INVERSE-DISTANCE INTERPOLATION BASED SET-POINT GENERATION METHODS FOR CLOSED-LOOP COMBUSTION CONTROL OF A CIDI ENGINE

Thesis

Presented in Partial Fulfillment of the Requirements for the Degree Master of Science in the Graduate School of The Ohio State University

By

Rajaram Seshu Maringanti, B.E. (Hons.)

Graduate Program in Mechanical Engineering

The Ohio State University

2009

Thesis Committee:

Professor Giorgio Rizzoni, Advisor
Dr. Shawn Midlam-Mohler, Co-advisor
Professor Steve Yurkovich
Professor Yann Guezennec
ABSTRACT

Closed-loop control of combustion is of great importance for conventional diesel engines in order to reduce the deterioration in engine performance and emissions caused by different sources of variability. Diesel engine emissions and performance are affected by variability in the combustion root-cause variables like air mass, residual mass, injection parameters etc. If combustion can be referenced based on the root-cause combustion variables, the engine performance and emissions can be improved. Such a referencing, however, leads to increased calibration effort for the combustion controller due to increase in the number of scheduling variables for generating such references as well as due to the increase in the number of references. Conventional methods for generating the set-points for the closed-loop combustion controller are not suitable as they become impractical as the number of scheduling variables increases. In this work, inverse-distance based interpolation methods have been developed to generate the set-points for the closed-loop combustion controller. The inverse-distance interpolation method provides the advantage of reduced calibration effort and can be extended to multiple-dimensional interpolation without significant increase in computational effort.

In this work, a closed-loop combustion control architecture has been developed for a heavy-duty diesel engine. The inverse-distance based calibration method has been demonstrated for a simplified version of the closed-loop combustion control architecture. The method involves an optimization approach that generates a scattered set of engine operating points where the engine should be calibrated. The inverse-distance interpolation method can interpolate with the scattered calibration data set to generate the set-points for the closed-loop combustion controller. The proposed method showed a great potential to improve the calibration effort when compared to conventional calibration methods that can be applied to the set-point generation problem. The inverse-distance based calibration approach resulted in a calibration effort
that is 5 to 25 times less than that of a conventional method and resulted in engine performance and emissions that are comparable to those of the conventional method.
DEDICATION

Dedicated to my parents and loved ones.
ACKNOWLEDGEMENTS

First of all, I would like to thank Cummins Engine Company for providing the monetary support which made this work possible. I would like to specially thank Ms. Joan Wills, Dr. John N Chi and Dr. Vivek Sujan of Cummins Engine Company for providing the opportunity to work on this project.

This work would not have been possible without the guidance of many people at the Center for Automotive Research in The Ohio State University. Many thanks to my co-adviser, Dr. Shawn Midlam-Mohler who has guided me on a day-to-day basis during the course of the project. His vast experience in the field of internal combustion engines was of immense help and he was always available with advice whenever it was needed. During this time, I learned many important things from him which helped me develop my technical skills as well as my personality.

My sincere thanks are also due to faculty members, research staff and students at the Center for Automotive Research. In particular, I thank Prof. Giorgio Rizzoni for giving me an opportunity to do my graduate research work at the center. My thanks are due to Prof. Steve Yurkovich and Prof. Yann Guezenec for their valuable feedback during my Master’s examination. My thanks are also due to the research staff members, Dr. Marcelo Canova and Dr. Fabio Chiara for their valuable technical advice during the course of the project. I would also like to thank graduate students Ming Fang, Jason Meyers and Chris Hoops who worked with me on the project and helped me whenever I was in need.

Last but not the least, my family and loved ones deserve many thanks from me for all the support that they extended to me during my graduate studies. Nothing would have been possible without their love and support.
VITA

June, 2004 ................................................. B.E. (Hons.) Mechanical Engineering,
Birla Institute of Technology and Science, Pilani, India

August, 2004 to July, 2007............................ Research Engineer,
Mercedes-Benz Research & Development India,
Bangalore, India

PUBLICATIONS


FIELD OF STUDY

Major Field: Mechanical Engineering
# TABLE OF CONTENTS

Abstract........................................................................................................... ii  
Dedication......................................................................................................... iv  
Acknowledgments........................................................................................... v  
Vita.................................................................................................................. vi  
List of Tables..................................................................................................... x  
List of Figures.................................................................................................... xii  

Chapters:

1 INTRODUCTION.......................................................................................... 1  

2 LITERATURE REVIEW .............................................................................. 6  

2.1 Introduction to Literature Review............................................................ 6  
2.2 Control of Diesel Engines ...................................................................... 6  
2.2.1 Fuel Injection System Control.............................................................. 7  
2.2.2 Air and EGR Control System............................................................... 8  
2.3 Closed-loop Combustion Control of Diesel Engines.............................. 12  
2.3.1 Cylinder Pressure-based Combustion Control................................... 18  
2.3.1.1 Heat Release Rate Based Combustion Metrics................................. 19  
2.3.1.2 Combustion Metrics Calculated Directly from Cylinder Pressure......... 25  
2.3.2 Control Methods for Combustion Control.......................................... 28  
2.3.3 Cylinder Pressure Data Processing for Combustion Control in Real-time 29  
2.3.4 Alternative Sensing Technologies for Closed-loop Combustion Control 33  
2.3.4.1 Ion current sensors ..................................................................... 33  
2.3.4.2 Accelerometers.......................................................................... 35  
2.3.4.3 Speed Sensors............................................................................ 36  
2.3.4.4 Torque Sensors........................................................................... 37  
2.4 Advancements in Diesel Engine Calibration Methods............................ 37  
2.4.1 Kernel based Interpolation Methods.................................................... 42  
2.4.1.1 Inverse-Distance based Interpolation Methods................................. 43  
2.4.1.2 Shortcomings of Inverse-Distance based Interpolation....................... 45  
2.5 Conclusions from the Literature Survey ................................................. 47  

3 TOOLS AND TECHNIQUES ........................................................................ 50
4 SENSITIVITY ANALYSIS ON A HIGH-FIDELITY DIESEL ENGINE MODEL

4.1 Introduction to Sensitivity Analysis .................................................................77
4.2 High Fidelity Diesel Engine Model ..................................................................77
  4.2.1 Validation of the High Fidelity Diesel Engine Model ....................................80
4.3 Sources of Variability in Engine Operation ......................................................86
  4.3.1 Design-Driven Sources of Variability ..........................................................86
  4.3.2 Component-Driven Variability .......................................................................92
  4.3.3 Environmental-Driven Variability .................................................................92
4.4 Effects of Engine Transience on Performance ................................................93
4.5 Sensitivity Analysis through Design of Experiments ........................................94
  4.5.1 Design of Experiments ..................................................................................96
4.6 ANOVA Analysis on the DOE Results ...............................................................98
  4.6.1 ANOVA analysis on ISFC and NOx ..............................................................101
  4.6.2 ANOVA Analysis on Combustion Metrics ....................................................104
  4.6.2.1 ANOVA Analysis on CA50 calculated from Gross Heat Release ...............105
  4.6.2.2 ANOVA analysis on CA50 calculated from Net Heat Release .................107
  4.6.2.3 ANOVA analysis on Rassweiler-Withrow Method .................................109
  4.6.2.4 Analysis of Maximum Pressure and Maximum Rate of Pressure Rise ........112
  4.6.2.5 ANOVA analysis on Combustion Pressure .................................................114
4.7 Conclusions from the ANOVA Analysis ........................................................117

5 CLOSED-LOOP COMBUSTION CONTROL ARCHITECTURE ..........................119
  5.1 Definition of the Closed-loop Combustion Control Problem .............................119
  5.2 Hypothetical Closed-loop Combustion Control Architecture ...........................121
  5.3 Closed-loop Combustion Control Architecture for Cummins Heavy-duty Diesel Engine ............123
LIST OF TABLES

Table                                      Page
Table 1: Calibration parameters of a diesel engine that has to meet 2010 emission regulations ............ 38
Table 2: Example of 1-way ANOVA........................................................................................................... 62
Table 3: One-way ANOVA Table.................................................................................................................. 63
Table 4: Mileage of different models of cars manufactured by different factories ................................. 65
Table 5: Summary of Cylinder-to-cylinder variability .............................................................................. 92
Table 6: 13-modes in terms of engine speed and load and the corresponding weighting factors........... 95
Table 7: Factors for Design of Experiments with Variation and Levels .................................................. 97
Table 8: Percentage Contribution of each factor to total variance ......................................................... 100
Table 9: Number of Neurons in the Hidden Layers for the Neural Networks Trained ......................... 151
Table 10: Factors chosen for DOE on the RGM ..................................................................................... 156
Table 11: Average Distance of Center of Mass for different values of $M$ ............................................ 164
Table 12: Comparison of Mean Center of Mass coordinates for different values of $M$ ...................... 169
Table 13: Comparison of Standard Deviations in Center of Mass coordinates for different values of $M$ ... 169
Table 14: Comparison of FTP Cycle integrated NOx and fuel consumption between the baseline and the conventional calibration case

Table 15: Comparison of FTP cycle integrated NOx and fuel consumption between the baseline and the inverse-distance method for exact FTP profiles

Table 16: Comparison of FTP cycle integrated NOx and fuel consumption between the baseline and the inverse-distance method for randomized FTP profiles
## LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Figure 1</td>
<td>Emission Standards for NOx and PM in the USA [DieselNet, 2007]</td>
<td>1</td>
</tr>
<tr>
<td>Figure 2</td>
<td>Diesel engine system schematic (after [1])</td>
<td>7</td>
</tr>
<tr>
<td>Figure 3</td>
<td>Common Rail Diesel Injection System (after [2])</td>
<td>8</td>
</tr>
<tr>
<td>Figure 4</td>
<td>Effect of Intake Manifold Conditions on Diesel Engine Performance and Emissions (after [4])</td>
<td>9</td>
</tr>
<tr>
<td>Figure 5</td>
<td>Effect of Cooled EGR on NOx and PM (after [4])</td>
<td>11</td>
</tr>
<tr>
<td>Figure 6</td>
<td>NOx and soot formation regions on φ-T map (after [6])</td>
<td>12</td>
</tr>
<tr>
<td>Figure 7</td>
<td>(a) Fuel Economy and Emissions as a function of PRM10 at zero EGR levels. (b) Economy and Emissions as a function of PRM10 at heavy EGR levels. (after [7])</td>
<td>14</td>
</tr>
<tr>
<td>Figure 8</td>
<td>(a) Desired change in heat release due to change in valve opening strategy (b) Intake and Exhaust valve opening profiles for desired change in heat release (after [9])</td>
<td>16</td>
</tr>
<tr>
<td>Figure 9</td>
<td>Sensitivity of NOx and soot to variability in operating conditions (after [13])</td>
<td>18</td>
</tr>
<tr>
<td>Figure 10</td>
<td>Effect of Heat Transfer on Combustion Heat Release (After [16])</td>
<td>20</td>
</tr>
<tr>
<td>Figure 11</td>
<td>Cylinder-cylinder variation of CA50 in different cases (after [18])</td>
<td>22</td>
</tr>
</tbody>
</table>
Figure 12: Comparison between heat release calculated from RW method and AHR method (after [20]) ..24

Figure 13: Comparison between Heat Release Rate and Diesel Pressure Ratio Model (after [20]) ..........25

Figure 14: Fired, Motoring and Combustion pressure traces of a conventional diesel engine (after [20]) ...26

Figure 15: (a) Center of gravity and secant length metrics (b) Correlation of center of gravity metric to diesel engine injection parameters (after [21])..........................................................27

Figure 16: Comparison of Available Time and the Computation Time for IMEP and ROHR calculations (after [23])..................................................................................................................31

Figure 17: (a) Heat release rate calculation from a noisy pressure signal (b) Performance of RW and DPR methods with a noisy pressure signal (after [20]) .................................................................32

Figure 18: Comparison of ion current and pressure signals in conventional diesel combustion (after [24]) 34

Figure 19: Comparison of cylinder pressure and accelerometer signals for a 4-cylinder engine (after [28])36

Figure 20: Design of Experiments Process (after [34]) ........................................................................39

Figure 21: NOx and Brake Thermal Efficiency Trade-off Using Model-Based Calibration Methods (after [33]) ........................................................................................................................................41

Figure 22: Illustration of the affect of even-valued exponent on the interpolated surface .................44

Figure 23: Illustration of the affect of odd-valued exponent on the interpolated surface ...............45

Figure 24. (a) Illustration of the ‘flat spots’ of the interpolated surface using original inverse-distance interpolation. (b) Illustration of the effect of replacing the interpolated value at the data points by their tangent corresponding planes ..................................................................................47
Figure 25: Multi-zone approach for combustion modeling ................................................................. 55

Figure 26: Finding the solution to $F(x) = 0$ using bisection method with $a_1$ and $b_1$ as the initial conditions. The solution is found by bisecting the search range in each step .............................................. 58

Figure 27: One-way ANOVA Table for the Shipment Problem generated by MATLAB ...................... 64

Figure 28: Two-way ANOVA results for the car mileage problem ....................................................... 66

Figure 29: N-way ANOVA analysis ....................................................................................................... 67

Figure 30: Structure of each neuron in an ANN ..................................................................................... 69

Figure 31: Different activation functions used in ANNs ........................................................................ 70

Figure 32: A multi-layer feed-forward network .................................................................................... 73

Figure 33: Genetic Algorithm Optimization Process ............................................................................. 75

Figure 34: Snapshot of Detailed GT-Power Engine Model ................................................................. 78

Figure 35: Engine Calibration Data Points ............................................................................................. 82

Figure 36: Comparison of Brake Torque ............................................................................................... 83

Figure 37: Comparison of Intake Manifold Pressure and Temperature .................................................. 83

Figure 38: Comparison of Charge Flow and EGR Flow ..................................................................... 84

Figure 39: Comparison of Turbine Inlet Pressure and Temperature ..................................................... 84

Figure 40: Comparison of Experimental and Predicted Indicated Specific Fuel Consumption ............. 85
Figure 41: Comparison of Experimental and Predicted NOx Concentration ........................................85
Figure 42: Comparison of Experimental and Predicted Soot .................................................................86
Figure 43: Intake Manifold Design .........................................................................................................88
Figure 44: Distribution of Fresh Air in Cylinders 1 to 6 ........................................................................90
Figure 45: Distribution of Residuals in Cylinders 1 to 6 .....................................................................90
Figure 46: Distribution of Total Trapped Mass in Cylinders 1 to 6 .........................................................91
Figure 47: Distribution of Charge Temperature in Cylinders 1 to 6 .....................................................91
Figure 48: Selection of available calibrated points closest to the 13-modes ............................................96
Figure 49: ANOVA Table for Mode 1 .....................................................................................................99
Figure 50: Sensitivity of NOx to engine root-cause variables ...............................................................101
Figure 51: Modes represented on engine load speed ..........................................................................102
Figure 52: Sensitivity of ISFC to engine root-cause variables ..............................................................103
Figure 53: Sensitivity of CA50_{Gross} of cylinder 1 on engine root-cause variables .........................106
Figure 54: Sensitivity of CA50_{Gross} of cylinder 6 on engine root-cause variables .........................107
Figure 55: Sensitivity of CA50_{Net} of cylinder 1 on engine root-cause variables ..........................108
Figure 56: Sensitivity of CA50_{Net} of cylinder 6 on engine root-cause variables ..........................109
Figure 57: Sensitivity of MFB50_{RW} of cylinder 1 on engine root-cause variables .........................111
Figure 58: Sensitivity of MFB50_{RW} of cylinder 1 on engine root-cause variables ........................................111

Figure 59: Cylinder Pressure traces at all the 11 modes .................................................................113

Figure 60: Rate of pressure rise in all the 11 modes ........................................................................113

Figure 61: Gross Heat Release Rates for all the 11 modes ...............................................................114

Figure 62: Comparison of actual cylinder pressure, compression pressure and combustion pressure ........115

Figure 63: Sensitivity of Maximum Combustion Pressure of cylinder 1 to engine root-cause variables ....116

Figure 64: Sensitivity of Combustion Pressure of cylinder 1 to engine root-cause variables ...............117

Figure 65: Hypothetical Closed-loop Combustion Control Architecture ...........................................121

Figure 66: Cummins Heavy-duty Diesel Engine Control Architecture .............................................125

Figure 67: Closed-loop Combustion Control Architecture for Cummins Heavy-duty Diesel Engine ........131

Figure 68: Depiction of Transient operation of an engine .................................................................133

Figure 69: Set-point generator output for the proposed control .........................................................134

Figure 70: Inputs and Outputs of the Set-Point Generator Scheme ...................................................134

Figure 71: Neighborhood definition for Inverse-Distance Interpolation .........................................139

Figure 72: Center of Mass Demonstration .......................................................................................140

Figure 73: Calibration Procedure for the Inverse-distance Set-Point Generation Method .................144

Figure 74: Comparison of EGR flow, MAP of RGM and DGM .......................................................148
Figure 75: Comparison of ISFC and IMEP of RGM and DGM ................................................................. 148

Figure 76: Structure of the Neural Network with 2 inputs, 4 hidden layers and 3 outputs ......................... 149

Figure 77: Error Histogram for Neural Network prediction of NOx ......................................................... 151

Figure 78: Error Histogram for Neural Network prediction of ISFC ...................................................... 152

Figure 79: Error Histogram for Neural Network prediction of IMEP ...................................................... 152

Figure 80: Simplified Combustion Control Architecture for Validation .................................................. 154

Figure 81: Normalized ISFC as a function of normalized MAP and EGR flow ....................................... 159

Figure 82: Normalized NOx as a function of normalized MAP and EGR flow ....................................... 159

Figure 83: Objective function values as a function of normalized MAP and EGR flow ......................... 160

Figure 84: Distance of Center of Mass (in %) as a function of the total number of calibrations, $M$ ........ 165

Figure 85: Histogram of Number of Neighbors for different values of $M$ .................................................. 166

Figure 86: Distribution of Coordinates of the Center of Mass for $M = 100$ ............................................ 167

Figure 87: Distribution of Coordinates of the Center of Mass for $M = 200$ ............................................ 167

Figure 88: Distribution of Coordinates of the Center of Mass for $M = 304$ ............................................ 168

Figure 89: Distribution of Coordinates of the Center of Mass for $M = 500$ ............................................ 168

Figure 90: Effect of Neighborhood bounds on Average Distance of Center of Mass ............................ 171
Figure 91: Comparison of Cumulative NOx emissions using exact FTP profiles for inverse-distance method
.........................................................................................................................................176

Figure 92: Comparison of total fuel consumption using exact FTP profiles for inverse-distance method.176

Figure 93: Comparison of cycle-integrated NOx emissions and fuel consumption for exact FTP profiles 177

Figure 94: Comparison of cycle-integrated NOx emissions and fuel consumption for randomized FTP profiles
.........................................................................................................................................177

Figure 95: A Transient Maneuver in the FTP cycle .................................................................179

Figure 96: Response of the conventional and inverse-distance set-point generators......................179

Figure 97: Normalized NOx emissions during the transients.....................................................180

Figure 98: Comparison of Calibration Effort for $p = 6$ ............................................................183

Figure 99: Comparison of Calibration Effort for $p = 8$ ............................................................184

Figure 100: Comparison of Calibration Effort for $p = 10$ ......................................................184
INTRODUCTION

Diesel engines are primarily used for heavy-duty transportation and commercial applications due to higher thermal efficiency compared to gasoline engines. Unfortunately, diesel engines have the disadvantage of increased nitric oxides (NOx) and particulate matter (PM) emissions which are directly linked to a wide range of environmental and health effects. As a result, governments across the world are tightening the regulations on emissions at a drastic rate. Figure 1 shows the regulations set by the Environmental Protection Agency of the United States on NOx and PM emissions for heavy-duty diesel engines. It can be observed that the NOx and PM emissions are expected to be almost negligible by the year 2010.

Figure 1: Emission Standards for NOx and PM in the USA [DieselNet, 2007]
The decrease in emissions like NOx and PM is often linked to decrease in the fuel efficiency of the diesel engines which is highly undesirable for engine manufacturers. Therefore, engine manufacturers are developing new technologies for curtailing emissions by not sacrificing the fuel efficiency of the engine. A variety of exhaust after-treatment devices are being used to meet the regulations. Diesel Particulate Filters (DPFs) are being used for removing the particulate matter and Selective Catalytic Reduction (SCR) catalysts are being used to curtail NOx emissions. The disadvantages with the after-treatment systems are their high costs and the increase in complexity of engine control. For example, the temperature of the exhaust gases plays a very important role in the efficiency of the exhaust after-treatment which demands a complex thermal management system in the engine. While the exhaust after-treatment systems are being continuously developed, the engine manufacturers are also looking at developing technologies that can modify the combustion process and reduce engine emissions considerably reducing the need for after-treatment systems. Exhaust gas recirculation (EGR) is one of the most effective methods to reduce NOx emissions as it reduces the peak temperature of combustion. But, hot EGR can increase the PM emissions as it results in reduction in air fuel ratio in the cylinders. Therefore, cooled EGR is being used as a strategy to decrease both NOx and PM emissions in modern diesel engines.

Engine manufacturers are also interested in alternative combustion modes like Homogenous Charge Compression Ignition (HCCI), Pre-mixed Charge Compression Ignition (PCCI), and Low Temperature Combustion (LTC) which have the potential to reduce both NOx and soot emissions considerably. Most of the alternative combustion modes rely on a homogenous and pre-mixed charge self-igniting in the combustion chamber leading to considerably reduced NOx and soot emissions. These combustion regimes are highly unstable and therefore have to be monitored using some kind of feedback from combustion. Due to the lack of a trigger for combustion like a fuel injection event or a spark event, it is necessary to have a direct or an indirect feedback to assess the progress of combustion. Closed-loop combustion control is therefore being developed by many researchers to gain more control authority on the alternative combustion modes.
Closed-loop combustion control can also be advantageous for conventional diesel combustion as it can help in reducing the dispersions in engine emissions and performance due to various sources of variability. Combustion in diesel engines gets affected due to various sources of variability like cylinder-to-cylinder variations, engine-to-engine variations, and variability in ambient conditions, fuel properties etc. The variability in combustion leads to deteriorated performance as the engine is calibrated assuming no variability. The effects of the variability on combustion can be compensated with the help of closed-loop combustion control. This helps in reducing the variability in engine performance and emissions from the calibrated values. The design of a combustion controller involves finding the appropriate combustion metrics that are well correlated to engine performance and emissions. This is because the technology for measuring emissions like NOx and PM has not matured enough for on-the-road vehicle application. Another important aspect of combustion control is the identification of the actuators that can control combustion. For a modern diesel engine, there are many possibilities for the actuators like EGR fraction, boost pressure, multiple injections etc. The choice of the actuator and combustion metrics depends on the objective of the combustion controller.

Feedback for combustion control can be achieved through direct or indirect means. The most direct method of obtaining combustion feedback is with the help of cylinder pressure measurements. Other direct methods like ion sensing probes are also possible but have limited range of operability for conventional combustion. Information about combustion progress can also be obtained through indirect means like measuring the vibrations of the cylinder block, fluctuations in engine speed caused by combustion etc. But these methods need a lot of processing and do not always yield reliable information about combustion. Cylinder pressure sensing is the most robust method to achieve combustion feedback. But, high costs have made cylinder pressure sensors prohibitive for use in production. However, engine manufacturers hope that the costs of high quality pressure sensors will come down with their increased use and developments in sensor technology. Therefore, cylinder-pressure based combustion control has become a primary focus of research in the recent times.
The tightened regulations on emissions and the competitive need to maintain high fuel efficiency are forcing the diesel engine manufacturers to explore new technologies like closed-loop combustion control. As a result, the number of degrees of freedom of diesel engines is increasing rapidly leading to tremendous increase in the calibration effort of the engines. Modern diesel engines have anywhere between 10-20 calibrated parameters which include parameters from the air handling unit, fuel injection system, exhaust after-treatment systems etc. For technologies like combustion control, the number of calibrated parameters increases further due to the need to optimize combustion in individual cylinders. The increase in the number of calibrated parameters also poses challenges in the form of increased memory requirements for the engine electronic control units. Most of the diesel engine control depends on calibration maps which are generated by an elaborate calibration procedure during the developmental phase of the engine. Conventional calibration techniques involve conducting experiments on the engine and optimizing its performance at each operating point. As the number of degrees of freedom increases, the number of operating points where experiments are conducted increases tremendously. Added to the normal operation of the engine, calibrations also need to be generated for specific operating conditions like changes in ambient conditions, rapid transients etc. The base calibration tables are corrected to address the special conditions in an ad-hoc way. Therefore, to reduce the calibration effort, various assumptions are made during calibration which makes them sub-optimal leading to deterioration in engine performance.

Engine manufacturers are looking at alternatives to conventional calibration methods. Model-based calibration is one of the alternatives that can help reducing the time spent in the test-cell during engine calibration. Models of the engine can be used to calibrate the engine in a simulation setup. In addition to the calibration effort, the engine manufacturers are also exploring methods to reduce the memory requirements for storing the calibrations on the engine ECUs. Conventional calibrations are in the form of multi-dimensional tables which pose memory and computational issues for engine ECUs. Therefore, scattered data interpolation methods that can handle interpolation in multiple dimensions without a great increase in computational complexity are of immense value. Another area of focus is to generate calibrations that are optimal for both steady-state and transient operation of engine. Modern diesel engines are expected to meet
emission and performance regulations both in steady-state and transient conditions. Improvements in various aspects of the conventional calibration techniques has a direct impact on expediting the deployment of advanced technologies like closed-loop combustion control.

In this study, concepts have been developed for closed-loop combustion control of a Cummins heavy duty diesel engine which operates in the conventional diesel combustion mode. The first part of the study deals with studying the sensitivity of diesel combustion to various sources of variability. This is done with the help of a high-fidelity model of the diesel engine. The sensitivity of engine emissions, performance and various combustion metrics have been identified during the sensitivity study. Depending on the sensitivity study, an architecture for closed-loop combustion control for the diesel engine has been proposed. The later part of the study focuses on a set-point generation scheme for the combustion controller. The scheme chosen for set-point generation aims at minimizing the calibration effort and improving the performance of the engine. A kernel based interpolation method has been developed to generate the set-points using a scattered set of calibrations. An optimization approach has been developed to identify the operating points where engine calibrations need to be performed. The kernel based set-point generation method has been compared to conventional calibration methods in terms of calibration effort and engine performance. In the end, conclusions are drawn with respect to the efficacy of the set-point generation method and how it can be improved and extended to the full-fledged closed-loop combustion control problem.
LITERATURE REVIEW

2.1 Introduction to Literature Review

This chapter discusses the outcome of a literature survey conducted on various topics on diesel engines that are relevant to the thesis. The focus of the thesis is on developing kernel based calibration methods for closed-loop combustion control of a conventional diesel engine. The goal of the literature survey is to identify publications in the areas of modern diesel engine control, closed-loop combustion control of engines and advancements in calibration methods for modern diesel engines. In the first part of the literature survey, publications on diesel engine control have been identified. The literature survey for this part is not very elaborate and its purpose is to identify the most important aspects of diesel engine control that are directly relevant to closed-loop combustion control. In the second part, recent works on application of closed-loop combustion control methods to different kinds of internal combustion engines have been reported. In the third part, publications have been studied that address the calibration methods that are being developed by the diesel engine community in order to meet the challenges of increased calibration effort of modern day diesel engines. In the end, conclusions are drawn regarding the state of research.

2.2 Control of Diesel Engines

Modern diesel engines use a variety of technologies to improve fuel efficiency and decrease emissions. These technologies include exhaust gas recirculation system (EGR), variable geometry turbocharger (VNT), common rail diesel injection system (CRDI) and a variety of exhaust gas after-treatment devices to achieve better fuel efficiencies while limiting emissions. The overall objective of such a diesel engine control is to provide the desired torque at the highest efficiency while meeting the exhaust gas and noise emission requirements. This requires an optimal coordination of the injection, EGR and air handling systems under stationary and dynamic conditions [1]. Figure 2 shows a schematic of the modern day diesel
engine system. The most important aspects of diesel engine control that are directly relevant to the focus of the thesis are fuel injection system control and air handling control. These aspects have been discussed in the next two sections.

Figure 2: Diesel engine system schematic (after [1])

2.2.1 Fuel Injection System Control

The fuel control system responds to the torque demand of the driver by controlling the fuel injection system to inject the appropriate amount of fuel. It also controls the injection timing and duration, injection pressure, and also the number of injections in order to minimize fuel consumption and exhaust gas emissions. Modern diesel engines use common rail diesel injection systems (CRDI) [2] and electronic unit injectors [3] for fuel injection. Figure 3 shows the picture of a common rail injection system. It consists of a low pressure fuel pump, a high pressure fuel pump with a pressure regulator valve, a common rail (or possibly two for “V” engines), injectors, pressure and flow limiters and an engine electronic control unit. The high pressure pump helps in pumping fuel into the common rail and the pressure regulator valve helps in maintaining the common rail pressure at a desired value. The common rail acts as a fuel accumulator and provides the injectors with high-pressure fuel. The electronic control unit (ECU) controls the common rail system to maintain the desired pressure levels in the common rail. The ECU also controls the amount of fuel injected and the injection timings. Most of the engine ECUs in the market today control the fuel system with the help of open loop feed-forward maps for injection parameters like rail pressure, injection amount and injection timings. Steady-state engine calibration maps are used for this feed-forward control. These
maps take the current speed of the engine and torque requirement as inputs to output the injection parameters. In addition, dynamic adaptation of the steady-state maps is employed during transients [4]. In the case of transients, the amount of fuel injected is limited to avoid the sudden drop in the air-fuel ratio which normally leads to high soot formation in diesel engines.

![Common Rail Diesel Injection System](image-url)

Figure 3: Common Rail Diesel Injection System (after [2])

### 2.2.2 Air and EGR Control System

The goal of air and EGR control path is to provide the diesel engine with the desired amount of fresh air and re-circulated exhaust gas. Intake manifold conditions have a strong effect on engine performance and emissions. Figure 4 shows the effect of intake manifold conditions on BSFC, NOx and soot formation of a diesel engine. Increase in intake manifold temperature typically increases BSFC, NOx and soot formations. Increase in BSFC and soot can be attributed to decreased air-fuel ratio and increase in NOx can be attributed to higher combustion temperatures. The second part of Figure 4 plots the effect of intake manifold pressure (keeping intake manifold temperature constant) on BSFC, NOx and soot. BSFC decreases with increase in boost pressure. NOx and soot exhibit opposite behaviors with increase in boost pressure. While NOx increases first and reaches a maximum and starts decreasing with increase in intake
manifold pressure, soot exhibits the opposite behavior. Intake manifold conditions are therefore crucial for achieving desired performance and emission levels. Most diesel engines accomplish air flow control with the help of a variable geometry turbocharger (VGT). Mass air flow and manifold pressure sensors in the intake manifold are commonly used as feedback to control the air flow.

![Graph showing the effect of intake manifold conditions on diesel engine performance and emissions](image)

Figure 4: Effect of Intake Manifold Conditions on Diesel Engine Performance and Emissions (after [4])

On the other hand, exhaust gas recirculation is done to reduce the peak cylinder temperature during combustion which helps in reduced nitrogen oxides emissions. Production EGR systems are either open loop or use feedback from the engine sensors to deliver the desired amount of EGR at a given operating point. Feedback systems provide EGR control during fairly steady-state conditions. Feed-forward schemes are necessary during heavy transients [1] due to delays and slow dynamics of the system. Most EGR
feedback controls rely on the air mass-flow sensors as feedback devices. Some EGR feedback systems rely on pressure sensors in intake and exhaust manifolds for feedback or even dedicated differential pressure flow meters in the EGR path. Although EGR helps in reducing NOx formation, it can increase the amount of soot formed due to lower oxygen concentrations and reduced density which leads to lower air-fuel ratios in the cylinder. For this reason EGR systems in modern day diesel engines use an EGR cooler to decrease the temperature which results in reduced intake manifold temperature. The use of EGR cooling also enables a more aggressive use of EGR at higher loads of engine without severely impacting the density of fresh air.

Figure 5 shows the effect of hot and cooled EGR on engine NOx and soot formation. It can be seen that cooled EGR can help in reducing both NOx and soot significantly. Due to the coupling between EGR and air flow, a coordinated control strategy is often required. In [5], the authors developed a coordinated EGR-VGT strategy with the help of a multivariable control design. Desired values of boost pressure and EGR ratio for a given speed and load of engine are determined from calibration maps such that fuel consumption can be minimized while keeping the emissions within the regulated limits.

