücksichtigung vorgegebener Kriterien sowie der Toleranzeinflüsse optimiert. Das optimierte Ergebnis wurde anschließend mit dem Ergebnis der Serienapplikation verglichen. Die Ergebnisse zeigen, dass die vorgegebenen Grenzwerte bezüglich NO_x -Rohemissionen, Kennfeldglättung und Fahrbarkeit eingehalten werden und die Verteilung der Ruß-Rohemissionen deutlich reduziert werden konnte.

8. Zusammenfassung und Diskussion

In dieser Arbeit wird eine neue Applikationsmethodik namens *Stochastic Engine Modeling and Optimization* für die datengetriebene Applikation von Verbrennungsmotoren vorgestellt, die eine robuste Applikation dieser unterstützt. Der Standardprozess in der Motorapplikation ist die Parametrierung eines goldenen, nominalen Motors und die Sicherstellung der Konformität des Motors durch die Anwendung von Sicherheitsfaktoren. Die Sicherheitsfaktoren berücksichtigen dabei Toleranzeinflüsse. Im Folgenden werden die gewonnenen Erkenntnisse und Ergebnisse dieser Arbeit kurz und knapp zusammengefasst. Darüber hinaus wird diskutiert, in welchen Bereichen noch Forschungsbedarf besteht.

Zunächst werden in Kapitel 2 Grundlagen vermittelt, die für den weiteren Verlauf der Arbeit wichtig sind. Anschließend wird in Kapitel 3 der Versuchsaufbau beschrieben, wobei auf die Versuchsträger, den Prüfstand und die Messdatenerfassung eingegangen wird. Der SEMO-Ansatz wird in Kapitel 4 vorgestellt und grob in Versuchsmethodik, Modellierung sowie Optimierung unterteilt. Hier wird auch auf den Stand der Technik eingegangen und die Anforderungen an die Methodik beschrieben. Das Kapitel schließt mit einer kritischen Bewertung und Abgrenzung der Methodik zum Stand der Technik.

In Kapitel 5 wird die Versuchsplanungsverfahren der SEMO-Methode beschrieben, welche für die Erstellung der stochastischen Modelle verwendet wird. Dabei wird auf die Versuchsaufbereitung, die Versuchsplanung und die Versuchsdurchführung eingegangen. Für die Versuchsdurchführung wird auf einen Ansatz namens *Online DoE with Constraint Modeling* zurückgegriffen. Um diesen Ansatz für die SEMO-Methode nutzbar zu machen, werden verschiedene Anpassungen bzw. Verbesserungen vorgenommen und eine modifizierter ODCM-Ansatz vorgestellt. Zudem erfolgt eine Diskussion zu weiteren Anpassungen, die zu einer noch effizienteren Versuchsdurchführung führen könnten.

In Kapitel 6 wird ein stochastisches Motormodell zur Vorhersage von Flottenemissionen vorgestellt, das den Kern des SEMO-Ansatzes bildet. Um die Leistungsfähigkeit des Modells besser einordnen zu können, werden zunächst die asymptotischen Eigenschaften dessen untersucht und diskutiert. Die Untersuchungen zeigen, dass der Modellfehler mit steigender Anzahl von Stützstellen und Stichproben abnimmt. Weiterhin wird gezeigt, dass ein optimales Stützstellen-Stichprobenverhältnis existiert, für das der Modellfehler minimal ist.

In Kapitel 7 wird die SEMO-Methode auf zwei Anwendungsfälle angewendet. Zunächst wird in Abschnitt 7.1 ein Reihen-Vierzylinder-Dieselmotor in einem festen Betriebspunkt optimiert. Dazu werden ein deterministisches und ein stochastisches Optimierungsszenario definiert und die Optimierungsergebnisse miteinander verglichen. Das stochastische Optimierungsszenario wird mit der SEMO-Methode gelöst. Für den deterministischen Fall werden nominale Modelle für Ruß, NO_x und das Verbrennungsgeräusch trainiert. Für den stochastischen Fall werden entsprechende Quantilmodelle trainiert. Die Ergebnisse zeigen, dass die Grenzwerte mit der SEMO-Methode eingehalten werden. Für das deterministische

Szenario mit Sicherheitsfaktoren ist dies nicht der Fall. Darüber hinaus wird mit der SEMO-Methode eine deutlich höhere Minderung der NO_x -Rohemissionen erreicht. Anschließend an den Vergleich der beiden Optimierungsszenarien wird die Effizienz der SEMO-Methode anhand eines numerischen Vergleichs mit der numerischen Lösung des stochastischen Problems gezeigt. SEMO-Methode ist wie erwartet um Größenordnungen weniger rechenintensiv als eine numerische Lösung des Problems. Zudem ist der Rechenaufwand der SEMO-Methode vergleichbar mit dem Rechenaufwand des deterministischen Szenarios. Zum Schluss wird in Abschnitt 7.2 die SEMO-Methode zur Optimierung eines V6-Dieselmotors in mehreren Betriebspunkten angewendet. Zunächst findet eine Validierung der stochastischen Modelle statt. Anschließend wird der Motor unter Berücksichtigung der Toleranzeinflüsse optimiert. Zur Validierung der Modelle werden verschiedene Motorkonfigurationen mit unterschiedlichen Toleranzlagen am dynamischen Prüfstand vermessen. Im Anschluss werden Modelle Ruß- und NO_x -Modelle für die verschiedenen Motorkonfigurationen trainiert und mit den Messdaten validiert. Die Ergebnisse zeigen eine sehr gute Übereinstimmung zwischen Messung und Vorhersage—sowohl für Ruß als auch für NO_x . Kleinere Diskrepanzen sind vor allem in dynamischen Situationen zu erkennen, aber kaum signifikant. Abschließend wird eine stochastische Optimierung des Motors unter Berücksichtigung aller Toleranzbeiträge durchgeführt. Hierzu wird eine gesonderte Kostenfunktion formuliert, die eine Emissionsoptimierung unter Einhaltung der Kriterien Fahrbarkeit und Kennfeldglättung anstrebt. Das Ergebnis der Optimierung wird mit dem Emissionsergebnis der Serienapplikation verglichen. Die Ergebnisse zeigen, dass die Ruß-Rohemissionen im Vergleich zur Serienapplikation unter Einhaltung der Grenzwerte deutlich reduziert werden konnten.

Mit den hier vorgestellten Ergebnissen wäre es denkbar, zusätzlich zu den betrachteten Bauteiltoleranzen weitere toleranzbehaftete Phänomene wie Straßenzustand, Verkehrsfluss oder andere Umwelteinflüsse in die Betrachtungen einzubeziehen. In [116] und [115] werden diese bereits bei der Optimierung berücksichtigt. Dies würde eine deutlich bessere Quantifizierung der gesamten inhärenten Streuung ermöglichen.

Darüber hinaus ist zu prüfen, inwieweit der SEMO-Ansatz auf Sonderbetriebsarten wie den DPF-Regenerationsbetrieb oder ganz allgemein auf den Ottomotor übertragbar ist. Auch hier spielen Toleranzeinflüsse eine große Rolle und beeinflussen die Robustheit der Anwendung.

A. Anhang

A.1. Wahrscheinlichkeitstheorie

Die Wahrscheinlichkeitstheorie ist ein Teilgebiet der Mathematik, das sich mit der Probabilistik beschäftigt, die den in der Natur vorkommenden Prozessen zugrunde liegt. Sie hinterfragt die Wahrscheinlichkeit mit der eine Zufallsvariable X die Werte seines Ergebnisraums Ω annimmt. Für einen Münzwurf z.B. bildet $\Omega = \{\text{Kopf}, \text{Zahl}\}$ die Menge der Wurfergebnisse ab, dessen Realisierungen mit $x_{\text{Kopf}} = 1$ und $x_{\text{Zahl}} = 0$ angegeben werden können. Die Wahrscheinlichkeit für ein Ereignis x wird durch $P(x)$ angegeben, wobei P das sogenannte Wahrscheinlichkeitsmaß ist. Im Folgenden sollen einige Begriffe der Wahrscheinlichkeitstheorie, die zum grundlegenden Verständnis der Arbeit beitragen, vorgestellt werden.

Prinzipiell wird zur Betrachtung von zufälligen Ereignissen von einem Zufallsexperiment ausgegangen, dessen mögliche Ergebnisse durch eine Menge Ω zusammengefasst werden können. I.d.R. interessiert uns jedoch nicht das konkrete Ergebnis für ein oder mehrere Ereignisse $\omega \in \Omega$, sondern ob diese eingetreten sind oder nicht. Um den Ereignissen Wahrscheinlichkeiten zuordnen zu können, werden diese in einem sog. Ereignissystem Σ aufgeführt. Dabei ist das Ereignissystem eine Menge von Teilmengen von Ω und stellt eine σ -Algebra dar.

Definition A.1 Ω sei eine nichtleere Menge und $\mathcal{P}(\Omega)$ die zugehörige Potenzmenge, also die Menge aller Teilmengen von Ω . Ein Mengensystem $\Sigma \subseteq \mathcal{P}(\Omega)$, ist eine σ -Algebra wenn folgendes gilt:

- 1. $\Omega \in \Sigma$.
- 2. Ist $S \in \Sigma$, dann ist auch das Komplement $S^c = \Omega \setminus S$ in Σ enthalten. Mit $\Omega \in \Sigma$ ist somit auch die Leere Menge \emptyset in Σ enthalten
- 3. Sind S_1, S_2, \dots, S_n mit $n \in \mathbb{N}$ in Σ enthalten, dann ist auch die Vereinigung $\bigcup_{i=1}^n S_i$ in Σ enthalten.

Eine Wahrscheinlichkeit ist somit ein Bild, welches mit einem Wahrscheinlichkeitsmaß P des Ergebnisraums in das Intervall $[0, 1]$ projiziert wird. Das Tripel (Ω, Σ, P) wird i.A. als Wahrscheinlichkeitsraum bezeichnet. Wenn Ω endlich oder abzählbar unendlich ist, entspricht Σ der Potenzmenge von Ω . Es gilt $P(\Omega) = 1$.

Wahrscheinlichkeitsdichtefunktion

Für den Fall, dass Ω eine überabzählbare Ergebnismenge ist, also im Falle der reellen Zahlen, lässt sich das Wahrscheinlichkeitsmaß durch eine Wahrscheinlichkeitsdichtefunktion definieren. In diesem Fall, wird als Ereignissystem statt der Potenzmenge der reellen Zahlen die Borelsche σ -Algebra gewählt, welche alle Intervalle von reellen Zahlen als Elemente enthält.

Definition A.2 (X, \mathcal{O}) sei ein topologischer Raum, wobei die Topologie \mathcal{O} das Mengensystem der offenen Mengen darstellt. Dann ist

$$\mathcal{B}(X) := \sigma(\mathcal{O}), \quad (\text{A.1})$$

wobei σ der σ -Operator ist.

