INNOVATIONS IN REPRESENTATION AND CALIBRATION OF RESIDUAL GAS FRACTION AND VOLUMETRIC EFFICIENCY IN A SPARK IGNITED, INTERNAL COMBUSTION ENGINE

A Thesis

Presented in Partial Fulfillment of the Requirements for

the Degree Master of Science

in the Graduate School of The Ohio State University

By

Jason Meyer

****

The Ohio State University
2008

Master’s Examination Committee: Approved by

Professor Yann Guezennec, Advisor
Professor Steve Yurkovich

____________________________
Advisor
Graduate Program in Mechanical Engineering
ABSTRACT

Accurate fuel control demands accurate predictions of fresh air mass. To meet this challenge, automobile manufacturers employ fresh-air prediction algorithms that are usually based on accurate volumetric efficiency (VE) maps. Calibrating these VE maps requires dense sets of experimental data that are costly to acquire. Moreover, new actuation technologies often add degrees of freedom to the air calibration problem, greatly increasing the burden placed on both physical dynamometer experiments as well as vehicle data collection. A data-rich engine mapping for a 4 degree of freedom engine (for example a conventional engine with independently variable intake and exhaust valve timing) can involve hundreds of thousands of experimental data points. This thesis offers a solution to this problem: use one-dimensional (1-D) engine simulation software together with intelligent sub-sampling of physical data to develop highly accurate simulators of the gas charge (fresh air and residuals) over the engine’s entire operating range. These simulators can then be exercised to generate the data-rich engine mapping that the VE and residual mapping requires.
One-dimensional simulators, like GT-Power\textsuperscript{1}, are able to capture the effect of one-dimensional wave dynamics throughout the intake and exhaust systems. Models with this level of detail are well suited for this calibration problem, because they have sufficient accuracy and execute quickly enough to be practical. This is in contrast to three- or four-dimension computational-fluid dynamics programs, which require considerably longer computation time. In addition to predicting trapped fresh air charge, one-dimensional simulators are also able to predict the mass of trapped residuals (for engines without external exhaust gas recirculation) because they are governed by the same wave dynamics and fluid dynamics as fresh air. In fact, a model designed to target an accurate volumetric efficiency prediction will also generate accurate estimates of the residual gas fraction. Even though it cannot be directly measured, the residual gas fraction is important because it influences combustion and spark timing requirements. This thesis will discuss a validated procedure of generating virtual VE and residual gas fraction maps.

The second focus of the thesis is the representation of this information. Because VE and residual gas fraction must be predicted every engine cycle for every cylinder, the engine control unit (ECU) constantly references this information. Storing these parameters which are dependant on the operating conditions is nontrivial. Both the computational complexity and the memory storage space are limited in an ECU. Therefore, the types of models that can be implemented in production vehicles are limited. It must be computationally simple, have a minimum number of calibration parameters and yet model the data with high accuracy. New methods of representing

\textsuperscript{1} GT-Power is a product in the GT-Suite software package and is trademarked by Gamma Technologies
both the volumetric efficiency and the residual gas fraction which meet the ECU limitations and have superior accuracy compared to the current practices are presented in this thesis.
DEDICATION

I dedicate this thesis to my parents
ACKNOWLEDGMENTS

I would like to thank my two advisors Dr. Guezennec and Dr. Yurkovich for providing guidance and support throughout my research. The insight and powertrain knowledge provided by Dr. Dudek and Layne Wiggins has been paramount to the success of my research and is appreciated. I would also like to thank fellow graduate students Sai Rajagopalan and Yiran Hu for their help. They have been my main teachers of optimization and abstract mathematical concepts. The staff at CAR especially Dr. Midlam-Mohler has been invaluable and I thank them.
VITA

May 19, 1985. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . Born – Centerville, OH

March 2007. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . BS Mechanical Engineering,
The Ohio State University.

April 2007 to Present. . . . . . . . . . . . . . . . . . . . . . . . Research Assistant
The Ohio State University.
Center for Automotive Research