The emission and fuel efficiency regulations on engines are being progressively tightened. Fuel and air handling control systems employed by most of the present day diesel engines are not sufficient to meet future emission and performance requirements. In order to meet the regulations, new technologies are being developed continuously. These new technologies demand for improved control of the system to achieve the objectives of reduced emissions and improved performance. One of the most important topics of study is the use of alternative diesel combustion modes. Alternative combustion modes, like low temperature combustion (LTC), Homogeneous Charge Compression Ignition (HCCI), and premixed charge compression ignition (PCCI), are being developed along with conventional diesel combustion to achieve emission regulations and desired performances. These alternative combustion modes help in operating the diesel engines in areas far from the high soot and high NOx formation areas. Figure 6 plots the operating regions of different alternative combustion modes in terms of their local equivalence ratio and local temperature [6]. It can be seen that there is a tremendous potential to decrease soot and NOx simultaneously by operating in these modes. But such an operation demands higher control authority –
most likely in the form of combustion feedback. It will be seen later that closed-loop combustion control becomes a necessity to enable these alternative combustion modes. On the other hand, closed-loop combustion control can help in improving emissions and performance of conventional diesel engines by operating closer to the NOx-PM trade-off boundaries. Improvement in the calibration processes of diesel engines is also an important area of development in order to meet future regulations [1]. Literature on industrial calibration practices and its developments will be dealt with in Section 2.4.

Figure 5: Effect of Cooled EGR on NOx and PM (after [4])
2.3 Closed-loop Combustion Control of Diesel Engines

Closed-loop combustion control (CLCC) involves controlling the combustion characteristics of an engine with the help of direct or indirect feedback from combustion. The advantages of closed-loop combustion control depend on its application. In spark ignited engines, CLCC has been used to widen the operating range to points close to knocking, pre-ignition and dilution limits of gasoline engines. In alternative combustion modes, like HCCI, PCCI, and LTC, closed loop combustion control is a necessity due to the lack of any combustion trigger. In the case of conventional diesel engines, the advantages of CLCC are not well established but the focus has been in the area of minimizing engine emission and performance variability resulting from various sources of variability. Closed loop combustion control has been explored to a great extent for application in spark-ignited engines and homogeneous charge compression ignition diesel engines (HCCI).

In [7], the authors demonstrated that, in spark ignited engines, feedback of cylinder pressure measurements can help in achieving multiple objectives like spark-timing and EGR control, cylinder-to-cylinder air-fuel ratio balancing, misfire detection, cold start control and knock detection. The authors define a combustion metric based on the ratio of in-cylinder pressure to the motoring pressure at different crank angle positions. The normalized ratio of the cylinder pressure to the motoring pressure at 10 degrees ATDC (PRM10) has a
strong correlation to the spark timing for Maximum Brake Torque (MBT). Another metric which has a good correlation to the dilution levels of the engine called DILPAR (dilution parameter) has been proposed by authors. DILPAR is derived from PRM10 and the final pressure ratio (FPR) which is the ratio of in-cylinder pressure to motoring pressure at the end of combustion (typically at 55 deg ATDC). The authors suggest two different control strategies of controlling either the PRM10 or DILPAR to achieve operation close to the dilution levels of the engine. For spark ignited engines, the MBT (Maximum Brake Torque) spark timing yields a PRM10 value of 0.55 under all operating conditions. Similarly, the DILPAR can also be controlled to a set-point which determines the spark timing for maximum brake torque for the dilution limit of the engine. In this work, the authors use closed-loop control of PRM10 with the help of spark advance as the control input. The PRM10 value has been maintained at 0.55 for all the normal operating conditions. When knock or pre-ignition are detected, the set-points for PRM10 have been reduced. Knock is detected with the help of a spark-plug boss sensor and partial burn cycles have been detected with the help of the FPR. The detection of the dilution limit is done by monitoring the co-variance in IMEP and FPR. With the help of closed-loop control, the authors were able to demonstrate engine operation close to the dilution limits of the engine and highest efficiency. Figure 7 plots the emissions and fuel efficiency of the spark-ignited engine as a function of PRM10 at zero and high EGR levels.
Closed-loop combustion control has been widely applied to Homogenous Combustion Charge Ignition (HCCI) engines by many authors. In HCCI engines, a homogeneous mixture of diesel fuel and air self-ignites in the cylinder. Because the homogeneous mixture auto-ignites, combustion will start more or less simultaneously through-out the whole cylinder [8]. A rapid heat release during the combustion helps in reduced NOx and improved thermal efficiency of HCCI engines. However, due to the lack of an easily identifiable trigger like fuel injection or spark, it is difficult to control the start of combustion in HCCI engines. Therefore, a feedback of combustion phasing becomes a necessity for control of HCCI engines. Closed-loop control of combustion phasing in HCCI engines has been successfully demonstrated by many
authors. In [8], the authors have demonstrated closed loop control of HCCI combustion for individual cylinders using dual fuels and inlet air pre-heating. Combustion control is achieved with the help of dual-fuels with different self-ignitability properties. The dual fuels used were n-heptane and isooctane. For the combustion timing control, crank angle at which 50\% of combustion heat release takes place, CA50, is used as the feedback variable. The authors used an estimation of the sensitivity of CA50 to fuel octane number to determine the ratio of the fuels to be used for control. Pre-heating of inlet air is done at low loads to maintain high combustion efficiency. However this is not the primary combustion controller. The author has demonstrated the performance of the closed-loop combustion timing control under various load and speed changing scenarios.

In [9], the authors have applied Variable Valve Actuation (VVA) techniques to control the combustion timing in HCCI engines. The authors used overlap between exhaust valve closing (EVC) and intake valve opening (IVO) to influence combustion timing. The overlap affects the temperature of the combustion chamber by trapping residual gases from the previous cycle thus affecting the combustion phasing. In the same study, the authors used intake valve closing (IVC) to affect the effective compression ratio of the cylinder thereby adjusting the pressure and temperature of the combustion chamber. The authors note that the IVC strategy is a stronger control strategy than the valve overlap strategy as heat-release shows more sensitivity to the IVC strategy than the overlap as shown in Figure 8. It can be seen that a smaller change in IVC produces the desired change in heat-release when compared to the EVC-IVO overlap. However, this phenomenon cannot be generalized for all the operating conditions and therefore the authors use both valve timing strategies at different operating conditions to control the combustion timing in HCCI engines. In this study the authors have chosen to control the start of combustion (SOC) calculated by the Arrhenius knock integral method with the help of the valve actuation strategy. Due to the computational limitations of calculating the SOC in real-time, the authors propose to control CA50 instead of SOC. A duration model is used to convert the CA50 into SOC using other inputs like air fuel ratio, mass of residuals \textit{etc.}
In [10], the authors have proposed a HCCI control strategy for gasoline engines based on variable valve actuation. The authors use IVC to control combustion timing and EVC to control the peak pressure of the cylinder. IVC helps in varying the effective compression ratio of the cylinder and EVC affects the residual content from the previous cycle in the cylinder. The authors were able to demonstrate a de-coupled controller in order to achieve the combustion timing and peak pressure control which is essential for the gasoline HCCI engine. The strength of this work lies in the de-coupling of the controllers which is a
common problem faced in multi-input multi-output (MIMO) systems. Similar applications of closed-loop combustion control can be found in pre-mixed diesel combustion [11] and low temperature combustion [12].

However, the application of closed loop combustion control for conventional diesel engines has not been extensively studied. In [13], the authors have proposed a coordinated combustion feedback control for conventional diesel combustion to minimize the dispersion in emissions for a conventional diesel engine. They have done a sensitivity analysis of the diesel engine emissions to various sources of variability like main timing, fuel quantity, fresh air, rail pressure etc. The results from the sensitivity analysis are shown in Figure 9. It can be observed from the figure that NOx shows high sensitivity to changes in main timing, boost pressure and intake manifold air flow (MAF). On the other hand particulate matter (soot) showed high sensitivity to injected fuel amount and MAF. It is therefore clear that a closed loop combustion control strategy based on injection timing and amount, and air handling system would be most effective in minimizing the dispersion in emissions. The authors propose a coordinated control strategy where engine states like CA50, IMEP, combustion instability (CIS), and rate of combustion are controlled through a coordinated control of the injection system and the air path. The fueling and air path set-points are adapted to ensure stable operation during high EGR operation and alternative combustion modes. In [14], the authors demonstrated closed-loop combustion control for a CRDI diesel engine with the help of an adaptive controller where a combustion metric which has a direct correlation with the start of combustion is used as feedback. Start of injection is again used as the control input to track the feedback variable. A radial-basis neural network model is used as a feed-forward generator for the control scheme. The neural network is adapted based on adaptation rules to cater to the change in engine behavior due to various sources of variability.
As explained before, closed-loop combustion control has been developed predominantly for application in spark ignited engines and alternative diesel combustion modes like HCCI. A complete evaluation of advantages of closed-loop combustion control for conventional diesel engines has not been seen in the literature. The purpose of this study is to develop and apply the concepts of closed-loop combustion control to conventional diesel combustion and to evaluate the need, challenges and potential of such an application. In the next sections, various aspects of closed-loop combustion control as addressed in the literature have been discussed.

### 2.3.1 Cylinder Pressure-based Combustion Control

Electronic cylinder pressure sensors and improved computational capabilities of micro-processors have facilitated the use of cylinder pressure for engine control and monitoring purposes. Therefore it has been widely used in optimum spark control, knock control, air fuel ratio control and misfire detection in spark-ignited engines [15]. Cylinder pressure versus crank angle data over the compression and expansion strokes of the engine operating cycle can be used to obtain quantitative information on the progress of combustion.
Therefore, cylinder pressure has a great potential in facilitating closed loop combustion control. A wide variety of feedback signals calculated from cylinder pressure measurements have been used in the literature. These feedback signals are either calculated directly from the cylinder pressure or from approximate heat release rate calculations. These feedback signals are dependent on the combustion regime of the diesel engines as the same feedback signal may not be suitable for different combustion recipes. The feedback signals act as a signature of combustion and are therefore termed as combustion metrics. The combustion metrics are correlated to engine performance and emissions which provide the feedback controller, information about the operation of the engine. This is done because direct in-vehicle measurement of engine performance variables like Brake Specific Fuel Consumption (BSFC) or engine emissions like NOx or soot is not possible.

2.3.1.1 Heat Release Rate Based Combustion Metrics

Heat release rate method models the rate of release of fuel’s chemical energy during the diesel engine combustion process. This method is the application of first law of thermodynamics to a quasi-static open system. The first law for such a system is given by Equation 1 where \( \frac{dQ}{dt} \) is the heat transfer rate across the system boundary into the system, \( p \frac{dV}{dt} \) is the rate of work transfer done by the system boundary displacement, \( m_i \) is the mass flow rate into the system, \( h_i \) is the enthalpy of the flux \( i \) entering or leaving the system and \( U \) is the energy of the material contained inside the system boundary.

\[
\frac{dQ}{dt} - p \frac{dV}{dt} + \sum_{i} m_i h_i = \frac{dU}{dt}
\]  

Equation 1
The above equation can be used to obtain the rate of release of fuel’s chemical energy. This method is called the heat release rate method and can be calculated from cylinder pressure data using Equation 2. Equation 1 is modified to give the rate of heat release in the crank angle domain. The heat release calculated from the Equation 2 is called the true heat release rate since it accounts for the heat transfer between the working fluid and cylinder walls. It is important to consider heat transfer during heat release rate analysis as it can affect the heat release significantly. Figure 10 from [16] plots the cumulative heat release rate curve of a diesel engine at mid-speed, mid-load operating point calculated with and without heat transfer and crevice effects. Although there are models available in the literature to accurately predict the heat transfer rates [17], the improved accuracy does not often justify the computational burden involved in calculating the heat release rate for CLCC in a real engine application. Therefore, apparent heat release rate (AHR) given in Equation 3 is used by most authors. Metrics derived from gross heat release rate and apparent heat release rate show similar sensitivity to engine emissions and performance. Another important simplification that most authors do is to assume a constant value for the specific heat ratio in the heat release equation.

Figure 10: Effect of Heat Transfer on Combustion Heat Release (After [16])
\[
\frac{dQ_{ch}}{d\theta} = \frac{1}{\gamma - 1} \left[ \gamma \cdot p \cdot \frac{dV}{d\theta} + V \cdot \frac{dp}{d\theta} \right] + \frac{dQ_{hu}}{d\theta} \tag{Equation 2}
\]

Normalized cumulative heat release can be calculated by integrating Equation 2 and normalizing the cumulative heat release. Authors in the literature use CA\(x\), which is the crank angle for \(x\)% of cumulative heat release as a feedback variable with CA50 being the most popular one. CA50 is the most robust metric that can be derived from the heat-release analysis as it has direct correlation to combustion efficiency and also least prone to noise. Metrics derived from early and late combustion are prone to noise since the signal to noise ratio (SNR) of the pressure signal is low in these regions.

In [18], the authors used CA50 as feedback in balancing cylinder-to-cylinder variations in IMEP of a HCCI engine. They compared the advantages of controlling CA50 in individual cylinders against a bank control strategy. Even fuel amount in each cylinder (case 1) and even fuel amount and inlet air temperature of all cylinders (case 3) are used as bank control strategies. In cylinder-by-cylinder control, CA50 of individual cylinders were controlled to be at 6 deg ATDC (After Top Dead Center) using inlet air heating for individual cylinders (case 2) and by using fuel amount in individual cylinders in case 4. In case 2, where the CA50 of individual cylinders were controlled with inlet air temperature of individual cylinders, the authors observed best performance in terms of NOx reduction and highest efficiency followed by case 4 where fuel amount is used for cylinder-to-cylinder control. Case 2 also resulted in least covariance in IMEP of individual cylinders indicating combustion stability. The authors have also shown that closed-loop combustion control with CA50 as feedback has the potential to respond to fast engine transients. They were able to adjust CA50 from 1 to 9° ATDC at 2000 RPM engine speed within 8 engine cycles by changing...
inlet air temperature. Figure 11 plots the cylinder-to-cylinder variation in CA50 for each case mentioned above.

![Cylinder-cylinder variation of CA50 in different cases (after [18])]({attachment:image.png})

In [9], the authors have used CA50 as feedback to control a HCCI engine through variable valve actuations. Although their control is based on start of combustion (SOC), they use CA50 as a feedback as it is more robust to noise in the cylinder pressure measurement. They use a duration model to convert CA50 into an SOC value and use it as a feed-forward component of their control in determining the intake valve closing position (IVC). In [19], the authors have identified CA50 as the most promising feedback control variable since it is insensitive to sensor noise and to the drift of measurement of absolute cylinder pressure as it occurs during the life-cycle of a sensor.

Other versions of combustion metrics which do not rely on heat release rate calculations can also be found in the literature. The Rassweiler-Withrow model is one such method which estimates the mass fraction of fuel burnt (MFB) during the combustion event. The method is based on the assumption that the change in pressure due to the piston motion and charge-to-wall heat transfer can be represented by polytropic processes. In this method, the pressure change during any crank angle interval is assumed to be made up of a pressure rise due to combustion $\Delta p_c$ and a pressure rise due to the volume change $\Delta p_v$, [20]. Since the pressure rise due to combustion is proportional to the mass of the fuel that burns, the MFB at a given crank
angle can be calculated by taking the ratio of the combustion pressure rise until that crank angle to the total combustion pressure rise during the entire combustion event as given by Equation 4. In the equation, the combustion event is limited between the Intake Valve Closing (IVC) and Exhaust Valve Opening (EVO). In this way the crank angle at $x\%$ mass fraction burnt can be calculated. The incremental heat release can be calculated based on the rise in the combustion pressure as shown in Equation 4. If the poly-tropic index $n$ is assumed to be equal to the ratio of specific heats and if the crank angle interval for the calculations are taken to be small, then the RW method can produce nearly identical results to the Apparent Heat Release Rate (AHR) equation as shown in Figure 12 [20].

\[
M_{FB\_RW}(\theta_i) = \frac{\int_{\theta_{\text{IVC}}}^{\theta_i} \Delta P_c(\theta) \, d\theta}{\int_{\theta_{\text{IVC}}}^{\theta_{\text{EVO}}} \Delta P_c(\theta) \, d\theta}
\]

\[
\Delta P_c(\theta_i) = p(\theta_i) - p(\theta_{i-1}) \left(\frac{V(\theta_{i-1})}{V(\theta_i)}\right)^n
\]

\[
\Delta Q(\theta_i) = \frac{V(\theta_i)}{n-1} \Delta P_c(\theta_i)
\]

Equation 4
Another model which approximates the mass fraction of fuel burnt (MFB) during combustion is the diesel pressure ratio model. This method was first proposed in SI engines where the fuel MFB percentage is calculated taken as the ratio of the pressure ratio (PR) to the maximum pressure ratio. The pressure ratio is defined as the ratio of the firing pressure to the motoring pressure. To extend the same concept to diesel combustion, a correction to the firing pressure and motored pressure can be made in the form of Firing Pressure coefficient (FPC) and Motored Pressure Coefficient (MPC) as shown in Equation 5 [20]. This is to account for the higher compression ratios and multiple injections that characterize diesel engines. With the help of FPC and MPC, the MFB curve calculated by the diesel pressure ratio method represents the cumulative heat release rate curve of combustion as shown in Figure 13.
\[
DPR(\theta) = \frac{P(\theta) + FPC}{P_{\text{mot}}(\theta) + MPC} - 1
\]

\[
MFB_{DPR}(\theta) = \frac{DPR(\theta)}{DPR_{\text{max}}}
\]

Equation 5

2.3.1.2 Combustion Metrics Calculated Directly from Cylinder Pressure

Many combustion metrics that are directly calculated from in-cylinder pressure traces can also be found in literature. In [10], the authors use crank angle at peak cylinder pressure and peak cylinder pressure as feedback variables to control cycle to cycle variations in a HCCI engine. The authors were able to derive a linearized model of peak cylinder pressure in a HCCI engine which takes into consideration the cyclic coupling in residual effected HCCI engines. The advantage of peak pressure is that it can be calculated from the cylinder pressure without pinning the measurement. However, crank angle at peak pressure is not a good approximation for CA50 for conventional diesel engines with multiple injection events. In [20], the authors recommend subtracting the motored pressure from the cylinder pressure trace and calculating the
maximum pressure from the resulting combustion pressure trace as a better approximation of CA50. Figure 14 demonstrates the process of calculating combustion pressure for conventional diesel combustion.

Figure 14: Fired, Motoring and Combustion pressure traces of a conventional diesel engine (after [20])

Another combustion metric commonly found in the literature is the maximum rate of pressure rise and crank angle at maximum rate of pressure rise. Due to the differentiation involved, this metric is more susceptible to noise in cylinder pressure signal. However this can be a good alternative to using maximum pressure for conventional diesel engine with multiple injection events as it can differentiate pressure rise due to compression and combustion. The rate of pressure rise due to combustion is often higher than that due to compression which makes crank angle at maximum rate of pressure rise a better metric for conventional diesel engines.

While combustion metrics related to maximum pressure and maximum rate of pressure rise are most common, some authors have proposed more unconventional metrics like center of gravity of pressure trace. In [21], the authors have used the center of gravity of pressure trace from the BDC of intake stroke till the BDC of the power stroke as a combustion metric. They have also proposed metrics like secant length at
various levels of pressure trace. Figure 15 plots the pressure trace and various combustion metrics proposed in [21]. The authors were able to find a unique relation between the center of gravity and the injection timing and quantity for a diesel engine.

Figure 15: (a) Center of gravity and secant length metrics (b) Correlation of center of gravity metric to diesel engine injection parameters (after [21])
2.3.2 Control Methods for Combustion Control

Multiple degrees of freedom in diesel engines make it a formidable task to choose a control input for closed loop combustion control. The choice of control input is largely determined by the combustion mode of the engine. In literature, a number of different control inputs have been used. Adjusting the injection parameters like injection timings and injection amount is the most common method for diesel engines without premixed charge. The advantage of injection system is that it provides the capability of taking a control action on a cycle by cycle basis. In [21], the authors have shown that injection timing and amount have significant influence on engine emissions and performance of diesel engines thus making them ideal candidates as control inputs. Variable valve actuation strategies have been widely used in the literature in HCCI engine applications which do not have an injection trigger like a spark or fuel injection. Shaver et.al., use Intake Valve Closing (IVC) as a method to control the effective compression ratio of the cylinder in order to control the combustion timing of residual-affected HCCI engines. In [9], The authors compare the IVC strategy with another valve opening strategy where the negative overlap between the exhaust valve closing (EVC) and intake valve opening (IVO) is varied. They found that the IVC strategy which varies the effective compression ratio is a stronger control strategy than the negative overlap strategy which affects the amount of trapped residual gases from the previous cycle. In [8], the authors have used a dual fuel strategy to control the combustion characteristics of a HCCI engine. They inject different proportions of n-heptane which is easily self-ignitable and iso-octane which is ignition resistant to control the combustion during HCCI operation. In [18], the authors have used inlet air heating as a method to balance cylinder-to-cylinder variations in a HCCI engine. Increasing inlet air temperature affected the combustion phasing by causing the auto-ignition temperature to be reached earlier and by decreasing the air-to-fuel ratio in the cylinder. By maintaining the combustion phasing across the cylinders to be equal using inlet air temperature, the authors reported improved break efficiency and reduced levels of NOx.

Due to the non-linear behavior of combustion in engines, it is highly difficult to apply conventional control design techniques to closed loop combustion control. Most of the control structures used in the literature are based on scheduled PI (Proportional-Integral) controllers which are most commonly used in industrial
applications. Most of the authors in the literature do not discuss the calibration or tuning methods for Proportion-Integral-Derivative (PID) controllers which suggests that conventional tuning techniques have been used. In [9], the authors suggest a model-based non-linear compensation to the PI control. The non-linear compensation acts as a strong feed-forward component which minimizes the control action needed by the PI controller. A linearized model for the peak cylinder pressure model of a HCCI engine was derived in [10] and a linear control technique (H₂ control) was applied. The authors also simplified the multi-input multi-output (MIMO) plant model into two single-input single-output (SISO) models and designed two linear controllers at two different time scales in order to minimize the effect of coupling between the two SISO models. In [22], the authors have used system identification techniques to derive MIMO transfer function models of the plant. They have applied Model Predictive Control (MPC) to the control problem. MPC has allowed them to specify constraints on the variable valve actuation mechanism which is their choice to control combustion timing in an HCCI engine. In [14], the authors used an adaptive feed-forward controller based on a radial basis function network along with a linear feedback controller. The adaptive feed-forward controller uses the output of the linear feedback controller to adapt the parameters of the radius basis function network. The feedback error learning method was applied for fast control of start of combustion during transient operation of the engine.

2.3.3 Cylinder Pressure Data Processing for Combustion Control in Real-time

Cylinder pressure based combustion control requires careful handling of cylinder pressure sensing and processing. Errors in the pressure signal processing can lead to errors in combustion metric calculations which may result in undesirable results. In the literature, a number of authors provide both qualitative and quantitative insights into pressure signal processing techniques for combustion control.

In [23], the authors discuss the hardware requirements for implementing a closed-loop combustion controller. They suggest a three layer hardware structure to carry out tasks of different complexity and time-scales. The first layer is a FPGA (Field Programmable Gate Array) board which calculates low-level sample by sample operations like digital filtering, angular referencing and other algebraic operations. This
layer delivers a true real-time execution. The second level of hardware uses the output from the FPGA to carry out calculations or operations that are done every cycle of each cylinder in an engine. Examples of such operations are calculation of CA50 or other combustion metrics. Therefore, hardware for the second layer should be capable of carrying out floating point arithmetic in real-time. The third layer is a high level controller which is supervisory in nature. Typical tasks of such a layer would be data monitoring and other data processing like statistical analysis over thousands of cycles of engine data. In [23], the authors also specify the software needs for combustion control. The pressure signal can be filtered for reducing combustion noise using digital filters which introduce minimum group delay into the signal like Butterworth and Bessel filters. The group delay introduced can be estimated and corrected for during signal processing. They also suggest sampling methods based on time domain and crank angle domain. Crank angle domain sampling would be advantageous because such a method would be independent of the engine speed. But, designing filters in crank angle domain may pose difficulties. Therefore, the authors suggest a time domain sampling at a frequency that provides sufficient crank angle based sampling for the highest speed of the engine. Down-sampling can be done at lower speeds to reduce data storage requirements. The authors also suggest methods to correct crank angular referencing of pressure signals to delays induced by filters and sensors. If not corrected, such delays can lead to significant errors in combustion metric calculations. The hardware architecture was evaluated by calculating the rate of heat release and IMEP for an eight cylinder high speed diesel engine. The pressure data is sampled at a frequency of 100kHz and the pressure is referenced to the crank angle which is sensed with a 60-2 crank angle encoder (every 6 deg). Operations on the crank-angle referenced pressure data are carried out on a FPGA between two successive crank teeth in order to calculate rate of heat release (ROHR) and IMEP by the end of the cycle. Figure 16 plots the available time to carry out computations between two successive crank teeth versus the actual computation time for calculating ROHR and IMEP as a function of speed. It can be seen that as speed increases, the ratio between available and actual time decreases and if extrapolated, the speed at which the ratio becomes one can be found out. The authors note that from this evaluation, they used only 50% of the gates, 70% of the memory thus leaving a lot of scope to implement other combustion control algorithms.
In [20], the authors compare the sensitivity of various combustion metrics to noise in cylinder pressure signal. The metrics based on heat release rate analysis are more susceptible to noise due to the derivative of pressure term in heat release calculations. On the other hand, metrics based on Rassweiler-Withrow (RW) and Diesel Pressure Ratio (DPR) methods are more robust to noise thus making them more suitable for real-time applications. Figure 17 shows the effect of a noisy pressure trace on heat release rate calculation and MFB calculation using the RW and DPR methods.
Figure 17: (a) Heat release rate calculation from a noisy pressure signal (b) Performance of RW and DPR methods with a noisy pressure signal (after [20])
2.3.4 Alternative Sensing Technologies for Closed-loop Combustion Control

Cylinder pressure sensor is the most fundamental variable that can be used to characterize combustion in internal combustion engines. However, high costs associated with cylinder pressure sensors have kept them away from use in production engines. As a result, various other technologies that can replace cylinder pressure sensors have been under study. Since detailed evaluation of each technology is out of the scope of this work, a brief overview of the alternative sensing technologies is provided in the following sections.

2.3.4.1 Ion current sensors

Ion current probes sense the current induced by ions formed due to chemical and thermal ionization during combustion. The following reactions represent the most significant chemical and thermal kinetics of combustion [24]. The main chemical ionization is represented by the first two equations and the third equation represents thermal ionization seen during combustion.

\[
CH + O \rightarrow CHO^+ + e^- \\
CHO^- + H_2 O \rightarrow H_2 O^- + CO \\
M + E_{\text{ion}} \leftrightarrow M^+ + e^- 
\]

Ion current sensors are usually mounted inside the cylinder in the place of the pressure sensor and measure the current due to the ions formed during combustion. This technique is more suitable for SI engines as a spark plug can be used as an ion sensing probe thus giving time resolved information of combustion [25]. As a result ion sensing probes have been used extensively in SI engine applications to do knock control and misfire detection. Efforts are now being directed to use ion probe sensors for air-fuel ratio (AFR) control and peak pressure control in SI engines. Application of ion current probes for diesel engine combustion control is still in the nascent stage and has many challenges. Several authors reported that the behavior of ion current signal matched with that of the pressure sensor [26-27]. But, the amplitude of the ion current signal is reported to be highly sensitive to the AFR and for very lean operations it results in low signal-to-
noise ratios thus making it difficult to characterize combustion. This is a problem with diesel engines and alternate combustion modes like HCCI which operate lean most of the time. The amplitude of the ion current signal is also reported to be weak at high speed and low load operations of the engine. Another known problem with the ion sensing probes for diesel combustion is the deposition of carbon between the electrodes [24]. The high carbon deposition hampers the ability to measure the rate of ionization of diesel combustion. The choice of ion current probe position in the combustion chamber can reduce the carbon deposition thus making it possible to measure the ion current signal. Also the effect of the multiple injections on the ion current signal has not been studied in the literature. Figure 18 compares cylinder pressure and ion current signals for conventional diesel combustion. It can be seen that ion current signal can be used to detect the start of combustion but needs heavy filtering to deduce combustion progress. This phenomenon limits the application of ion current sensors to conventional diesel engines.

![Figure 18: Comparison of ion current and pressure signals in conventional diesel combustion (after [24])]
In [24], the authors applied ion sensing to a HCCI engine and reported that the fast reaction in HCCI was detected by this method. Effects of various factors like EGR on ion-current signals are studied and it is reported that high EGR levels might weaken the signal thus making it difficult to use in highly diluted combustion regimes.

2.3.4.2 Accelerometers

Accelerometers mounted on the engine can record the vibrations of the engine due to combustion and knocking. They have been used widely to predict knocking in SI engines. Accelerometer signals filtered using low pass filters can carry the signature of combustion from the cylinders. Figure 19 plots compares the pressure signals from a 6 cylinder engine and the accelerometer signals mounted on the cylinder head at two different speeds. Attempts have been made to reconstruct cylinder pressures using structure-borne vibration signals. In [28], the authors propose a signal model that incorporates a linear time invariant transfer function to describe the effect of pressure signal on the observed structure-borne sound. The in-cylinder pressure signal is parameterized as three components caused by compression and 2 stages of combustion. This model of the pressure is obtained offline with training data. The parameters of the pressure traces for each window of combustion are then found from the accelerometer signals by using an estimation-maximization algorithm. Due to the necessary filtering operation which involves information loss in time domain, it leads to inaccurate prediction of crank angle degree (CAD) of the start of combustion. This method of using structure borne sound to predict combustion could be more difficult in large engines with more than 4 cylinders due to the overlapping of many combustion events in a given window. Also, in order to isolate combustion events from individual cylinders, some authors proposed multiple accelerometers. The signal from the nearest accelerometer can be used to isolate the combustion of a given cylinder. For combustion metrics beyond Start of Combustion (SOC), it seems unlikely that this sensing technique is promising due to the difficulty in reconstructing pressure.
2.3.4.3 Speed Sensors

Fluctuations in crank speed have been used to monitor the combustion events in the engine. A crank shaft dynamics model which accounts for crank angle dependent moments of inertia, friction and torsional effects is needed to relate speed fluctuations to combustion events in a given cylinder. In [29], the authors have used stochastic models to estimate torque and pressure from fluctuations in engine speed. It is not clear from the literature that the correlation between speed fluctuations and cylinder pressure would be sufficient to calculate heat release rate which is required for calculating combustion metrics in most cases.

In [30], the authors report the use of crank speed fluctuations and pressure from one cylinder to reconstruct the pressures in all the other cylinders. The available pressure of the key cylinder was used for adaptive calibration of the system, determination of the absolute torque value as well as for modeling of the pressure signal. In order to estimate the pressure trace for cylinders without a pressure sensor, a parametric pressure model was developed and it was combined with the torque estimation. This approach becomes difficult for
engines with more than four cylinders because of overlap of multiple combustion events. In such cases, cylinder pressure sensors in multiple cylinders may be a solution. However, the capability of this method in predicting cylinder-to-cylinder differences in combustion is questionable as uniformity is assumed in constructing pressure traces for cylinders without pressure sensors. This method of using a limited number of pressure sensors to improve the prediction of cylinder pressure could be useful in reducing the costs associated with using pressure sensors in every cylinder.

2.3.4.4 Torque Sensors

Attempts have been made to use crank shaft torque sensors to isolate torques generated by each cylinder and correlate them to combustion. This method also needs an accurate model of the crankshaft dynamics which accounts for crank angle dependent moments of inertia, friction and torsional effects to accurately predict the torques contributed by individual cylinders. In [31], the authors proposed a model to calculate the individual torques and used a metric called 50% Torque Ratio to predict the start of combustion. Similar to the case of speed fluctuation sensing, it is tough to use torque to estimate cylinder pressures with overlapping combustion events. Additionally, the high cost of torque sensors make them prohibitive to use in production.

2.4 Advancements in Diesel Engine Calibration Methods

Calibration is the process of finding operating parameters for the control algorithm of an engine which generally maximize its performance and reduce emissions under various operating conditions. In traditional calibration, this may or may not be a rigorous optimization process and often time has a strong human element in the process. The output of such a process is conventionally a set of calibration maps which specify the operating parameters of the engine at various combinations of engine speeds and loads. The complexity of generating the calibration maps increases exponentially with increase in the number of degrees of freedoms of the engine system. An additional degree of freedom typically increases the mapping time and the size of the calibration tables by a factor of 2 (for two position devices) and 3-7 (for continuously variable devices) [32]. Due to the number of degrees of freedom in modern diesel engines,
conventional calibration methods are no longer attractive because of the need to reduce design and development time and costs. Table 1 table gives an idea of the number of parameters that need to be calibrated in modern diesel engines [33]. If conventional methods are used where each calibration parameter is swept in its operating range and experiments are done at every combination of calibration parameters, it is in a very large calibration effort which is impractical by any means.