Das Wahrscheinlichkeitsmaß P ist in diesem Fall für $B \in \mathcal{B}(\mathbb{R})$ durch

$$P(B) := \int_B f(x) dx \quad (\text{A.2})$$

gegeben, wobei $f(x)$ die Wahrscheinlichkeitsdichtefunktion ist. Für die Wahrscheinlichkeitsdichtefunktion gilt, dass sie positiv und integrierbar ist. Zudem ist sie normiert in dem Sinne, dass

$$\int_{-\infty}^{\infty} f(x) dx = 1 \quad (\text{A.3})$$

gilt.

Verteilungsfunktion

Definition A.3 X sei eine multivariate Zufallsvariable auf einem Wahrscheinlichkeitsraum (Ω, Σ, P) mit Ergebnisraum Ω , Sigma-Algebra Σ und Wahrscheinlichkeitsmaß P . Dann ist ihre Verteilungsfunktion $F: \mathbb{R} \rightarrow [0, 1]$ mit $\lim_{x \rightarrow -\infty} F(x) = 0$ und $\lim_{x \rightarrow \infty} F(x) = 1$ als

$$F_X(x) = P(X \leq x). \quad (\text{A.4})$$

definiert, wobei F eine monotone stetige Funktion ist.

Bedingte Wahrscheinlichkeiten

Die bedingte Wahrscheinlichkeit gibt die Wahrscheinlichkeit des Eintretens eines Ereignisses x_A unter der Bedingung an, dass ein zweites Ereignis x_B bereits eingetreten ist. Hierbei schränkt die Bedingung x_B i.d.R. die Eintrittswahrscheinlichkeit von x_A ein.

Definition A.4 x_A und x_B seien beliebige Ereignisse mit $P(x_B) > 0$, so ist die Wahrscheinlichkeit von x_A unter der Bedingung x_B definiert als

$$P(x_A | x_B) = \frac{P(x_A \cap x_B)}{P(x_B)}, \quad (\text{A.5})$$

wobei $P(x_A \cap x_B)$ die Wahrscheinlichkeit ist, dass x_A und x_B gemeinsam auftreten. Für den Fall, dass die beiden Ereignisse unkorreliert sind, ist die Wahrscheinlichkeit gilt $P(x_A | x_B) = P(x_A)$.

Gemeinsame Verteilung von Zufallsvariablen

Die gemeinsame Verteilung mehrerer Zufallsvariablen $X = \{X_1, X_2, \dots, X_n\}$, die auf einem Wahrscheinlichkeitsraum definiert sind, gibt die Wahrscheinlichkeit an, dass jede Zufallsvariable X_1, X_2, \dots, X_n in einen bestimmten Bereich fällt. Im Bereich der reellen Zahlen spricht man auch von einer gemeinsamen Wahrscheinlichkeitsdichteverteilung. Für den Fall n unabhängiger Zufallsvariablen gilt die Produktregel und die gemeinsame Wahrscheinlichkeitsdichteverteilung ist durch

$$f(x) = \prod_{i=1}^n f(x_i) \quad (\text{A.6})$$

gegeben, wobei. Sind die Zufallsvariablen korreliert bzw. voneinander abhängig, so gilt für die gemeinsame Wahrscheinlichkeitsdichte gemäß der Produktregel

$$f(x) = f(x_1) \prod_{i=2}^n f(x_i | x_{i-1}, \dots, x_1), \quad (\text{A.7})$$

Normalverteilung

Die Normalverteilung, auch Gaußverteilung genannt, ist die in der Stochastik am häufigsten auftretende Wahrscheinlichkeitsverteilung. Sie ist der Namensgeber des Gaußprozesses und spielt daher eine besondere Rolle im Rahmen dieser Arbeit. Eine stetige Zufallsvariable X hat eine Normalverteilung, wenn sie folgende Wahrscheinlichkeitsdichte hat:

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right), \quad -\infty < x < \infty. \quad (\text{A.8})$$

Hierbei ist μ der Erwartungswert der und σ^2 die Varianz. Man schreibt auch $X \sim \mathcal{N}(\mu, \sigma^2)$. Hierdurch wird ausgedrückt, dass die Zufallsvariable X normalverteilt ist mit Erwartungswert μ und Varianz σ^2 .

Summe unabhängiger Normalverteilungen

Die Summe zweier unabhängig normalverteilter Zufallsvariablen ist wieder normalverteilt.

Definition A.5 Seien X_1 und X_2 zwei unabhängige normalverteilte Zufallsvariablen mit $X_1 \sim \mathcal{N}(\mu_1, \sigma_1)$ und $X_2 \sim \mathcal{N}(\mu_2, \sigma_2)$, dann ist auch die Summe $X_1 + X_2$ normalverteilt mit

$$X_1 + X_2 \sim \mathcal{N}(\mu_1 + \mu_2, \sigma_1, \sigma_2) \quad (\text{A.9})$$

Evidenz

Die Evidenz bzw. Marginal Likelihood ist eine Wahrscheinlichkeitsfunktion bei der eine oder mehrere Parameter ausintegriert worden sind. Im Kontext bayesianischer Statistik ist sie wie folgt definiert.

Definition A.6 $\text{vec} \mathbf{x}_i \sim P(\mathbf{x}_i | \mathbf{w})$ seien unabhängig und identisch verteilte Zufallsvariablen unter \mathbf{w} mit $\mathbf{X} = (\mathbf{x}_1, \dots, \mathbf{x}_N)$. Der Parameter \mathbf{w} sei an sich durch die Verteilung $P(\mathbf{w} | \lambda)$ beschrieben. So ist die Evidenz $P(\mathbf{X} | \lambda)$ durch ausintegrieren von \mathbf{w} durch

$$P(\mathbf{X} | \lambda) = \int_{\mathbf{w}} P(\mathbf{X} | \mathbf{w}) P(\mathbf{w} | \lambda) d\mathbf{w} \quad (\text{A.10})$$

gegeben.

Randverteilung mehrdimensionaler Normalverteilungen

Definition A.7 Sei $X \sim \mathcal{N}(\mu, \Sigma)$ eine normalverteilte multivariate Zufallsvariable. Weiterhin seien $X_1 \sim \mathcal{N}(\mu_1, \Sigma_{11})$ und $X_2 \sim \mathcal{N}(\mu_2, \Sigma_{22})$ unabhängige Randverteilungen von X , so gilt

$$X = \begin{bmatrix} X_1 \\ X_2 \end{bmatrix} \sim \mathcal{N} \left(\begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}, \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix} \right), \quad (\text{A.11})$$

wobei $\Sigma_{12} = \Sigma_{21} = \text{cov}(X_1, X_2)$ ist.

Bedingte Normalverteilungen

Definition A.8 Sei X eine normalverteilte Zufallsvariable mit Randverteilungen X_1 und X_2 . So gilt für die bedingte Normalverteilung $X_1 | X_2$ unter $X_2 = x_2$

$$X_1 | X_2 = x_2 \sim \mathcal{N} \left(\mu_1 + \Sigma_{12} \Sigma_{22}^{-1} (x_2 - \mu_2), \Sigma_{11} - \Sigma_{12} \Sigma_{22}^{-1} \Sigma_{21} \right). \quad (\text{A.12})$$

Bayes-Theorem

Das Bayes-Theorem, auch Satz von Bayes genannt, ist eines der wichtigsten Sätze der Wahrscheinlichkeitstheorie und ist nach dem englischen Mathematiker Thomas Bayes benannt, der ihn erstmals 1763 in [9] als einen Spezialfall postuliert hat. Das Theorem liefert

einen fundamentalen Zusammenhang zwischen zwei bedingten Wahrscheinlichkeiten und ist definiert als

$$P(x_A | x_B) = \frac{P(x_B | x_A) \cdot P(x_A)}{P(x_B)} \quad (\text{A.13})$$

Im der Bayes'schen Modellierung wird dieser Zusammenhang zur Schätzung der sogenannten A-Posteriori-Wahrscheinlichkeit eines Zufallsereignisses genutzt. Diese gibt die Wahrscheinlichkeit einer unbekannten Größe, welche als Zufallsvariable behandelt wird, in Abhängigkeit von den aus einem Experiment bzw. vorliegenden Daten gewonnenen Erkenntnissen an. Die A-Posteriori-Wahrscheinlichkeit der Modellgewichte wird demnach als

$$P(\mathbf{w} | \mathbf{y}, \mathbf{X}) = \frac{P(\mathbf{y} | \mathbf{X}, \mathbf{w}) \cdot P(\mathbf{w})}{P(\mathbf{y} | \mathbf{X})}. \quad (\text{A.14})$$

geschrieben, wobei \mathbf{w} die Gewichte, \mathbf{X} und \mathbf{y} die Eingangsparameter und zugehörigen Beobachtungen der Trainingsdaten darstellen. Hierbei ist $P(\mathbf{w})$ die A-Priori-Wahrscheinlichkeit über den Modellparametern und $P(\mathbf{y} | \mathbf{X}, \mathbf{w})$ die sogenannte *Likelihood*. Der Term unter dem Bruch ist die sogenannte *Marginal Likelihood* bzw. die Evidenz und stellt eine normalisierende Konstante dar, welche durch

$$P(\mathbf{y} | \mathbf{X}) = \int P(\mathbf{y} | \mathbf{X}, \mathbf{w}) P(\mathbf{w}) d\mathbf{w} \quad (\text{A.15})$$

gegeben ist.

Likelihood-Funktion

In der Statistik gibt die Likelihood-Funktion \mathcal{L} die Anpassungsgüte eines statistischen Modells an einen Datenstichprobe an. Sie wird aus der gemeinsamen Wahrscheinlichkeitsverteilung der Stichprobe gebildet, stellt jedoch eine Funktion der Modellparameter dar. So können im Bereich der Modellierung mit Hilfe der Maximum-Likelihood-Methode unter Hinzunahme der Likelihood-Funktion optimale Modellparameter gefunden werden. Die Likelihood-Funktion ist definiert als

$$\mathcal{L}(\mathbf{w}) = \prod_{i=1}^n P(y_i | \mathbf{x}_i, \mathbf{w}). \quad (\text{A.16})$$

Optimale Parameter für \mathbf{x} erhält man durch Minimieren der Likelihood bzw. durch das Maximieren der negativen Likelihood. Aus Optimierungsgründen wird i.d.R. der Logarithmus der negativen Likelihood maximiert. Diese besitzt auf Grund der Monotonie ihr Maximum am Selben Punkt, ist jedoch weniger rechenintensiv. Das zugehörige Optimierungsproblem lässt sich also

$$\hat{\mathbf{w}} = \arg \max_{\mathbf{w} \in \Theta} (-\log \mathcal{L}(\mathbf{w})) \quad (\text{A.17})$$

notieren.

A.2. Algorithmen

Generierung von Sobol-Folgen

Zunächst werden für jede Dimension d der zu generierenden Folge primitive Polynome der Form

$$p_j(x) = x^{s_j} + a_{1,j}x^{s_j-1} + \dots + a_{s_j-1,j}x + 1 \quad (\text{A.18})$$

mit $j = 1, \dots, d$ gebildet, wobei a einen Koeffizienten und s_j die Ordnung des primitiven Polynoms darstellen. Aus den hieraus errechneten Koeffizienten werden anschließend d Mengen ungerader natürlicher Zahlen $\mathcal{M}_j = \{m_{1,j}, m_{2,j}, \dots, m_{\lceil \log_2 N \rceil, j}\}$ generiert, wobei $0 < m_{i,j} < 2^i$ gilt und N die Anzahl der zu generierenden Glieder darstellt. Für $i \leq s_j$ werden die d kleinsten ungeraden ganzen Zahlen gewählt. Und für $i > s_j$ gilt

$$m_{i,j} = 2^1 a_{1,j} m_{i-1,j} \oplus \dots \oplus 2^{s_j-1} a_{s_j-1,j} m_{i-s_j+1,j} \oplus 2^{s_j} m_{i-s_j,j} \oplus m_{i-s_j,j}, \quad (\text{A.19})$$

wobei \oplus einen bitweisen Exklusiv-Oder-Operator darstellt. Aus den ermittelten Werten $m_{i,j}$ wird die Richtungszahl (engl.: *direction number*)

$$v_{i,j} = \frac{m_{i,j}}{2^i} \quad (\text{A.20})$$

ermittelt, wobei $v_{i,j} \in \{0, 1\}$ ist. Mit Hilfe der Richtungszahl ist die k te Zahl der Sobol-Folge durch

$$x_{k,j} = b_1 v_{1,j} \oplus b_2 v_{2,j} \oplus \dots \oplus b_i v_{i,j} \quad (\text{A.21})$$

gegeben, wobei b_i eine bitweise Darstellung von i ist und $b_i \neq 0$ gilt. Die finale Umwandlung in eine Zahl im Intervall $[0, 1]$ erfolgt dann durch

$$y_{k,j} = \frac{x_{k,j}}{2^{b_j}}. \quad (\text{A.22})$$

Nomenklatur

Physikalische Größen

Formelzeichen	Einheit	Bezeichnung
α	°KW	Kurbelwinkel
V	m ³	Volumen
V_h	m ³	Hubvolumen
$X_{0,5}$	°KW n. OT	0,5 %-Umsatzpunkt
X_5	°KW n. OT	5 %-Umsatzpunkt
X_{50}	°KW n. OT	Umsatzschwerpunkt
x_{RG}	%	Abgasrückhalterate

Abkürzungen und Indizes

Abkürzung	Bezeichnung
SUV	Sport-Utility-Vehicle
HFM	Heißfilm-Luftmassenmesser
ODCM	Online DoE with Constraint Modelling
WLTC	Worldwide harmonized Light-duty vehicles Test Cycle
NO_x	Stickoxid
UHC	Unverbrannte Kohlenwasserstoffe
LNT	NO _x -Speicherkatalysator (Lean NO _x Trap)
GP	Gauß-Prozess (Gaussian process)
SE	quadratisch exponentieller (squared exponential)
ARD	automatische Relevanzbestimmung (automatic relevance determination)
MAE	mittlere absolute Fehler (mean absolute error)
RE	relative Fehler (relative error)
PE	prozentualer Fehler (percentage error)
MSE	mittlere quadratische Fehler (mean squared error)

RMSE	Root-Mean-Square-Error
NRMSE	normalisierte Root-Mean-Square-Error
R²	Bestimmtheitsmaß (coefficient of determination)
CV	Kreuzvalidierung (cross-validation)
LOOCV	Leave-One-Out-Kreuzvalidierung (leave-one-out cross-validation)
DoE	statistischen Versuchsplanung (design of experiments)
VTG	variable Turbinengeometrie
AGR	Abgasrückführung
LLK	Ladeluftkühler
DOC	Dieseloxydationskatalysator
SCR	selektive katalytische Reduktion
DPF	Dieselpartikelfilter
SDPF	SCR beschichteter DPF
m_L	Frischlufthmasse
p₂	Ladedruck
ψ_{SW}	Drallklappenstellung
PCR	Commonraildruck
m_{PE1}	Menge der ersten Voreinspritzung
m_{PE2}	Menge der zweiten Voreinspritzung
m_{NE}	Menge der Nacheinspritzung
φ_{HE}	Ansteuerbeginn der Haupteinspritzung
X_{AGR}	Abgasrückführtrate
MC	Monte-Carlo
QMC	Quasi-Monte-Carlo
SLSQP	Sequential-Least-Squares-Programming
FSN	Schwärzungszahl (Filter Smoke Number)
WC-Modell	Worst-Case-Modell
BC-Modell	Best-Case-Modell
m_{Ges}	eingespritzte Gesamtkraftstoffmasse
n_{Mot}	Motordrehzahl
ψ_{AGR,ND}	Niederdruckanteil der Abgasrückführtrate
SEMO	Stochastic Engine Modeling and Optimization

Abbildungsverzeichnis

2.1. Von links nach rechts: Modell mit Unteranpassung. Gut angepasstes Modell. Modell mit Überanpassung.	14
3.1. Schematischer Aufbau des Vierzylindermotors.	18
3.2. Schematischer Aufbau des V6-Motors.	19
4.1. Aufbau eines Motorsteuergeräts.	24
4.2. Schematische Darstellung der Motorparameter über der Drehzahl und der eingespritzten Kraftstoffmasse.	25
4.3. Motorbetriebsraum über Drehzahl und Last.	26
5.1. Ablauf einer klassischen DoE. Schwarzer Kreis: Ausgangspunkt für die Festlegung der Versuchsraumgrenzen, schwarze Kreuze: geplante Messungen, hellblaue Punkte: (vermessene) fahrbare Messpunkte, rote Kreuze: (abgebrochene) nicht fahrbare Messpunkte, hellrote Fläche: (unbekannter) nicht fahrbarer Versuchsraum, schwarz gestrichelt eingekreiste Fläche: nicht fahrbare Messpunkte, die zu einer möglichen Beschädigung des Motors führen können.	36
5.2. Schematischer Ablauf einer stationären Messung am Motorprüfstand [68]. Dargestellt sind drei beispielhafte Messgrößen mit unterschiedlichen Einschwingzeiten. Aufgetragen sind auch die für die Messung angenommene Einschwingzeit und die Messdauer.	37
5.3. Initialer Versuchsplan aufgespannt über der Motordrehzahl und der Gesamteinspritzmenge.	39
5.4. ODCM-Verfahren. Schwarze Kreuze: geplante Messungen, dunkelblaues Quadrat: sicherer Punkt, hellblaue Punkte: fahrbare Messpunkte, rote Kreuze: nicht fahrbare Messpunkte, schwarze Raute: übersprungener Punkt, hellroter Bereich: nicht fahrbarer Versuchsraum.	41
5.5. Modifiziertes Online-DoE-Verfahren. Schwarze Kreuze: geplante Messungen, dunkelblaues Quadrat: sicherer Punkt, hellblaue Punkte: fahrbare Messpunkte, rote Kreuze: nicht fahrbare Messpunkte, schwarze Raute: übersprungener Punkt.	43
5.6. NRMSE, Spezifität und Selektivität zweier GP mit unterschiedlichem Kernel über der Anzahl an Trainingsdaten, die für die Modellbildung verwendet wurden. Die Trainingsdaten für das jeweilige Modell werden zufällig aus der Gesamtheit der Trainingsdaten initialisiert. Schwarz durchgezogene Linie: isotroper SE Kernel, Blau gestrichelte Linie: ARD SE Kernel.	44

5.7. NRMSE, Spezifität und Selektivität zweier GP mit unterschiedlichem Kernel über der Anzahl an Trainingsdaten, die für die Modellbildung verwendet wurden. Die Trainingsdaten für das jeweilige Modell werden zufällig aus der Gesamtheit der Trainingsdaten (mit Stichprobenwiederholung) initialisiert. Schwarz durchgezogene Linie: isotroper SE Kernel, Blau gestrichelte Linie: ARD SE Kernel.	45
6.1. Schematische Darstellung des Modellbildungsprozesses.	50
6.2. Qualitative Darstellung des Einspritzmusters mit zwei Voreinspritzungen und einer Nacheinspritzung um den oberen Totpunkt [84], mit Genehmigung von Springer Nature.	56
6.3. NRMSE der Quantilmodelle mit 256, 512 und 1024 Stützstellen sowie der MC und QMC-Approximationen über der Stichprobenanzahl N_{samp} . Die Schwärzungszahl, gemessen an der ersten Probenahmestelle vor dem DOC, wird modelliert. Die Unterschreitungswahrscheinlichkeit des Quantils beträgt $p = 0.95$. Der NRMSE wird für 100 Testpunkte berechnet, die sich von den Trainingspunkten unterscheiden. Zur Bestimmung der Box-Plots werden die Experimente 50 Mal wiederholt [85].	60
6.4. NRMSE der Quantilmodelle über dem Verhältnis der Anzahl der Stützstellen zur Anzahl der Stichproben $\frac{N_{grid}}{N_{samp}}$ für $N_{grid} \cdot N_{samp} = 2^{17}$ (Kreis), $N_{grid} \cdot N_{samp} = 2^{18}$ (Kreuz), $N_{grid} \cdot N_{samp} = 2^{19}$ (Plus) und $N_{grid} \cdot N_{samp} = 2^{20}$ (Raute) Funktionsaufrufe des Nominalmodells, die zur Berechnung der Quantilwerte notwendig sind. Es wird die Schwärzungszahl modelliert, die an der ersten Probenahmestelle vor dem DOC gemessen wurde. Die Unterschreitungswahrscheinlichkeit am Quantil beträgt $p = 0.95$. Der NRMSE wird für 100 Testpunkte berechnet, die sich von den Trainingspunkten unterscheiden. Die gestrichelten Linien stellen Geraden konstanter Stützstellenanzahl dar [85].	62
7.1. Ergebnisse der NO_x -Rohemissionen und der Schwärzungszahl für das deterministische und das stochastische Optimierungsszenario. Deterministisches Szenario: nominales Ergebnis (schwarzer Kreis), quasi zufällige Realisierungen (blaue Kreise), Wahrscheinlichkeitsdichteverteilung (blaue durchgezogene Linie). Stochastisches Szenario: nominales Ergebnis (schwarzes Quadrat), quasi zufällige Realisierungen (rote Quadrate), Wahrscheinlichkeitsdichteverteilung (rot gestrichelte Linie) [85].	71
7.2. Ergebnisse der NO_x -Rohemissionen und des Verbrennungsgeräuschs für das deterministische und stochastische Optimierungsszenario. Deterministisches Szenario: nominales Ergebnis (schwarzer Kreis), quasi zufällige Realisierungen (blaue Kreise), Wahrscheinlichkeitsdichteverteilung (blaue durchgezogene Linie). Stochastisches Szenario: nominales Ergebnis (schwarzes Quadrat), quasi zufällige Realisierungen (rote Quadrate), Wahrscheinlichkeitsdichteverteilung (rote gestrichelte Linie) [85].	72
7.3. Motordrehzahl, Motordrehmoment sowie kumulierte und absolute NO_x -Rohemissionen im WLTC (Messung und Vorhersage) [84].	80
7.4. Motordrehzahl, Motordrehmoment, sowie kumulierte und absolute Ruß-Rohemissionen im WLTC (Messung und Vorhersage) [84].	81