<table>
<thead>
<tr>
<th>ENGINE SYSTEM</th>
<th>Sub-System</th>
<th>Calibration Parameter</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fuel Injection</td>
<td>Pilot &amp; Main</td>
<td>• Injection</td>
</tr>
<tr>
<td></td>
<td>Split Main</td>
<td>• Timing</td>
</tr>
<tr>
<td></td>
<td>Post-injection</td>
<td>• Injection Pressure</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Rate Shaping</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Fuel Quantity per injection</td>
</tr>
<tr>
<td>Air Management</td>
<td>Turbocharger(s)</td>
<td>• VGT setting</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• VNT setting</td>
</tr>
<tr>
<td></td>
<td>Valve actuation</td>
<td>• VVA</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• VCP</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Camless valve actuation</td>
</tr>
<tr>
<td></td>
<td>Air Flow</td>
<td>• Throttle</td>
</tr>
<tr>
<td></td>
<td>Management</td>
<td>• Intake valve actuation</td>
</tr>
<tr>
<td></td>
<td>Intercooler</td>
<td>• Intercooler bypass valve</td>
</tr>
<tr>
<td>Exhaust Gas</td>
<td>Flow</td>
<td>• EGR valve position</td>
</tr>
<tr>
<td>Recirculation</td>
<td>management</td>
<td></td>
</tr>
<tr>
<td>Afttreatment Control</td>
<td>SCR</td>
<td>• Urea injection rates</td>
</tr>
<tr>
<td></td>
<td>Lean NOx</td>
<td>• Rich AFR operation for regeneration</td>
</tr>
<tr>
<td></td>
<td>catalysis</td>
<td></td>
</tr>
<tr>
<td></td>
<td>NOx absorbers</td>
<td></td>
</tr>
<tr>
<td></td>
<td>DPF Regeneration</td>
<td>• Exhaust temperature and possibly flow control</td>
</tr>
</tbody>
</table>

Table 1: Calibration parameters of a diesel engine that has to meet 2010 emission regulations

Several new calibration techniques are therefore being explored to minimize the cost and time required for engine calibration. Many of these approaches deal with statistical methods, such as Design of Experiments (DOE). Design of Experiments is a statistical technique which starts with identifying the response variables of a system and the all the possible factors that influence the response variables. A model of the response variables as a function of factors is assumed in the beginning and experiments are done such that the
parameters of this model are identified. This identification is often done in an iterative fashion where the model structure is adapted to the results of the experiments. In this model identification, the degree of influence of each factor and the interactions between the factors on the response variable are identified and the model is accordingly corrected to remove unnecessary terms. Experiments are also repeated such that experimental errors are not learned by the model as normal behavior. After identifying the model with sufficient accuracy, the model is used for optimization of the response variables or any objective function constituting the response variables. Figure 20 gives a schematic of the design of experiments as shown by [34]. This method significantly reduces the number of experiments compared to the conventional experimentation methods where all the combinations of the factors are evaluated in experiments.

![Figure 20: Design of Experiments Process (after [34])](image)

On the other hand, model-based calibration methods are also being explored to reduce engine test cell operation time. Artificial Neural Network [35-36] based engine models have been explored by some authors. The disadvantage of these models is the requirement of enormous training data which often defeats the purpose of reducing the time and effort required for calibration. Model-based techniques based on physics and accepted empirical models of engine systems are the most promising for reducing the
calibration efforts. 1-D cycle simulation tools like WAVE, GT-Power are being used extensively to simulate engine behavior [37]. This technique requires calibration and validation of the engine models for representing a specific engine. The test cell data requirement for such a validation is still less because of the confidence one can have on physics-based modeling techniques used by the models. This confidence will not be present when using black box models like artificial neural networks.

In [37], the authors demonstrated an integrated model-based calibration environment which automates the generation of an input data set for calibration, running simulations and optimizing the outputs to find the best calibration dataset for a modern V6 engine equipped with two-step VVA and intake cam phasing. They use interpolation techniques to find the throttle position required for each load and also remove unrealistic combinations of input parameters which are practically not feasible in real engine simulations thus saving some simulation time. The optimization of the cam phasing and the two-step VVA is done using a parsing system which parses the entire simulation data based on the constraints provided and sorts them based on the optimization function. The optimization function can be a weighted function to accommodate multi-objective optimization.

All the methods discussed above help in steady-state engine mapping. However, steady-state engine calibration is not sufficient for applications where the emission standards have to be met in a transient drive cycle like the Federal Test Procedure (FTP). Therefore, efforts are also being focused on developing methods that can enable calibration of engines for transient performance. Dynamic model based engine calibration methods can prove to be helpful for transient calibration of engines. In [33], the authors first designed a number of transient cycles that represent typical field operation of an engine. Transient experiments on the engine were done using the transient cycles for the purpose of data collection. This process requires a base calibration table to do the initial experiments on the engine but it can be relatively crude. The data collected was then brought into a simulation environment where dynamic neural networks are developed for emissions and performance variables of the engine. The engine calibration optimization process was done offline with the help of the dynamic neural network models. The optimization was done on an FTP cycle such that fuel consumption is minimized across the cycle while limiting NOx and PM
emissions. The authors reported a considerable improvement in engine emissions and fuel consumption both in simulation as well as in experiments by using the off-line model-based calibration method by using different objective functions used in calibration. Figure 21 shows the results that the authors have shown in terms of NOx and brake thermal efficiency during an FTP drive cycle. The baseline simulation corresponds to the results from the base-calibration tables. Other points on the plot represent the results from different simulations with varying objective functions used during calibration.

![Figure 21: NOx and Brake Thermal Efficiency Trade-off Using Model-Based Calibration Methods (after [33])]({})

As the number of independent variables in engine control increase, so do the dimension and size of calibration tables. Engine electronic control units have a hard limit on their memory and processing capabilities and therefore storing multi-dimensional calibration tables for engine control is an unattractive option. Data interpolation techniques have been developed by a few authors to reduce the number of data...
points that have to be stored for each calibration parameter. In [32], the authors have applied interpolation techniques based on inverse-distance interpolation to determine the coefficients to estimate cylinder air charge of a variable cam operated spark ignited engine. Inverse-distance interpolation helps in reducing the amount of data that needs to be stored when compared to conventional look up table interpolation methods which require uniformly gridded data. The inverse-distance method allows the use of sparse data to interpolate the calibration maps thus making it attractive for engine control applications.

In this study, a combination of model-based calibration, black-box models, and interpolation techniques have been used to optimize the performance of closed-loop combustion control. Detailed and reduced order GT-Power models of the engine have been used to do model-based calibration. Details of GT-Power modeling are given in the next chapter. Black-box neural network models have been used as an alternatives to detailed combustion models in order to predict emissions and performance during controller development. These black-box methods are discussed in detail in the next chapter. Interpolation based on inverse distance weighting scheme is developed to generate the set points for controller tracking. The calibration data for interpolation is generated by optimizing engine performance on transient drive-cycles. The next section deals with the literature on inverse-distance based interpolation and its application to engine calibration.

2.4.1 Kernel based Interpolation Methods

Kernel-based methods are non-parametric methods that interpolate using a symmetric kernel function $K(u)$ and compute the interpolated value $Y$ at some $x$ by using the following equation:

$$
Y(x) = \frac{\sum_{i=1}^{N} K(x - X_i) \cdot Y_i}{\sum_{i=1}^{N} K(x - X_i)}
$$

Equation 6
Where, \( X_i \) and \( Y_i \) are known calibrated values and \( x \) is the value at which \( Y \) needs to be found. The choice of kernel function depends on the application. Various kernel functions have been used in the literature ranging from parabolic functions, inverse distance and Gaussian kernels. Among the various kernel functions, inverse-distance function is highly suitable for engine mapping because the interpolated value passes through the measurements. Also, the computational effort required for inverse-distance method is comparatively less than that required for other kernel functions. In [32], the authors have used inverse distance as the kernel function for interpolation of volumetric efficiency maps of a SI engine and demonstrated its advantages. In the next section, various aspects of inverse-distance interpolation methods have been discussed in order to evaluate their application to diesel engine calibration.

### 2.4.1.1 Inverse-Distance based Interpolation Methods

Inverse distance based interpolation techniques were first proposed by Shepard [38] as a solution to interpolation of irregularly spaced data. Since then, inverse-distance interpolation methods have been referred to as Shepard methods. Shepard proposed inverse-distance interpolation as a solution to one and two dimensional interpolation of scattered data points. He defined the interpolation function \( f(P) \) formed by the individual point \( z_i \) as described in Equation 7 where \( d_i \) is the Cartesian distance between the independent variables of point \( P \) and those of the point \( z_i \), \( u \) is the exponent on the distance, \( N \) is the number of data points available for interpolation.

\[
f(P) = \sum_{i=1}^{N} \frac{1}{d_i^u} z_i \quad \text{when } d_i \neq 0
\]

\[
f(P) = \frac{1}{\sum_{i=1}^{N} d_i^u} \quad \text{Equation 7}
\]

\[
f(P) = z_i \quad \text{when } d_i = 0
\]

It can be observed from the equation that the nearest point to \( P \) has the highest weight in the interpolation. When the point \( P \) coincides on an existing \( z_i \), the weight corresponding to \( z_i \) becomes infinite and the
interpolation results in $z_i$ itself. The behavior of the interpolation function is affected by the value of the exponent $u$ on the distance term in the equation. Shepard proved with the help of partial derivatives of the interpolation function that the surface formed by inverse distance interpolation would be differentiable only if $u \geq 1$. For values of $u < 1$, the surface becomes discontinuous as $P(x,y)$ approaches $z_i$. The effect of $u$ on the interpolation output has been demonstrated with the help of a 1-D interpolation example as shown in Figure 22 and Figure 23. Even values of $u$ produce smooth interpolation output as shown in Figure 22. But, as the value of $u$ increases (being even), the interpolation output flattens at the data points. For odd values of $u$, the interpolation output has slope discontinuities as shown in Figure 23. Therefore, for engine mapping, even values of $u$ are more suitable. The most commonly used exponent is two, since it serves the purpose of general surface mapping and description and can also be computed easily as it removes the square root from the Euclidean distance equation [38].

![Influence of $u$ on Inverse-distance interpolation](image.png)

Figure 22: Illustration of the affect of even-valued exponent on the interpolated surface
2.4.1.2 Shortcomings of Inverse-Distance based Interpolation

Although the inverse-distance method is a simple interpolation method with useful properties, it has several shortcomings:

1. When the number of data points is large, the calculation of distances from every data point becomes cumbersome. So, the method in its original form can become impractical for interpolating large data sets. (Restricting the calculation to only nearby points can remedy this limitation.)

2. Since the interpolation is based only on distance and does not consider the directionality of the data points, the results of the interpolation can sometimes be counter-intuitive. Also, the surface generated by the interpolation is bounded by the data points provided. Therefore, care should be
taken in not using the inverse-distance method for extrapolation of data. (Selecting a proper
spacing of data to avoid extrapolation can address this limitation.)

3. Also, it can be observed from Figure 22 that the surface generated by the inverse-distance based
interpolation has zero directional derivatives at the data points when \( u > 1 \). This is an undesirable
constraint on the interpolated surface.

To deal with the shortcoming of interpolating large data sets, Shepard suggests two different methods of
eliminating distant data points during the interpolation. To do such elimination, either (1) an arbitrary
distance criterion like searching data points within a radius \( r \) can be used or (2) an arbitrary number of \( n \)
nearest data points can be used for interpolation. Both the methods have their own drawbacks as choosing
an arbitrary radius may result in finding no data points or an unmanageable number of data points to be
found and constraining the interpolation to \( n \) data points may require ranking between points that are
equally distant from a given point. Shepard therefore suggests a combination of the two methods. He
suggests having bounds on the number of data points to be used. Also, he suggests finding a radius \( r \) to be
used for this search that is based on the density of the data points. The radius of neighborhood for search in
has to be selected through the knowledge of the characteristics of the surface and the density of the data
points. Another method could be to increase the radius of search until a good number of data points are
found.

To include the directional information to the interpolation function, Shepard suggests modified weighting
of each point depending on its orientation with respect to other data points. But, the extension of this
method to higher dimensions would not be straight-forward and would involve an increase in the
computational complexity of the method. In order to deal with the zero directional derivatives of the
interpolated surface at the data points, Shepard suggests finding an increment to the function values at the
data points such that the interpolated surface would achieve desired partial derivatives at a given data point.
But, in most cases the partial derivatives at the data points are not available and need to be estimated.
Methods for estimating the partial derivatives are given by [39]. Figure 24 illustrates the effect of adding estimated partial derivatives to an inverse-distance based interpolation surface at the data points [40].

![Figure 24](image.png)

Figure 24. (a) Illustration of the ‘flat spots’ of the interpolated surface using original inverse-distance interpolation. (b) Illustration of the effect of replacing the interpolated value at the data points by their tangent corresponding planes

### 2.5 Conclusions from the Literature Survey

In the first part of this work, a survey of literature on diesel engine control and closed-loop combustion control has been conducted. Closed-loop combustion control has been widely applied to spark-ignited engines for various applications like spark-timing and EGR control, cylinder-to-cylinder air-fuel ratio balancing, misfire detection, cold start control and knock detection. Closed-loop combustion control has also been widely applied to alternative diesel combustion modes like HCCI and other low temperature combustion regimes. In these combustion modes, closed-loop combustion control is a necessity due to the lack of a identifiable trigger for combustion. For conventional diesel combustion, closed-loop combustion control has been applied by a few authors. The primary advantage of closed-loop combustion control for
conventional diesel engine is to reduce the dispersions in emissions and fuel consumptions caused due to many sources of variability described in Chapter 4. In the literature survey, many combustion metrics derived from cylinder pressure that can function as feed-back variables for closed-loop combustion control have been identified. The combustion metrics can be derived either from the heat release analysis or directly from the cylinder pressure. The choice of the combustion metric is highly dependent on the objective of the closed-loop combustion controller and the actuator used to achieve the objective. Fuel injection parameters are primary actuators for combustion regimes where there is no pre-mixed charge. The advantage of using fuel injection parameters is the fast response times which makes cycle-by-cycle control of combustion possible. Variable Valve Actuation strategies are widely used for combustion regimes like HCCI, LTC where there is no combustion trigger. The effective compression ratio and the amount of trapped residuals in the cylinders can be controlled through variable valve actuation strategies. Other control methods that have been used are dual fuels with opposing self-ignitability properties and inlet air pre-heating. As far as the controller design is concerned, most authors in the literature preferred Proportional-Integral control over systematic design of controllers. The non-linear behavior of combustion makes it difficult to apply conventional controller design methods. Strong model-based feed-forward component has often been used as a non-linear compensator which minimizes the control actions of the feed-back controller. Unconventional methods like model predictive control have been used by a few authors. Literature on real-time implementation challenges for closed-loop combustion control has been reviewed during the survey. A hierarchical structure for hardware with varying level of computation power in each level has been proposed by some authors. The lowest level of the hardware is a processor with limited computation power which can carry out simple tasks like digital filtering of cylinder pressure signals, crank angle referencing etc. At the middle level, cycle level tasks can be carried out which need a higher level of computation power like combustion metric calculation. At the highest level, tasks like data monitoring and analysis can be conducted. These tasks are conducted at a frequency of thousands of cycles but the computational power required for this level is very high. The authors also discuss about the sampling requirements for combustion control. Crank angle based sampling is the ideal method but poses
problems in designing filters in crank-angle domain. Therefore, many authors suggest time domain sampling with sufficient crank angle resolution for the maximum speed of the engine.

In the second part of the literature survey, the advancements in engine calibration methods have been studied. Conventional calibration methods used for calibrating diesel engines are highly cumbersome and time consuming. For an ideal closed-loop combustion controller, the calibration effort increases even more because of the need to calibrate individual cylinders. In the survey, recent advancements to calibration efforts have been studied. Model based calibration is one of the methods that can transfer the calibration effort to simulation and minimize the engine calibration effort in the test-cell. Models of the engine can be derived from commercial engine modeling software like GT-Power and black-box models like neural networks. Another goal of calibration is to reduce the size of calibration tables in the engine control unit. Scattered data interpolation methods are desirable for reducing the number of data points needed for calibrating engines and the size of calibration tables. In particular, literature on inverse-distance interpolation method has been studied due to its potential as a simple and robust scattered data interpolation method. Inverse-distance based interpolation methods are among the most viable methods for extension of interpolation to higher dimensions which makes them attractive for application in diesel engine calibration.

In this study, a closed-loop combustion control architecture has been developed for a heavy-duty diesel engine. The inverse-distance based interpolation method is applied to generating the feed-forward control commands for closed-loop combustion controller. The inverse-distance interpolation method has been validated in simulation using a GT-Power model of the engine and neural network models for engine emissions and performance. A comparison study has been conducted to compare the inverse-distance method against the conventional calibration approach.
TOOLS AND TECHNIQUES

3.1 Tools and Techniques used in the Study

In this study, a modeling and simulation based approach has been taken to evaluate various aspects of closed-loop combustion control of conventional diesel engines. Model-based approaches have the advantage of fast evaluation of new technologies in a simulation framework and thus reduce design cycle times. In this study, various tools and techniques have been used to enable modeling and simulation based evaluation of closed loop combustion control. The tools include commercial software for engine system development, algorithm development and simulation. It is often required to use multiple tools in conjunction with one another as every tool is developed with a special functionality. For example, in this study, GT-Power, a commercial engine modeling tool is used for modeling and simulation of diesel engine system whereas MATLAB/Simulink is used for control and algorithm development. In order to combine the strengths of both tools, a co-simulation framework of GT-Power and Simulink is used in the study. The co-simulation framework helps by simultaneously using both MATLAB/Simulink and GT-Power for modeling various aspects of the system.

The term ‘technique’ refers to specific method used for modeling different functionalities of a system. The modeling technique used for a given problem is dependent on the application of the model. For example, if the goal of the study depends on detailed modeling of certain aspects of the system, a modeling technique or method that can address the behavior of the system in a detailed fashion is needed. On the other hand, if the goal is to develop a controller for the system, a reduced-order model which describes the system behavior in a simple way may be sufficient. Lumped parameter or zero-dimensional modeling is a classic example of reduced-order modeling technique. Both lumped parameter and higher dimensional modeling are physics-based modeling techniques and only differ in the degree in which they describe a model’s behavior with the lumped model being a discretized version of a scale that is truly continuous. On the other
hand, it is sometimes essential to have a black-box model which only describes the input-output relationship of a system without any knowledge of the physics of the model. This is the case if the system exhibits behavior that cannot be easily derived from physics. Such black-box models can be both parametric and non-parametric. Neural networks and curve-fitting are examples of parametric modeling techniques. As the name suggests, these parameters for the black-box models need to be identified with the help of training or curve-fitting methods. Such methods therefore require input-output pairs of the system either generated from another model or from experiments for identifying the parameters of the black-box models. On the other hand, there are non-parametric methods that do not rely on training. Kernel-based methods described in the previous chapter are examples of such non-parametric methods. They still require input-output pairs but do not require any kind of training or identification. In this study, both parametric and non-parametric methods have been used.

Design of Experiments (DOE), sensitivity analysis and optimization are common to the calibration phase of engines. In this study GT-Power’s DOE tool has been used to conduct full-factorial design of experiments on the diesel engine model. For sensitivity analysis, Analysis of Variance (ANOVA) has been applied to the experiments. For optimization, various methods have been evaluated and due to the non-linear behavior of diesel engines, Genetic Algorithms have been used. Each of the tools and techniques mentioned in this section are explained in detail in this chapter.

### 3.2 GT-Power for Engine Systems Simulation

GT-Power is one of the leading software tools available commercially for engine and vehicle system design and development. It provides in-built libraries to model various components of engine systems including manifolds, valves, injection systems, turbochargers and all types of internal combustion engines. GT-Power is designed for both steady-state and transient simulations of engine and vehicle systems. It uses one-dimensional gas dynamic equations to model gas flow and heat transfer. Apart from flow and thermal modeling, it is also possible to model mechanical, electrical and control systems with GT-Power. GT-Power also provides models of varying complexity to facilitate simple and advanced modeling of engine
systems. For example, for diesel combustion, GT-Power gives the option to model a diesel engine with a mean value engine model and also with a detailed DI-Jet diesel combustion model to predict emissions like NOx and soot.

The flow model of GT-Power involves solving the Navier Stokes equations in one dimension along the flow direction. Navier Stokes equations give the equations for conservation of mass, momentum and energy of a fluid. The whole system is discretized into small volumes connected by boundaries. Scalar flow variables like pressure, temperature etc. are assumed to be the same for every volume and vector flow variables like mass flux, velocity etc. are calculated for the boundaries of each discretized volume. Flow splits are also modeled as volume elements but are designed to solve the momentum equation in three dimensions. Heat transfer from the fluid to pipe walls, friction losses and pipe geometries are also considered in flow modeling. The accuracy of the system depends on the discretization level and the ability to approximate the system as 1-D rather than 3-D. However, high discretization can lead to small step sizes during simulation leading to longer simulation times. The discretization level therefore is usually determined by the type of simulations. For general performance analysis of engines, discretization lengths of 0.4 times the bore of the cylinder are recommended. But, for higher frequency operation like speeds greater than 6000 RPM, discretization lengths that are half of those for general performance analysis have to be used.

Models in GT-Power are divided into templates, objects and parts. A template is a generic structure that is used to define an object. For example, there is a pipe template with attributes like diameter, length etc. An object is a template with specific values for each attribute. A part is an instance of an object which is the basic building block of a system. Multiple parts can be created from an object which share common attributes. It is also possible to override the values of the attributes in a part to make it different from other parts created from the same object. Objects in GT-Power are divided into components, connections and reference objects. Components are objects that represent physical parts of a system like a pipe, cylinder etc. Connections are objects that typically connect two components. Examples of a flow connection are orifice connections, valve connections etc. Connections can also be mechanical (inertial) and also signal-based.
(sensors, actuators). Reference objects are collections of data that are referred by other objects. Reference objects are typically used to store data that are used in several different places. For example, properties of air are stored as a reference object so that different parts can access these properties. Another use of reference objects is to provide different options to model physical processes like combustion, heat transfer etc.

3.2.1.1 Engine Simulation using GT-Power

GT-Power provides various inbuilt components and connections to model various parts of an engine. It provides different types of cylinder valve models, like cam-driven and solenoid valves. On the other hand, it provides different specialized pipe models to model cylinder intake and exhaust ports. These pipes are different from regular pipe models as they also model fuel accumulation, evaporation and their transport into the engine cylinder. In-cylinder flow models are used in conjunction with combustion and heat transfer models used in the cylinder. For example, a multi-zone flow model is used when using a predictive combustion model. In addition, swirl and tumble caused by the fluids entering the cylinder through intake valves can also be modeled in a simplified manner. In-cylinder heat transfer can be modeled using various methods like Woshni’s method, the Hohenberg method etc.

Fuel Injection of different types can be modeled with GT-Power using different injection templates. For diesel engines, direct injection can be modeled by connecting an injection connection to the cylinder of the engine. It is possible to model multiple injection profiles by defining the injection amount, injection timing and rate shape for each injection. This type of connection is often used to model multiple injections in diesel engines or gasoline direct injection engines. It is also possible to model port-injections and injections at constant fuel-air ratios (for SI gasoline engines).

In GT-Power, combustion is calculated in either predictive or non-predictive fashion. In the non-predictive combustion models, a burn rate is specified by the user. In this case, fuel and air burn in the prescribed way without depending on cylinder variables like pressure, residual fraction etc. This kind of combustion modeling is useful in studying parameters that are not highly dependent on the burn rate, like the effect of
intake runner design on the volumetric efficiency of the engine. Use of a non-predictive combustion model can reduce the simulation times significantly. In predictive combustion models, the fuel burn rate is calculated using the appropriate inputs like cylinder pressure, temperature, equivalence ratio, residual fraction etc. Predictive combustion is useful in studying engine variables that are significantly dependent on burn rate like the effect of injection timings on heat release rate of a diesel engine. Simulation of predictive combustion models is substantially slower, therefore, they are avoided in cases where a complex combustion model is not required.

Another important combustion modeling concept is the use of two-zone and a multiple zone approach to combustion. In two-zone approach, at the start of combustion the cylinder is divided into two zones, burned and unburned. All the contents of the cylinder at the start of combustion are in the unburned zone. The burn rate of the combustion determines the rate at which the contents of the cylinder are moved from the unburned zone to the burned zone as combustion progresses. As the contents are transferred from the unburned zone to the burned zone, a chemical equilibrium calculation determines the concentrations of products of combustion (N₂, O₂, H₂O etc). This is a specific form of the two-zone approach that GT-Power implements. A multi-zone approach is used to model the combustion phenomenon more accurately. DI Jet combustion model is an example of a multi-zone predictive combustion model used to model Direct-Injection Diesel engine combustion. This model is particularly useful in predicting burn rates and NOx emissions of DI Diesel Engine. This approach divides the cylinder into multiple zones. The total injection fuel mass is divided into a number of axial and radial zones as shown in Figure 25. As the fuel is injected into the cylinder, a new axial slice is formed. Each zone is further divided into sub-zones for liquid fuel, unburned vapor fuel, entrained air and burned gases.
Immediately after injection, the zone is 100% liquid fuel. As the zone moves into the cylinder, it “entrains” air and the fuel begins to evaporate, thus forming the unburned subzone. The mass of the entrained air causes the velocity of the zone to decrease because momentum of the zone is conserved. The outer zones entrain air more quickly than the inner zones, thus decreasing their velocity more quickly and resulting in less penetration distance as can be seen in the figure immediately above. From the masses of vapor fuel and entrained air in each unburned subzone, the zonal fuel to air ratio is known. The zonal temperature is calculated taking into account the temperature of the injected fuel, entrained air temperature, and the effects of the evaporating fuel. When the combination of cylinder pressure, zonal temperature, and fuel-to-air ratio becomes combustible, the fuel in the zone ignites, further changing the temperature and composition. All products of combustion will be created in the burned subzone. GT-Power is also capable of modeling exhaust emissions and knock. GT-Power predicts 11 products of combustion (N2, O2, CO2, CO, H2O, H2, H, O, OH, NO, N) using chemical equilibrium equations. NOx calculations are done using extended Zeldovich mechanism. GT-Power is also capable of predicting soot in diesel combustion but with limited accuracy. The soot prediction can only be used for qualitative studies.

GT-Power also gives the capability to model a mean-valued engine model based on calibration maps. The mean value engine model uses maps for the cylinder air flow and the efficiency of the engine. These maps are either defined as tables or calculated by an external model. Mean value engine models are useful in decreasing the simulation time but they do not give detailed cylinder variables like crank-resolved cylinder pressure, emissions etc. In general, the maps needed for mean value cylinder model are derived from a
detailed combustion model by conducting design of experiments or directly from experimental data. GT-Power provides an inbuilt DOE tool which can be used to design DOEs on detailed GT-Power models.

3.2.1.2 Optimization using GT-Power

In GT-Power, it is possible to optimize engine system parameters to achieve a desired performance of a given variable. An example of such an optimization is to find the optimal intake runner length to maximize the volumetric efficiency of an engine. GT-Power gives various methods to carry out such an optimization ranging from design of experiments to direct optimization techniques. Some of these techniques are discussed in the following sections.

3.2.1.3 Design of Experiments for Optimization

Design of experiments (DOE) is a method used to design simulation runs for every possible combination of input parameters to be optimized. In this method, a discrete set of values for each input parameter is defined based on the input parameter’s operation range. The engine simulations are then run at combinations of different values of input parameters. These combinations can be both full-factorial or reduced factorial. In full-factorial DOEs every combination of input parameters is run as a separate case. Reduced factorial DOEs reduce the number of combinations by eliminating the unrealistic combinations of input parameters. Design of experiments is a robust method to optimize the input parameters to achieve desired performance. GT-Power provides a post-processing tool to process the outputs of the DOEs. Using the post-processing tool it is possible to define the response variables in the engine simulation that have to be optimized. The tool provides the capability to fit response surfaces, statistical analysis (ANOVA) and carry out constrained optimization. It is possible to define weighted objective functions in cases where two response variables have to be optimized. For example, one may want to minimize NOx emissions for a diesel engine by varying the amount of EGR content without significantly increasing the fuel consumption (ISFC). In such cases a weighted objective function can be defined with appropriate weights for NOx and ISFC during the optimization process. Design of experiments has its advantages as it is more flexible for carrying out different optimizations using the same set of experiments. However, the process consumes lot
of simulation and post-processing time for carrying out optimization studies and therefore not very suitable
optimization purposes where multiple input parameters have to be optimized. The number of experiments
to be designed increases exponentially with increase in input parameters. Therefore, GT-Power also
provides other techniques for engine optimization.

3.2.1.4 Direct Optimization Methods

Direct optimization methods are iterative methods used to optimize input parameters to achieve a given
objective. The input parameters to be used for optimization along with their ranges can be specified. The
objective function should be specified as a cycle-averaged value of a response variable of the engine
system simulation. The objective function has to be a uni-modal function within the range of the input
parameters. It is also possible to specify up to two constraints on the optimization. There are two
algorithms available in GT-Power to do the direct optimization.

1. Discrete-grid Optimization: In this type of optimization, the search range for each input parameter
   is shrunk by half every iteration of the algorithm. This method is also called as bisection method
   and it is robust yet simple method to find the solution of the optimization problem. The bisection
   algorithm is often used in root-finding problems and it is initialized with two initial guesses for
   solution between which the solution should lie. The algorithm then shrinks the search range of the
   algorithm by half in every iteration. This method is relatively slow compared to other optimization
   methods like gradient descent methods. The convergence of this method is dependent on the initial
   guesses and this method also expects the objective function to have no local minima. Figure 26
   illustrates the optimization methodology of the bisection method.
2. Brent’s Optimization Method: Brent’s method for optimization combines the bisection algorithm with other optimization methods like the golden search method and parabolic method to increase the convergence rate of optimization. This method requires fewer function evaluations compared to the discrete-grid optimization method thus making it faster in convergence. This method is particularly useful when there are more than 2 independent variables to be optimized. However, this method is not as robust as the discrete-grid method as it may settle at local extrema during the optimization.

GT-Power also provides the option to use an external optimizer for doing the optimization. This is done by running the GT-Power model iteratively and extracting the results from the simulation using a data extraction utility routine called ‘gb2csv’. The response variable or the objective function value extracted can be used by the external optimizer to modify the solution. The new solution of input parameters are fed
to GT-Power model to get the results for the next iteration and this process can be carried out until the optimization goal is reached.

3.3 MATLAB/Simulink for Modeling and Simulation

MATLAB is a technical computing environment from Mathworks Inc. which can be used for computation, visualization and programming of problems that can be represented in a mathematical form. Typical applications of MATLAB include algorithm development, modeling, simulation, data analysis and visualization. MATLAB can also be used for data acquisition and prototyping purposes. The basic data element of MATLAB is an array or a matrix from which it derives its name. Therefore, MATLAB is a very convenient tool to deal with problems involving matrix and vector operations. MATLAB also has a vast number of mathematical functions ranging from elementary functions like sum, trigonometric functions to more sophisticated functions like matrix inverse, eigen-values, fourier transforms etc. These functions make MATLAB a very handy tool to do fast programming when compared to programming languages like C. It is possible to implement scripts in MATLAB similar to C-code to implement control flow, functions and data structures. With the help of MATLAB it is possible to do data visualization with the help of plots in two and three dimension. In addition, MATLAB also offers toolboxes to do specific tasks like optimization, control system development, neural network training etc. These toolboxes are nothing but a collection of MATLAB scripts for various algorithms specific to an area or methodology.

3.3.1 Simulink for Dynamic Simulations

Simulink is a MATLAB based tool developed for modeling and simulation of dynamic systems. It is a graphical block-based tool instead of a programming language. With the help of Simulink, it is possible to model a dynamic system represented in the form of mathematical equations and simulate the dynamic behavior of the system and analyze it. Simulink has a library of blocks which can be used to model a dynamic system. Examples of such block-sets are sinks, sources, linear and non-linear components, connectors etc. For example, if a system’s dynamic behavior can be represented in the form of a transfer function, it can be simulated with the help of a step input block (source) and a transfer function block
(linear component). A scope block (sink) can be used to see the system’s response to a step-input. Simulink can simulate a system in continuous time, discrete time and also in hybrid mode. It can also simulate multirate systems with blocks that are executed at different rates in the system. Simulink also consists of an inbuilt set of numerical solvers that are used to solve the differential, algebraic equations of a system. The choice of the solver depends on whether the system is to be simulated in continuous time or discrete time and the complexity of the dynamic system to be simulated.