7.5. Temperatur vor Turbolader im WLTC [84].	83
7.6. Motordrehzahl, Motordrehmoment und kumulierte NO _x -Rohemissionen im WLTC (Messung und Vorhersage) für die verschiedenen Konfigurationen [84].	86
7.7. Motordrehzahl, Motordrehmoment und absolute NO _x -Rohemissionen im WLTC (Messung und Vorhersage) für die verschiedenen Konfigurationen [84].	87
7.8. Motordrehzahl, Motordrehmoment und kumulierte Ruß-Rohemissionen im WLTC (Messung und Vorhersage) für die verschiedenen Konfigurationen [84].	88
7.9. Motordrehzahl, Motordrehmoment und absolute Ruß-Rohemissionen im WLTC (Messung und Vorhersage) für die verschiedenen Konfigurationen [84].	89
7.10. Schematische Darstellung des Regelkreises. \mathbf{z}_{Soll} und \mathbf{z}_{Ist} sind der Soll- und der Ist-Betriebspunkt zu einem bestimmten Zeitschritt. $\Delta\mathbf{z}$ ist die zugehörige Betriebspunkttoleranz (Haupteinspritzmengenabweichung) und $\Delta\mathbf{x}$ die Parametertoleranz [84].	93
7.11. WLTC-Simulationsergebnisse mit Kerndichteschätzungen für einen nominalen Motor (orange) und 200 Motoren mit zufälligem Toleranzverhalten (blau). Dargestellt sind die Ergebnisse für die NO _x - und Ruß-Rohemissionen [84].	94

Tabellenverzeichnis

3.1. Spezifikationen des Vierzylinder-Forschungsmotors.	17
3.2. Spezifikationen des Sechszylinder-V-Motors.	19
3.3. Messdatenerfassung und -sensorik.	21
3.4. Übersicht der Analyseeinheiten des Horiba MEXA-ONE.	21
5.1. Modellparameterübersicht.	39
5.2. Variationsbereich der Modellparameter.	40
6.1. Betriebsraumgrenzen des DoE-Plans [85].	57
6.2. NRMSE und R^2 der Trainingsdaten sowie NRMSE und R^2 mit der LOOCV-Methode für das nominale Ruß-Rohemissionsmodell [85].	58
6.3. Parametertoleranzen des Reihen-Vierzylinder-Dieselmotors [85].	59
7.1. Sicherheitsfaktoren für das deterministische Optimierungsszenario [85].	67
7.2. NRMSE und R^2 der Trainingsdaten sowie NRMSE und R^2 mit der LOOCV-Methode für die nominalen Motormodelle (Reihen-Vierzylinder-Dieselmotor) [85].	67
7.3. NRMSE und R^2 der Trainings- und Testdaten für die Quantilmodelle [85].	69
7.4. Optimierungsergebnisse für das deterministische und das stochastische Optimierungsszenario. Dargestellt sind die Ergebnisse für die NO_x - und Ruß-Rohemissionen sowie für das Verbrennungsgeräusch. Aus den mit jeweils 10^6 QMC-Stichproben approximierten Ausgangsverteilungen wurden die verschiedenen Statistiken der jeweiligen Optima berechnet [85].	70
7.5. Optimierungsergebnisse für das stochastische Optimierungsszenario sowie für den Brute-Force-Ansatz. Dargestellt sind die Ergebnisse für die NO_x - und Ruß-Rohemissionen sowie für das Verbrennungsgeräusch. Aus den mit jeweils 10^6 QMC-Stichproben approximierten Ausgangsverteilungen wurden die verschiedenen Statistiken der jeweiligen Optima berechnet. [85]	74
7.6. Rechenaufwand für die deterministische und stochastische Optimierung sowie für den Brute-Force-Ansatz. Angegeben ist die Rechenzeit in Sekunden.	75
7.7. Liste der, durch das ODCM-Verfahren, überwachten Grenzwerte [84].	77
7.8. Betriebsraumgrenzen des DoE-Versuchsplans für die Untersuchungen am Sechszylinder-Motor [84].	78
7.9. NRMSE und R^2 der Trainingsdaten sowie NRMSE und R^2 mit der LOOCV-Methode für das nominale Ruß-Rohemissionsmodells (V6-Dieselmotor) [84].	79
7.10. Parametertoleranzen des V6-Dieselmotors [84].	90

Literaturverzeichnis

- [1] ANDREW WILSON ; HANNES NICKISCH: Kernel Interpolation for Scalable Structured Gaussian Processes (KISS-GP). In: FRANCIS BACH (Hrsg.) ; DAVID BLEI (Hrsg.): *Proceedings of the 32nd International Conference on Machine Learning* Bd. 37, PMLR, Juni 2015, S. 1775–1784. – URL <http://proceedings.mlr.press/v37/wilson15.html>
- [2] ANTONOV, I.A. ; SALEEV, V.M.: An economic method of computing LPT-sequences. In: *USSR Computational Mathematics and Mathematical Physics* 19 (1979), Januar, Nr. 1, S. 252–256. – URL <https://www.sciencedirect.com/science/article/pii/0041555379900855>. – ISSN 0041-5553
- [3] ASMUSSEN, S. ; GLYNN, Peter W.: *Stochastic Modelling and Applied Probability*. Bd. 57: *Stochastic Simulation: Algorithms and Analysis*. Springer, New York, NY, 2007. – URL <https://link.springer.com/book/10.1007/978-0-387-69033-9>. – ISSN: 0172-4568. – ISBN 978-0-387-30679-7
- [4] AVERBECK, Bruno B.: Theory of Choice in Bandit, Information Sampling and Foraging Tasks. In: *PLOS Computational Biology* 11 (2015), März, Nr. 3, S. e1004164. – URL <https://doi.org/10.1371/journal.pcbi.1004164>. – Publisher: Public Library of Science
- [5] AVL LIST GMBH: *Emission Measurement Instruments - AVL Smoke Meter*. AVL List GmbH, 2018. – URL <https://www.avl.com/documents/10138/885893/AVL+Smoke+Meter+415SE+English>
- [6] AYINDE, Babajide O. ; INANC, Tamer ; ZURADA, Jacek M.: Redundant feature pruning for accelerated inference in deep neural networks. In: *Neural Networks* 118 (2019), Oktober, S. 148–158. – URL <https://doi.org/10.1016/j.neunet.2019.04.021>. – ISSN 0893-6080
- [7] BATES, R. A. ; WYNN, H. P.: Modelling Feasible Design Regions Using Lattice-based Kernel Methods. In: *Quality and Reliability Engineering International* 20 (2004), März, Nr. 2, S. 135–142. – URL <https://doi.org/10.1002/qre.624>. – Zugriffsdatum: 2021-03-09. – Publisher: John Wiley & Sons, Ltd. – ISSN 0748-8017
- [8] BAUER, Frank ; PEREVERZEV, Sergei ; ROSASCO, Lorenzo: On regularization algorithms in learning theory. In: *Journal of Complexity* 23 (2007), Februar, Nr. 1, S. 52–72. – URL <https://www.sciencedirect.com/science/article/pii/S0885064X06000781>. – ISSN 0885-064X

- [9] BAYES, Thomas ; PRICE, null: LII. An essay towards solving a problem in the doctrine of chances. By the late Rev. Mr. Bayes, F. R. S. communicated by Mr. Price, in a letter to John Canton, A. M. F. R. S. In: *Philosophical Transactions of the Royal Society of London* 53 (1763), Januar, S. 370–418. – URL <https://doi.org/10.1098/rstl.1763.0053>. – Zugriffsdatum: 2021-08-08. – Publisher: Royal Society
- [10] BEATRICE, Carlo ; NAPOLITANO, Pierpaolo ; GUIDO, Chiara: Injection Parameter Optimization by DoE of a Light-duty Diesel Engine Fed by Bio-ethanol/RME/Diesel Blend. In: *Applied Energy* 113 (2014), S. 373–384. – Place: Amsterdam, Netherlands Publisher: Elsevier. – ISSN 0306-2619
- [11] BENGIO, Yoshua ; BERGSTRA, James: Slow, Decorrelated Features for Pretraining Complex Cell-like Networks. In: BENGIO, Y. (Hrsg.) ; SCHUURMANS, D. (Hrsg.) ; LAFFERTY, J. (Hrsg.) ; WILLIAMS, C. (Hrsg.) ; CULOTTA, A. (Hrsg.): *Advances in Neural Information Processing Systems* Bd. 22, Curran Associates, Inc., 2009, S. 99–107
- [12] BERGER, Benjamin: *Modeling and Optimization for Stationary Base Engine Calibration*. Munich, Germany, mediaTUM, PhD Thesis, 2012. – URL <http://mediatum.ub.tum.de?id=1108936>
- [13] BERGER, Benjamin ; RAUSCHER, Florian: Robust Gaussian Process Modeling For Engine Calibration. In: *IFAC Proceedings Volumes* 45 (2012), Nr. 2, S. 159–164. – URL <https://www.sciencedirect.com/science/article/pii/S1474667016306619>. – Place: Amsterdam, Netherlands Publisher: Elsevier. – ISSN 1474-6670
- [14] BERGER-TAL, Oded ; NATHAN, Jonathan ; MERON, Ehud ; SALTZ, David: The Exploration-Exploitation Dilemma: A Multidisciplinary Framework. In: *PLOS ONE* 9 (2014), April, Nr. 4, S. e95693. – URL <https://doi.org/10.1371/journal.pone.0095693>. – Publisher: Public Library of Science
- [15] BISHOP, Christopher M. ; JORDAN, M. (Hrsg.) ; KLEINBERG, J. (Hrsg.) ; SCHÖLKOPF, B. (Hrsg.): *Pattern Recognition and Machine Learning*. 1. New York : Springer Science+Business Media, Januar 2006 (Information Science and Statistics). – URL <https://www.springer.com/de/book/9780387310732>. – ISBN 978-0-387-31073-2
- [16] BLUM, Avrim L. ; LANGLEY, Pat: Selection of relevant features and examples in machine learning. In: *Relevance* 97 (1997), Dezember, Nr. 1, S. 245–271. – URL [https://doi.org/10.1016/S0004-3702\(97\)00063-5](https://doi.org/10.1016/S0004-3702(97)00063-5). – ISSN 0004-3702
- [17] BOX, G. E. P. ; COX, D. R.: An analysis of transformations. In: *Journal of the Royal Statistical Society: Series B (Methodological)* 26 (1964), Nr. 2, S. 211–243. – URL <https://rss.onlinelibrary.wiley.com/doi/abs/10.1111/j.2517-6161.1964.tb00553.x>. – Publisher: John Wiley and Sons

- [18] BROCHU, Eric ; BROCHU, Tyson ; FREITAS, Nando de: A Bayesian Interactive Optimization Approach to Procedural Animation Design. In: *Proceedings of the 2010 ACM SIGGRAPH/Eurographics Symposium on Computer Animation*. Madrid, Spain : Eurographics Association, 2010 (SCA '10), S. 103–112. – URL <https://dl.acm.org/doi/10.5555/1921427.1921443>
- [19] BÄHR, Johannes ; ERKER, Paul: Der Übergang ins digitale Zeitalter. In: *Bosch: Geschichte eines Weltunternehmens*. 1. München : Verlag C.H.BECK Literatur - Sachbuch - Wissenschaft, 2013, S. 452–468. – URL <https://doi.org/10.17104/9783406639845-452>. – Zugriffsdatum: 2021-11-02. – ISBN 978-3-406-63984-5
- [20] CAFLISCH, Russel E.: Monte Carlo and Quasi-Monte Carlo Methods. In: *Acta Numerica* Bd. 7. Cambridge, UK : Cambridge University Press, 1998, S. 1–49
- [21] CASTAGNÉ, M. ; BENTOLILA, Y. ; CHAUDOYE, F. ; HALLÉ, A. ; NICOLAS, F. ; SINOQUET, D.: Comparison of Engine Calibration Methods Based on Design of Experiments (DoE). In: *Oil & Gas Science and Technology - Rev. IFP* 63 (2008), Nr. 4, S. 563–582. – URL <https://doi.org/10.2516/ogst:2008029>
- [22] CATENI, Silvia ; COLLA, Valentina ; VANNUCCI, Marco: A method for resampling imbalanced datasets in binary classification tasks for real-world problems. In: *Neurocomputing* 135 (2014), Juli, S. 32–41. – URL <https://www.sciencedirect.com/science/article/pii/S0925231213011429>. – ISSN 0925-2312
- [23] CHALUPKA, Krzysztof ; WILLIAMS, Christopher K. I. ; MURRAY, Iain: A Framework for Evaluating Approximation Methods for Gaussian Process Regression. In: *J. Mach. Learn. Res.* 14 (2013), Februar, Nr. 1, S. 333–350. – Publisher: JMLR.org. – ISSN 1532-4435
- [24] CHARLES J. STONE: Optimal Global Rates of Convergence for Nonparametric Regression. In: *The Annals of Statistics* 10 (1982), Dezember, Nr. 4, S. 1040–1053. – URL <https://doi.org/10.1214/aos/1176345969>
- [25] CHAWLA, Nitesh V.: Data Mining for Imbalanced Datasets: An Overview. In: MAIMON, Oded (Hrsg.) ; ROKACH, Lior (Hrsg.): *Data Mining and Knowledge Discovery Handbook*. Boston, MA : Springer US, 2010, S. 875–886. – URL https://doi.org/10.1007/978-0-387-09823-4_45. – ISBN 978-0-387-09823-4
- [26] CHAWLA, Nitesh V. ; BOWYER, Kevin W. ; HALL, Lawrence O. ; KEGELMEYER, W. P.: SMOTE: Synthetic Minority over-Sampling Technique. In: *Journal of Artificial Intelligence Research* 16 (2002), Juni, Nr. 1, S. 321–357. – URL <https://doi.org/10.1613/jair.953>. – ISSN 1076-9757
- [27] CHISARI, Corrado: *SOBOL - The Sobol Quasirandom Sequence*. 2011. – URL https://people.sc.fsu.edu/jburkardt/py_src/sobol/sobol.html
- [28] DAMIANOU, Andreas ; LAWRENCE, Neil D.: Deep Gaussian Processes. In: CARVALHO, Carlos M. (Hrsg.) ; RAVIKUMAR, Pradeep (Hrsg.): *Proceedings of the Sixteenth International Conference on Artificial Intelligence and Statistics* Bd. 31. Scottsdale, Arizona, USA : PMLR, April 2013, S. 207–215. – URL <http://proceedings.mlr.press/v31/damianou13a.pdf>

- [29] DEFLORIAN, Michael: *Versuchsplanung und Methoden zur Identifikation zeitkontinuierlicher Zustandsraummodelle am Beispiel des Verbrennungsmotors*. München, Technische Universität München, Doctoral Thesis, 2011. – URL <https://mediatum.ub.tum.de/1083567>
- [30] DEFLORIAN, Michael ; KLÖPPER, Florian ; RÜCKERT, Joachim: Online Dynamic Black Box Modelling and Adaptive Experiment Design in Combustion Engine Calibration. In: *6th IFAC Symposium on Advances in Automotive Control* 43 (2010), Juli, Nr. 7, S. 703–708. – URL <https://doi.org/10.3182/20100712-3-DE-2013.00068>. – Zugriffsdatum: 2021-02-18. – ISSN 1474-6670
- [31] DENIL, Misha ; SHAKIBI, Babak ; DINH, Laurent ; RANZATO, Marc\textquotesingle Aurelio ; FREITAS, Nando de: Predicting Parameters in Deep Learning. In: BURGESS, C. J. C. (Hrsg.) ; BOTTOU, L. (Hrsg.) ; WELLING, M. (Hrsg.) ; GHAHRAMANI, Z. (Hrsg.) ; WEINBERGER, K. Q. (Hrsg.): *Advances in Neural Information Processing Systems* Bd. 26, Curran Associates, Inc., 2013, S. 2148–2156. – URL <https://proceedings.neurips.cc/paper/2013/file/7fec306d1e665bc9c748b5d2b99a6e97-Paper.pdf>
- [32] DIDCOCK, Nico ; JAKUBEK, Stefan ; HAMETNER, Christoph ; KEUTH, Nikolaus: Data based modelling of driveability for internal combustion engines. In: *Engineering Optimization* 50 (2018), Oktober, Nr. 10, S. 1755–1771. – URL <https://doi.org/10.1080/0305215X.2017.1417397>. – Publisher: Taylor & Francis. – ISSN 0305-215X
- [33] DOOREN, Stijn van ; BALERNA, Camillo ; SALAZAR, Mauro ; AMSTUTZ, Alois ; ONDER, Christopher H.: Optimal Diesel engine calibration using convex modelling of Pareto frontiers. In: *Control Engineering Practice* 96 (2020), März, S. 104313. – URL <https://doi.org/10.1016/j.conengprac.2020.104313>. – Zugriffsdatum: 2021-02-24. – ISSN 0967-0661
- [34] DU, Xiaoping ; CHEN, Wei ; WANG, Yu: Most Probable Point-Based Methods. In: SINGHEE, Amith (Hrsg.) ; RUTENBAR, Rob A. (Hrsg.): *Extreme Statistics in Nanoscale Memory Design*. Boston, MA : Springer US, 2010, S. 179–202. – URL https://doi.org/10.1007/978-1-4419-6606-3_7. – ISBN 978-1-4419-6606-3
- [35] ERIKSSON, David ; DONG, Kun ; LEE, Eric ; BINDEL, David ; WILSON, Andrew G.: Scaling Gaussian Process Regression with Derivatives. In: BENGIO, S. (Hrsg.) ; WALLACH, H. (Hrsg.) ; LAROCHELLE, H. (Hrsg.) ; GRAUMAN, K. (Hrsg.) ; CESA-BIANCHI, N. (Hrsg.) ; GARNETT, R. (Hrsg.): *Advances in Neural Information Processing Systems* Bd. 31, Curran Associates, Inc., 2018, S. 6867–6877. – URL <https://proceedings.neurips.cc/paper/2018/file/c2f32522a84d5e6357e6abac087f1b0b-Paper.pdf>
- [36] ETEMADI, N.: An elementary proof of the strong law of large numbers. In: *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete* 55 (1981), Februar, Nr. 1, S. 119–122. – URL <https://doi.org/10.1007/BF01013465>. – ISSN 1432-2064

- [37] EUROPEAN COMMISSION: Commission Regulation (EC) No 692/2008 of 18 July 2008 implementing and amending Regulation (EC) No 715/2007 of the European Parliament and of the Council on type-approval of motor vehicles with respect to emissions from light passenger and commercial vehicles (Euro 5 and Euro 6) and on access to vehicle repair and maintenance information (Text with EEA relevance)/Text with EEA relevance. In: *Official Journal of the European Union* 51 (2019), September, S. 1–136. – URL <http://data.europa.eu/eli/reg/2008/692/2019-09-01>. – ISSN 1725-2555
- [38] FAWCETT, Tom: An introduction to ROC analysis. In: *ROC Analysis in Pattern Recognition* 27 (2006), Juni, Nr. 8, S. 861–874. – URL <https://www.sciencedirect.com/science/article/pii/S016786550500303X>. – ISSN 0167-8655
- [39] FRAZIER, Peter I.: Bayesian Optimization. In: *Recent Advances in Optimization and Modeling of Contemporary Problems*. INFORMS, Oktober 2018 (TutORials in Operations Research), S. 255–278. – URL <https://pubsonline.informs.org/doi/abs/10.1287/educ.2018.0188>. – Section: 11
- [40] FRIEDRICH, Christian ; AUER, Matthias ; STIESCH, Gunnar: Model Based Calibration Techniques for Medium Speed Engine Optimization: Investigations on Common Modeling Approaches for Modeling of Selected Steady State Engine Outputs. In: *SAE International Journal of Engines* 9 (2016), Oktober, Nr. 4, S. 1989–1998. – URL <https://doi.org/10.4271/2016-01-2156>. – Place: Warrendale, PA, USA Publisher: SAE International. – ISSN 19463936, 19463944
- [41] GALANTI, Silvio ; JUNG, Alan: Low-Discrepancy Sequences. In: *The Journal of Derivatives* 5 (1997), August, Nr. 1, S. 63. – URL <http://jod.pm-research.com/content/5/1/63.abstract>
- [42] GELBART, Michael A. ; SNOEK, Jasper ; ADAMS, Ryan P.: Bayesian Optimization with Unknown Constraints. In: *Proceedings of the Thirtieth Conference on Uncertainty in Artificial Intelligence*. Quebec City, Quebec, Canada : AUAI Press, 2014 (UAI'14), S. 250–259. – URL <https://dl.acm.org/doi/10.5555/3020751.3020778>. – event-place: Quebec City, QC, Canada. – ISBN 978-0-9749039-1-0
- [43] GNEITING, Tilmann: Compactly Supported Correlation Functions. In: *Journal of Multivariate Analysis* 83 (2002), November, Nr. 2, S. 493–508. – URL <https://www.sciencedirect.com/science/article/pii/S0047259X01920561>. – ISSN 0047-259X
- [44] GRAMACY, Robert B. ; LEE, Herbert K. H.: Bayesian Treed Gaussian Process Models With an Application to Computer Modeling. In: *Journal of the American Statistical Association* 103 (2008), September, Nr. 483, S. 1119–1130. – URL <https://doi.org/10.1198/016214508000000689>. – Publisher: Taylor & Francis. – ISSN 0162-1459
- [45] GSCHWEITL, Kurt ; PFLUEGL, Horst ; FORTUNA, Tiziana ; LEITHGOEB, Rainer: Increasing the efficiency of model-based engine applications through the use of CAMEO online DoE toolbox. In: *ATZ worldwide* 103 (2001), Juli, Nr. 7, S. 17–20. – URL <https://doi.org/10.1007/BF03226800>. – Zugriffsdatum: 2021-02-25. – ISSN 2192-9076