Just like MATLAB toolboxes for specific applications, there exist Simulink block-sets for specific applications. Examples of such block sets are Simulink Control Design, Signal Processing Block Set, Parameter Estimation Block Set etc. Simulink also provides tools to do rapid prototyping. Algorithms developed in Simulink can be converted into codes to run on target electronic control units (ECUs). Simulink provides Real-time Workshop to do this conversion of Simulink models into target-specific codes. These codes can be used in conjunction with other real-time simulation tools like dSPACE to carry out rapid prototyping of the algorithms developed.

### 3.3.2 GT-Power and Simulink Interface

GT-Power is a powerful tool to model and simulate complex engine systems. On the other hand, Simulink is a tool that is highly useful in the development of model-based control algorithms and rapid prototyping. In order to exploit the strengths of both the tools, it is useful if they can be used together in a co-simulation.

GT-Power provides an object called ‘SimulinkHarness’ which can be used for this purpose. Any GT-Power control signal like sensor and actuator signals can be exchanged between GT-Power and Simulink for this purpose. The general practice is to send sensor signals to the Simulink control algorithm which outputs the necessary control action in the form of an actuator signal which is again fed back to the GT-Power model. The following figure shows a GT-Power Model embedded into a Simulink model diagram.
3.4 Numerical Tools Used in the Study

This section describes various numerical tools used in this study for modeling and analysis purposes. Proprietary codes of the numerical methods from MATLAB have been used for this study. The following sections describe briefly about the methods and tools used for the study.

3.4.1 Analysis of Variance (ANOVA) for Data Analysis

ANOVA analysis is a statistical method used to determine if differences exist among two or more groups of data. It does this by comparing the means of the groups to see if they are statistically different. The output of ANOVA analysis is typically a table containing Sum of squares, Degrees of Freedom, Mean Sum of Squares, F statistic and p-value. Each of these terms are explained while explaining the different types of analysis that can be done using ANOVA.

1. One way ANOVA: This method analyzes the effect of one characteristic or factor on one dependent variable. An example of such an analysis is to compare whether the bacteria count in different shipments of milk are different. Here the characteristic is shipment number and the dependent variable is bacteria count. The following example of one way ANOVA is taken from MATLAB. In Table 2: Example of 1-way ANOVA, the data from a study of bacteria count from different shipments are given. The bacteria count from six randomly chosen cartons of a given shipment is given in the columns
The idea of the analysis of variance is to take a summary of variability in the observations and partition it into different sources. The first step is to calculate the Total Sum of Squares ($SS_{Total}$). It is given by the sum of squares of deviations of all observations from the grand mean given by Equation 8. The value of the sum of squares for the above data set is also given.

$$SS_{Total} = \sum (Y_{i,j} - \bar{Y})^2 = 1360.17$$  

The next step is to calculate the Sum of Squares ‘within’ each group given by the following formula. This gives an estimate of the variability of the observations within a group. This is sometimes referred to as $SS_{Error}$. The sum of squares within each shipment for the above data is also given below.

$$SS_{Within} = \sum (Y_{i,j} - \bar{Y}_j)^2 = 557.17$$  

<table>
<thead>
<tr>
<th>Shipment No</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bacteria Count ($Y_{ij}$)</td>
<td>24</td>
<td>14</td>
<td>11</td>
<td>7</td>
<td>19</td>
</tr>
<tr>
<td></td>
<td>15</td>
<td>7</td>
<td>9</td>
<td>7</td>
<td>24</td>
</tr>
<tr>
<td></td>
<td>21</td>
<td>12</td>
<td>7</td>
<td>4</td>
<td>19</td>
</tr>
<tr>
<td></td>
<td>27</td>
<td>17</td>
<td>13</td>
<td>7</td>
<td>15</td>
</tr>
<tr>
<td></td>
<td>33</td>
<td>14</td>
<td>12</td>
<td>12</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td>23</td>
<td>16</td>
<td>18</td>
<td>18</td>
<td>20</td>
</tr>
<tr>
<td>Mean ((\bar{Y}_j))</td>
<td>23.8333</td>
<td>13.3333</td>
<td>11.6667</td>
<td>9.1667</td>
<td>17.8333</td>
</tr>
<tr>
<td>Sample Variance</td>
<td>36.1667</td>
<td>12.6667</td>
<td>14.2667</td>
<td>25.3667</td>
<td>22.9667</td>
</tr>
<tr>
<td>Grand Mean of all data ((\bar{Y}))</td>
<td>15.1667</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 2: Example of 1-way ANOVA
In the next step, the sum of squares for the observations of all the groups is calculated. This is different from $SS_{total}$ as the mean of each group is used instead of the each sample for the calculation. The formula for the same and the value for the above data set are given below. This sum of squares is often referred to as $SS_{Among}$.

$$SS_{Among} = \sum (\bar{Y}_j - \bar{Y})^2 = 803$$

Equation 10

It can be observed from the above three equations that $SS_{Total} = SS_{Among} + SS_{Within}$. Associated with each sum of squares is a degree of freedom. In general, one starts with as many degrees of freedom as there are observations and loses one degree of freedom for every sample mean calculated. For the $SS_{Total}$ there is one grand sample average, therefore there are $N - 1$ degrees of freedom where $N$ is the total number of observations. There are $n_i - 1$ degrees of freedom within each group where $n_i$ is the number of observations in each group. Therefore, there are $\sum n_i - 1 = N - k$ degrees of freedom for $SS_{Within}$ where $k$ is the number of groups. That leaves $k - 1$ degrees of freedom for $SS_{Among}$. The sum of squares divided by the associated number of degrees of freedom gives the mean sum of squares. The sum of squares, degrees of freedom and mean sum of squares for the above data set can be summarized as follows.

<table>
<thead>
<tr>
<th>Source of Variation</th>
<th>Sum of Squares</th>
<th>Degrees of Freedom</th>
<th>Mean Sum of Squares</th>
</tr>
</thead>
<tbody>
<tr>
<td>Among</td>
<td>803</td>
<td>4</td>
<td>200.75</td>
</tr>
<tr>
<td>Within</td>
<td>557.17</td>
<td>25</td>
<td>22.287</td>
</tr>
<tr>
<td>Total</td>
<td>1360.17</td>
<td>29</td>
<td></td>
</tr>
</tbody>
</table>

Table 3: One-way ANOVA Table
Each mean square in the above table captures different elements of variability. The mean square within is also sometimes referred to as mean square error. Within sample variability is attributed to the sampling error. The mean square among contains some of this error variability but also variability due to differences among group means. To compare the significance of the differences among the group means, a ratio

\[ F = \frac{MS_{\text{Among}}}{MS_{\text{Within}}} \]

is defined. Values of \( F \) close to 1 can indicate that the differences in means can be attributed to natural or random variability. Values of \( F \) much larger than 1 indicate a significant variability of means between the groups. The values of \( F \) which signify this variability can be read out from a table of the \( F \) distribution. This tabulation assumes that the original observations are normally distributed with a common error variance. From the \( F \) distribution, a \( p \)-value can be calculated depending on the critical \( F \) value. The \( p \)-value is usually calculated with a confidence interval of 95% that the variability seen is because of random variability (null hypothesis). A high \( p \)-value represents that the null-hypothesis is true and a very small \( p \)-value represents that the samples in each group are indeed different. Figure 27 gives the output of MATLAB’s \texttt{anova1} script for the above problem. It can be observed that the \( p \)-value in this case is very small indicating that the mean of bacteria count in each shipment were indeed different.

![Figure 1: One-way ANOVA](image)

**Figure 27:** One-way ANOVA Table for the Shipment Problem generated by MATLAB

1. Two-way ANOVA: A two way ANOVA analysis is different from the one-way in that the groups in two way ANOVA have two defining characteristics instead of one. The analysis for a two way ANOVA is the same except that the sum of squares is calculated both for rows as well as columns of
data and also for the interactions between the two factors. The following table gives the output of two-way ANOVA analysis on the effect of car model and the factory on the mileage of a car. There are three models of cars and two factories. Three cars of the same model from each factory are selected for the study. The data is given in Table 4. The result of the ANOVA analysis on the data using *anova2* function in MATLAB is given in Figure 28: Two-way ANOVA results for the car mileage problem. It can be observed that the model and the factory have a significant influence on the variability in means. This can be concluded from the p-values for the rows and columns from Figure 28. The interactions are not significant as explained by the p-values which also say that the observed result is likely even without the effect of interactions.

<table>
<thead>
<tr>
<th>Mileage (mi/gallons)</th>
<th>Model A</th>
<th>Model B</th>
<th>Model C</th>
</tr>
</thead>
<tbody>
<tr>
<td>Factory 1</td>
<td>33.300</td>
<td>34.500</td>
<td>37.400</td>
</tr>
<tr>
<td></td>
<td>33.400</td>
<td>34.800</td>
<td>36.800</td>
</tr>
<tr>
<td></td>
<td>32.900</td>
<td>33.800</td>
<td>37.600</td>
</tr>
<tr>
<td>Factory 2</td>
<td>32.600</td>
<td>33.400</td>
<td>36.600</td>
</tr>
<tr>
<td></td>
<td>32.500</td>
<td>33.700</td>
<td>37.000</td>
</tr>
<tr>
<td></td>
<td>33.000</td>
<td>33.900</td>
<td>36.700</td>
</tr>
</tbody>
</table>

Table 4: Mileage of different models of cars manufactured by different factories
Figure 28: Two-way ANOVA results for the car mileage problem

2. N-way ANOVA: When the number of characteristics or factors under study is more than two, n-way ANOVA analysis can be used. The same concepts used for the two-way ANOVA can be extended to n-way ANOVA. And the ANOVA analysis outputs the significance of a single factor on the dependent variable and all the possible interactions between the factors. The n-way ANOVA analysis can be done using MATLAB’s \texttt{anovan} function. With the help of this function, it is possible to define the factors and the interactions that need to be considered for a given data set. Figure 29 gives the ANOVA analysis of a dataset to find the effect of a car’s origin country, number of cylinders and manufacturing date on the mileage. In this analysis only the effect of the interaction between the origin and the manufacturing date is selected for study along with the primary factors.
ANOVA analysis is a powerful statistical tool to do variability analysis. In this study, ANOVA analysis has been used to analyze the effect of various engine variables on engine emissions and performance. The data for the analysis is obtained from a design of experiments done on a virtual engine on GT-Power. The results of the ANOVA analysis done in the study will be discussed in detail in Chapter 3.

3.4.2 Artificial Neural Networks

Artificial Neural Network (ANN) is a parametric black-box modeling technique used to represent input-output behavior of systems. They are highly versatile in nature and can learn complex input-output relationships like highly non-linear behavior exhibited by some systems. ANNs are often used to substitute complex physics-based models in a system for reducing the computational complexity.

An artificial neural network is a mathematical structure that emulates the behavior of a biological neural network. It usually consists of a large number of highly interconnected processing elements also called as neurons. An artificial neural network, like the human brain, learns by example. Therefore, artificial neural networks find application in pattern recognition and data classification. Just like a biological neural network which can learn to carry out complex functions, an artificial neural network can also learn complex behavior in data which conventional models are not capable of.
In the human brain, a typical neuron collects signals from others through a host of fine structures called *dendrites*. The neuron sends out spikes of electrical activity through a long, thin strand known as an *axon*, which splits into thousands of branches. At the end of each branch, a structure called a *synapse* converts the activity from the axon into electrical effects that inhibit or excite activity from the axon into electrical effects that inhibit or excite activity in the connected neurons. When a neuron receives excitatory input that is sufficiently large compared with its inhibitory input, it sends a spike of electrical activity down its axon. Learning occurs by changing the effectiveness of the synapses so that the influence of one neuron on another changes. An artificial neural network is very similar to a biological neural network. It consists of multiple interconnected neurons. In the process of learning, the artificial neural network learns when to fire a particular neuron. An artificial neural network usually has an input layer, a layer of neurons and an output layer. The number and the type of the neuron depend on the complexity of data that the network has to learn. Each neuron can be viewed as an activation function that takes a particular value depending on the inputs.

In general an artificial neural network has the following elements as represented in the form of a diagram in Figure 30.

1. A set of processing units (neurons,)
2. A state of activation $y_k$ for every unit, which is equivalent to the output of the unit
3. Connections between the units. Generally each connection is defined by a weight $w_{jk}$ which determines the effect which the signal of unit $j$ has on unit $k$
4. A propagation rule, which determines the effective input $v_k$ of a unit from its external inputs;
5. An activation function $\Psi_k$, which determines the new level of activation based on the effective input $v_k(t)$ and the current activation $y_k(t)$ (i.e., the update)
6. An external input (aka bias, offset) $b_k$ for each unit

68
7. A method for information gathering (the learning rule)

8. An environment within which the system must operate, providing input signals and -- if necessary -- error signals.

![Figure 30: Structure of each neuron in an ANN](image)

Each unit or neuron receives inputs from its neighbors or external sources and uses this to calculate an output which is propagated to other neurons. A neural network can have a number of such neurons acting parallel on their inputs and calculating the outputs. As explained above, a neural network will have an input layer which takes inputs from external sources, an output layer which sends data out of the neural network and hidden layers which consist of neurons whose inputs and outputs are hidden inside the neural network itself. The input to the activation function of a given neuron is nothing but the weighted sum of all the inputs connected to the neuron plus a bias. This value $v_k$ is passed through the activation function. Often the activation function is a non-decreasing threshold function of its input. The threshold function can be hard-limiting (sgn) or linear or smoothly limiting function (sigmoid) as shown in Figure 31.
3.4.2.1 Neural Network Topologies

In the previous section, the properties of the basic element of a neural network (neuron) were discussed. In this section, the pattern of connection between the neurons and the data flow are discussed. Neural networks can be broadly classified into two groups based on their topology:

1. Feed-forward networks: Neural networks where the data propagation from input to output is strictly feed-forward are called feed-forward networks. Such networks can have multiple layers of neurons and there is not feedback of the output of the neurons to the inputs. These networks are most suitable for modeling systems that are static in nature.

2. Recurrent Networks: In these networks, the previous state or outputs of a neuron are fed back as inputs along with new inputs to calculate the new outputs. These networks are useful in modeling the dynamic behavior of a system.

In this study, only feed-forward neural networks have been used and therefore the discussion will be limited to feed-forward networks.

3.4.2.2 Neural Network Training Algorithms

As discussed above a neural network has to be trained to respond in a particular set of inputs in a particular fashion. This training involves learning of the weights associated with each connection between the
neurons and the biases associated with each neuron. The simplest neural network has one or more input neuron and an output neuron which are connected through a series of weights and a bias as shown in the following figure.

Rosenblatt first proposed this structure and called it a perceptron. The output of the above neural network structure can be represented by Equation 11 where $\Psi$ is an activation function. If $\Psi$ is a hard-limiting activation function, this structure of neural networks can be used for data classification purposes as the output of the neural network can be either 1 or -1.

$$y = \Psi\left(\sum_{i=1}^{2} w_i x_i + \theta\right)$$  \hspace{1cm} \text{Equation 11}$$

The learning rule for the weights and bias of a single layer perceptron is given by the [41] where $d(x)$ is the desired output of the perceptron and $x_i$ is the $i^{th}$ input of the perceptron. This weights and biases are updated only when the perceptron does not respond correctly for a given input.

$$\Delta w_i = d(x) x_i$$  \hspace{1cm} \text{Equation 12}$$

$$\Delta \theta = d(x)$$
This training algorithm was further generalized by Widrow and Hoff using the least mean square learning procedure which is generally used in data fitting for linear models. The output of the least mean square error minimization algorithm modified the learning rule as follows where $\delta$ is the error in the output of the neural network and $\gamma$ is a constant of proportionality. Therefore, this learning rule is also called as the delta rule.

$$\Delta w_j = \gamma \delta_j x_j$$  \hspace{1cm} \text{Equation 13}$$

However, Widrow and Hoff have given the delta rule only for a single layer neural network model with a hard-limiting activation function. Rumelhart, Hinton and Williams extended this delta learning rule to a multi-layer feed-forward network. A multi-layer feed-forward network is shown in Figure 32. This method of learning the weights for a multi-layer rule is also called as Back-Propagation algorithm. The central idea of this algorithm is that the errors in the output layer of the multi-layer networks are back-propagated to the hidden layers to update the weights and biases of the hidden layers. It can be proved that the weights updation rule will remain the same as Equation 13 for the output layer of a multi-layer feed-forward network. But for the weights pertaining to the hidden neurons should be updated by using the back-propagation rule. The updation rule is given by Equation 14 where $h$ represents a hidden neuron and $o$ represents an output neuron. This rule requires the derivative of the activation function to be calculated at its input value. Several improvements have been made to the back-propagation algorithm to make it reliable. The effectiveness of the back-propagation algorithm depends on the number of iterations one can afford to minimize the error, the number of learning samples available for training and the number of hidden units.
\[ \Delta w_{bh} = \gamma_b \delta_{h} x_j \]

\[ \delta_{h} = F'(v_h) \sum_{a=1}^{N} \delta_{a} w_{ha} \]

Equation 14

\[ \text{Figure 32: A multi-layer feed-forward network} \]

In this study feed-forward neural networks have been used to model the non-linear behavior of diesel engines. MATLAB’s neural network toolbox has been extensively used for the training purposes. MATLAB provides a highly powerful toolbox with training algorithms like the back-propagation algorithm for training feed-forward neural networks.

3.4.3 Genetic Algorithms

Genetic Algorithms are a class of Evolutionary Algorithms that make use of evolutionary concepts like inheritance, mutation, selection and cross-over to solve optimization and search problems. They find
application in lot of fields including engineering, economics and computational science. The basic requirements for implementing a genetic algorithm are as follows

1. A genetic representation of the pool of probable solutions to a problem. This genetic representation is known as the chromosome or the population and the individual members of this solution are called the phenotypes or individuals. The population evolves from one generation to the next generation by using inheritance, mutation, selection, and cross-over techniques just like any biological creature evolves.

2. For the population to evolve, a selection criterion has to be identified which determines the fitness of a particular individual of the population. This criterion is also called as the fitness function. The evolution process has to take place such that the fitness value of each individual in a given population becomes better from generation to generation. Finally, the solution of the problem is the individual with the best fitness value which is predefined.

3.4.3.1 Optimization Process Using Genetic Algorithms

The process of optimization through genetic algorithms can be summarized by the flow chart given in Figure 33. The process consists of four important elements

1. Initialization: The initialization stage consists of creating an initial population which is randomly generated depending on any constraints on the individuals. This initial population evolves through generations by selection, mutation and cross-over of the individuals of the population.
2. Selection: Selection is the stage where the individuals with the best fitness values are selected for reproduction. The fitter the individual, the more number of times it will be selected to reproduce. Different methods are used for the selection process. In a tournament method of selection, a tournament is conducted between the individuals of the population and the winner is selected. Another method commonly used for selection is Roulette wheel selection also called as fitness proportionate selection. In this selection, a fitness probability is assigned to each individual of a population which is nothing but the ratio of the individual's fitness value to the sum of fitness values of all individuals. A Roulette wheel is simulated with these probability values. The individuals with higher probability are more likely to be selected for reproduction.
3. **Cross-over:** In this stage, characteristics of two parents are crossed-over to produce two offsprings. The most common way of doing cross-over is to randomly find a locus in the parents and interchange the sub-sequences before and after the locus to form two off-springs. Examples of crossover methods are one-point cross over, two-point crossover and heuristic. This process emulates biological recombination between two single chromosome organisms.

4. **Mutation:** In this process, the characteristics of an individual are randomly mutated. For example, an individual with a genetic binary representation of 00001000 can be randomly mutated at the second position to 01001000. A common method used for mutation is to generate a random variable for every bit of an individual and decide whether the bit will be mutated or not. Uniform and Gaussian distributions are normally used in mutation process. This process also closely mimics biological mutation.

In this study, genetic algorithms are used to solve optimization problems in engine calibration and design of experiments. The Genetic Algorithm Toolbox from MATLAB and customized Fortran codes of genetic algorithm have been used to do the optimization in the study.
4.1 Introduction to Sensitivity Analysis

This chapter describes the sensitivity analysis done on a high fidelity GT-Power diesel engine model to identify the sensitivity of diesel engine performance and emissions to disturbances in operating conditions. The sensitivity of various combustion metrics to disturbances in engine operating variables is also identified in this process. Firstly, a design of experiments (DOE) on a GT-Power based diesel engine model is conducted with root-cause variables of the engine as factors and engine emissions, fuel consumption and combustion metrics as the response variables. Analysis of Variance (ANOVA), a statistical approach described in Chapter 2, has been used to identify the significance of each engine root cause variable on the response variables. The results from the sensitivity analysis have been used to identify various possibilities of designing the closed-loop combustion control architecture. A viable architecture is then chosen in order to demonstrate closed-loop combustion control on the heavy-duty diesel engine under consideration. The following sections describe in detail the high-fidelity diesel engine model, design of experiments and the ANOVA analysis conducted on the DOE results.

4.2 High Fidelity Diesel Engine Model

For doing the sensitivity analysis, a high fidelity diesel engine model developed in GT-Power has been used. The model was originally developed and calibrated by Cummins. GT-Power is a tool widely used for engine modeling and simulation and has the ability to model complex engine configurations. As discussed in Chapter 2, GT-Power also has the capability to model combustion to an acceptable level of accuracy needed for sensitivity studies of this nature. The engine considered for study and modeled in GT-Power is a six-cylinder heavy duty diesel engine equipped with a turbo-charger, cooled exhaust gas recirculation loop and an intercooler. A DI Jet model developed by Yoshizaki et.al and Morel et.al. which uses a multi-zone
approach to model the diffusive phenomenon of conventional diesel combustion is used in the diesel engine model. A snapshot of the GT-Power model of the engine is shown in Figure 34.

![Figure 34: Snapshot of Detailed GT-Power Engine Model](image)

The main components of the diesel engine model are described below.

1. Cylinders: The cylinders are modeled using the ‘EngCylinder’ object. The combustion in cylinders has been modeled using the DI Jet Model. For the sensitivity analysis, cylinders 1 and 6 have been chosen to run as independent cylinders with detailed combustion models whereas cylinders 2-5 have been designated as ‘slave’ cylinders. By using the ‘slave’ attribute for a cylinder, the combustion rate calculated in the ‘master’ cylinder (here cylinder 1) will be imposed rather than being calculated. This helps in reducing the computation time associated with using a detailed combustion model like the DI Jet model. A limitation of this approach is that the effect of
cylinder-to-cylinder combustion variability cannot be obtained for cylinders marked as slave. Since cylinders 1 and 6 were found to be on the extremes of cylinder-to-cylinder variability, detailed combustion models in the two cylinders can provide the necessary information on the maximum effects of cylinder-to-cylinder variability. Heat transfer inside the cylinder is modeled using the Woshni Heat Transfer Correlation proposed by Woshni. Woshni’s correlation models the convective heat transfer \( \dot{Q}_{ht} \) between the combustion gases and the cylinder walls by Equation 15 where \( h_c \) is the convective heat transfer coefficient, \( A \) is the heat transfer area, \( T_{gas} \) and \( T_{wall} \) are instantaneous gas and wall temperatures, \( B \) is the bore of the cylinder, \( p \) is the instantaneous gas pressure, \( T \) is the instantaneous gas temperature, \( w \) is the cylinder gas velocity. The heat transfer calculation is important for the heat release rate analysis of combustion which is critical in calculating combustion metrics.

\[
\dot{Q}_{ht} = h_c A(T_{gas} - T_{wall}) \\
h_c (W / m^2.K) = 3.26 B (m)^{-0.2} p (kPa)^{0.8} T (K)^{-0.55} w (m/s)^{0.8}
\]

Equation 15

2. Injectors: The injectors are modeled using a ‘InjMultiProfConn’ object which allows the modeling of multiple pulses of injection. The total mass of fuel to be injected in each pulse, the injection timings (in crank angle degrees) and the injection profile can be specified in the injector model. The simulated diesel engine had three injection pulses, pilot, main and post, which were modeled using this injector model.

3. Engine Cranktrain: The ‘EngineCrankTrain’ object has been used to model the cranktrain. This object specifies the attributes of an engine's cranktrain and models the crank-slider mechanisms and crankshaft which translate the torques generated by individual cylinders into the crankshaft output torque.
4. Intake and Exhaust Manifolds: The intake and exhaust manifolds have been modeled using ‘Pipe’ and ‘FlowSplit’ objects. The flow dynamics in the pipes are modeled in one-dimension along the flow direction by solving the Navier Stokes Equations. In the case of ‘FlowSplit’ objects, the momentum equation is solved in all the three dimensions. Exact geometric specifications of the manifolds can be specified using the Pipe and FlowSplit models.

5. Exhaust Gas Recirculation: Exhaust Gas from the exhaust manifold is recirculated into the intake manifolds. The exhaust gas recirculation flow is controlled through a valve modeled as an orifice using the ‘Orifice’ connection. The exhaust gas is cooled before entering the intake manifold. The heat transfer between the exhaust gas and the coolant is modeled by assuming a constant coolant temperature. The constant coolant temperature is specified as the wall temperature for the exhaust gas recirculation pipes.

6. Variable Geometry Turbocharger and Intercooler: The variable geometry turbocharger is modeled with the help of a ‘compressor’, ‘turbine’ and ‘FreeShaft’ models. Compressor maps and turbine maps can be specified in the models. The variable geometry command for the turbine is given as a control input to the turbine model. Turbine maps are specified for a discrete set of variable geometry commands. The maps for geometry other than those specified will be generated through interpolation. The shaft model simulates the turbocharger dynamics. The intercooler is modeled as a simple heat exchanger with variable efficiency. The efficiency of the heat-exchanger varies with the mass flow rate of air through the intercooler.

4.2.1 Validation of the High Fidelity Diesel Engine Model

The detailed GT-Power model has been calibrated by Cummins for optimal engine performance and emissions at 149 engine operating points. Figure 35 represents the calibration data points on the engine torque-speed plane. The calibration procedure involves identifying the set-points for engine air handling system, injection and cooling systems which result in optimal engine performance. The engine parameters that have been calibrated at each of the 149 operating parameters are as follows
1. Pilot, Main and Post injection quantities

2. Pilot, Main and Post injection timings

3. EGR valve position for optimal EGR fraction

4. Variable Geometry Turbocharger (VGT) position for Boost Pressure

5. Rail Pressure

6. EGR Coolant Temperature

7. Charge Coolant Temperature

These calibrated data points have been simulated on the GT-Power model and the outputs have been compared against experimental data from the engine test cell. For the sake of comparison, the actual data has been normalized for confidentiality reasons. Engine variables like brake torque, engine intake and exhaust flow variables predicted by the detailed GT-Power model are compared with the test cell data as shown in Figure 36. It can be observed that the GT-Power model predicts the engine performance with good accuracy. Engine torque is predicted within ±5% error. Other flow variables like intake manifold pressure and temperature (Figure 37), EGR and charge flows (Figure 38) and turbine inlet pressure and temperature (Figure 39) are predicted within ±10% error in most cases.
Figure 40 compare the Indicated Specific Fuel Consumption, NOx concentration and soot prediction of the high fidelity DI-Jet combustion model. The accuracy of prediction of ISFC and NOx is agreeable since the right trends in prediction are sufficient for studies of this nature. Soot prediction by the combustion model is highly inaccurate and does not catch the trends which leads to omitting soot as a response variable for the sensitivity analysis. However, NOx-soot tradeoff is of great importance to diesel engines and the omission of soot in this study is only because of inadequacy of an accurate prediction model.
Figure 36: Comparison of Brake Torque

Figure 37: Comparison of Intake Manifold Pressure and Temperature
Figure 38: Comparison of Charge Flow and EGR Flow

Figure 39: Comparison of Turbine Inlet Pressure and Temperature
Figure 40: Comparison of Experimental and Predicted Indicated Specific Fuel Consumption

Figure 41: Comparison of Experimental and Predicted NOx Concentration
4.3 Sources of Variability in Engine Operation

The goal of the sensitivity analysis is to identify the sensitivity of the engine’s performance to various sources of variability in the operation of the engine. Calibration of a modern day diesel engine is a formidable task due to its high number of degrees of freedom. Due to the increased number of calibrated parameters, it is inevitable for engine calibrators to make assumptions to simplify the calibration process. These assumptions can surface out as disturbances to ideal engine operation and may lead to deterioration in engine performance. In the following sections, various sources of variability have been identified and their possible effects on the engine performance have been described.

4.3.1 Design-Driven Sources of Variability

Variability induced by virtue of the design of the engine can be called design-driven variability. Non-uniform distribution of fresh air and residual gases amongst the cylinders of a multi-cylinder engine is an example of design driven variability because it is a natural consequence of flow dynamics of the intake
manifold. In an engine with in-line cylinder arrangement, depending on the intake manifold design, the trapped mass compositions between the cylinders can be different leading to different combustion conditions. Although, care is taken by engine designers during the design of the intake manifold to minimize the maldistribution, it is practically impossible to completely eliminate it. Conventionally, during the engine calibration the cylinders are assumed to have the same ratio of fresh air and residuals and the engine is calibrated for engine-out emissions and performance rather than for individual cylinders. This assumption is made to eliminate the need to calibrate individual cylinders’ performance. By considering such variability induced by design, it is possible to further improve engine-out emissions and performance by optimizing the performance of each cylinder. For example, the differences in the cylinder conditions can be compensated with the help of different injection parameters for each cylinder.

In order to study the cylinder-to-cylinder variability, a study on the detailed GT-Power model has been conducted where the in-cylinder conditions of all the cylinders have been observed for different operating conditions. Since the GT-Power model is a deterministic model, it can be used only for analyzing the systematic variability induced by design. Figure 43 shows the design of the intake manifold for the heavy-duty diesel engine under study. The flow distribution between the cylinders can be expected to be symmetric about cylinder number three where the runners of the intake manifold diverge to the remaining cylinders.
To quantify the variability in cylinder conditions, the GT-Power model has been simulated for all the calibration points shown in Figure 35. The following variables for each cylinder have been extracted from the simulation results and a cylinder-to-cylinder mass distribution study has been carried out.

1. Trapped Fresh Air at Cycle Start
2. Trapped Residuals at Cycle Start
3. Trapped Mass at Cycle Start
4. Average Temperature of Incoming Gases

The variation in each of the above variables from the mean value for each simulated case has been plotted as histograms in Figure 44. The summary of the cylinder to cylinder variability is given by Table 5. It can be observed that the variation in the cylinder conditions can be high. The maximum variability between any two cylinders (in this case always the 1st and 6th cylinders because of inline cylinder configuration) indicates significant differences in the composition of the extreme cylinders. In particular, it can be observed that the cylinder 1 receives a higher percentage of fresh air and lower percentage of residual fraction. It is vice-versa for cylinder 6. Higher fresh air leads to more oxygen content and lower residuals leading to higher combustion temperature both of which contribute to higher NOx levels in cylinder 1 and vice-versa in cylinder 6. On the other hand, higher amount of fresh air increases combustion efficiency leading to better fuel consumption in cylinder number one and vice-versa in cylinder number six. These results undermine the assumption of uniform cylinder conditions during combustion while calibrating for engine performance. By considering the cylinder-to-cylinder variability, it is possible to optimize the performance of individual cylinders thus improving over-all engine performance. Closed-loop combustion control can be designed to see cylinder-to-cylinder variability as a disturbance to optimal engine performance and compensate for it. Another advantage of considering cylinder-to-cylinder variations is balancing torque of the cylinders thus improving the overall drivability of the vehicle.
Figure 44: Distribution of Fresh Air in Cylinders 1 to 6

Figure 45: Distribution of Residuals in Cylinders 1 to 6
Figure 46: Distribution of Total Trapped Mass in Cylinders 1 to 6

Figure 47: Distribution of Charge Temperature in Cylinders 1 to 6
<table>
<thead>
<tr>
<th>Variable Name</th>
<th>Standard deviation (cylinder-to-cylinder)</th>
<th>Maximum deviation from mean cylinder in all the cases</th>
<th>Maximum variability between any two cylinders in all the cases</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trapped Fresh Air</td>
<td>±2.7%</td>
<td>±4%</td>
<td>7%</td>
</tr>
<tr>
<td>Trapped Residuals</td>
<td>±5.7%</td>
<td>±8%</td>
<td>14%</td>
</tr>
<tr>
<td>Trapped Mass</td>
<td>±1.3%</td>
<td>±3%</td>
<td>6.5%</td>
</tr>
<tr>
<td>Average Intake Temperature</td>
<td>±1.4%</td>
<td>±6 K</td>
<td>12 K</td>
</tr>
</tbody>
</table>

Table 5: Summary of Cylinder-to-cylinder variability

4.3.2 Component-Driven Variability

Component-to-component variability is another disturbance that is usually neglected during the calibration phase of an engine. It is assumed that the calibration done on a particular engine will be optimal for all the other engines of the same type and make. However, component-to-component variability due to manufacturing tolerances and component aging can lead to differences in engine performance. This can lead to variability in performance from engine to engine or component to component within the same engine. For example, all the fuel injectors in a given engine are not exactly the same and may have differences in their injection characteristics due to manufacturing defects or aging. Component driven variabilities may affect the engine performance by affecting the root-cause variables of the engine. Root cause variables are the engine variables like air and EGR flows, fuel injection parameters etc. which characterize engine operating conditions. Therefore, it is sufficient to study the sensitivity of engine performance to variability in the root-cause variables to study the effect of component-to-component variability.