- [46] GUTJAHR, Tobias ; KRUSE, Thomas ; HUBER, Thorsten: Advanced modeling and optimization for virtual calibration of internal combustion engines. In: *NDIA Ground Vehicle Systems Engineering and Technology Symposium*. Novi, MI, USA : NDIA Michigan Chapter, 2017, S. 1–9
- [47] GUYON, Isabelle ; ELISSEEFF, André: An Introduction of Variable and Feature Selection. In: *Journal Machine Learning Research* 3 (2003), Januar, S. 1157–1182. – URL <https://www.jmlr.org/papers/volume3/guyon03a/guyon03a.pdf>. – Zugriffsdatum: 2021-01-03
- [48] HAAS, Simon-Florian: *Experimentelle und theoretische Untersuchung homogener und teilhomogener Dieselmotorenverfahren*. Stuttgart, Universität Stuttgart, Doctoral Thesis, 2007. – URL <http://dx.doi.org/10.18419/opus-4110>. – Zugriffsdatum: 2021-12-24
- [49] HAASE, Dirk: *Ein neues Verfahren zur modellbasierten Prozessoptimierung auf der Grundlage der statistischen Versuchsplanung am Beispiel eines Ottomotors mit elektromagnetischer Ventilsteuerung (EMVS)*. Dresden, Technische Universität Dresden, Doctoral Thesis, 2004
- [50] HAFNER, Michael ; ISERMANN, Rolf: Multiobjective optimization of feedforward control maps in engine management systems towards low consumption and low emissions. In: *Transactions of the Institute of Measurement and Control* 25 (2003), März, Nr. 1, S. 57–74. – URL <https://doi.org/10.1191/0142331203tm074oa>. – Zugriffsdatum: 2021-02-24. – Publisher: SAGE Publications Ltd STM. – ISSN 0142-3312
- [51] HAFNER, Michael ; SCHÜLER, Matthias ; ISERMANN, Rolf: Einsatz schneller neuronaler Netze zur modellbasierten Optimierung von Verbrennungsmotoren. In: *MTZ - Motortechnische Zeitschrift* 61 (2000), November, Nr. 11, S. 798–805. – URL <https://doi.org/10.1007/BF03227317>. – Zugriffsdatum: 2021-02-16. – ISSN 2192-8843
- [52] HAN, Song ; POOL, Jeff ; NARANG, Sharan ; MAO, Huizi ; GONG, Enhao ; TANG, Shijian ; ELSER, Erich ; VAJDA, Peter ; PALURI, Manohar ; TRAN, John ; CATANZARO, Bryan ; DALLY, William J.: DSD: Dense-Sparse-Dense Training for Deep Neural Networks. In: *International Conference on Learning Representations*, OpenReview.net, 2017, S. 1–13. – URL https://openreview.net/pdf?id=HyoST_9x1
- [53] HARTMANN, B. ; HEUSER, P. ; KLOPPENBURG, E. ; DIENER, R.: Online-methods for engine test bed measurements considering engine limits. In: BARGENDE, M. (Hrsg.) ; REUSS, Hans-Christian (Hrsg.) ; WIEDEMANN, J. (Hrsg.): *16th International Stuttgart Symposium*. Wiesbaden : Springer Fachmedien Wiesbaden, 2016, S. 1251–1264. – ISBN 978-3-658-13255-2
- [54] HARTMANN, Benjamin: *Lokale Modellnetze zur Identifikation und Versuchsplanung nichtlinearer Systeme*. Siegen, Universität Siegen, Doctoral Thesis, 2013. – URL <https://dspace.ub.uni-siegen.de/handle/ubsi/786>
- [55] HASTIE, Trevor ; TIBSHIRANI, Rob ; FRIEDMAN, Jerome: The elements of statistical learning: data mining, inference, and prediction. (2009), Januar

- [56] HENSMAN, James ; FUSI, Nicol{\o} ; LAWRENCE, Neil D.: Gaussian processes for big data. In: *Uncertainty in Artificial Intelligence - Proceedings of the 29th Conference, UAI 2013*. Bellevue, WA, USA : AUAI Press, Juli 2013 (UAI'13), S. 282–290. – URL <https://arxiv.org/abs/1309.6835>
- [57] HOFFMANN, Stefan ; SCHROTT, Michael ; HUBER, Thorsten ; KRUSE, Thomas: Model-based Methods for the Calibration of Modern Internal Combustion Engines. In: *MTZ worldwide* 76 (2015), Nr. 4, S. 24–29. – Place: Berlin & Heidelberg, Germany Publisher: Springer. – ISSN 2192-9114
- [58] HORNIK, Kurt ; STINCHCOMBE, Maxwell ; WHITE, Halbert: Multilayer feedforward networks are universal approximators. In: *Neural Networks* 2 (1989), Januar, Nr. 5, S. 359–366. – URL <https://www.sciencedirect.com/science/article/pii/0893608089900208>. – ISSN 0893-6080
- [59] JAMES, Gareth ; WITTEN, Daniela ; HASTIE, Trevor ; TIBSHIRANI, Robert: Statistical Learning. In: JAMES, Gareth (Hrsg.) ; WITTEN, Daniela (Hrsg.) ; HASTIE, Trevor (Hrsg.) ; TIBSHIRANI, Robert (Hrsg.): *An Introduction to Statistical Learning: with Applications in R*. New York, NY : Springer New York, 2013, S. 15–57. – URL https://doi.org/10.1007/978-1-4614-7138-7_2. – ISBN 978-1-4614-7138-7
- [60] JANKOV, Kristian: *Beitrag zur automatisierten Steuerkennfeld-Applikation bei Fahrzeug-Dieselmotoren*. Berlin, Technische Universität Berlin, Doctoral Thesis, 2008. – URL <http://dx.doi.org/10.14279/depositonce-1955>
- [61] JEROME SACKS ; WILLIAM J. WELCH ; TOBY J. MITCHELL ; HENRY P. WYNN: Design and Analysis of Computer Experiments. In: *Statistical Science* 4 (1989), November, Nr. 4, S. 409–423. – URL <https://doi.org/10.1214/ss/1177012413>
- [62] KATEHAKIS, Michael N. ; VEINOTT, Arthur F.: The Multi-Armed Bandit Problem: Decomposition and Computation. In: *Mathematics of Operations Research* 12 (1987), Mai, Nr. 2, S. 262–268. – URL <https://doi.org/10.1287/moor.12.2.262>. – Zugriffsdatum: 2021-02-18. – Publisher: INFORMS. – ISSN 0364-765X
- [63] KOHAVI, Ron ; JOHN, George H.: Wrappers for feature subset selection. In: *Relevance* 97 (1997), Dezember, Nr. 1, S. 273–324. – URL [https://doi.org/10.1016/S0004-3702\(97\)00043-X](https://doi.org/10.1016/S0004-3702(97)00043-X). – Zugriffsdatum: 2021-02-03. – ISSN 0004-3702
- [64] KOLMANOVSKY, Ilya ; SIVERGUINA, Irina ; LYGOE, Bob: Optimization of Powertrain Operating Policy for Feasibility Assessment and Calibration: Stochastic Dynamic Programming Approach. In: *Proceedings of the American Control Conference* Bd. 2. New York City, NY, USA : IEEE, 2002, S. 1425–1430. – ISBN 0-7803-7298-0
- [65] KRAFT, Dieter: *Deutsche Forschungs- und Versuchsanstalt für Luft- und Raumfahrt Köln: Forschungsbericht*. Bd. 88: *A Software Package for Sequential Quadratic Programming*. Cologne, Germany : Wiss. Berichtswesen d. DFVLR, 1988

- [66] KRÜGER, Michael ; BAREISS, Stefan ; KUFFERATH, Andreas ; NABER, Dirk ; RUFF, Daniel ; SCHUMACHER, Herbert: Further optimization of NO_x emissions under the EU 6d regulation. In: BARGENDE, Michael (Hrsg.) ; REUSS, Hans-Christian (Hrsg.) ; WAGNER, Andreas (Hrsg.) ; WIEDEMANN, Jochen (Hrsg.): *19. Internationales Stuttgarter Symposium*. Berlin & Heidelberg, Germany : Springer, 2019, S. 876–895. – ISBN 978-3-658-25939-6
- [67] KUDER, Johannes ; KRUSE, Thomas: Parameteroptimierung an Ottomotoren mit Direkteinspritzung. In: *MTZ - Motortechnische Zeitschrift* 61 (2000), Juni, Nr. 6, S. 378–384. – URL <https://doi.org/10.1007/BF03226577>. – Zugriffsdatum: 2021-02-16. – ISSN 2192-8843
- [68] KÖTTER, Hinrich ; SEQUENZ, Heiko: Stationäre Motorvermessung mit verschiedenen Methoden und Modellen. In: ISERMANN, Rolf (Hrsg.): *Elektronisches Management motorischer Fahrzeugantriebe: Elektronik, Modellbildung, Regelung und Diagnose für Verbrennungsmotoren, Getriebe und Elektroantriebe*. 1. Wiesbaden : Vieweg+Teubner, 2010, S. 130–166. – URL https://doi.org/10.1007/978-3-8348-9389-5_6. – ISBN 978-3-8348-9389-5
- [69] LANGOUËT, Hoël ; MÉTIVIER, Ludovic ; SINOQUET, Delphine ; TRAN, Quang-Huy: Optimization for Engine Calibration. In: *EngOpt 2008 : International Conference on Engineering Optimization*. Rio de Janeiro, Brazil : E-Papers Servicos Ed. Ltda., 2008, S. 1–5. – URL <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.618.7080&rep=rep1&type=pdf>
- [70] LIEBL, Johannes ; MUNK, Friedrich ; HOHENNER, Hans ; LUDWIG, Bernhard: Die Steuerung der neuen BMW Valvetronic-Motoren. In: *MTZ - Motortechnische Zeitschrift* 62 (2001), Juli, Nr. 7, S. 516–527. – URL <https://doi.org/10.1007/BF03227347>. – ISSN 2192-8843
- [71] MACKAY, David J. C.: Bayesian Interpolation. In: SMITH, C. R. (Hrsg.) ; ERICKSON, Gary J. (Hrsg.) ; NEUDORFER, Paul O. (Hrsg.): *Maximum Entropy and Bayesian Methods: Seattle, 1991*. Dordrecht : Springer Netherlands, 1992, S. 39–66. – URL https://doi.org/10.1007/978-94-017-2219-3_3. – ISBN 978-94-017-2219-3
- [72] MACKAY, David J. C.: *Gaussian Processes - A Replacement for Supervised Neural Networks?* 1997. – URL <http://www.inference.org.uk/mackay/gp.pdf>
- [73] MARTINEZ-CANTIN, Ruben ; FREITAS, Nando de ; BROCHU, Eric ; CASTELLANOS, José ; DOUCET, Arnaud: A Bayesian exploration-exploitation approach for optimal online sensing and planning with a visually guided mobile robot. In: *Autonomous Robots* 27 (2009), August, Nr. 2, S. 93–103. – URL <https://doi.org/10.1007/s10514-009-9130-2>. – ISSN 1573-7527
- [74] MATHWORKS: *Model-Based Calibration Toolbox - Getting Started Guide*. Natick, MA, USA : The MathWorks, Inc., 2020. – URL https://de.mathworks.com/help/pdf_doc/mbc/mbc_gs.pdf. – Zugriffsdatum: 2021-02-25. – Note: R2020b

- [75] MATTHIES, Hermann G.: QUANTIFYING UNCERTAINTY: MODERN COMPUTATIONAL REPRESENTATION OF PROBABILITY AND APPLICATIONS. In: IBRAHIMBEGOVIC, Adnan (Hrsg.) ; KOZAR, Ivica (Hrsg.): *Extreme Man-Made and Natural Hazards in Dynamics of Structures*. Berlin & Heidelberg, Germany : Springer, 2007 (NATO Security through Science), S. 105–135
- [76] MAURYA, Rakesh K.: Knocking and Combustion Noise Analysis. In: *Reciprocating Engine Combustion Diagnostics: In-Cylinder Pressure Measurement and Analysis*. Berlin & Heidelberg, Germany : Springer, 2019, S. 461–542. – ISBN 978-3-030-11954-6
- [77] MCKAY, M. D. ; BECKMAN, R. J. ; CONOVER, W. J.: A Comparison of Three Methods for Selecting Values of Input Variables in the Analysis of Output from a Computer Code. In: *Technometrics* 21 (1979), Nr. 2, S. 239–245. – URL <http://www.jstor.org/stable/1268522>. – Place: London, England Publisher: Taylor & Francis. – ISSN 00401706
- [78] MELLO, Tito Homem-de: On Rates of Convergence for Stochastic Optimization Problems Under Non-Independent and Identically Distributed Sampling. In: *SIAM Journal on Optimization* 19 (2008), Januar, Nr. 2, S. 524–551. – URL <https://doi.org/10.1137/060657418>. – Zugriffsdatum: 2022-05-13. – Publisher: Society for Industrial and Applied Mathematics. – ISSN 1052-6234
- [79] METROPOLIS, Nicholas ; ULAM, S.: The Monte Carlo Method. In: *Journal of the American Statistical Association* 44 (1949), Nr. 247, S. 335–341. – URL <https://www.tandfonline.com/doi/abs/10.1080/01621459.1949.10483310>. – Place: London, England Publisher: Taylor & Francis
- [80] MICHALIS TITSIAS: Variational Learning of Inducing Variables in Sparse Gaussian Processes. In: DAVID VAN DYK (Hrsg.) ; MAX WELLING (Hrsg.): *Proceedings of the Twelfth International Conference on Artificial Intelligence and Statistics* Bd. 5, PMLR, April 2009, S. 567–574. – URL <http://proceedings.mlr.press/v5/titsias09a.html>
- [81] MILLO, Federico ; ARYA, Pranav ; MALLAMO, Fabio: Optimization of automotive diesel engine calibration using genetic algorithm techniques. In: *Energy* 158 (2018), September, S. 807–819. – URL <https://www.sciencedirect.com/science/article/pii/S0360544218311095>. – ISSN 0360-5442
- [82] MOLINARO, Annette M. ; SIMON, Richard ; PFEIFFER, Ruth M.: Prediction error estimation: a comparison of resampling methods. In: *Bioinformatics* 21 (2005), Mai, Nr. 15. – URL <https://doi.org/10.1093/bioinformatics/bti499>. – Zugriffsdatum: 2021-02-23. – ISSN 1367-4803
- [83] MONTOMOLI, Francesco ; CARNEVALE, Mauro ; D’AMMARO, Antonio ; MASSINI, Michela ; SALVADORI, Simone: *Overview of Uncertainty Quantification Methods*. 1. Berlin & Heidelberg, Germany : Springer, 2015 (SpringerBriefs in Applied Sciences and Technology). – ISSN: 2191-530X Publication Title: Uncertainty Quantification in Computational Fluid Dynamics and Aircraft Engines

- [84] MOURAT, Kadir ; ECKSTEIN, Carola ; KOCH, Thomas: Application of stochastic design optimization to a passenger car diesel engine to reduce emission spread in a vehicle fleet. In: *Automotive and Engine Technology* 6 (2021), Juni, Nr. 1, S. 99–112. – URL <https://rdcu.be/ciD0d>. – ISSN 2365-5135
- [85] MOURAT, Kadir ; ECKSTEIN, Carola ; KOCH, Thomas: A stochastic design optimization methodology to reduce emission spread in combustion engines. In: *Automotive and Engine Technology* 6 (2021), Juni, Nr. 1, S. 15–29. – URL <https://doi.org/10.1007/s41104-020-00073-y>. – ISSN 2365-5135
- [86] NEAL, Radford M.: *Lecture Notes in Statistics*. Bd. 118: *Bayesian Learning for Neural Networks*. 1. Berlin & Heidelberg, Germany : Springer, 1996. – URL <https://rd.springer.com/book/10.1007/978-1-4612-0745-0>. – ISBN 978-1-4612-0745-0
- [87] NELLES, Oliver: Local Linear Neuro-Fuzzy Models: Fundamentals. In: NELLES, Oliver (Hrsg.): *Nonlinear System Identification: From Classical Approaches to Neural Networks and Fuzzy Models*. Berlin, Heidelberg : Springer Berlin Heidelberg, 2001, S. 341–389. – URL https://doi.org/10.1007/978-3-662-04323-3_12. – ISBN 978-3-662-04323-3
- [88] NELLES, Oliver ; FINK, A ; BABUSKA, Robert ; SETNES, M: Comparison of Two Construction Algorithms for Takagi-Sugeno Fuzzy Models. In: *International Journal of Applied Mathematics and Computer Science* 10 (2000), Januar, Nr. 4, S. 835–855. – URL <https://www.amcs.uz.zgora.pl/?action=paper&paper=931>. – publisher: Walter de Gruyter. – ISSN 1641-876X
- [89] OBERKAMPF, William L. ; HELTON, Jon C. ; JOSLYN, Cliff A. ; WOJTKIEWICZ, Steven F. ; FERNON, Scott: Challenge problems: uncertainty in system response given uncertain parameters. In: *Reliability Engineering & System Safety* 85 (2004), Nr. 1, S. 11–19. – URL <http://www.sciencedirect.com/science/article/pii/S0951832004000493>. – ISSN 0951-8320
- [90] PANDEY, Varun ; DOOREN, Stijn van ; RITZMANN, Johannes ; PLA, Benjamín ; ONDER, Christopher: Variable smoothing of optimal diesel engine calibration for improved performance and drivability during transient operation. In: *International Journal of Engine Research* (2020), Juni. – URL <https://doi.org/10.1177/1468087420918801>. – Zugriffsdatum: 2021-03-07. – Publisher: SAGE Publications. – ISSN 1468-0874
- [91] PLUMLEE, Matthew ; TUO, Rui: Building Accurate Emulators for Stochastic Simulations via Quantile Kriging. In: *Technometrics* 56 (2014), Dezember, Nr. 4, S. 466–473. – URL <https://www.tandfonline.com/doi/full/10.1080/00401706.2013.860919>. – Place: London, England Publisher: Taylor & Francis. – ISSN 1537-2723
- [92] PRESS, William H. ; TEUKOLSKY, Saul A. ; VETTERLING, William T. ; FLANNERY, Brian P.: *Numerical Recipes 3rd Edition: The Art of Scientific Computing*. 3. USA : Cambridge University Press, 2007. – URL <https://doi.org/10.5555/1403886>. – ISBN 0-521-88068-8