4.3.3 Environmental-Driven Variability

This category of variability refers to the effect of engine environmental conditions on its performance. All the external systems that the engine has to interact with become its environment and impose boundary
conditions on the operation of the engine. Variations in engine environmental conditions can be in the form of variations in ambient conditions (temperature and pressure), variations in fuel properties due to differences in quality of fuels and variations in boundary conditions imposed by an after-treatment system on the engine like back-pressure from a diesel particulate filter. Variations in ambient conditions are often taken into consideration in the conventional calibration processes and engine calibration maps are corrected for changes in ambient conditions. As a result, there is still a lot of scope for improvement in engine performance by accounting for environmental disturbances in a robust way. It has to be noted that variations in engine environmental conditions affect all the cylinders equally in an engine and therefore are much easier to deal with when compared to cylinder-to-cylinder or component driven variabilities. Apart from the variations in fuel properties, all other environmental variations affect the charge conditions of the engine. Therefore, a sensitivity of engine performance to environmental sources of variability can be studied with the help of the root-cause variables. In this study, the effect of variability in fuel properties has not been studied as the GT-Power model lacks the modeling capability to simulate such effects.

4.3.4 Effects of Engine Transience on Performance

In this section, the effect of transient operation of engine on its performance is discussed. The conventional practice in engine calibration is to calibrate the engine’s performance and emissions for steady-state operation. Therefore this calibration can be expected to give optimal results in the field only during steady-state operation of the engine. However, during transient operations, engine performance can be far from optimal if the same steady-state calibrations are used. This problem is compounded by the existence of sub-systems which exhibit different dynamic behavior. Fuel injection systems in an engine have the smallest time constants and are capable of changing injection parameters very rapidly. During load changes, the fuel injection system can supply the necessary change in fuel quantity by the next engine cycle. Therefore load changes in engines take place very rapidly. A change in engine fueling amount may change the engine speed depending on vehicle load. Speed changes are usually slower compared to changes in the load. The slowest dynamics in the engine can be expected from the air handling sub-system. For example, a change request in boost pressure requires the turbocharger to operate at higher speed resulting in ‘turbo-lag’ and
this lag is compounded by the intake manifold volume filling dynamics. Although engine calibrations are tested for their performance on transient drive-cycles like the FTP, they are not necessarily optimized for such transient performance. A classic example of transient operation that can lead to undesired engine performance is during load changes of an engine. Load changes of an engine are driven by the driver’s command and are almost instantaneous in nature. However, the boost and EGR changes associated with a load change are not instantaneous in nature due to their slower dynamics. Most engine controllers schedule different calibrated parameters only as a function of engine speed and load assuming that the rest of the system would have reached steady-state. Therefore, during such transient events, the engine performance may not be optimal as the dynamics of the entire system are not taking into consideration during the scheduling of control parameters. In modern engines, discrepancies like these are corrected for by correcting the calibration maps for transient operation of the engine. The engine calibrations therefore have to go through multiple corrections which leads to increase in the calibration effort without a significant increase in the performance due to the ad-hoc nature of the corrections. As a result, the engine performance is not necessarily optimal for the operating conditions. In this study, kernel-based interpolation methods are used to improve the transient performance of the engine by scheduling the set-points and feed-forward component of closed-loop combustion control based on the dynamic variables of the engine. This procedure will be discussed in detail in Chapter 4.

4.4 Sensitivity Analysis through Design of Experiments

In the previous section, various disturbances that can potentially affect engine performance have been identified. All the disturbances except the variabilities in fuel properties affect the root-cause variables of the engine which leads to loss of performance. In this section, a design of experiments developed for the root-cause variables of a diesel engine has been discussed along with the analysis of results. Analysis of Variance (ANOVA) has been applied to the results from design of experiments to identify the key factors that influence engine performance, like Indicated Specific Fuel Efficiency and NOx emissions. For this analysis, steady-state operating points closest to the 13-mode test have been chosen from the calibrated data set available for the GT-Power model of the heavy duty diesel engine. The 13-mode test is a standard in the
heavy-duty engine industry which includes 13 steady-state operating points of the engine characterized by percentage of nominal speeds and loads. Each steady-state point is assigned a weight and the emissions are calculated as a weighted sum of engine-out emissions at each steady-state point. Steady-state points with higher expected frequency of operation are assigned higher weights and vice-versa. Table 6 gives the 13 modes for a heavy-duty diesel engine along with the speeds and loads each mode represents and the weighting factor associated with each mode. Figure 48 plots the calibrated points closest to the eleven out of thirteen modes for the heavy-duty engine under study. Two modes at engine idle have been ignored in this analysis due to the unavailability of calibration at engine idle conditions.

<table>
<thead>
<tr>
<th>Mode</th>
<th>Speed (% of nominal)</th>
<th>Load (% of nominal)</th>
<th>Weighting factor</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>idle</td>
<td>-</td>
<td>0.410/2</td>
</tr>
<tr>
<td>2</td>
<td>40</td>
<td>20</td>
<td>0.037</td>
</tr>
<tr>
<td>3</td>
<td>40</td>
<td>40</td>
<td>0.027</td>
</tr>
<tr>
<td>4</td>
<td>idle</td>
<td>-</td>
<td>0.410/2</td>
</tr>
<tr>
<td>5</td>
<td>60</td>
<td>20</td>
<td>0.029</td>
</tr>
<tr>
<td>6</td>
<td>60</td>
<td>40</td>
<td>0.064</td>
</tr>
<tr>
<td>7</td>
<td>80</td>
<td>40</td>
<td>0.041</td>
</tr>
<tr>
<td>8</td>
<td>80</td>
<td>60</td>
<td>0.032</td>
</tr>
<tr>
<td>9</td>
<td>60</td>
<td>60</td>
<td>0.077</td>
</tr>
<tr>
<td>10</td>
<td>60</td>
<td>80</td>
<td>0.055</td>
</tr>
<tr>
<td>11</td>
<td>60</td>
<td>95</td>
<td>0.049</td>
</tr>
<tr>
<td>12</td>
<td>80</td>
<td>80</td>
<td>0.037</td>
</tr>
<tr>
<td>13</td>
<td>60</td>
<td>5</td>
<td>0.142</td>
</tr>
</tbody>
</table>

Table 6: 13-modes in terms of engine speed and load and the corresponding weighting factors
4.4.1 Design of Experiments

The factors for the design of experiments should be chosen such that the root-cause variables that affect engine operation are included. Any variabilities described in the previous sections, except for changes in fuel properties, affect engine performance through the root cause variables. Root-cause variables, as the name suggests, are the basic necessities for combustion like fresh air, residuals and injection parameters. Other variables, like swirl induced by the cylinder intake port design, can affect the mixing conditions inside the cylinder and therefore can also be called a root-cause variable.

Table 7 gives the root-cause variables chosen for the study along with the actuators used to induce variation in each of the variables. The root cause variables are perturbed from their nominal values by an amount shown in the third column of the table. The variation in the root-cause variables is indicative of the effect of various variabilities discussed earlier as observed by Cummins. The number of levels for each factor of the
DOE is indicated in the third column of the table. The center of DOE is taken as the base calibration available for each actuator command and the actuator command is varied such that the desired variation is achieved in the root-cause variable. A full-factorial DOE with the factors, variation and levels described in Table 7 is set-up in the GT-Power model of the diesel engine under study. Three levels have been assigned for each factor which gives $3^6 = 729$ cases for each mode for a full-factorial DOE. In order to simulate cylinder-to-cylinder variation in combustion due to variability in the root-cause variables, DI Jet combustion models are chosen to run for cylinders 1 and 6 and cylinders 3-5 have been chosen as slave cylinders in the GT-Power model. This results in less computational burden compared to running DI Jet combustion models on all the six cylinders and at the same time gives an estimate of the largest effect of cylinder-to-cylinder variation since cylinders 1 and 6 are the extremes of this phenomenon. The DOE cases for all the modes have been simulated and the results from the DOE have been analyzed to extract the sensitivity of engine performance to the root-cause variables.

<table>
<thead>
<tr>
<th>Root Cause Variable</th>
<th>Actuator to change root cause variable in the DGM</th>
<th>Variation in parameters from calibrated value, Levels</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rail Pressure</td>
<td>Rail Pressure in the Common Rail</td>
<td>±20 bar, 3 Levels</td>
</tr>
<tr>
<td>Start of Injection</td>
<td>Injector Command</td>
<td>±0.5 CAD, 3 Levels</td>
</tr>
<tr>
<td>Total Fuel Amount</td>
<td>Injector Command</td>
<td>±5%, 3 Levels</td>
</tr>
<tr>
<td>EGR Flow</td>
<td>EGR Valve Command</td>
<td>±5% variation in EGR fraction, 3 Levels</td>
</tr>
<tr>
<td>Boost Pressure</td>
<td>Turbocharger Command</td>
<td>±5% variation in Boost Pressure, 3 Levels</td>
</tr>
<tr>
<td>Swirl Ratio</td>
<td>Swirl Ratio Specification in Cylinder Model</td>
<td>±0.2, 3 Levels</td>
</tr>
</tbody>
</table>

Table 7: Factors for Design of Experiments with Variation and Levels
4.4.2 ANOVA Analysis on the DOE Results

Analysis of Variance has been discussed in Chapter 3. It is a method that is used to compare the means of two or more groups of data where each group is characterized by one or more factors. Clearly, for this study, the number of factors is six and therefore N-way ANOVA analysis has been used. Before doing the ANOVA analysis, response variables have to be chosen for the experiment. For the present sensitivity analysis, the response variables pertain to engine performance, emissions and combustion metrics. Indicated specific fuel consumption (ISFC), engine-out NOx emissions and combustion metrics identified during literature survey have been chosen as the response variables for the ANOVA analysis. The results from the DOE of every mode are extracted using the GT-Power’s post-processing tool GT-Post.

MATLAB has been used to carry out the ANOVA analysis. MATLAB provides inbuilt codes to perform one, two and n-way ANOVA analysis on a given set of data. The output of the ANOVA analysis is a table consisting of the sum of squares, number of degrees of freedom, mean sum of squares, F and p values for every main factor in the ANOVA analysis as described in Chapter 3. Figure 49 shows the ANOVA table for NOx sensitivity in mode 1 as output by MATLAB. In generating this table, a first level of interaction between the factors is also considered for the sensitivity analysis. The significance of each factor and interactions can be identified by looking at the probability (p-value) associated with each factor. However, since the data in this study is generated by a deterministic GT-Power model of the engine, the term probability makes little sense. For this reason, the ratio of Sum of Squared Errors of each factor to the Total Sum of Squared Errors is taken as the significance indicator for each factor. This significance is represented in the form of a percentage and it signifies the percentage contribution of each factor or interaction to the total variance in the data and is represented by Equation 16. The percentage contribution of each factor to the total variance in the data for the ANOVA analysis shown in Figure 49 is tabulated in Table 8. It can be observed that the percentage contributions of all the factors and interactions add up to hundred. The error in the ANOVA table represents the unexplained variance in the data.
% Variance Contribution = \frac{SSE_{\text{Factor}}}{SSE_{\text{Total}}} \times 100

\begin{table}[h]
\centering
\begin{tabular}{lrrrr}
\hline
Source & Sum Sq. & d.f. & Mean Sq. & F & Prob>F \\
\hline
EGR & 333196.2 & 2 & 166596.1 & 9315.98 & 0 \\
Boost & 480394.4 & 2 & 241197.2 & 12199.46 & 0 \\
Fuel Qty & 81537.8 & 2 & 40768.9 & 2035.04 & 0 \\
SOI & 384766.7 & 2 & 192381.3 & 9602.99 & 0 \\
Rail P & 60303.7 & 2 & 30151.9 & 1505.07 & 0 \\
Swirl & 28848.1 & 2 & 14424 & 720 & 0 \\
EGR*Boost & 242 & 4 & 60.5 & 3.02 & 0.0175 \\
EGR*Fuel Qty & 723.9 & 4 & 181 & 9.03 & 0 \\
EGR*SOI & 424.1 & 4 & 106 & 5.29 & 0.0003 \\
EGR*Rail P & 129.1 & 4 & 32.3 & 1.61 & 0.1697 \\
EGR*Swirl & 60.9 & 4 & 15.2 & 0.76 & 0.5819 \\
Boost*Fuel Qty & 61046.3 & 4 & 15261.6 & 761.8 & 0 \\
Boost*SOI & 109 & 4 & 27.2 & 1.36 & 0.2464 \\
Boost*Rail P & 372.2 & 4 & 93.1 & 4.64 & 0.0111 \\
Boost*Swirl & 102.2 & 4 & 25.6 & 1.28 & 0.2781 \\
Fuel Qty*SOI & 461.6 & 4 & 115.4 & 5.76 & 0.0001 \\
Fuel Qty*Rail P & 134.1 & 4 & 33.5 & 1.67 & 0.1543 \\
Fuel Qty*Swirl & 259.2 & 4 & 64.9 & 3.22 & 0.0121 \\
SOI*Rail P & 33.7 & 4 & 8.4 & 0.42 & 0.7942 \\
SOI*Swirl & 39.7 & 4 & 9.9 & 0.49 & 0.7335 \\
Rail P*Swirl & 70.2 & 4 & 17.6 & 0.88 & 0.4778 \\
Error & 13142 & 656 & 20 & & \\
Total & 1454392.8 & 728 & & & \\
\hline
\end{tabular}
\caption{Analysis of Variance}
\end{table}

Constrained (Type III) sums of squares.

Figure 49: ANOVA Table for Mode 1
<table>
<thead>
<tr>
<th>Factor</th>
<th>% Contribution to Total Variance</th>
</tr>
</thead>
<tbody>
<tr>
<td>EGR</td>
<td>22.91</td>
</tr>
<tr>
<td>Boost</td>
<td>33.58</td>
</tr>
<tr>
<td>Fuel Qty</td>
<td>5.61</td>
</tr>
<tr>
<td>SOI</td>
<td>26.46</td>
</tr>
<tr>
<td>Rail P</td>
<td>4.15</td>
</tr>
<tr>
<td>Swirl</td>
<td>1.98</td>
</tr>
<tr>
<td>EGR*Boost</td>
<td>0.02</td>
</tr>
<tr>
<td>EGR*Fuel Qty</td>
<td>0.05</td>
</tr>
<tr>
<td>EGR*SOI</td>
<td>0.03</td>
</tr>
<tr>
<td>EGR*Rail P</td>
<td>0.01</td>
</tr>
<tr>
<td>EGR*Swirl</td>
<td>0.00</td>
</tr>
<tr>
<td>Boost*Fuel Qty</td>
<td>4.20</td>
</tr>
<tr>
<td>Boost*SOI</td>
<td>0.01</td>
</tr>
<tr>
<td>Boost*Rail P</td>
<td>0.03</td>
</tr>
<tr>
<td>Boost*Swirl</td>
<td>0.01</td>
</tr>
<tr>
<td>Fuel Qty*SOI</td>
<td>0.03</td>
</tr>
<tr>
<td>Fuel Qty*Rail P</td>
<td>0.01</td>
</tr>
<tr>
<td>Fuel Qty*Swirl</td>
<td>0.02</td>
</tr>
<tr>
<td>SOI*Rail P</td>
<td>0.00</td>
</tr>
<tr>
<td>SOI*Swirl</td>
<td>0.00</td>
</tr>
<tr>
<td>Rail P*Swirl</td>
<td>0.00</td>
</tr>
<tr>
<td>Error</td>
<td>0.90</td>
</tr>
</tbody>
</table>

Table 8: Percentage Contribution of each factor to total variance
4.4.2.1 ANOVA analysis on ISFC and NOx

The sensitivity of ISFC and NOx on the factors chosen for the Design of Experiments is found using the ANOVA analysis. The significance of each factor is calculated as the percentage contribution of each factor to the total variance in the DOE data for each of the 11 modes. As a first step, ANOVA analysis is done considering the factors and first-level of interactions between the factors. Only interactions between the factors which showed significant levels compared to the main factors were chosen for the ANOVA result plots. The percentage contribution of each factor and interactions to the total variance is calculated as explained in the previous section. This exercise is done for all the 11 modes chosen for the analysis. The ANOVA results for engine out NOx and engine ISFC are plotted as horizontal bar charts in Figure 50 and Figure 52 respectively.

![Sensitivity of NOx to Engine Root-cause Variables](image)

Figure 50: Sensitivity of NOx to engine root-cause variables
The significance of the each DOE factor on engine out NOx emissions for each mode is plotted in Figure 50. The contribution of each factor is represented by a unique color in the plot. The area shaded in white color represents the unexplained variance in the data. The model chosen for NOx is largely linear with only one interaction between fuel quantity and boost. The other interactions were found to be insignificant in the preliminary analysis. It can be observed from Figure 50 that among the six factors chosen for the DOE with their expected variability, EGR flow, boost, SOI and fuel quantity contribute most to the total variability of NOx in each mode. Also, the relative significance of each factor varies with the mode. Figure 51 plots the speed and load associated with each mode on the engine load-speed plane for reference. It can be seen that in all the modes either EGR or boost show the maximum influence on NOx for a 5% variability from their nominal values. For the modes (7, 8, 10 and 11) at low loads of engine, boost shows a large influence on NOx. Start of Injection exhibits comparable significance in the NOx variability followed by fuel quantity. Rail pressure and swirl ratio show least significance among the main factors. Also, it was observed that the interaction between boost and fuel quantity showed appreciable significance and in some modes it was comparable to that of rail pressure and swirl. It can therefore be concluded that variability observed in NOx can be compensated by changing EGR flow, boost, injection timing or quantity.

Figure 51: Modes represented on engine load speed
The significance of each DOE factor on engine ISFC is plotted in Figure 52. It can be observed that boost pressure, fuel quantity and injection timings have the highest contributions to ISFC variability. The effect of 5% change in EGR flow on ISFC is almost negligible in all the modes. Boost pressure by far has the maximum contribution to the variability in all the modes followed by injection timing and injection quantity. The effect of interactions is completely absent in the case of ISFC which suggests that the influence of the factors is additive in nature. Therefore, boost pressure, injection timings and quantity are the key factors of influence as far as ISFC is concerned. It can therefore be concluded from the sensitivity analysis on NOx and ISFC that variability in air handling system has the highest effect on NOx and ISFC followed by the injection parameters. In other words, it can be said that air handling system has higher controllability on NOx and ISFC compared to the fueling system. However, control of air handling system...
is a challenge because of the high coupling between the EGR and VGT systems and slow dynamics. On the other hand fuel injection system offers very fast response times in compensating for the dispersions in NOx and ISFC.

4.4.2.2 ANOVA Analysis on Combustion Metrics

In this section, the results of ANOVA analysis on various combustion metrics are discussed. In Chapter 2, various combustion metrics used in the literature have been discussed in detail. For this study, only combustion metrics that can be calculated from cylinder pressure have been selected. The purpose of this exercise is to find the sensitivity of the combustion metrics to variability in the six factors chosen for the DOE. Clearly, combustion metrics that exhibit high sensitivity to the factors can be monitored during combustion in order to identify the existence of variability in the operating conditions. For the sensitivity analysis, the following combustion metrics have been chosen

1. Crank Angle at 50\% Gross Heat Release - CA_{50_{\text{Gross}}}

2. Crank Angle at 50\% Net Heat Release - CA_{50_{\text{Net}}}

3. Crank Angle at 50\% Mass Fraction Burnt based on Rassweiler-Withrow Method (MFB_{50_{\text{RW}}})

4. Maximum Cylinder Pressure (P_{\text{max}})

5. Crank Angle at Maximum Cylinder Pressure (CA_{P_{\text{max}}})

6. Maximum Rate of Pressure Rise (dp/d\theta_{\text{max}})

7. Crank Angle at Maximum Rate of Pressure Rise (CA_{dp/d\theta_{\text{max}}})

8. Maximum Combustion Pressure (P_{\text{comb,max}})

For doing the ANOVA analysis, the pressure traces from each DOE case for all the 11 modes have been extracted from cylinder number one and six of the GT-Power model. As explained earlier, for the DOE,
DIJet combustion models have been used only for cylinders one and six in the GT-Power model to reduce computation time. This method can give an estimate of the maximum effect of cylinder-to-cylinder variability on combustion because cylinders one and six are the extremes of this phenomenon.

4.4.2.2.1 ANOVA Analysis on CA50 calculated from Gross Heat Release

Heat release rate method models the rate of release of fuel’s chemical energy during the diesel engine combustion process. Gross heat release method calculates the total chemical energy release during combustion which includes the rate at which work is done on the piston and also the heat transfer between the hot combustion gases and cylinder walls. Therefore, the gross heat release method gives much more accurate information on the rate of release of diesel fuel’s energy during combustion than the apparent heat release method discussed in Chapter 2. However, this method requires the calculation of heat transfer rate during combustion which is usually accomplished by empirical models like Woshni’s correlation. Due to the increased computational burden of calculating the heat transfer, gross heat release rate method may not be suitable for applications like closed loop combustion control where such calculations have to be done on an engine cycle to cycle basis. In this study however, CA50 calculated from gross heat release rate is used as a baseline metric as it is probably the most accurate of all the metrics chosen for the study. The gross heat release rate can be calculated from crank-resolved pressure data using the following equation. CA50 is the crank angle at which 50% of total heat release takes place.

\[
\frac{dQ_{ch}}{d\theta} = \frac{1}{\gamma - 1} \left[ \gamma \cdot p \cdot \frac{dV}{d\theta} + V \cdot \frac{dp}{d\theta} \right] + \frac{dQ_{hu}}{d\theta}
\]

Equation 17

Figure 53 plots the results from ANOVA analysis for CA50 of cylinder 1 calculated for all the 11 modes chosen in the study. It can be observed that boost and injection timing (SOI) show maximum influence on CA50 based on the mode. They are followed by fuel quantity. The effect of fuel quantity is predominant at
modes at high loads (mode 1, 4 and 9). EGR flow has minimal effect on the metric and for modes with low loads (mode 7, 10 and 11) the effect of EGR increases. Variability in rail pressure and swirl has least effect on the metric among the main factors of the DOE. The interactions between the factors did not show any significance and were therefore excluded from the ANOVA analysis plots. Figure 54 plots the ANOVA plot for CA50\textsubscript{Gross} calculated for cylinder 6. It can be observed that no noticeable difference can be observed in the significance of the factors on the metric for cylinder 6. This indicates that the variabilities in the factors affect both cylinders one and six identically.

![Sensitivity of CA50 Gross to Engine Root-cause Variables](image)

Figure 53: Sensitivity of CA50\textsubscript{Gross} of cylinder 1 on engine root-cause variables
4.4.2.2 ANOVA analysis on CA50 calculated from Net Heat Release

As explained earlier, the computational burden involved in calculating gross heat release rate makes it an undesirable method to calculate CA50 for applications like CLCC. Therefore, CA50 calculated from net heat release rate of combustion is often used as a substitute to CA50\textsubscript{Gross}. The only difference in these two methods is that heat transfer rate during combustion is not accounted for in the heat release calculation. As a result, the net heat release rate is also called as apparent heat release rate. This reduces the accuracy at which one can predict the true CA50 value which is nothing but CA50\textsubscript{Gross}. However, it gives a close approximation of the true CA50 with much less computation burden. Equation 18 gives the net heat release rate calculation from crank-resolved cylinder pressure data.
\[
\frac{dQ_{\text{net}}}{d\theta} = \frac{dQ_{\text{ch}}}{d\theta} - \frac{dQ_{\text{at}}}{d\theta} = \frac{1}{\gamma - 1}\left[\gamma \cdot p \cdot \frac{dV}{d\theta} + V \cdot \frac{dp}{d\theta}\right]
\]

Equation 18

Figure 55: Sensitivity of CA50\text{Net} of cylinder 1 on engine root-cause variables plots the significance of each DOE factor on CA50\text{Net} calculated for cylinder 1. The significance level of each factor on CA50\text{Net} is very similar to that of CA50\text{Gross} seen in the previous section. Figure 56 plots the CA50\text{Net} calculated for cylinder 6.

Figure 55: Sensitivity of CA50\text{Net} of cylinder 1 on engine root-cause variables
4.4.2.2.3 ANOVA analysis on Rassweiler-Withrow Method

Rassweiler-Withrow method is an approximation of the apparent heat release rate method and calculates the mass fraction of fuel burnt during the combustion. The method is based on the assumption that the change in pressure due to the piston motion and charge-to-wall heat transfer can be represented by polytropic processes. In this method, the pressure change during any crank angle interval is assumed to be made up of a pressure rise due to combustion $\Delta p_c$ and a pressure rise due to the volume change $\Delta p_v$. Since the pressure rise due to combustion is proportional to the mass of the fuel that burns, the MFB at a given crank angle can be calculated by taking the ratio of the combustion pressure rise until that crank angle to the total combustion pressure rise during the entire combustion event as given by Equation 19. In the calculation of MFB50, the polytropic index $n$ is approximated by the ratio of specific heats of the...
combustion gases (γ). This means that the compression process is adiabatic and therefore, MFB50 is an approximation of the apparent heat release method.

\[ MFB_{RW}(\theta_i) = \frac{\int_{\theta_{nc}}^{\theta_i} \Delta P_c(\theta)}{\Delta P_c(\theta_{nc})} \]

\[ \Delta P_c(\theta_i) = p(\theta_i) - p(\theta_{i-1}) \left( \frac{V(\theta_{i-1})}{V(\theta_i)} \right)^n \]

\[ \Delta Q(\theta_i) = \frac{V(\theta_i)}{n-1} \Delta P_c(\theta_i) \]

Figure 57 plots the sensitivity of 50% of mass fraction burnt (MFB50) of cylinder 1 calculated from the Rassweiler-Withrow Method. Comparing with the sensitivity of CA50 calculated from heat release analysis, it can be observed that the effect of injection timing on MFB50 increases in all the modes. Boost pressure also shows high significance followed by fuel quantity and EGR flow. The significance of EGR however reduces which makes MFB50 less sensitive to variabilities in EGR. Therefore, MFB50 can also act as a substitute for CA50 calculated from gross and net heat release methods. Unlike CA50 calculation, there is no derivative involved in MFB50 calculation which means that it can be more robust to noise in cylinder pressure signal. Figure 58 plots the sensitivity of MFB50 of cylinder 6. The sensitivity is identical to that of cylinder 1.
Figure 57: Sensitivity of MFB50$_{RW}$ of cylinder 1 on engine root-cause variables

Figure 58: Sensitivity of MFB50$_{RW}$ of cylinder 1 on engine root-cause variables

111
4.4.2.4 Analysis of Maximum Pressure and Maximum Rate of Pressure Rise

Maximum cylinder pressure and maximum rate of pressure rise are also commonly used as combustion metrics for closed-loop combustion control. For large diesel engines with high compression ratios, the maximum cylinder pressure is often observed at the TDC of the cylinder caused by the compression. Similarly, the maximum rate of pressure rise can also be caused purely by the compression stroke (and may be partly by pilot injection if it exists) and may occur before the TDC of the cylinder. Therefore, maximum pressure and rate of pressure rise are not convenient metrics for conventional diesel combustion as they may not reflect the complete effect of combustion. The same phenomenon was observed for the heavy-duty diesel engine under consideration in this study. Figure 59 plots the cylinder pressure traces for all the eleven modes at their nominal operating conditions. It can be observed that the maximum pressure in all the modes is exactly at the TDC of the cylinder.

Figure 60 plots the rate of pressure rise for all the modes at their nominal operating conditions. It can be seen that the maximum rate of pressure rise happens before the TDC. For comparison, the gross heat release rates for all the eleven modes are plotted in Figure 61. It can be seen that the release of fuel’s chemical energy starts to happen only after the TDC. Thus, maximum pressure and rate of pressure rise in all the modes do not account for the effects of combustion on the cylinder pressure. In Chapter 2, it was discussed that for conventional diesel engines, the pressure rise due to combustion also called as combustion pressure would be a convenient metric. The combustion pressure can be obtained by subtracting the motoring pressure from the actual cylinder pressure. The next section describes the calculation of combustion pressure and its sensitivity to the root-cause variables.
Figure 59: Cylinder Pressure traces at all the 11 modes

Figure 60: Rate of pressure rise in all the 11 modes
4.4.2.2.5 ANOVA analysis on Combustion Pressure

The combustion pressure is calculated using an approximation of the motoring pressure trace by using the Rassweiler-Withrow method. It is assumed that the rise in pressure due to compression can be approximated by an adiabatic compression process. The combustion pressure is the difference between the actual cylinder pressure and the approximated motoring trace. Equation 20 shows the calculation of the combustion pressure where $\gamma$ is the ratio of specific heats of the combustion gases. For the calculation, $\gamma$ calculated by GT-Power as a function of temperature and gas composition has been used. In practice, a constant value for $\gamma$ can be used. The value of $\gamma$ during combustion varies between 1.25-1.35. Figure 62 plots the virtual pressures due to compression/expansion and combustion along with the actual cylinder pressure for comparison. It can be observed pressure rise due to compression is positive and that due to expansion is negative. The rise in pressure due to combustion occurs well after the TDC of the cylinder. The sum of both compression and combustion pressures results in the actual cylinder pressure signal.

Figure 61: Gross Heat Release Rates for all the 11 modes
\[ p_r(\theta_i) = p(\theta_{i-1}) \left( \frac{V(\theta_{i-1})}{V(\theta_i)} \right)^y \]
\[ \Delta p_c(\theta_i) = p(\theta_i) - p_r(\theta_i) \quad \text{Equation 20} \]

Figure 62: Comparison of actual cylinder pressure, compression pressure and combustion pressure

The combustion pressure calculated can be a good combustion metric for closed-loop combustion control. For the ANOVA analysis, the maximum combustion pressures for all the modes have been extracted. Figure 63 plots the sensitivity of maximum combustion pressure to the engine root cause variables. It can be observed that fuel quantity has the highest significance on the maximum combustion pressure. Maximum cylinder pressures are strongly correlated to the torque output of the engine and therefore it can be expected that they are also strongly correlated to fuel quantities. Boost and injection timings also show high significance on the maximum combustion pressure followed by EGR. Figure 64 plots the sensitivity of
combustion pressure for cylinder 6. The combustion pressure shows sensitivity identical to that of cylinder 1.

Figure 63: Sensitivity of Maximum Combustion Pressure of cylinder 1 to engine root-cause variables
4.5 Conclusions from the ANOVA Analysis

In this chapter, various disturbances that can deteriorate engine performance have been discussed. The disturbances affect the combustion through the root-cause variables of the engine. A design of experiments with the engine root-cause variables has been designed and performed on the GT-Power model of the diesel engine. The results from the DOE have been used to do a sensitivity analysis on engine performance, emissions and combustion metrics. From the six factors chosen as the engine root-cause variables, it was observed that EGR flow, boost pressure, fuel quantity and injection timings have the most significance on the variability on ISFC, NOx and combustion metrics chosen in the study. The sensitivity information can be used to design the closed loop combustion control architecture. It was observed that engine ISFC and NOx had higher sensitivity to air handling system when compared to the fuel injection system. However, the faster response times of fuel injection systems can be an advantage to correct for dispersions rapidly.
On the other hand, combustion metrics showed sensitivity to both fuel injection system parameters and boost pressure. Appropriate combustion metric can be chosen for the closed-loop control architecture depending on its correlation to engine NOx and ISFC and also its sensitivity to the engine root-cause variables. It was observed that the combustion metrics $\text{CA50}_{\text{Gross}}$ and $\text{CA50}_{\text{Net}}$ showed sensitivities to all the main engine root-cause variables. Therefore, they are ideal candidates for combustion metrics as the effects of variability in the root-cause variables can be readily seen in these combustion metrics. On the other hand, MFB50$_{\text{RW}}$, which is an approximation of the apparent heat release rate method does not show significant sensitivity to variability in EGR flow. Therefore, it cannot completely identify the variability in NOx contributed by variability in EGR flow. It can therefore be a good metric where there is another system which can take care of variability in EGR. Finally, combustion pressure showed high sensitivity to fuel injection parameters making it an ideal metric to identify variability in fuel quantity and changes in engine performance due to variability in fuel quantity. For example, combustion pressure can be a good metric if the goal of the CLCC is to control noise and drivability which are correlated to cylinder pressure directly. The sensitivity analysis can therefore play a significant role in identifying the closed-loop combustion control architecture. The actuators and the feedback variables can be chosen depending on the goal of the closed-loop control. In Chapter 4, the goal of the closed-loop combustion controller for the heavy duty diesel engine under study is established and the combustion control architecture is derived with the help of the results from the sensitivity analysis.
5.1 Definition of the Closed-loop Combustion Control Problem

It is a known fact that in diesel engines, there is a clear trade-off between emissions and engine performance. It was discussed in Chapter 2 that there are two significant trade-offs, one between NOx and soot and another between NOx and ISFC in diesel engines. In addition to emissions and performance, noise and drivability are other factors that have to be considered while choosing the operating parameters of a diesel engine. The operating parameters of the engine should be carefully chosen in order to satisfy a multi-variable objective function. The methodology of choosing the operating parameters is not well-established and varies from manufacturer to manufacturer. Ideally, the operating parameters have to be chosen in an engine calibration process using an objective function as shown in Equation 21. The weights $\alpha$’s in the equation are chosen depending on the relative importance of each variable at a given operating condition. For example, for city driving conditions, engine out emissions and noise could be more important. Similarly, for high-way driving conditions, Brake Specific Fuel Consumption (BSFC) and drivability may be more important than the other parameters. The relative importance between the variables is often dictated by the local regulations on engine performance. However, the method in which manufacturers calibrate engines is not as systematic as the ideal method. Some of the calibration is still done in an ad-hoc way to overcome the challenges of doing a multi-variable optimization of engine performance.

$$f_{opt} = \alpha_1 BSFC + \alpha_2 Noise + \alpha_3 Drivability + \alpha_4 PM + \alpha_5 NOx$$

Equation 21

In Chapter 4, the assumptions made during engine calibration process and their effects on engine-out NOx and ISFC have been discussed. The following are the most important sources that can lead to deterioration in engine performance and emissions.
1. Cylinder-to-cylinder variability

2. Component-driven variability

3. Environmental driven variability

4. Transient operation of the engine

It was seen that whatever be the source of variability, it finally affects engine performance through the combustion root-cause variables. Combustion root-cause variables are the engine operating variables that are basic to the combustion phenomenon. The sensitivity analysis study discussed in Chapter 4 was aimed at identifying key-factors from the root-cause variables that have high influence on engine out NOx and ISFC.

The overall goal of closed-loop combustion control (CLCC) is to minimize the effects of changes in engine operating conditions on engine performance, emissions, drivability, and noise by direct or indirect feedback from combustion. As discussed in Chapter 2, there are multiple choices available for combustion feedback. Cylinder pressure and ion-current probes are examples of sensors that can provide direct feedback. The disadvantages of ion-current probes for application to conventional diesel combustion control have been discussed in Chapter 2. On the other hand, indirect feedback can be provided by sensors that are not placed directly in the combustion chamber like accelerometers, crank speed sensors and torque sensors. However, indirect sensing is not a viable option because of its complexity and reduced capability to give reliable information on combustion. Cylinder pressure sensing is the most viable technique for closed-loop combustion control because of the direct correlation of cylinder pressure to the combustion phenomenon. Despite their high costs, cylinder pressure sensors are still being evaluated as the prime enablers of closed-loop combustion control due to their potential to give high quality combustion feedback. For this study, it is assumed that cylinder pressure sensors are available for providing combustion feedback without going into the details of their economic viability.
5.2 Hypothetical Closed-loop Combustion Control Architecture

In this section, a hypothetical architecture for closed-loop combustion control is proposed. The hypothetical architecture is an example of how combustion control can be implemented in an ideal scenario. The hypothetical architecture provides a framework to analyze various challenges of combustion control. Figure 65 shows the hypothetical architecture for closed-loop combustion control. As the name suggests, the architecture is impractical for implementation in a vehicle due to various aspects which are discussed below. The reasons which contribute to the hypothetical nature of the control architecture are used to define a simplified architecture that is more feasible to implement in a real-world scenario.

![Hypothetical Closed-loop Combustion Control Architecture Diagram]

Figure 65: Hypothetical Closed-loop Combustion Control Architecture

The hypothetical architecture reflects the objective of the closed-loop combustion control in an ideal fashion. The set of feed-back variables represented by \( Y \) represent the variables of the objective function shown in Equation 21 like NOx, PM, ISFC, Drivability and Noise. The hypothesis is that these variables
can be directly measured in a vehicle. In practice, it is not practical to measure some of the feedback variables like NOx and PM in a production vehicle. This is because of the costs associated with sensors that can provide reliable measurements. Therefore, it is necessary to replace these variables with combustion metrics that correlate well to the variables of interest. Variables like engine noise and drivability are qualitative terms for which there are no standard definitions. Quantitative metrics that can represent noise and drivability of the engine can be derived based on way in which they are defined. Therefore, for each variable or a group of variables of the objective function of the closed-loop combustion controller, a combustion metric that closely correlates to the variable(s) needs to be defined. For practical implementation, the combustion metric has to be derived from the combustion variable that is sensed to achieve combustion control.

The controller in the architecture is assumed to have a feed-back and feed-forward component. The feedback combustion controller corrects for disturbances to the normal operation of the engine caused by various sources of variability discussed before. Both the feed-forward and feedback components sum up to produce the desired values for the combustion root-cause variables \( U \) for each cylinder of the engine. Ideally, the set \( U \) should consist of all the possible root-cause variables that affect combustion. In practice, one would want to generate the desired values only for the root-cause variables that have significant impact on the objective of the closed-loop combustion controller. A sensitivity analysis similar to the one described in Chapter 4 can be used to find the root-cause variables that significantly impact combustion.

The set-point generator in the architecture shown in Figure 65 is a block that generates optimal set-points for the combustion control feed-back variables and the feed-forward components of the controller. The set-points can be generated by optimizing the objective function described in Equation 21 at different operating conditions of the engine. The optimal operating conditions which result in the minimization of the objective function are the feed-forward components of the control. As explained earlier, parameters of the objective function may vary with vehicle driving conditions. The optimization may be done in real-time if the number of independent variables to be optimized is manageable. Online optimization becomes infeasible if the number of independent variables increases as the computational burden to do such a task also increases.
In such a situation the optimization can be done offline for a predefined set of operating conditions and the feed-forward components and the set-points can be generated using interpolation techniques depending on the present operating conditions.

The outputs $U$ of the closed-loop combustion controller are the desired combustion root-cause variables for the engine. The root-cause variables are basic necessities of combustion like air and residual conditions and fuel injection parameters. Root-cause variables like fresh air and residual content have to controlled with the help of a feedback controller. On the other hand, the desired values for fuel injection parameters like fuel injection amount and injection timings can be commanded from the fuel injectors in a feed-forward fashion. The engine controller shown in Figure 65 accomplishes the feed-back control of some of the root-cause variables with the help of an inner feedback loop. Due to the existence of cascaded feedback loops in the combustion control architecture, it becomes extremely difficult to ensure the stability and robustness of the controllers.

5.3 Closed-loop Combustion Control Architecture for Cummins Heavy-duty Diesel Engine

It is clear that the design of closed-loop combustion control architecture is not a trivial task. In order to design an architecture that is feasible for real-world implementation, various aspects need to be considered. The design process for the architecture is tightly coupled with the objective of combustion control, sensing capabilities and tolerance to increased calibration effort. The architecture should also be able to integrate seamlessly into the existent engine control architecture. In this study, the closed-loop combustion control architecture has been designed for a Cummins heavy-duty diesel engine. The architecture has been designed such that it can be seamlessly integrated into an existent engine control architecture. Also, various other assumptions and constraints have been applied to deduce a feasible architecture. The details of this process are discussed in this section.
5.3.1 Traditional Engine Control Architecture of Cummins Heavy-duty Diesel Engines

The heavy-duty diesel engine under consideration is equipped with a variable geometry turbocharger with an intercooler, a cooled EGR loop and common rail injection system capable of multiple injections. An abstraction of the control architecture traditionally used by the engine is depicted in Figure 66: Cummins Heavy-duty Diesel Engine Control Architecture. The control architecture is based on a map based algorithm which converts driver’s demand for torque into set-points for the air handling unit and fuel injection sub-system. The air handling unit includes the variable geometry turbocharger and the exhaust gas recirculation loop. Boost pressure and residual gas fraction in the intake manifold of the engine are controlled to track their set-points with the help of feedback. Boost pressure is measured with a manifold pressure sensor and the residual fraction is estimated with the help of pressure sensors in the intake and exhaust manifolds. On the other hand, the injection system receives the set-points for rail pressure, injection timings (pilot, main and post) and injection quantities (pilot, main and post) as inputs. The rail pressure is controlled with a rail pressure control valve with the help of a feedback controller. The set-points for injection timings and quantities are converted into appropriate inputs to the fuel injectors such that the desirable quantity is injected at the desirable time. Traditionally, the same injection commands are given to all the cylinders despite the existence of cylinder-to-cylinder variations. It has to be noted that the architecture has feedback for the air handling unit and rail pressure and there is no feedback available from the fuel injectors. The amount of fuel injected is limited by a fuel limiter during heavy transient operations. This is done to reduce particulate emissions during transients to compensate for the lack of fresh air due to turbocharger lag and intake manifold dynamics. The combustion inside the cylinder is represented with a cylinder block in the control architecture. The combustion is a function of the outputs from the air handling and fuel injection sub-systems. The output from the cylinder is torque which is delivered to the wheels. The driver closes the loop by changing the accelerator position in the case of a change in the desired torque.
The traditional control architecture has been a tested method and therefore used by many engine manufacturers. It yields desired results when the engine is operating at steady-state. But during transient operation of the engine, the design of the architecture poses issues that may lead to deterioration in engine performance. Consider a scenario when the driver demands an increase in the torque and depresses the accelerator pedal. The torque controller of the engine immediately caters to the torque request by increasing the amount of fuel injected into the cylinders. The increase in the fuel quantity happens almost in an instantaneous fashion. The calibration maps of the controller generate set-points for EGR fraction, boost pressure, injection timings and rail pressure based on the current engine speed and load. It was noted in Chapter 4 that the air handling system has the slowest response time in the engine whereas the injection events can be changed on a cycle-by-cycle basis. Due to this fact, the set-points for injection parameters are
achieved instantaneously whereas the desired boost and EGR levels lag behind by several engine cycles due to slower dynamics. As a result, during the transients the engine operates at an undesirable combination of injection parameters, EGR fraction and boost which may lead to deterioration in engine performance and emissions. For an engine that operates on a transient drive cycle, the loss in engine performance and increase in emissions during transients can be significant. The primary reason for not scheduling injection parameters considering the lag in the air handling unit is increased calibration effort and the need to store multi-dimensional calibration tables.

Figure 66 shows a very high-level picture of the engine control architecture depicting only the aspects directly relevant to closed-loop combustion control. In reality, the engine control unit also monitors and controls various other engine sub-systems like the engine cooling and lubrication system, auxiliaries and exhaust after-treatment system. The closed-loop combustion controller has to be designed keeping the traditional engine control architecture in mind. Due to tremendous efforts already invested in the infrastructure required to develop and implement traditional engine control architectures, it would be highly undesirable to redefine it completely. A combustion control architecture that can be integrated into an already existent engine control unit can also expedite its deployment into production engines. A combustion controller that also addresses the shortcomings of a traditional control architecture would be an ideal solution to this problem. Therefore, certain assumptions have been made in order to develop a combustion control architecture for the Cummins heavy-duty diesel engine.

5.3.2 Constraints on Closed-loop Combustion Control

From Figure 66 it is clear that the combustion controller has the option of controlling the set-points for the air handling unit and the fuel injection parameters for affecting the root-cause variables of combustion in the cylinders. The constraints described in this section aim at identifying the sub-set of actuators that can be used for combustion control. The primary motivation behind the constraints is to minimize the integration issues between the combustion controller and the traditional control architecture. The constraints identified are as follows
1. The closed-loop combustion controller does not modify the set-points for the air handling unit of the engine. This assumption has been made to avoid the complexities in designing two cascaded feed-back controllers as described in the section where the hypothetical control architecture was discussed. Also, due to slow dynamics involved with the air handling unit, the response time of a combustion controller based on adjusting the boost levels and residual fraction would be very slow. However, the combustion controller can use the outputs from the air handling control unit in taking a decision. This aspect can be particularly useful for addressing the short-coming of the traditional control system during transient operation of the engine and compensate for the deterioration of engine performance and emissions due to the dynamics in the air handling unit.

2. The closed-loop combustion controller should not directly impact the torque production of the engine. This is to make sure that the driver receives the torque that he demands. In the existent engine control architecture, the driver request for torque is converted into the total amount of fuel injected into the cylinder. This means that the closed-loop combustion controller should not alter the total fuel injected into all the cylinders.

3. The feedback variable chosen for closed-loop combustion control should be a cylinder pressure based combustion metric that is correlated well to the objective function chosen for combustion control. The combustion metrics should also be highly sensitive to changes in the root-cause variables. This is important because the closed-loop combustion controller does not intend to close the loop on all the root-cause variables. Also, the computational complexity of calculating the combustion metrics from cylinder pressure should be minimal.

4. The addition of closed-loop combustion control to the engine control architecture should not considerably increase the effort required to calibrate the engine.
5.3.3 Choice of Actuators

The first assumption in the previous section eliminates boost pressure and EGR fraction as feasible control actuators for closed-loop combustion control. This is actually a constraint that considerably reduces the effectiveness of the closed-loop combustion control since boost and EGR have high influence on engine emissions and performance as seen in the sensitivity analysis. However, EGR and boost are already controlled with the help of feedback which makes it unlikely to expect variabilities in them since the feedback controllers try to correct for the error. On the other hand, slow response times of boost and EGR compared to the fuel injection system make them undesirable if CLCC aims at cycle-by-cycle control of combustion which is desirable during transient operation of the engine. Lastly, the existence of CLCC can let the engine calibrators calibrate EGR levels closer to the engine dilution limits without the fear of poor operation which gives the opportunity to reduce the emissions further.

The second assumption in the previous section restricts the use of fuel quantity for closed-loop combustion control. However, it does not completely eliminate fuel quantity from the set of actuators since it only places a constraint on the total fuel quantity injected into the cylinders. It is therefore possible to deliver different amounts of fuel quantities to each cylinder by keeping the total fuel quantity the same. Such a kind of control is highly desirable to balance the torques produced by each cylinder which is directly connected to the drivability of the vehicle. At a cylinder level, the fuel quantity can be distributed between the multiple injections. However, design of such a control structure could be highly complex as the number of control variables increases tremendously.

Among the other root-cause variables that directly affect combustion are injection timings, rail pressure and swirl factor of the inlet ports of the engine. Swirl has been eliminated since it cannot be actively controlled in the engine under study. Controlling the rail pressure set-point would introduce the problem of cascaded control loops since rail pressure is already controlled with the help of feedback. Moreover, it was observed in Chapter 4, that engine performance and emissions are more sensitive to injection timing than rail pressure. Injection timing is therefore another choice for the set of actuators. The pilot, main and post
injection timings of each cylinder can be controlled individually but implementing such a controller can be highly complex due to the multi-input multi-output nature of the problem. An alternative method is to move all the three injection timings together.

5.3.4 Choice of Feedback Variables

The feedback variables are combustion metrics that substitute the need to measure actual emissions and performance parameters in the engine. Therefore, it is clear that the choice of the combustion metric depends on its correlation to the emission or performance parameter that it represents. Also, the combustion metric must be highly sensitive to changes in the root-cause variables; particularly to the actuator chosen for controlling the combustion metric. In the previous section, it was observed that change in injection timings and restricted use of fuel quantity are the two actuating possibilities available for closed-loop combustion control of the heavy-duty diesel engine under study. From Chapter 4, it was observed that combustion metrics derived directly from heat release or from an approximation of heat release showed higher sensitivity to injecting timings compared to those derived directly from cylinder pressure. Therefore, CA50_{Gross}, CA50_{Net}, MFB50_{RW} are more suitable as feedback variables for the control loop with injection timing as the actuator. Of the three, MFB50_{RW} showed deteriorated sensitivity to variability in EGR making it less desirable to use as feedback since EGR has very high significance as far as NOx and PM are concerned. CA50_{Gross} and CA50_{Net} are both suitable for feedback since they show sensitivity to all the root-cause variables. The choice of the combustion metric between CA50_{Gross} and CA50_{Net} is determined by the availability of the computational resources as CA50_{Gross} introduces extra computational burden for calculating the heat transfer rate in the engine. Due to the high correlation of heat release to combustion, this control loop is more suitable to address the dispersion in emissions and fuel consumption of the engine.

The other variables of interest for combustion control are engine noise and drivability. As explained earlier, these two terms are qualitative with no standard definitions. For example, noise from the cylinders can be correlated to the maximum rate of pressure rise in the cylinders. Drivability can be defined as the fluctuations in crank shaft torque which is the result of imbalance in torques contributed by the individual
cylinders to the crank shaft. When the torque contribution by individual cylinders is balanced, it can lead to better drivability of the vehicle as the fluctuations in the total torque would be minimal. The torque contributed by each cylinders can be correlated to the IMEP of the cylinder, maximum cylinder pressure etc. In Chapter 4, it was observed that the metrics like maximum combustion pressure are sensitive to the total fuel quantity injected into the cylinders. Therefore, fuel quantity is more suitable as an actuator for noise and drivability control aspects of combustion control provided they are defined as explained above.

5.3.5 Closed-loop Combustion Control Architecture

In the last three sections, various assumptions have been made to develop the framework for closed-loop combustion control architecture. The assumptions have been made to enable a smooth integration of the architecture into the existing control architecture of the heavy-duty diesel engine under consideration. The set of feasible actuators and feedback variables have been discussed based on the assumptions. Figure 67 shows the new engine architecture for the heavy-duty diesel engine along with closed-loop combustion controller. There primary differences between this architecture and the hypothetical architecture area as follows

1. The architecture shown in Figure 67 has two feedback loops that are complementary and not cascaded as in the hypothetical control architecture. The closed-loop combustion controller and the engine controller work in tandem complementing each other. The closed-loop combustion controller modifies the injection parameters of the individual cylinders and the engine controller controls the air handling unit and the rail pressure. However, there are some interactions between both the loops. For example, the total fuel that has to be injected is determined by the engine controller and it acts as a constraint on the closed-loop combustion controller which can only change the distribution of the total fuel between the cylinders by keeping the total amount same. It has to be noted that the closed-loop combustion controller modifies the fuel injection parameters whose response time on combustion is almost instantaneous. On the other hand, the engine controller controls the air and residual system which is many times slower than the fuel injection system. The closed-loop combustion controller should be
designed in such a way that it schedules fueling parameters considering the time lag in the air handling unit. This aspect can be taken care at the set-point generator level of the closed-loop combustion control. The set-point generation problem is formalized in the next section and the rest of the study is aimed at developing methodologies for the set-point generation problem.

2. The second difference between the proposed control architecture and the hypothetical one is the feedback variables. The feedback variables in the hypothetical control architecture have been replaced by combustion metrics. The choice of the combustion metrics are chosen depending on the objective of the combustion controller and the computational capability of the combustion metric calculator. Therefore, they have not been explicitly represented in the combustion control architecture. Possible candidates for the combustion metrics have been discussed in the previous section.

Figure 67: Closed-loop Combustion Control Architecture for Cummins Heavy-duty Diesel Engine

\[ U = \{Q_{\text{main},i}, Q_{\text{pilot},i}, Q_{\text{post},i}, \text{SOI}_{\text{main},i}, \text{SOI}_{\text{pilot},i}, \text{SOI}_{\text{post},i}\} \quad (i = \text{cylinder number}) \]

\[ Y = [\text{CM}_1, \text{CM}_2, \ldots] \]

\[ V = [P_{\text{boost}}, \text{EGR Fraction}, \text{Total Fuel}, \text{Rail Pressure etc}] \]

FF – Feed-Forward

SP – Set-Point

CM – Combustion Metric
5.4 Set-point Generation for Closed-loop Combustion Control

In the last section, the architecture for the closed-loop combustion control of heavy duty diesel engine under study has been proposed. The combustion controller has feedback and feed-forward components and a set-point generator that generates the set-points for the combustion metrics and the feed-forward actuator commands. It is a known fact that combustion is a non-linear phenomenon and conventional control methods are difficult to develop for combustion control. It is therefore important to have a strong feed-forward component that minimizes the control effort of the closed-loop controller. The feed-forward component can be viewed as a non-linear compensator for the combustion controller which can eliminate the need to develop a non-linear feedback control method. The difficulty of developing a non-linear feedback controller that is stable and robust across the entire operating engine is non-trivial and the success of such an effort is not guaranteed. On the other hand, with the help of a strong feed-forward component, it would be sufficient to develop a linear control method for the feedback control which is much more straightforward. Therefore, the focus of this study has been shifted to the feed-forward component generation for the combustion control. It has to be noted that the set-points for the combustion metrics can be generated by operating the individual cylinders with the feed-forward actuator commands. Therefore, the problem of set-point generation boils down to generating the feed-forward components of the control. The feed-forward component generation will be referred to as set-point generation from this point onwards.

Another important aspect of the set-point generation problem is generating the feed-forward components and set-points for combustion metrics during engine transients. Transients in the engine are initiated by a change in driver’s torque request. In order to meet the torque request, the torque controller injects an equivalent amount of fuel that produces the change in torque in the immediately next engine cycle. The engine speed changes depending on the vehicle and road load on the engine. During this transient phase, a traditional set-point generation would have generated set-points for the injection parameters assuming that the air handling unit has reached the new steady-state. The air handling unit, however, due to its slower dynamics requires a number of cycles to provide the required boost and EGR levels that are optimal for this new operation. This phenomenon is depicted by Figure 68. It can be observed that the gross fuel quantity
changes instantaneously from operating point (OP) 1 to 2, whereas the boost pressure and EGR take several cycles to reach their set-points. In a traditional engine control method, calibrations would be available only at the steady-state values of EGR and boost. The calibration approach that is proposed in this study aims at adding data points for the transient operation of the engine as represented by the ‘stars’ in the figure.

By using the present values of EGR fraction and boost pressure as inputs, the set-point generator can generate the set-points and the feed-forward components during the engine transients. The outputs of the set-point generator for the proposed control are depicted in Figure 69. The inputs and outputs of the set-point generator block are shown in Figure 70.

Figure 68: Depiction of Transient operation of an engine.
Figure 69: Set-point generator output for the proposed control

Figure 70: Inputs and Outputs of the Set-Point Generator Scheme
5.5 Methods for Set-point Generation

The goal of the set-point generator block to generate optimal set-points for the combustion metrics and optimal feed-forward components of the combustion controller at a given operating condition of the engine. This process involves an optimization of the objective function that represents the goal of the combustion controller. As the number of independent variables that represent the operating conditions of the engine increases, the calibration effort to generate the set-points also increases rapidly. The independent variables for the set-point generation include engine speed, load, EGR fraction and boost. The set-point generator should output six injection parameters (3 SOIs and 3 quantities) for each cylinder which comes to 36 parameters (the engine has 6 cylinders). A conventional method to generate the set-points is to carry out experiments on the engine at different combinations of the independent variables and optimize the injection parameters at each operating condition. For example, if the operating range of each independent variable is divided into 10 equally spaced points, the number of data points at which each injection parameter has to be optimized would be $10^4$. Such a task should be done for every injection parameter for each cylinder. Many of the operating conditions may not be even encountered during the actual operation of the engine. The resulting calibration effort is enormous and highly impractical. With the help of engineering intuition and calibration experience, the calibration effort can be reduced by many folds. However, the calibration effort would still be high and can lead to increased development costs. The result of such a calibration effort would be 4D tables for each injection parameter of the cylinders which could be 10000 data points for each injection parameter. The number of points for 36 injection parameters for all the six cylinders would be $36*10^4$. The memory required to store so many data points is enormous and cannot be managed by the present-day engine control units. Therefore a set-point generation scheme which has the following characteristics is highly desirable

1. Requires less calibration effort

2. Capable of dealing with multiple dimension interpolation

3. Less computation burden
Inverse-distance interpolation method discussed in Chapter 2 has all the desirable characteristics enumerated above. Inverse-distance interpolation method facilitates scattered data interpolation. Therefore, using such a method can eliminate the need to use a uniform grid of data points during calibration and therefore reduces the number of points at which engine calibration has to be carried out. The calibration points can be chosen from the operating conditions that the engine is operated frequently in the field. The inverse-distance method can also be easily extended to interpolation in multiple dimension which serves the purpose of generating the set-points using an input space that consists of four variables. The method has some shortcomings, however, when they are addressed adequately, it can serve as a very simple yet robust interpolation technique. The following section deals with the details of how the inverse-distance method is applied to the set-point generation problem for combustion control.

5.5.1 Inverse-distance Interpolation based Set-point Generation

The inverse-distance interpolation technique interpolates based on a normalized weighting method. The weights are calculated based on the reciprocal of distances of a point to the available measurements. Because of the weighting scheme based on inverse-distances, the closest measurements have the highest influence on the interpolation output. The interpolated output can be represented by Equation 22 where \( z_i \) represents the measurements available, \( d_i \) is the Cartesian distance between the independent variables of point \( P \) and those of the point \( z_i \), \( u \) is the exponent on the distance, \( N \) is the number of measurements available for interpolation.

\[
f(P) = \frac{\sum_{i=1}^{N} \frac{1}{d_i^u} z_i}{\sum_{i=1}^{N} \frac{1}{d_i^u}} \tag{Equation 22}
\]

The method’s capability to interpolate using a scattered set of measurements makes it a potential method to reduce the calibration effort for the set-point generations scheme. Also, the extension of the method to
higher dimensions is straightforward without increasing the computational complexity considerably. For higher dimensions, only the calculation of the Cartesian distances is affected which is now calculated in the n-D space of independent variables. The computational complexity increases linearly with the increase in the number of dimensions. The number of mathematical operations required for the inverse-distance method with \( u = 2 \) where \( N \) points are used for interpolation in \( \alpha \) dimensions is as follows.

Number of Additions = \( N \cdot (\alpha + 1) - 2 \)

Number of Subtractions = \( N \cdot \alpha \)

Number of Multiplications = \( N \cdot (\alpha + 1) \)

Number of Divisions = \( N + 1 \)

It can be seen that the number of mathematical operations required are all functions of \( N \), the number of points used for interpolation. Therefore, if \( N \) increases, the interpolation can become computationally cumbersome. It is therefore desirable to restrict \( N \) to a value that can still yield reliable interpolation output. In Chapter 2, it was noted that the exponent \( u \) can be used to define the shape of the interpolation surface. Higher even values of \( u \) make the surface flatter near the measurements. It was also observed that for any value of \( u > 1 \), the directional derivatives become zero at the measurements leading to undesired constraint on the interpolation surface. However, for lower values of \( u \), this effect can be ignored. For ease in computation, the values of \( u \) can be selected as even numbers (with \( u = 2 \) being the simplest) which can eliminate the need to calculate the square roots while calculating the Cartesian distances. Calculating square roots are computationally expensive in an engine control unit. Another known shortcoming of the interpolation function is lack of directionality in the calculation. This is because the interpolation is based only on the distances and does not take into consideration the orientation of the measurements while doing the interpolation. Also, the inverse-distance method is incapable of extrapolating out of the bounds of the available measurements. In this study, the shortcomings have been addressed so that the inverse-distance interpolation method can be used for the set-point generation method.
5.5.1.1 Definition of a Neighborhood for Interpolation

When the available number of measurements or calibrations is very high, the inverse-distance interpolation becomes computationally cumbersome. The number of data points influences the number of all the four arithmetic operations as shown in the previous section. In practice, the number of data points used for interpolation can be picked from a neighborhood near the present operating point. This is similar to doing a piece-wise interpolation and only data points close to the present operating point will be used in the interpolation. The boundaries of the neighborhood are defined by the density of calibrations available and can be varied in each dimension.

The neighborhood for the set-point generation problem should be defined in the 4D space of engine speed, gross fuel amount, EGR fraction and boost pressure. In addition to defining the neighborhood bounds, it is also essential to define the number of data points used for interpolation. When the interpolation is being done at a region of the input space where there is a high density of calibrations, it is possible that the number of neighbors may be very high. To avoid such a problem, the available measurements can be ranked based on their distances and the $N$ nearest calibrations can be used for interpolation. On the other hand, when the number of calibrations available are less than a threshold value, $N_{\text{thres}}$, the bounds of the neighborhood can be increased iteratively until the required number of data points are available for interpolation. An example of a neighborhood is shown in Figure 71 where the number of independent variables is taken as three as anything beyond three cannot be visualized. The operating point at which the interpolation is being done is at the center of the box and represented by a ‘star’. The points represented by ‘plus’ are the values of independent variables at which calibrations are available. In this case, the bounds of the independent variables (X1, X2 and X3) are taken as ±100% from the present operating point for simplicity of representation.
5.5.1.2 Overcoming the Directionality Constraint of Inverse-Distance Method

The inverse-distance method does not consider the direction of the measurement while doing the interpolation. This may lead to a bias in the interpolation output. For example, at a given operating point of an engine, if the EGR fractions of all the calibration points are concentrated only in one half of the neighborhood (either positive or negative), the interpolation output will be biased towards EGR fraction values either higher or lower to the present operating point. The worst case of this phenomenon is when such a bias exists in all the dimensions of the interpolation. This phenomenon can be owed to the fact that the inverse-distance interpolation cannot extrapolate beyond the bounds of the calibration data. Therefore, care should be taken during the calibration phase so that all the possible operating ranges of the engine are covered. To overcome this constraint, a metric closely related to the center of mass of solid bodies is defined. The metric is explained with the help of Figure 72. Consider a simple case where the inverse-
distance method is used for two-dimensional interpolation. In Figure 72, the origin represents the present operating point in two dimensions and the points \((\Delta X_i, \Delta Y_i)\) represent the calibrations available for interpolation. The point \(C\) \((\Delta X_c, \Delta Y_c)\) represents the coordinates of the center of mass of all the available calibration points calculated using Equation 23. In the case of solids, the mass of each particle contributes to the center of mass. But, in this case, the inverse-distance of each point acts as the mass of the point in the center of mass calculation of Equation 23. The center of mass therefore is closest to points with highest ‘mass’, i.e., to the points closest to the operating point.

Figure 72: Center of Mass Demonstration
Ideally, for accurate interpolation, the center-of-mass should coincide with the operating point. This is only possible when the operating point is right on top of an existing calibration or when the available calibrations are oriented to the operating point in such a way that the center of mass coincides with the point. At best, it would be desirable to have a calibration data set that would minimize the distance of every operating point to the center of mass of the neighboring calibrations. However, to do such a minimization it is required to know the engine operating profile beforehand which is again impractical. However, one can define a set of operating profiles where the engine will operate most in its application and generate a calibration data set. An optimization procedure based on genetic algorithms is used to carry out the optimization. The output of the optimization is the set of input variables (here speed, fuel, EGR fraction and boost pressure) where the engine calibration should be performed to ensure accurate interpolation during that particular driving conditions. It should be noted that, this optimization requires the profiles for EGR fraction and boost pressure of the engine during the drive-cycle which poses challenges as such profiles cannot be generated without calibrating the air handling system of the engine. In practice, the profiles for EGR fraction and boost can be generated from a crude calibration of the air handling unit. Another method is to use the calibrations from an engine whose specifications match closely with those of the engine under consideration. In the next section, the genetic algorithm based optimization procedure has been explained.

### 5.5.2 Genetic Algorithm based Optimization

Genetic algorithm based optimization procedure can be used to find an optimal set of engine operating points represented by engine speed, gross fuel amount, EGR fraction and boost pressure that represent engine operation during a particular drive cycle. Genetic algorithms and their applications have been

\[
\Delta X^*_c = \frac{\sum_{i=1}^{n} \frac{1}{d_i} \cdot \Delta X_i}{\sum_{i=1}^{n} \frac{1}{d_i}}, \quad \Delta Y^*_c = \frac{\sum_{i=1}^{n} \frac{1}{d_i} \cdot \Delta Y_i}{\sum_{i=1}^{n} \frac{1}{d_i}}
\]

Equation 23
discussed in Chapter 3. They are ideal for this problem because of their ability to solve optimization problems involving a high number of variables. If $M$ operating conditions have to be generated, each represented by engine speed, load, EGR fraction and boost pressure, there are effectively $M\times 4$ variables in the optimization problem. The ability of genetic algorithms to converge without a good initial guess of the solution is also a very desirable feature for optimization problems similar to the one at hand.