- [93] QUIÑONERO-CANDELA, Joaquin ; RASMUSSEN, Carl E.: A Unifying View of Sparse Approximate Gaussian Process Regression. In: *J. Mach. Learn. Res.* 6 (2005), Dezember, S. 1939–1959. – Publisher: JMLR.org. – ISSN 1532-4435
- [94] RAIDT, Benedikt: *Versuchs- und Applikationsmethode für die Lasterfassung turboaufgeladener Ottomotoren*. München, Technische Universität München, Dissertation, 2020. – URL <https://mediatum.ub.tum.de/doc/1515232/1515232.pdf>
- [95] RASMUSSEN, Carl ; GHAHRAMANI, Zoubin: Occam’s Razor. In: LEEN, T. (Hrsg.) ; DIETTERICH, T. (Hrsg.) ; TRESP, V. (Hrsg.): *Advances in Neural Information Processing Systems* Bd. 13, MIT Press, 2001, S. 294–300. – URL <https://proceedings.neurips.cc/paper/2000/file/0950ca92a4dcf426067cfd2246bb5ff3-Paper.pdf>
- [96] RASMUSSEN, Carl E. ; WILLIAMS, C. K. I.: *Gaussian Processes for Machine Learning*. Cambridge, MA, USA : MIT Press, 2006 (Adaptive Computation and Machine Learning). – URL <http://www.gaussianprocess.org/gpml/chapters/RW.pdf>
- [97] REŞİTOĞLU, İbrahim A. ; ALTINIŞIK, Kemal ; KESKIN, Ali: The pollutant emissions from diesel-engine vehicles and exhaust aftertreatment systems. In: *Clean Technologies and Environmental Policy* 17 (2015), Januar, Nr. 1, S. 15–27. – URL <https://doi.org/10.1007/s10098-014-0793-9>. – Zugriffsdatum: 2021-02-25. – ISSN 1618-9558
- [98] ROY, Christopher J. ; OBERKAMPF, William L.: A comprehensive framework for verification, validation, and uncertainty quantification in scientific computing. In: *Computer Methods in Applied Mechanics and Engineering* 200 (2011), Nr. 25–28, S. 2131–2144. – URL <http://www.sciencedirect.com/science/article/pii/S0045782511001290>. – ISSN 0045-7825
- [99] SAMMUT, Claude (Hrsg.) ; WEBB, Geoffrey I. (Hrsg.): *Encyclopedia of Machine Learning and Data Mining*. 2. Boston, MA, USA : Springer, 2017. – URL <https://doi.org/10.1007/978-1-4899-7687-1>. – ISBN ISBN 978-1-4899-7685-7
- [100] SCHILLINGER, Mark ; HARTMANN, Benjamin ; JACOB, Martin: Dynamic Safe Active Learning for Calibration. In: RÖPKE, Carsten (Hrsg.) ; GÜHMANN, Clemens (Hrsg.): *International Conference on Calibration Methods and Automotive Data Analytics*. Tübingen, Germany : Expert Verlag, 2019, S. 258–277
- [101] SCHILLINGER, Mark ; MOURAT, Kadir ; HARTMANN, Benjamin ; ECKSTEIN, Carola ; JACOB, Martin ; KLOPPENBURG, Ernst ; NELLES, Oliver: Modern Online DoE Methods for Calibration – Constraint Modeling, Continuous Boundary Estimation, and Active Learning. In: RÖPKE, Carsten (Hrsg.) ; GÜHMANN, Clemens (Hrsg.): *Automotive Data Analytics, Methods and Design of Experiments (DOE)*. Tübingen, Germany : Expert Verlag, 2017, S. 329–344

- [102] SCHREIBER, Alexander: Dynamische Motorvermessung mit verschiedenen Methoden und Modellen. In: ISERMANN, Rolf (Hrsg.): *Elektronisches Management motorischer Fahrzeugantriebe: Elektronik, Modellbildung, Regelung und Diagnose für Verbrennungsmotoren, Getriebe und Elektroantriebe*. 1. Wiesbaden : Vieweg+Teubner, 2010, S. 167–199. – URL https://doi.org/10.1007/978-3-8348-9389-5_7. – ISBN 978-3-8348-9389-5
- [103] SCHÄUFFELE, Jörg ; ZURAWKA, Thomas: Methoden und Werkzeuge in der Entwicklung. In: *Automotive Software Engineering*. 6. Springer Vieweg, 2016 (ATZMTZ), S. 207–326. – URL <https://doi.org/10.1007/978-3-658-11815-0>. – ISBN 978-3-658-11814-3
- [104] SEQUENZ, Heiko: *Emission Modelling and Model-based Optimisation of the Engine Control*. VDI-Verlag, 2013 (Berichte aus dem Institut für Automatisierungstechnik und Mechatronik der TU Darmstadt). – URL <https://tuprints.ulb.tu-darmstadt.de/3948/>. – ISBN 978-3-18-522208-5
- [105] SHAHRIARI, B. ; SWERSKY, K. ; WANG, Z. ; ADAMS, R. P. ; FREITAS, N. de: Taking the Human Out of the Loop: A Review of Bayesian Optimization. In: *Proceedings of the IEEE* 104 (2016), Nr. 1, S. 148–175. – URL <https://ieeexplore.ieee.org/document/7352306>. – ISSN 0018-9219
- [106] SIEBERTZ, Karl ; BEBBER, David van ; HOCHKIRCHEN, Thomas: Versuchspläne. In: *Statistische Versuchsplanung: Design of Experiments (DoE)*. 2. Berlin, Heidelberg : Springer Vieweg, 2017 (VDI-Buch), S. 27–59. – URL https://doi.org/10.1007/978-3-662-55743-3_2. – Zugriffsdatum: 2021-02-15. – ISBN 978-3-662-55743-3
- [107] SOBOL, Ilya M.: On the Distribution of Points in a Cube and the Approximate Evaluation of Integrals. In: *USSR Computational Mathematics and Mathematical Physics* 7 (1967), Nr. 4, S. 86–112. – Place: Amsterdam, Netherlands Publisher: Elsevier. – ISSN 0041-5553
- [108] SOBOL, Ilya M. ; ASOTSKY, Danil ; KREININ, Alexander ; KUCHERENKO, Sergei: Construction and Comparison of High-Dimensional Sobol’ Generators. In: *Wilmott* 2011 (2011), Nr. 56, S. 64–79. – URL <https://onlinelibrary.wiley.com/doi/abs/10.1002/wilm.10056>. – Place: Hoboken, NJ, USA Publisher: John Wiley and Sons
- [109] STONE, M.: Cross-Validatory Choice and Assessment of Statistical Predictions. In: *Journal of the Royal Statistical Society. Series B (Methodological)* 36 (1974), Nr. 2, S. 111–133. – URL <https://rss.onlinelibrary.wiley.com/doi/abs/10.1111/j.2517-6161.1974.tb00994.x>
- [110] SUNG, Alexander ; KLÖPPER, Florian ; MITTERER, Alexander ; WACHTMEISTER, Georg ; ZELL, Andreas: Modellbasierte Online-Optimierung in der Simulation und am Motorenprüfstand. In: *MTZ - Motortechnische Zeitschrift* 68 (2007), Januar, Nr. 1, S. 42–48. – URL <https://doi.org/10.1007/BF03225444>. – ISSN 2192-8843
- [111] THE GPY AUTHORS: *GPY: A Gaussian process framework in python*. 2012. – URL <http://github.com/SheffieldML/GPy>

- [112] TIETZE, Nils: *Model-based Calibration of Engine Control Units Using Gaussian Process Regression*. Darmstadt, Technische Universität Darmstadt, Doctoral Thesis, 2015. – URL <https://tuprints.ulb.tu-darmstadt.de/id/eprint/4572>. – Zugriffsdatum: 2021-10-03
- [113] VEHTARI, Aki ; TOLVANEN, Ville ; MONONEN, Tommi ; WINTHER, Ole: Bayesian Leave-One-Out Cross-Validation Approximations for Gaussian Latent Variable Models. In: *Journal of Machine Learning Research* 17 (2016), Juni, S. 3581–3618. – URL <http://jmlr.org/papers/volume17/14-540/14-540.pdf>. – Place: Brookline, MA, USA Publisher: Microtome
- [114] VIRTANEN, Pauli ; GOMMERS, Ralf ; OLIPHANT, Travis E. ; HABERLAND, Matt ; REDDY, Tyler ; COURNAPEAU, David ; BUROVSKI, Evgeni ; PETERSON, Pearu ; WECKESER, Warren ; BRIGHT, Jonathan ; WALT, Stéfan J. van der ; BRETT, Matthew ; WILSON, Joshua ; JARROD MILLMAN, K. ; MAYOROV, Nikolay ; NELSON, Andrew R. J. ; JONES, Eric ; KERN, Robert ; LARSON, Eric ; CAREY, CJ ; POLAT, İlhan ; FENG, Yu ; MOORE, Eric W. ; VANDERPLAS, Jake ; LAXALDE, Denis ; PERKTOLD, Josef ; CIRMANN, Robert ; HENRIKSEN, Ian ; QUINTERO, E. A. ; HARRIS, Charles R. ; ARCHIBALD, Anne M. ; RIBEIRO, Antônio H. ; PEDREGOSA, Fabian ; MULBREGT, Paul van ; SCIPY 1.0 CONTRIBUTORS: SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. In: *Nature Methods* 17 (2020), S. 261–272. – URL <https://rdcu.be/b08Wh>
- [115] WASSERBURGER, A. ; DIDCOCK, N. ; HAMETNER, C.: Efficient real driving emissions calibration of automotive powertrains under operating uncertainties. In: *Engineering Optimization* (2021), Oktober, S. 1–18. – URL <https://doi.org/10.1080/0305215X.2021.1989589>. – Publisher: Taylor & Francis. – ISSN 0305-215X
- [116] WASSERBURGER, A. ; HAMETNER, C. ; DIDCOCK, N.: Risk-averse Real Driving Emissions Optimization Considering Stochastic Influences. In: *Engineering Optimization* (2020), S. 1–17. – Place: London, England Publisher: Taylor & Francis
- [117] WILLIAMS, C. K. I.: Prediction with Gaussian Processes: From Linear Regression to Linear Prediction and Beyond. In: JORDAN, Michael I. (Hrsg.): *Learning in Graphical Models*. Berlin & Heidelberg, Germany : Springer, 1998, S. 599–621. – ISBN 978-94-011-5014-9
- [118] WILLIAMS, C. K. I. ; RASMUSSEN, Carl E.: Gaussian Processes for Regression. In: TOURETZKY, David S. (Hrsg.) ; MOZER, Michael C. (Hrsg.) ; HASSELMO, Michael E. (Hrsg.): *Proceedings of the 8th International Conference on Neural Information Processing Systems*. Cambridge, MA, USA : MIT Press, 1996 (NIPS'95), S. 514–520. – URL <http://dl.acm.org/citation.cfm?id=2998828.2998901>. – Backup Publisher: Max-Planck-Gesellschaft
- [119] WINTRICH, Thomas ; ROTHE, Stefan ; BUCHER, Kilian ; HITZ, Hans-Jürgen: Diesel Injection System with Closed-loop Control. In: *MTZ worldwide* 79 (2018), Nr. 9, S. 54–59. – URL <https://rd.springer.com/article/10.1007/978-3-331-3018-0062-0>. – Place: Berlin & Heidelberg, Germany Publisher: Springer. – ISSN 2192-9114

- [120] WONG, Pak K. ; GAO, Xiang H. ; WONG, Ka I. ; VONG, Chi M.: Online Extreme Learning Machine Based Modeling and Optimization for Point-by-point Engine Calibration. In: *Neurocomputing* 277 (2018), Februar, S. 187–197. – Place: Amsterdam, Netherlands Publisher: Elsevier. – ISSN 0925-2312
- [121] WOOLDRIDGE, Jeffrey M.: The Simple Regression Model. In: *Introductory Econometrics: A Modern Approach*. 7. Boston, MA, USA : Cengage Learning, 2018, S. 20–65. – ISBN 978-1-337-55886-0
- [122] YOON, Jaehong ; HWANG, Sung J.: Combined Group and Exclusive Sparsity for Deep Neural Networks. In: PRECUP, Doina (Hrsg.) ; TEH, Yee W. (Hrsg.): *Proceedings of the 34th International Conference on Machine Learning* Bd. 70. International Convention Centre, Sydney, Australia : PMLR, August 2017, S. 3958–3966. – URL <http://proceedings.mlr.press/v70/yoon17a.html>