The objective function for the genetic algorithm based optimization is based on the center of mass metric discussed in the previous section. At any given operating point, the distance between the center of mass of all the available calibration data points (within a neighborhood) and the operating point should be minimized for the inverse-distance interpolation method to give a reliable interpolation output. Therefore, for the optimization, the objective function may be chosen such that the distance of the center of mass is minimized. The root mean squared (RMS) value of the distance of the center of mass at every operating point of the engine is a good example of such an objective function shown in Equation 24 where $d_c$ represents the distance between the center of mass and the present operating point, $n$ represents the total number of samples in chosen drive cycle. In practice, the objective function can be modified with the help of penalty functions to achieve desired optimization results.

$$f(x) = \sqrt{\frac{1}{n} \sum_{i=1}^{n} d_{c,i}^2}$$

Equation 24

The parameters that have to determined before the optimization are as follows

1. The dynamic engine operating profiles for engine speed, gross fuel quantity, EGR fraction and boost pressure that represent typical engine operating conditions.

2. Number of data points to be identified, $M$

3. The bounds of the neighborhood for searching the neighbors during the optimization procedure

4. The maximum number of calibrations within a neighborhood used for interpolation denoted by $m$
As the number of data points increase, the interpolation output should ideally improve. The ideal number of data points can be determined based on the trade-off between the calibration effort and improved interpolation outputs. The bounds of the neighborhood should be as close to the operating point as possible. But, with less number of data points, the neighborhood has to be broader. As the number of data points increase, it is possible to narrow down the neighborhood bounds. It is possible that at certain operating conditions, the neighborhood may not contain any neighbors. The occurrence of such a situation can be minimized by adding a penalty to the fitness function of the optimization in such an event.

In this study, a Fortran implementation of the genetic algorithm has been used. The algorithm is an improved version of an open-source Fortran implementation. The use of Fortran enables fast execution of the algorithm.

5.6 Step-by-Step Approach for Inverse-distance based Set-point Generation

In the previous section, the inverse-distance method and various aspects of the method have been discussed in detail. The relevance of the inverse-distance interpolation method to the combustion control set-point generation has been established. In this section, the calibration procedure for the inverse-distance interpolation method has been described with the help of a flow diagram shown in Figure 73.

The calibration procedure for the inverse-distance method involves two stages. In the first stage, a set of different operating conditions representative of the engine operation in field is generated. The operating conditions are represented by engine speed, gross fuel amount, EGR fraction and boost pressure. These operating conditions are used in the genetic algorithm based optimization to generate operating points of the engine where calibration should be performed. To generate the representative operating conditions, it is required to have some form of calibrations for the engine. However, these calibrations need not be aggressive and optimal and a crude calibration that represents the actual engine operating conditions within an acceptable error is sufficient. The conventional calibration procedure can be toned down to generate this crude calibration for the engine. For generating the engine speed, a vehicle model which models the road load, vehicle load and engine crank shaft dynamics can be used. The gross fuel amount, EGR fraction and
boost pressure can be generated with the help of the crude calibration. The next step in the first stage of the calibration procedure is the genetic algorithm based optimization discussed in the previous section. The genetic algorithm based optimization procedure is central to this set-point generation method. The goal of this optimization is to generate operating points of the engine by minimizing the center of mass metric. The inputs to the genetic algorithm are the total number of operating points to be generated ($M$), the neighborhood bounds to calculate the center of mass metric and a set of operating points of the engine generated in the first stage of calibration procedure. The output of this step is a set of $M$ operating points where the injection parameters are optimized.

![Diagram of Calibration Procedure for the Inverse-distance Set-Point Generation Method](image)

**Figure 73: Calibration Procedure for the Inverse-distance Set-Point Generation Method**

In the second stage of the calibration procedure, the outputs of the set-point generator are calibrated for engine performance and emissions. For the closed-loop combustion controller, it was seen that the injection parameters are the suitable feed-forward components. Therefore, the injection parameters like injection timings and quantities are optimized at each of the operating points generated by the genetic algorithm. The objective function used for the optimization needs to be specified for this stage. This calibration can be model-based or can be done in an experimental setup. The output of this stage is the calibrations at operating conditions where the engine will be operated on the engine. It has to be noted that these
calibrations are scattered in the entire operating region of the engine. The inverse-distance interpolation method can now be used to generate the set-points for CLCC during the normal vehicle operation which is represented by the final block of Figure 73.

In this chapter, the closed-loop combustion control architecture for the heavy-duty diesel engine under study has been developed. Firstly, a hypothetical control architecture was discussed and it was simplified to arrive at a feasible design. The set of actuators and feedback variables were identified for closed-loop combustion control. A step-by-step approach for implementing the inverse-distance interpolation method to generate the set-points for the closed-loop combustion controller has been proposed. In the next chapter, the set-point generation scheme has been validated in a simulation setup and compared to a conventional set-point generation method.
VALIDATION OF KERNEL-BASED SET-POINT GENERATION FOR CLOSED-LOOP COMBUSTION CONTROL

6.1 Validation of the Inverse-distance Method for Set-point Generation

The goal of this chapter is to implement the inverse-distance set-point generation approach for the closed-loop combustion controller. The step-by-step approach for the implementation as proposed in the previous chapter has been applied to set-point generation for a simplified combustion control architecture. Various aspects of the step-by-step approach have been addressed in the implementation. The engine performance and emissions which resulted from the inverse-distance set-point generation method have been compared to those generated from a conventional calibration approach. For an apples-to-apples comparison of the conventional calibration approach and the inverse-distance approach, both the approaches have been implemented using a reduced GT-Power model of the engine. Neural network models have been used to predict engine performance and emissions during the validation procedure. In the end, the calibration effort required for the inverse-distance interpolation based set-point generation has been compared to that of conventional method with the help of some assumptions.

6.2 Reduced GT-Power Model

A reduced GT-Power model of the heavy-duty diesel engine developed by Cummins has been used for the comparison of a conventional calibration and inverse-distance based approaches for the set-point generation. The reduced GT-Power model (RGM) has been chosen because of its faster simulation times compared to the detailed GT-Power model (DGM) of the engine described in Chapter 3. Due to the reduced computational time, the RGM enables a faster evaluation of both the approaches in a simulation setup.

The RGM is a highly simplified version of the detailed GT-Power model. The cylinder in the RGM is modeled using a mean-value engine cylinder model which uses mapped functions for engine volumetric
efficiencies and distributions of fuel energy. These mapped functions are derived from a detailed GT-Power model. The RGM models the behavior of one cylinder and assumes all the other cylinders behave similarly and therefore does not predict cylinder-to-cylinder variability in combustion. Due to this limitation, the RGM can only be used for set-point generation for a single cylinder. For this study, this is sufficient to establish the efficacy of the set-point generation methodology. The RGM cannot generate crank angle resolved cylinder data which is required to calculate many combustion metrics. So, the set-point generation has been limited to the generation of the feed-forward components. As noted earlier, the set-points for the combustion metrics can be generated by operating the cylinders with the feed-forward injection parameters. The RGM does not have the capability to predict emissions like NOx and soot. Therefore, alternative methods are needed to predict emissions in the RGM. In this study, neural network models have been used to predict engine out NOx and ISFC of the engine. Details of the neural network modeling have been provided in the next section.

The intake and exhaust manifolds of the engine in the RGM are modeled using lumped parameter models. The lumped parameter modeling technique is also referred to as zero-dimensional modeling technique as it models the components as one control volume as opposed to discrete flow elements. Although the simplified modeling approach improves the computation times, it leads to inaccuracies in the model predictions. The RGM has been validated against the outputs of the detailed GT-Power model to ascertain the model’s accuracy. Figure 74 and Figure 75 compare the model predictions against the DGM model predictions. EGR flow, intake manifold pressure (MAP), ISFC and IMEP predictions of the RGM have been compared against the DGM’s outputs. It can be observed that the RGM predicts the variables within ±10 % error which is suitable for this analysis. Although, RGM models ISFC and IMEP, neural network models trained on data generated from the detailed GT-Power model have been used during the comparison study.
Figure 74: Comparison of EGR flow, MAP of RGM and DGM

Figure 75: Comparison of ISFC and IMEP of RGM and DGM
6.3 Neural Networks for Engine Emissions and Performance

Neural network models have been developed for engine-out NOx concentration and the engine performance variables, ISFC and IMEP. The neural networks have been used in the comparison study between the conventional and inverse-distance methods. For the development of the neural networks, MATLAB’s Neural Network toolbox has been used. In all the neural networks, feed-forward architecture described in Chapter 3 has been used with one hidden layer of neurons. Non-linear tan sigmoid activation functions are used for the hidden layer and linear activation functions are used for the output layer of the neural network. The back propagation algorithm described in Chapter 3 has been used to train the neural networks. Figure 76 shows the structure of the neural networks for a case with 2 inputs, 4 hidden neurons and 3 outputs. In this study, the same structure has been maintained for all the neural networks trained and only the number of neurons have been changed to arrive at the best performing neural network.

The data for training the neural networks was generated from various design of experiments conducted on the detailed GT-Power model of the engine. The data from the design of experiments on DGM discussed in Chapter 3 has been used in conjunction with other DOEs that were conducted as a part of a larger research
effort with Cummins. All the DOEs were carried out on the 11-modes of the engine discussed in Chapter 4. The DOEs provided a good set of training data for the neural networks to predict the effects of engine operating conditions like EGR fraction, boost pressure and injection timings on engine out NOx, ISFC and IMEP. However, due to the limited set of data used for training, the neural networks can only be used to predict trends in the engine emissions and performances. The neural networks for NOx, ISFC and IMEP have been trained with the following inputs.

1. Engine Speed

2. Gross fuel quantity

3. Main Injection Timing

4. Pilot Injection Timing

5. EGR fraction

6. Boost Pressure

7. Intake Manifold Temperature

The effect of post injection timing and rail pressure on the variables has been ignored because of unavailability of training data to capture the effects. Table 9 gives the number of neurons in the hidden layer for each neural network. Figure 77-79 plot the error in neural network predictions for the three variables as compared to the training data in the form of histograms.
<table>
<thead>
<tr>
<th>Variable</th>
<th>Number of Neurons</th>
</tr>
</thead>
<tbody>
<tr>
<td>NOx</td>
<td>40</td>
</tr>
<tr>
<td>ISFC</td>
<td>40</td>
</tr>
<tr>
<td>IMEP</td>
<td>30</td>
</tr>
</tbody>
</table>

Table 9: Number of Neurons in the Hidden Layers for the Neural Networks Trained

![Error Histogram](image)

Figure 77: Error Histogram for Neural Network prediction of NOx
Figure 78: Error Histogram for Neural Network prediction of ISFC

Figure 79: Error Histogram for Neural Network prediction of IMEP
6.4 Simplified Closed-loop Combustion Control Architecture for Validation

The reduced GT-Power model enables a faster evaluation of the validation approach that has been proposed. But, the reduced modeling capability limits the ability to evaluate all the aspects of the closed-loop combustion control architecture proposed in Chapter 5. Therefore a simplified combustion control architecture has been developed for the validation of the set-point generation using the inverse-distance approach. The concepts developed for the simplified control architecture can be directly extended to a full-fledged control architecture. The following assumptions have been made in developing a simplified control architecture.

The objective of the simplified closed-loop combustion controller has been simplified to minimize the variability only in engine out NOx and ISFC of the engine. Other variables of interest like soot, noise and drivability have been excluded from the objective of the combustion controller. The combustion metric, CA50, calculated from net heat release rate of combustion is chosen as the feed-back variable for the simplified controller because it is correlated well to NOx and ISFC. As noted in the previous chapter, injection timing is the preferable actuator for controlling a combustion metric like CA50. So, the simplified closed-loop combustion controller adjusts the injection timings such that the set-point for CA50 is achieved. For simplicity, the closed-loop controller moves the three injection timings as a whole from their nominal values instead of moving them individually. For this study, the feedback loop is ignored because the RGM cannot generate the crank angle resolved cylinder pressure. The set-point generation method is demonstrated only to generate the feed-forward components of the injection timings at various engine operating conditions. As it was noted before, the set points of the combustion metric (here CA50) can be generated by operating the individual cylinders at the nominal injection parameters (feed-forward component) without any disturbance to the root-cause variables. Figure 80 shows the simplified combustion control architecture used in this validation study. The feedback loop has been shown in the architecture although it was ignored in the study. It can be observed that the plant is the mean cylinder model in the reduced GT-power model of the engine. There is only one feed-back to the combustion controller which is CA50. The output of the controller is the set of injection timings which are the same for
all the cylinders. The objective function for the set-point generation is a weighted sum of NOx and ISFC whose weights can be adjusted based on the relative importance of each variable. The goal of the set-point generator is to generate the set-points for the combustion metric and the feed-forward injection timings that minimize the objective function at a given operating condition. The operating conditions are represented by engine speed, gross fuel, EGR fraction and boost pressure. The engine controller controls the air handling unit to achieve the set-points for EGR fraction and boost pressure. The engine controller also determines the gross fuel amount depending on the driver’s torque request.

Figure 80: Simplified Combustion Control Architecture for Validation


6.5 Set-point Generation using Conventional Calibration Method

The goal of this effort is to emulate the conventional calibration methodology for the set-point generation problem. The conventional method is used to calibrate the fuel injection system and also the air handling unit of the RGM. The calibrations generated in this effort for the air handling system of the engine will also be used in the evaluation of the inverse-distance method. In the conventional calibration procedure, a design of experiments is conducted on the engine with the engine operating conditions as the factors. Traditionally, only engine speed and load are used to represent the engine operating conditions. Load of the engine is represented by the gross fuel quantity injected into the cylinders. For every combination of engine speed and load, the air handling unit and the fuel injection system are calibrated to minimize the objective function chosen for the optimization. The objective function may be a weighted sum of engine performance and emissions as shown in Equation 21 in Chapter 5. The output of a conventional engine calibration procedure is a set of 2-D calibration maps for the root-cause variables of the engine like the EGR fraction, boost pressure, injection timings, rail pressure etc. as functions of engine speed and load.

In order to emulate the conventional calibration procedure, a design of experiments has been done on the reduced GT-Power model. The goal of the calibration procedure is to generate calibration tables for EGR fraction, boost pressure and injection timings. Rail pressure has been omitted from the calibration procedure and a base engine calibration table has been used for it. Therefore, the factors for design of experiments have been chosen as engine speed, gross fuel quantity, EGR valve position (to vary EGR fraction), variable geometry turbocharger vane position (to vary boost pressure) and injection timings. The following table gives the levels of each factor in the DOE.
<table>
<thead>
<tr>
<th>DOE factor</th>
<th>Number of Levels</th>
</tr>
</thead>
<tbody>
<tr>
<td>Engine Speed</td>
<td>19</td>
</tr>
<tr>
<td>Engine Load (fuel quantity)</td>
<td>16</td>
</tr>
<tr>
<td>EGR valve position</td>
<td>21</td>
</tr>
<tr>
<td>VGT position</td>
<td>21</td>
</tr>
<tr>
<td>Injection Timings</td>
<td>21</td>
</tr>
<tr>
<td>(All injections moved together)</td>
<td></td>
</tr>
</tbody>
</table>

Table 10: Factors chosen for DOE on the RGM

It can be observed that the DOE has approximately 2.8 million (19x16x21x21x21) operating conditions at which engine emissions and performance have to be evaluated. In practice, it is impossible to calibrate an engine at so many points. The actual methods used by engine calibrators to reduce the number of data points are still ad-hoc in nature. With the help of engineering intuition and calibration experience, the calibrators may be able to reduce the number of points. For example, a coarser grid may be used for some parameters and interpolation methods can be used to fill the missing data if it is known from experience that the parameters behave smoothly in the missing regions. Engine calibrators also recycle calibrations from an engine that has similar specifications to the engine being calibrated. In this way, it is possible to have good starting points for the calibration which can reduce the number of experiments considerably. The actual effort invested by the engine calibrators would be many times less than that of the ideal method. If the number of points evaluated for EGR, boost and injection timings are taken as 10 instead of 21, the actual calibration effort would be 9 times less and can be done by evaluating 304000 operating points, which is still cumbersome to carry out in an engine test cell. For this study, the DOE is conducted on the RGM at all the operating points of the DOE due to lack of good starting points for EGR and boost pressure to reduce the number of experiments. The calibration effort, however has been reduced with the help of some assumptions.
The evaluation of 2.8 million engine operating conditions is a highly time-consuming task even in a simulation setup. The reduced GT-Power model takes approximately 15 seconds to simulate the steady state engine behavior for each operating condition. Simulation of 2.8 million cases takes approximately 16 months if it is done on the RGM itself which is impractical by any means. In order to speed up the simulation time, the DOE is split between GT-Power and MATLAB. The RGM is used to simulate the dynamic behavior of the engine which is primarily exhibited by the air handling unit. The fuel injection parameters affect combustion in an almost instantaneous fashion. Therefore the effect of injection timings on the engine emissions and performance can be evaluated using static models as long as the rest of the engine sub-systems have reached steady state. The DOE has been partitioned between GT-power and MATLAB where GT-power is used to simulate the dynamic behavior and MATLAB is used for the static modeling. The advantage of the method is the computation time can be reduced significantly (here by a factor of 21, the number of levels in injection timings). Therefore, the entire DOE can be done within a month’s time. However, there are some disadvantages with this method. The effect of injection timings is not entirely instantaneous since there are effects of cycle-to-cycle coupling. The change in injection timings can change the combustion temperature and affect the cylinder intake conditions through the exhaust recirculation. However the cycle-to-cycle coupling has pronounced effects only if the injection timings are changed by large amounts. In the DOE, a nominal injection timing profile is obtained from the base calibrations of the engine that the RGM represents. The nominal injection timing profile is used for the part of DOE done on the RGM in GT-Power. For the static DOE in MATLAB, the nominal injection timings are moved only by a maximum of 2 crank angle degrees from the nominal values. The DOE where the dynamics in EGR fraction and boost are simulated is carried out first followed by the static DOE on injection timings in MATLAB. In order to evaluate the emissions and performance in MATLAB, neural network models of engine emissions and performance discussed in the previous section were used.

6.5.1 Optimization of the DOE Results

The design of experiments for the RGM has been done by splitting the experiments between GT-Power and MATLAB as explained earlier. Using the results from the design of experiments, the RGM has been
calibrated for engine performance and emissions. For every combination of speed and load, optimal operating conditions like EGR fraction, boost pressure and injection timings have been found out. For the optimization, the objective function shown in Equation 25 has been used. Equal weights are given to NOx and ISFC during the optimization. Normalized values of ISFC and NOx have been used in the objective function. The other variables like soot, noise and drivability have been omitted from the objective function due to the modeling limitations of the RGM. The constraint on IMEP is to ascertain the generation of desired engine torque at every operating point.

\[ f_{opt} = \frac{1}{2} \cdot ISFC_{Norm} + \frac{1}{2} \cdot NOx_{Norm} \]

Equation 25

Subject to \( IMEP = IMEP_{des} \)

Figure 81-83 show the normalized ISFC, normalized NOx and the objective function values as function of normalized MAP and EGR flow. The plots are generated for a mid speed, mid load operating point at nominal injection timings. The operating conditions which satisfy the IMEP constraint within ±5% of desired IMEP value have been shown in the plot.
Figure 81: Normalized ISFC as a function of normalized MAP and EGR flow

Figure 82: Normalized NOx as a function of normalized MAP and EGR flow
The EGR fraction, boost pressure and injection timings corresponding to the minimum value of the objective function (Figure 83) where the IMEP constraint is met are chosen as the calibrated values. The result of the calibration process is the set-points for EGR fraction, boost pressure and injection timings for every combination of engine speed and load. It has to be noted that the result of the traditional calibration method is set-points based only on engine speed and load. The transients in EGR fraction and boost pressure do not have any effect on the injection timings which is a short-coming of the traditional calibration method described in the previous chapter. The calibration tables generated for injection timings are used for the set-point generation of the closed-loop combustion control architecture. The calibrations generated for EGR fraction and boost pressure are used as set-points for the air handling unit of the RGM. The engine performance and emissions that result by using the conventional calibration tables are compared with those generated when inverse-distance method is used for generating the set-points for the injection timings. The next section deals with the inverse-distance based set-point generation method.
6.6 Set-Point Generation for the Simplified Combustion Control Architecture Using Inverse-Distance Method

At the end of Chapter 5, a step-by-step approach has been proposed for the inverse-distance based set-point generation. In this section, the step-by-step approach has been applied to the set-point generation of the simplified closed-loop combustion control architecture. In the inverse-distance method EGR fraction and boost pressure along with engine speed and gross fuel amount are used to represent the operating conditions of the engine. This is to improve the transient performance of the engine as explained in the previous chapter. For the simplified combustion control architecture, the scope of the set-point generation is to generate feed-forward injection timings and set-points for the combustion metric, CA50. The focus of this study, however, has been limited to the generation of the feed-forward injection timings for the mean-value cylinder model of the RGM. Therefore, the outputs of the set-point generator are the injection timings. For this analysis, only pilot and main injection timings have been controlled with the feed-forward controller and the post injection timing has been generated with the help of a base calibration table.

6.6.1 Stage1: Genetic Algorithm Optimization for the Generation of Calibration Points

In this stage, the genetic algorithm based optimization approach proposed in the previous chapter is used to generate operating points in the input space of the set-point generator where calibration of the RGM will be carried out. The goal of the optimization is to minimize the distance of the center of mass of neighboring calibrations and every operating point of the engine. This process is necessary for the inverse-distance interpolation to give reliable interpolation outputs as discussed in the previous chapter. However, it is impossible to carry out this optimization for every possible operating point of the engine. Therefore, this optimization can be carried out on various drive-cycles that represent typical operating conditions of the engine in the field. In this study, the drive-cycle has been chosen as the Federal Test Procedure (FTP). For the optimization, the operating profiles of all the input variables of the set-point generator should be available. For this study, the FTP profile for engine speed and gross fuel amount have been taken from a vehicle test where the vehicle was equipped with the diesel engine under study. The profiles for EGR
fraction and boost pressure have been taken from the calibration tables generated for RGM using the conventional method. Such input profiles are not usually known for engines unless the engine has been calibrated before. In practice, a crude calibration can be performed to generate the input profiles for various driving conditions. Profiles generated using crude calibration tables may not represent the actual operating conditions of the engine and are therefore prone to some error. To emulate this problem, a random error of ±10% has been injected into the FTP profiles generated using the calibration tables generated during the conventional calibration approach. The optimization has been repeated for both exact and randomized FTP profiles to study the effect of error on the calibrations generated.

Genetic algorithms have been used for the optimization procedure. Genetic algorithms are global optimization methods that are highly suitable for optimization problems involving a large number of variables. The fitness function chosen for the optimization has been shown in Equation 26. This is a modified version of the fitness function shown in Equation 24 in the previous chapter. In Equation 26, $d_c$ represents the distance between the center of mass of the calibrated neighbors and the present operating point and $n$ represents the total number of samples in the FTP drive cycle. The terms $ES_c$, $Q_c$, $EGR_c$ and $BP_c$ represent the normalized center of mass coordinates for engine speed, fuel quantity, EGR fraction and boost pressure. The first term is the root-mean-squared (RMS) error of the center of mass distances in the entire FTP cycle and the second term consists of the average of center of mass coordinates at each operating point. The second term helps in minimizing the bias of the center of mass coordinates during the optimization. The third term is a penalty term which is a function of the number of calibration points that have not been used for the center of mass calculation during the entire FTP cycle. The penalty function is included to minimize the number of unused calibration points in the optimization results. Also, in the optimization, a penalty is added for operating points of the FTP drive cycle which do not have any neighbors within the defined neighborhood. For such operating points, the center of mass coordinates are assigned a high value which acts as a penalty for the objective function.
As described in the previous chapter, the following parameters also have to be identified in this step before carrying out the genetic algorithm based optimization.

1. The total number of operating points that have to be generated by the optimization denoted by $M$
2. The bounds of the neighborhood for inverse-distance interpolation
3. The maximum number of calibrations within a neighborhood used for interpolation denoted by $m$

The values for the above parameters cannot be chosen arbitrarily as they strongly influence the optimization results. All the three parameters together determine the distribution of calibration points in the input space of the set-point generator. In the following sections, the importance of each parameter is described and appropriate values for each parameter have been arrived at.

6.6.1.1 Effect of Total Number of Operating Points on Center of Mass

The total number of operating points, $M$, that is chosen to represent the FTP drive cycle determines the calibration effort of the inverse-distance method as well as the quality of interpolation output. An appropriate value for $M$ is that which results in desired engine performance and emissions and also minimizes the calibration effort. To analyze the effect of $M$, the optimization procedure has been carried out for different values of $M$. For these experiments, the neighborhood bounds and the maximum number of points within the neighborhood used for interpolation have been fixed. The bounds of the neighborhood have been chosen as $\pm 15\%$ of the range of each variable in the input space. For calculating the center of mass for the fitness function, a maximum of ten nearest neighbors within the neighborhood have been used ($m = 10$). The optimizations have been carried out separately using both exact and randomized FTP profiles as inputs to analyze the effect of the error in the input profiles on the optimization.
The FTP cycle chosen for the optimization has 3354 operating points after removing the operating conditions at idle. If the genetic algorithm is asked to pick 3354 calibration points, it will pick exactly all the operating points of the FTP. The center of mass distance for such a case would be zero as all the calibrations would lie right on top of each operating point. But, the calibrations generated would be unsuitable for any other driving conditions, which is undesirable. In the experiments, the total number of calibration points, $M$, has been chosen as 100, 200, 304, 500, 1000 for each optimization case. The genetic algorithm optimization for each value of $M$ has been carried out by keeping all the neighborhood bounds and the maximum number of calibrations used for center of mass calculation the same across the cases. For the center of mass representation, the present operating point is taken as the origin. The center of mass coordinates are normalized by the maximum value of the input variable and represented as percentages of engine speed, gross fuel, EGR fraction and boost pressure. The distance of the center of mass is therefore represented in percentage. Table 11 tabulates the average distance of the center of mass for each value of $M$, both for exact FTP profiles and randomized FTP profiles.

<table>
<thead>
<tr>
<th>Number of points ($M$)</th>
<th>Average distance (%) (Exact FTP Profiles)</th>
<th>Average distance (%) (Randomized FTP profiles)</th>
</tr>
</thead>
<tbody>
<tr>
<td>100</td>
<td>7.4769</td>
<td>7.5585</td>
</tr>
<tr>
<td>200</td>
<td>6.4774</td>
<td>6.5221</td>
</tr>
<tr>
<td>304</td>
<td>5.7279</td>
<td>6.0464</td>
</tr>
<tr>
<td>500</td>
<td>5.0781</td>
<td>5.2935</td>
</tr>
<tr>
<td>1000</td>
<td>4.0571</td>
<td>4.0578</td>
</tr>
</tbody>
</table>

Table 11: Average Distance of Center of Mass for different values of $M$

It can be observed from the table that the average distance of the center of mass decreases with the increase in the number of operating points chosen to represent the FTP cycle. For the cases with less number of calibration points, the genetic algorithm has to distribute the available calibration points such that the center of mass is minimized for the whole FTP cycle. With the increase in the number of available calibration
points, the genetic algorithm has the freedom to place them closer to the operating conditions of the FTP drive cycle. The center of mass has been plotted as a function of the number of calibration points in Figure 84. It can be seen that with the increase in the number of points, there are diminishing returns in terms of the average distance of the center of mass. It can also be observed that the average distance of center of mass has not been impacted significantly by the random error injected into the input profiles. In a way, the bounds used for the neighborhood also act as a tolerance for the error in the FTP profiles. If the error in the FTP profiles is within the bounds used for the optimization, the inverse-distance method would still find a good number of neighbors within the bounds for interpolation. Figure 85 plots the histogram of number of neighbors for the first four values of $M$. It can be seen that the number of FTP operating points that have more than 10 neighbors increase significantly with the increase in the $M$.

![Figure 84: Distance of Center of Mass (in %) as a function of the total number of calibrations, $M$](image)
Figure 85: Histogram of Number of Neighbors for different values of $M$

Figure 86-89 plot the histograms for the center of mass coordinates for each value of $M$ where exact FTP profiles have been used. It can be seen from the histograms that the spread of the coordinates reduces with increase in $M$. Table 12 and Table 13 compare the means and standard deviations of the center of mass for different values of $M$. The means and standard deviations of the experiments with randomized FTP profiles are also tabulated. All the means of the coordinates are within 1%. The standard deviations decrease with increase in $M$ as indicated by the spread in the histograms. The means and standard deviations in the cases where exact FTP profiles are used are very similar to those where randomized FTP profiles are used indicating a good tolerance of the center of mass to the error in the input profiles. Increase in the total number of calibration points reduces the average distance of the center of mass. Reduced center of mass improves the interpolation output of the inverse-distance method but nothing can be said about its effect on engine performance. The effect of center of mass on engine performance and emissions is studied later in the chapter which can be used to arrive at an appropriate value for $M$. The choice of $M$ depends on the value that yields desirable engine performance and also minimizes the calibration effort.
Figure 86: Distribution of Coordinates of the Center of Mass for $M = 100$

Figure 87: Distribution of Coordinates of the Center of Mass for $M = 200$
Figure 88: Distribution of Coordinates of the Center of Mass for $M = 304$

Figure 89: Distribution of Coordinates of the Center of Mass for $M = 500$
<table>
<thead>
<tr>
<th>Number of Points</th>
<th>Speed Coordinates</th>
<th>Fuel Coordinates</th>
<th>EGR Coordinates</th>
<th>Boost Pressure Coordinates</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Exact FTP</td>
<td>Randomized FTP</td>
<td>Exact FTP</td>
<td>Randomized FTP</td>
</tr>
<tr>
<td>100</td>
<td>0.2756</td>
<td>0.0698</td>
<td>0.0332</td>
<td>-0.2094</td>
</tr>
<tr>
<td>200</td>
<td>0.3994</td>
<td>-0.0819</td>
<td>0.2895</td>
<td>-0.3118</td>
</tr>
<tr>
<td>304</td>
<td>0.2235</td>
<td>0.1414</td>
<td>0.4365</td>
<td>0.6817</td>
</tr>
<tr>
<td>500</td>
<td>0.4897</td>
<td>0.1414</td>
<td>0.4574</td>
<td>0.6134</td>
</tr>
<tr>
<td>1000</td>
<td>0.0186</td>
<td>0.0210</td>
<td>0.5264</td>
<td>0.5317</td>
</tr>
</tbody>
</table>

Table 12: Comparison of Mean Center of Mass coordinates for different values of $M$

<table>
<thead>
<tr>
<th>Number of Points</th>
<th>Speed Coordinates</th>
<th>Fuel Coordinates</th>
<th>EGR Coordinates</th>
<th>Boost Pressure Coordinates</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Exact FTP</td>
<td>Randomized FTP</td>
<td>Exact FTP</td>
<td>Randomized FTP</td>
</tr>
<tr>
<td>1000</td>
<td>2.3403</td>
<td>2.3409</td>
<td>1.9822</td>
<td>1.9826</td>
</tr>
</tbody>
</table>

Table 13: Comparison of Standard Deviations in Center of Mass coordinates for different values of $M$
6.6.1.2 Effect of Neighborhood Bounds on Center of Mass

In order to reduce the computational burden for the inverse-distance method, a neighborhood is defined and only calibrations within the neighborhood are used for interpolation. There is a clear trade-off between the neighborhood bounds and the total number of calibration points chosen for the inverse-distance method. If the total number of calibrations are less, the neighborhood should be expanded to ensure that a minimum number of calibrations are available for interpolation at every given operating point. On the other hand, expanding the neighborhood beyond a certain point is not desirable as calibrations far away from the present operating condition may be used for interpolation which may deteriorate the performance of the engine. The genetic algorithm uses the neighborhood bounds and the total number of calibration points to distribute the calibrations across the operating range of the given drive cycle. If the bounds are too small, then there is a possibility that the resultant calibrations will become overly specific to the driving cycle for which the calibrations have been optimized. Such a calibration data set will be intolerant to errors in the input drive cycle profiles. Therefore, the bounds used for the neighborhood should also be dependent on the confidence one has on the correctness of the input driving profiles used in the genetic algorithm.

To study the effect of bounds on the center of mass, experiments were conducted where the bounds have been varied by keeping the other parameters of the method the same. The values for the bounds are taken as ±15%, ±10% and ±5% of the range each input variable. The optimizations have been done for $M = 304$ and $m = 10$. Figure 90 plots the effect of neighborhood bounds on the average distance of the center of mass. As expected, the center of mass increases with increase in the neighborhood bounds. The choice of the bounds is tied up to their influence on engine performance and emissions. In this study, this effect of neighborhood bounds on engine performance and emissions has not been studied and ±15% has been used as bounds for the neighborhood when studying the effect of inverse-distance method on engine performance and emissions.

The third parameter of the inverse-distance method is the maximum number of calibrations ($m$) that are used for interpolation in a given neighborhood. This value of this parameter is directly connected to the
computational burden of the inverse-distance interpolation. This value is therefore dependent on the processing power of the hardware on which inverse-distance method is implemented. Due to the lack of information on such hardware, this value has been chosen as 10 for this study. Therefore, for the evaluation of effect of inverse-distance method on engine out emissions and performance, the total number of calibration points \( (M) \) is the only parameter that has been varied.

![Bar Chart](chart.png)

Figure 90: Effect of Neighborhood bounds on Average Distance of Center of Mass

6.6.2 Stage 2: Optimization of Injection Timings for the Calibrations Generated by Genetic Algorithm

This is the second stage of the calibration where the injection timings at the operating points generated by the genetic algorithm are calibrated for engine performance and emissions. The injection timings are optimized at each of the operating points to minimize the objective function chosen for closed-loop combustion control. The optimization procedure is similar to the one used for the conventional calibration.
method where equal weights have been given for both ISFC and NOx concentrations with a constraint on the IMEP. Since the values for EGR fraction and boost pressure have already been generated by the genetic algorithm, it is not required to optimize them. This step has been implemented for all the operating points generated by the genetic algorithm for different values of $M$. The calibrations generated in this stage are used in the inverse-distance interpolation method to generate the set-points for CLCC. The algorithm for the interpolation has been implemented in MATLAB. The inverse-distance interpolation method to generate the injection timing ($SOI$) is shown in Equation 27 where $d$ is the Cartesian distance of the present operating point and an operating point for which calibration is available in the 4-D Cartesian coordinate system constituted by engine speed, gross fuel quantity, EGR fraction and boost pressure.

$$SOI_k = \frac{\sum_{i=1}^{10} \frac{1}{d_i^2} \cdot SOI_i}{\sum_{i=1}^{10} \frac{1}{d_i^2}}$$

Equation 27

At a given operating point, the inverse-distance interpolation algorithm finds neighbors from the calibrations that are within the neighborhood whose bounds are defined by ± 15% of the range of each input variable. The algorithm then sorts the calibration points based on distance and uses the nearest 10 points for the interpolation. This algorithm is repeated at every engine operating point and the injection timings are calculated.

6.7 Comparison of Engine Performance and Emissions between the Inverse-distance and Conventional Set-Point Generation Methods

The inverse-distance interpolation and conventional calibration methods have been implemented as algorithms in Simulink. FTP simulations have been done on the reduced GT-Power model of the engine using both conventional and inverse-distance methods to generate set-points for the injection timings. For
calculating the emissions and performance during the FTP drive cycle, neural network models of NOx and ISFC have been used. In the case of inverse-distance method, the RGM has been simulated on a FTP drive-cycle with different number of calibration points (M) for the set-point generation. The other parameters of the inverse-distance method like the neighborhood bounds and maximum number of neighbors used for interpolation have not been varied for cases with different values of M. The neighborhood bounds are taken as ±15% of the range of each input variable. A maximum number of 10 neighbors within the neighborhood have been used for interpolation to minimize the computational complexity. For the inverse-distance method, calibrations generated while using exact and randomized FTP profiles in the genetic algorithm optimization stage have been used in the comparison study. For comparing the engine performance and emissions, the FTP drive cycle has also been simulated by using the calibrations for the injection timings generated using the conventional calibration method.

All the FTP simulations done differ only in the method and the calibrations used for generating the injection timings. The engine torque output during the FTP cycle simulations is approximately equal due to the IMEP constraint used during the calibration procedure. For boost and EGR fraction, the calibration tables generated during the conventional calibration method have been used in all the simulations. Therefore, any difference in the engine performance and emissions between the simulations can be attributed to the differences in injection timing profiles used during the simulations. The engine performance and emissions generated by each method are compared and the comparison study has been reported in the next section.

The conventional and inverse-distance methods have to be compared against an ideal case to evaluate the performance and emissions in each case. The ideal case should result in ideal performance and emissions of the engine for the FTP drive cycle. For such a case, the objective of the calibration has to be met at every operating condition of the FTP drive cycle. This is only possible if injection timings have been optimized at every operating point of the FTP drive cycle. Such a calibration set will result in the optimal trade-off between NOx and ISFC during the FTP drive cycle. Even though, such a calibration set will not produce optimal results for any other driving cycle, it provides for an ideal case for the comparison of the
performance of conventional and inverse-distance methods for the FTP drive cycle. To generate the calibrations for the ideal case, the injection timings have been optimized at every operating condition of the FTP drive cycle to minimize the engine out NOx and ISFC using the objective function given by Equation 25. The optimal values of cumulative NOx and fuel consumption during this process have been considered as ‘ideal’ for the comparison study.

For each FTP simulation, the NOx concentration and engine ISFC have been calculated by the neural network models of NOx and ISFC. It has to be noted that the neural networks are good at predicting the trends rather than the absolute values of the variables. Therefore, importance has to be given to the trends in these results rather than the absolute numbers. The cycle integrated engine out NOx emissions and the fuel consumption in each simulation have been calculated for each FTP simulation. For the sake of comparison, the NOx emissions and the fuel consumption have been normalized by those obtained from the ideal FTP case where every operating point on the FTP drive cycle was optimized. Table 14 tabulates cycle-integrated NOx emissions and fuel consumption for the conventional calibration method normalized by the corresponding values of the ideal FTP case. Table 15 tabulates the normalized values of cycle-integrated NOx and fuel consumption for different cases of the inverse-distance method when exact FTP profiles have been used in the genetic algorithm and Table 16 tabulates those when randomized FTP profiles have been used. The cumulative NOx and fuel consumption for different cases of the inverse-distance method have also been plotted in Figure 91 and Figure 92. The values for the ideal and conventional calibrations have been represented as horizontal lines in the plots. The ideal case is also equivalent to the inverse-distance method where the number of calibrations are chosen as the total number of operating conditions in FTP. For the inverse-distance method, with the increase in the value of $M$, cumulative NOx first decreases and then increases. For fuel consumption, the opposite trend has been observed. As the value of $M$ increases, the values tend to 1, which represents the ideal FTP case. The ideal case represents the optimal trade-off between ISFC and NOx which is determined by the objective function used during the engine calibration. The NOx and fuel consumption of different cases have also been compared in Figure 93 and Figure 94.
### Table 14: Comparison of FTP Cycle integrated NOx and fuel consumption between the baseline and the conventional calibration case

<table>
<thead>
<tr>
<th></th>
<th>Ideal FTP</th>
<th>Conventional Calibration</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normalized Cumulative NOx emissions</td>
<td>1.00</td>
<td>1.0072</td>
</tr>
<tr>
<td>Normalized Cumulative Fuel Consumption</td>
<td>1.00</td>
<td>0.9785</td>
</tr>
</tbody>
</table>

### Table 15: Comparison of FTP cycle integrated NOx and fuel consumption between the baseline and the inverse-distance method for exact FTP profiles

<table>
<thead>
<tr>
<th></th>
<th>Ideal</th>
<th>M = 100</th>
<th>M = 200</th>
<th>M = 304</th>
<th>M = 500</th>
<th>M = 1000</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normalized Cumulative NOx emissions</td>
<td>1.00</td>
<td>1.0074</td>
<td>0.9523</td>
<td>0.9487</td>
<td>0.9536</td>
<td>0.9780</td>
</tr>
<tr>
<td>Normalized Cumulative Fuel Consumption</td>
<td>1.00</td>
<td>1.0091</td>
<td>1.0161</td>
<td>1.0285</td>
<td>1.0259</td>
<td>1.0168</td>
</tr>
</tbody>
</table>

### Table 16: Comparison of FTP cycle integrated NOx and fuel consumption between the baseline and the inverse-distance method for randomized FTP profiles

<table>
<thead>
<tr>
<th></th>
<th>Ideal</th>
<th>M = 100</th>
<th>M = 200</th>
<th>M = 304</th>
<th>M = 500</th>
<th>M = 1000</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normalized Cumulative NOx emissions</td>
<td>1.00</td>
<td>0.9899</td>
<td>1.001</td>
<td>1.0170</td>
<td>0.9806</td>
<td>0.9791</td>
</tr>
<tr>
<td>Normalized Cumulative Fuel Consumption</td>
<td>1.00</td>
<td>1.0143</td>
<td>1.0074</td>
<td>1.0046</td>
<td>1.0137</td>
<td>1.0160</td>
</tr>
</tbody>
</table>
Figure 91: Comparison of Cumulative NOx emissions using exact FTP profiles for inverse-distance method

Figure 92: Comparison of total fuel consumption using exact FTP profiles for inverse-distance method
Figure 93: Comparison of cycle-integrated NOx emissions and fuel consumption for exact FTP profiles

Figure 94: Comparison of cycle-integrated NOx emissions and fuel consumption for randomized FTP profiles
Unlike the conventional calibration method, the inverse-distance method uses EGR fraction and boost pressure as inputs to the set-point generator. This is done to improve the transient performance of the engine and to reduce spikes in NOx. The injection timings compensate for the dynamics in EGR fraction and boost pressure reducing the spikes in NOx emissions during transients in the case of inverse-distance method. In order to demonstrate this effect, a transient load change event from the FTP simulation has been studied. The behavior of the engine variables during the transient event have been plotted in Figure 95. It can be see that the gross fuel injected has been ramped almost instantaneously and the corresponding changes in engine speed, boost pressure and EGR are much slower and take several seconds to change. Figure 96 plots the response of the set-point generator for the conventional and inverse-distance methods. The main injection timing in the conventional case reaches steady state when engine speed and fuel quantity reach steady-state as it is scheduled based on those two variables. On the other hand, the injection timing in the case of inverse-distance method reaches steady state much slower and keeps up with the dynamics in EGR and boost pressure. For generating the injection timing command during the transient phase, the inverse-distance method uses calibrations whose operating conditions were similar to that of the transient conditions in FTP. The resulting normalized NOx between the two methods is shown in Figure 97. It can be observed that during the transients, the instantaneous difference in NOx between the two methods is significant with the inverse-distance method giving smaller spikes.

The steady-state performance of the inverse-distance and the conventional set-point generation methods are expected to be similar. However, it has to be noted that the steady-state behavior of the inverse-distance method is still dependent on the parameters like the total number of calibrations used for the method. As the calibration data used for the interpolation changes, it is possible that the inverse-distance method results in different steady-state behavior of the engine as a different set of calibration points may be used to determine the interpolation outputs during the steady-state.
Figure 95: A Transient Maneuver in the FTP cycle

Figure 96: Response of the conventional and inverse-distance set-point generators
6.8 Comparison of Calibration Effort Between the Inverse-Distance and Conventional Set-Point Generation Methods

In this section, the calibration effort required for the inverse-distance method has been compared against that required for conventional method of set-point generation. The example of set-point generation for the simplified combustion control architecture has been used for this comparison. The set-point generator for the simplified combustion controller takes the present operating conditions of the engine to generate injection timings for the cylinders. The engine operating conditions are represented by engine speed, gross fuel amount, EGR fraction and boost pressure. It is assumed that the same injection timing commands are used for all the cylinders and all injections of a cylinder are moved simultaneously during the calibration procedure. Therefore at every operating condition of the engine, the injection timings are optimized to meet the objective of the calibration.

Figure 97: Normalized NOx emissions during the transients
In the case of conventional calibration, the engine is calibrated for every engine speed and gross fuel quantity using a uniform grid based approach. At every engine speed and gross fuel quantity, the air and fuel injection parameters are optimized. If the calibration is done at \( s \) engine speeds and \( f \) fuel quantities, and \( p \) points are used for each of the remaining three dimensions, it results in a calibration effort of evaluating \( s^*f^*p^3 \) engine test cell experiments. As described previously, it is difficult to quantify the calibration effort of a conventional method exactly as calibrators use ad-hoc methods to minimize the effort. With the help of prior calibration experience, it is possible to generate good starting points for calibrating the engine at each speed and load and therefore reduce the value of \( p \). The methods that manufacturers use to generate these starting points are well guarded secrets. It was learned that, for the engine that has been used for this study, the number of points evaluated for every speed and load are in the order of few hundreds. For the engine under study, the typical grid size of engine speed and fuel quantity is 19x16 which amounts to 304 engine speed and load points. Therefore, the total calibration effort for the conventional calibration approach can be approximated as \( 19^*16^*p^*p^*p = 304^*p^3 \) operating points to be evaluated.

For the inverse-distance interpolation method, it is necessary to have a crude calibration of the engine which is used to generate operating profiles for the input variables that are used in the genetic algorithm optimization. However, the engine need not be aggressively calibrated for generating these base calibration tables. Methods that are already used to reduce the calibration effort of conventional methods can be used more aggressively to generate these base calibration tables. The factor by which the calibration effort is reduced for generating the base calibrations when compared to the full-fledged conventional method varies depending on the engine and also the calibrator’s experience. For example, it could be possible to use a coarser grid for EGR points if the calibrator has the knowledge that the calibration surfaces are smooth for variations in EGR. For the calibration effort comparison, this factor of reduction, denoted by \( k \), has been taken as a variable and the calibration effort has been compared for different values of \( k \). It is also assumed that the number of points in all the 5 dimensions can be reduced by the same factor. Therefore the number of engine experiments that should be done to generate the base calibration table for the inverse-distance
method are \((19*16*p*p*p)/k^5 = (304*p^3)/k^5\). Once the base calibration table has been generated, the operating profiles for engine speed, fuel amount, EGR fraction and boost pressure can be generated for a given drive cycle. These operating profiles can be used in the genetic algorithm optimization to generate \(M\) operating points that represent the chosen drive cycle for the inverse-distance interpolation method. At each of the \(M\) operating points represented by engine speed, gross fuel, EGR fraction and boost pressure, the injection timings have to be optimized. The optimization of injection timings is common to both conventional and inverse-distance methods. Assuming the same \(p\) number of injection timings (used in conventional calibration method) are evaluated at each of the \(M\) operating points, the total calibration effort for the inverse-distance method would be \((304*p^3)/k^5 + p*M\). In this expression, the value of \(p, k\) and \(M\) are variable. To compare the calibration effort between the conventional and inverse-distance methods, the values of \(p, k\) and \(M\) are varied. The ratio of the number of experiments in conventional method to that of the inverse-distance method can be used as a measure to compare the calibration effort between the two methods. The ratio is represented as \(\beta\) which is given by Equation 28 where \(p\) is the number of points evaluated for each input variable EGR, boost pressure and injection timings, \(k\) is the reduction factor that is used to generate the base calibration table for the inverse-distance method and \(M\) is the total number of operating points found using the genetic algorithm.

\[
\beta = \frac{304, p^3}{\left(\frac{304, p^3}{k^5}\right) + p*M}
\]

Equation 28

The values of \(p, k\) and \(M\) have been varied to find the calibration effort ratio (\(\beta\)) at each combination of the three factors. The value of \(\beta\) for \(p = 6\) has been plotted in Figure 98 for different values of \(k\). The value \(k = 1\) represents no reduction in the grid size from the conventional calibration method to generate the crude calibration table for inverse-distance method. For this case, the calibration effort for the inverse-distance method is higher than conventional calibration case for reasons that are obvious. In practice, the value of \(k\) will always be higher than 1 as the base calibration tables required for the inverse-distance method can be
crude. It can be seen that, for the case where \( k = 2 \), the calibration effort for the inverse-distance method can reduce significantly by a factor of 5 to 25 depending on the value of \( M \) when compared to the conventional method. Figure 99 and Figure 100 compare the calibration efforts for \( p = 8 \) and \( p = 10 \). The calibration effort required for the inverse-distance method reduces further with increase in \( p \) for \( k > 1 \). Depending on the actual value of \( p \) and \( k \) for a given engine, the reduction in calibration effort can be read out from these figures.

![Comparison of Calibration Effort (p = 6)](image)

**Figure 98:** Comparison of Calibration Effort for \( p = 6 \)
Figure 99: Comparison of Calibration Effort for $p = 8$

Figure 100: Comparison of Calibration Effort for $p = 10$
6.9 Conclusions from the Validation

In this chapter, the inverse-distance interpolation method has been demonstrated as a method to generate the set-points for a simplified closed-loop combustion controller. A reduced GT-Power model of engine and neural network based engine performance and emission models have been used to study the effect of inverse-distance method on engine performance and emissions. For comparison, a conventional calibration method has also been implemented on the RGM. The engine performance and emissions from the inverse-distance method and the conventional calibration method have been compared. The inverse-distance method resulted in engine performance and emissions that closely match with that of the conventional method. The cycle-integrated NOx emissions of the inverse-distance method have reduced when compared to the conventional method and on the other hand fuel consumption results for the inverse-distance method were higher when compared to the conventional method. It was seen that, as the value of $M$ increases, the NOx and fuel consumption of the inverse-distance method tend towards the ideal FTP case. The inverse-distance set-point generation method also resulted in better transient performance of the engine when compared to the conventional method resulting in reduced NOx spikes during transients.

The calibration efforts required for the inverse-distance method have been compared to that of a conventional method with the help of some assumption. From this analysis, it can be concluded that the inverse-distance method has a high potential to reduce the calibration effort when compared to conventional calibration methods used by engine manufacturers. Depending on the values of $p$, $k$ and $M$, the calibration effort can be reduced by a factor of 5 to 25 for the set-point generation of the simplified combustion controller. The parameters $p$, $k$ depend on the calibration practices followed by the manufacturer. For the calibration generated for the inverse-distance method to represent all the possible engine driving conditions, the value of $M$ should be high. By simulating the engine performance and emissions over the proposed driving cycle for different values of $M$, it is possible to choose the value that meets the emissions and performance objectives.
CONCLUSIONS

In this study, concepts have been developed for the application of closed-loop combustion control to conventional diesel combustion. Closed-loop combustion control is a promising technology that can help in reducing the variability in engine emissions and performance caused by design driven, component driven and environmental driven sources of variability. With the tightening of emission regulations, the variability in engine emissions caused by sources of variability is becoming significant. To keep the engines compliant with the emission regulations, manufacturers are forced to calibrate the engines beyond the regulatory limits of emissions to allow for some margin for the emission performance to decline because of variability effects. This often leads to deterioration in engine performance, like higher fuel consumption, which is highly undesirable for the engine manufacturers. With the help of closed-loop combustion control, the variability in engine emissions can be minimized and the engine can be calibrated closer to the regulatory limits of emissions. As a result, the engine can also be calibrated for higher fuel efficiency without the fear of non-compliance in emissions.

In Chapter 4 of the thesis, various sources of variability in a diesel engine that can affect engine performance and emissions have been identified. The effects of the variability are studied with the help of a high-fidelity diesel engine model of a Cummins heavy-duty diesel engine. The variability in engines can be classified as design-driven, component driven and environmental driven. Cylinder-to-cylinder variability in combustion is a design driven variability which can lead to significant differences in combustion of each cylinder. The variability was found to be particularly large between cylinders 1 and 6, as they were situated at either extreme of the intake manifold. Although the variability caused by design cannot be completely eliminated, it is possible to compensate for the effect of such variability on combustion by using different injection parameters for each cylinder. Component driven variability in combustion can be caused by differences in components from engine-to-engine due to manufacturing tolerances and component aging. Such a variability affects combustion in cylinders through the root-cause variables of combustion. Root-
cause variables are the basic necessities of combustion like fresh air, residuals and fuel injection parameters. Environmental driven variability can be caused by variation in the boundary conditions of the engine like ambient conditions, fuel properties and conditions imposed by an after-treatment system. Environmental driven variability also affects combustion through the root-cause variables. So, in order to study the sensitivity of diesel combustion to various sources of variability, a design of experiment has been developed where the root-cause variables of the engine have been perturbed from their nominal values. For this analysis, EGR fraction, boost pressure, fuel injection amount, injection timings, rail pressure and swirl have been chosen as the root-cause variables. Engine-out NOx and Indicated Specific Fuel Consumption (ISFC) have been taken as combustion performance variables for the sensitivity analysis. Analysis of Variance Method (ANOVA) has been used to identify the significance of each root-cause variable on the response variables. From the ANOVA analysis for engine out NOx and ISFC, it was identified that EGR fraction, boost pressure, fuel quantity and injection timings are key factors which can contribute significantly to the variability in NOx and ISFC. EGR fraction and boost pressure showed the highest effect on NOx compared to the fuel injection system parameters. In the case of ISFC, boost pressure showed highest significance. Therefore NOx and ISFC have shown more sensitivity to variability in the air-handling unit compared to that in the fuel injection system.

A sensitivity analysis has also been done on various combustion metrics chosen during the literature survey. Among those chosen, the combustion metrics derived from combustion heat release rate analysis showed sensitivity to variability to all the key-factors like EGR fraction, boost pressure, fuel quantity and injection timings. This makes them ideal candidates for combustion feedback since they can be used to sense variability in engine out NOx and ISFC caused by the key factors. Combustion metrics like maximum pressure and maximum rate of pressure rise are not suitable for conventional diesel engines as they often occur before the heat release of combustion due to high compression ratios. Therefore, metrics like maximum combustion pressure have been derived which separate the effect of combustion from the cylinder pressure measurement by subtracting the motoring trace from the firing trace. In the sensitivity analysis, maximum combustion pressure showed considerable sensitivity to variability in the gross fuel
quantity. Thus, metrics derived from combustion pressure can be used to sense variability in fuel quantity. Such metrics will also be good candidates for monitoring engine noise and drivability. From the sensitivity analysis study, the key factors that influence engine performance and emissions have been identified. Also, various combustion metrics that can be used to achieve different goals of closed-loop combustion control have been identified. This information has been used in designing a closed-loop combustion control architecture.

In Chapter 5, the overall goal and the architecture of the closed-loop combustion controller has been defined. Firstly, a hypothetical control architecture has been proposed that served as a framework for identifying various challenges associated with the design of a closed-loop combustion control architecture. With the help of some assumptions, the closed-loop combustion control architecture that can be seamlessly integrated into the existing engine control architecture of Cummins heavy-duty diesel engine has been proposed. The proposed combustion controller closes the loop on fuel injection parameters of the engine and does not modify the set-points of the air handling unit nor can it interfere with the engine torque controller. The set of possible actuators and feedback variables have been identified with the help of the assumptions. The emphasis of the rest of the work in the thesis is on generating set-points for the closed-loop combustion controller.

Set-point generation refers to generating the feed-forward components and the set-points for the combustion metrics for the closed-loop combustion controller. A strong feed-forward component is an important part of the combustion control problem due to the non-linearity in combustion phenomenon. The proposed set-point generator creates the combustion references by taking engine speed, load, EGR fraction and boost pressure as inputs. Dynamic variables, like EGR fraction and boost pressure, have been included in the inputs to generate appropriate references during engine transients. Due to the higher number of independent variables and outputs, conventional calibration methods are not suitable for the set-point generation problem as they result in very high calibration effort. Therefore, a calibration approach based on kernel based interpolation methods has been proposed as a solution with the goal of reducing calibration effort required for the set-point generation problem.
The calibration method chosen for the set-point generation is based on inverse-distance based interpolation. The inverse-distance method uses a normalized weighting approach to interpolate between the calibrations of the engine where the weights are determined based on inverse-distance. As a result, the calibration closest to an operating point has the maximum influence on the interpolation output and vice-versa. The inverse-distance method facilitates the use of scattered calibration data for set-point interpolation and it can be extended to multiple dimensional interpolation without a significant increase in the computational effort. However, the method has some shortcomings which have been addressed in the study. In order to reduce the computational burden in interpolation, the inverse-distance method interpolates using only a set of neighboring points within a pre-defined neighborhood from the present operating point. The bounds of the neighborhood can be defined based on the density of the calibration points. Another shortcoming of the inverse-distance method is that the method is not capable of extrapolation beyond the bounds of the calibrated data. To overcome this problem, the ‘center of mass metric’ has been developed. For the inverse-distance interpolation to generate reliable outputs, the distance of the center of mass of all the neighboring calibrations and the present operating point should be close to zero.

A two-stage calibration approach has been proposed to implement the inverse-distance interpolation method for the set-point generation. In the first-stage, a crude calibration is generated for the engine which is used to generate the operating conditions of the engine during different driving scenarios. The operating conditions thus generated are used in a genetic algorithm based optimization procedure. The goal of the optimization is to generate a pre-defined number of engine operating points which minimize the distance between the center of mass of neighboring calibrations and every expected operating point of the engine. The output of the optimization approach is a set of engine operating conditions defined by the input space of the set-point generator (here engine speed, gross fuel amount, EGR fraction and boost pressure) where the engine should be calibrated for emissions and performance. Inverse-distance interpolation can be used on such calibrations to generate set-points during the actual operation of the engine.

In Chapter 6, the inverse-distance method has been applied to the set-point generation problem for a simplified closed-loop combustion control problem. The objective of the closed-loop combustion control
has been limited to minimizing NOx and ISFC of the engine and injection timings are used as the actuator variables for achieving combustion control. A reduced GT-Power model (RGM) of the diesel engine under study has been used to evaluate the set-point generation method. The reduced GT-Power model facilitates a quick evaluation of the inverse-distance method. However, it is incapable of predicting cylinder-to-cylinder variations in combustion and crank-resolved cylinder pressure. Therefore, the set-point generation problem has been simplified to generating feed-forward injection timing commands that are the same for all the cylinders. The inverse-distance based calibration method has been compared to a conventional calibration method both in terms of calibration effort and engine performance. For an apples-to-apples comparison of engine performance and emissions, both the methods have been implemented in simulation. In a conventional calibration approach, only engine speed and load are used as the independent variables of the engine. In order to emulate a conventional calibration procedure, a design of experiments (DOE) has been conducted at each combination of the engine speed and load where the EGR fraction, boost pressure and injection timings of the engine have been varied. The EGR fraction, boost pressure and injection timings at each engine speed-load combination have been optimized for NOx and ISFC based on an equally weighted objective function. For the optimization, neural network models of NOx and ISFC have been used. The outputs of the conventional calibration approach are 2-D calibration tables for EGR fraction, boost pressure and injection timings with engine speed and load as inputs.

The two-stage calibration approach for the inverse-distance interpolation based calibration method has also been implemented for doing the comparison study. For the inverse-distance approach, the input space for the set-point generation consists of engine speed, load, EGR fraction and boost pressure. For the first stage of the calibration, the calibration tables for EGR fraction and boost pressure generated from the conventional approach have been used to generate the set of operating conditions during a FTP drive cycle operation of the engine. These operating conditions have been used in the genetic algorithm based optimization such that the center of mass is minimized. Data sets containing 100, 200, 304, 500 and 1000 operating points have been generated from the genetic algorithm based optimization approach. In the second stage of the calibration, the injection timings at all the operating points generated in the
optimization approach have been optimized to minimize engine out NOx and ISFC. The output of the inverse-distance calibration approach is a set of calibrations scattered in the 4-D input space of the set-point generator.

Simulations of the FTP cycle have been carried out using the calibrations generated in conventional and inverse-distance based calibration methods. Linear interpolation has been used to interpolate the calibrations generated in the conventional approach and inverse-distance interpolation method has been used for the calibrations generated in the inverse-distance based calibration method. An ideal FTP case, where calibrations are available at each operating point of the FTP drive cycle, has been used as the baseline for comparing the performance and emissions between the conventional and inverse-distance methods. The inverse-distance set-point generation method resulted in reduced NOx spikes during transients of the engine when compared to the conventional method. This is because of the inclusion of EGR fraction and boost pressure in the input space of the inverse-distance based set-point generation method. Therefore, the FTP cycle-integrated NOx emissions of the inverse-distance method have reduced when compared to the conventional method. The fuel consumption results for the inverse-distance method were higher when compared to the conventional method. It was seen that, as the number of calibrations generated in the inverse-distance method increased, the NOx and fuel consumption of the inverse-distance method approached those of the ideal FTP case.

The calibration efforts required for the inverse-distance method have been compared to that of the conventional method with the help of some assumptions. The inverse-distance method has shown a high potential to reduce the calibration effort when compared to conventional calibration. The inverse-distance approach is a two-stage calibration approach. However, the calibration required for the first stage can be crude and can be derived from a conventional calibration approach with reduced number of levels in each of the input variables. A factor of reduction, $k$, has been defined to represent the degree of crudeness in the calibration when compared to a conventional method. Another parameter for the inverse-distance method is the number of calibrations generated from the optimization method, $M$. Depending on the values of $k$ and $M$, the reduction in the calibration effort for the inverse-distance method is of the order of $1/5^{th}$ to $1/25^{th}$ of
that required for the conventional calibration method. This is a significant improvement which can result in reduced engine development costs and duration for engine manufacturers.

The inverse-distance interpolation method has shown considerable promise in simulation as a method for generating the set-points of the closed-loop combustion controller. The method would be useful to reduce the calibration effort for a full-fledged closed-loop combustion controller for which conventional calibration would be highly impractical. The next logical step for the work is to validate the method in an experimental setup of the engine. The implementation of the inverse-distance interpolation algorithm on engine control unit may be a challenging task. The challenging part of such an implementation would be to implement the search algorithm to search for calibrations in the neighborhood. In this work, the effect of parameters of the inverse-distance method, like the neighborhood bounds, on engine performance and emissions has not been studied. The ideal way of finding the appropriate parameters for the inverse-distance method is to tie the genetic algorithm optimization to engine performance and emission models and let the algorithm choose the parameters and the corresponding calibrations based on engine performance and emissions. The center of mass metric can be used as a constraint or a penalty in the optimization. However, such an optimization problem would be complex to formulate and the validity of the calibrations generated by such a process depends on the accuracy of the models used to predict engine performance and emissions.
BIBLIOGRAPHY


### NOMENCLATURE

#### ACRONYMS

<table>
<thead>
<tr>
<th>ACRONYM</th>
<th>FULL FORM</th>
</tr>
</thead>
<tbody>
<tr>
<td>AFR</td>
<td>Air Fuel Ratio</td>
</tr>
<tr>
<td>ATDC</td>
<td>After Top-Dead Center</td>
</tr>
<tr>
<td>BSFC</td>
<td>Brake Specific Fuel Consumption</td>
</tr>
<tr>
<td>CAD</td>
<td>Crank Angle Degree</td>
</tr>
<tr>
<td>CAx</td>
<td>Crank Angle at x% of heat release</td>
</tr>
<tr>
<td>CIDI</td>
<td>Compression Ignition Direct Injection</td>
</tr>
<tr>
<td>CLCC</td>
<td>Closed Loop Combustion Controller</td>
</tr>
<tr>
<td>CRDI</td>
<td>Common Rail Diesel Injection</td>
</tr>
<tr>
<td>DOE</td>
<td>Design of Experiments</td>
</tr>
<tr>
<td>EGR</td>
<td>Exhaust Gas Recirculation</td>
</tr>
<tr>
<td>EVC</td>
<td>Exhaust Valve Closing</td>
</tr>
<tr>
<td>EVO</td>
<td>Exhaust Valve Opening</td>
</tr>
<tr>
<td>FTP</td>
<td>Federal Test Procedure</td>
</tr>
<tr>
<td>HCCI</td>
<td>Homogenous Charge Compression Ignition</td>
</tr>
<tr>
<td>IMEP</td>
<td>Indicated Mean Effective Pressure</td>
</tr>
<tr>
<td>ISFC</td>
<td>Indicated Specific Fuel Consumption</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Description</td>
</tr>
<tr>
<td>--------------</td>
<td>--------------------------------------------------</td>
</tr>
<tr>
<td>IVC</td>
<td>Intake Valve Closing</td>
</tr>
<tr>
<td>LTC</td>
<td>Low Temperature Combustion</td>
</tr>
<tr>
<td>MAF</td>
<td>Manifold Air Flow</td>
</tr>
<tr>
<td>MFB</td>
<td>Mass Fraction Burnt</td>
</tr>
<tr>
<td>MIMO</td>
<td>Multi-Input-Multi-Output</td>
</tr>
<tr>
<td>NOx</td>
<td>Nitrogen Oxides (NO + NO$_2$)</td>
</tr>
<tr>
<td>PCCI</td>
<td>Pre-mixed Charge Compression Ignition</td>
</tr>
<tr>
<td>PM</td>
<td>Particulate Matter</td>
</tr>
<tr>
<td>PI</td>
<td>Proportional-Integral (controller)</td>
</tr>
<tr>
<td>PID</td>
<td>Proportional-Integral-Derivative (controller)</td>
</tr>
<tr>
<td>SI</td>
<td>Spark Ignition</td>
</tr>
<tr>
<td>SISO</td>
<td>Single-Input-Single-Output</td>
</tr>
<tr>
<td>SOI</td>
<td>Start of Injection</td>
</tr>
<tr>
<td>TDC</td>
<td>Top Dead Center</td>
</tr>
<tr>
<td>VVA</td>
<td>Variable Valve Actuation</td>
</tr>
</tbody>
</table>