FIELDS OF STUDY

Major Field: Mechanical Engineering
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Title</th>
<th>Page #</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABSTRACT</td>
<td>ii</td>
</tr>
<tr>
<td>DEDICATION</td>
<td>v</td>
</tr>
<tr>
<td>ACKNOWLEDGMENTS</td>
<td>vi</td>
</tr>
<tr>
<td>VITA</td>
<td>vii</td>
</tr>
<tr>
<td>LIST OF TABLES</td>
<td>x</td>
</tr>
<tr>
<td>LIST OF FIGURES</td>
<td>xi</td>
</tr>
<tr>
<td>LIST OF ABBREVIATIONS</td>
<td>xiii</td>
</tr>
<tr>
<td>CHAPTER 1</td>
<td>1</td>
</tr>
<tr>
<td>1.1 Introduction</td>
<td>1</td>
</tr>
<tr>
<td>CHAPTER 2</td>
<td>8</td>
</tr>
<tr>
<td>2.1 Cylinder Air Mass Predicting and Volumetric Efficiency</td>
<td>8</td>
</tr>
<tr>
<td>2.1.1 Steady-state Detection</td>
<td>10</td>
</tr>
<tr>
<td>2.1.2 Hardware Limitation and Robustness</td>
<td>10</td>
</tr>
<tr>
<td>2.2 Trapped Residual Gas Mass</td>
<td>11</td>
</tr>
<tr>
<td>2.2.1 Existing Physics-based Residual Gas Models</td>
<td>17</td>
</tr>
<tr>
<td>2.3 Engine Modeling Software</td>
<td>21</td>
</tr>
<tr>
<td>2.3.1 Design using Engine Models</td>
<td>21</td>
</tr>
<tr>
<td>2.3.2 Virtual Representations using Engine Models</td>
<td>22</td>
</tr>
<tr>
<td>2.4 GT-Power</td>
<td>23</td>
</tr>
<tr>
<td>2.4.1 How GT-Power Generates a Solution</td>
<td>25</td>
</tr>
<tr>
<td>2.5 Calibration Techniques</td>
<td>27</td>
</tr>
<tr>
<td>2.5.1 Least Squares</td>
<td>27</td>
</tr>
<tr>
<td>2.5.2 Stepwise Linear Regression</td>
<td>28</td>
</tr>
<tr>
<td>2.5.3 Gradient-based Nonlinear Optimization</td>
<td>29</td>
</tr>
<tr>
<td>2.5.4 Genetic Algorithms</td>
<td>30</td>
</tr>
<tr>
<td>2.6 Model Sensitivity</td>
<td>31</td>
</tr>
<tr>
<td>Table #:</td>
<td>Description</td>
</tr>
<tr>
<td>---------</td>
<td>-----------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Table 1</td>
<td>Distribution of VE and Distribution of VE Relative to Parked VE Table</td>
</tr>
<tr>
<td>Table 2</td>
<td>Comparison between Error of Engine Speed Warping Model with and without the Warping Term</td>
</tr>
<tr>
<td>Table 3</td>
<td>VE Error Comparison between Park Table Reference only and the Full Quadratic VE Function</td>
</tr>
<tr>
<td>Table 4</td>
<td>VE Error Comparison between Full Quadratic and Reduced Quadratic VE Functions</td>
</tr>
<tr>
<td>Table 5</td>
<td>VE Error of the VE Function Compared to Vehicle Data</td>
</tr>
<tr>
<td>Table 6</td>
<td>Volumetric Efficiency Comparison between Vehicle Data and Reduced Quadratic VE Function Calibrated from Vehicle Data</td>
</tr>
<tr>
<td>Table 7</td>
<td>Residual Gas Fraction Error Statistics for MIT Function</td>
</tr>
<tr>
<td>Table 8</td>
<td>Residual Gas Fraction Error Statistics for Senecal Function</td>
</tr>
<tr>
<td>Table 9</td>
<td>Error Statistics for Inverse MAP Residual Gas Model</td>
</tr>
<tr>
<td>Table 10</td>
<td>Effect of Barometric Pressure on the Inverse MAP Residual Gas Fraction Model</td>
</tr>
</tbody>
</table>
**LIST OF FIGURES**

<table>
<thead>
<tr>
<th>Figure #:</th>
<th>Description</th>
<th>Page #</th>
</tr>
</thead>
<tbody>
<tr>
<td>Figure 1</td>
<td>Overlap Factor</td>
<td>18</td>
</tr>
<tr>
<td>Figure 2</td>
<td>Object-Oriented Map of GT-Power Model</td>
<td>25</td>
</tr>
<tr>
<td>Figure 3</td>
<td>Illustration of Component Discritization [17]</td>
<td>26</td>
</tr>
<tr>
<td>Figure 4</td>
<td>Definition of Sensitivity to Defined Parameter</td>
<td>33</td>
</tr>
<tr>
<td>Figure 5</td>
<td>Example Sensitivity Study Showing the Effect of Wall Temperature on VE</td>
<td>35</td>
</tr>
<tr>
<td>Figure 6</td>
<td>Randomly Sampled, Space-filling Dataset to be used for the Model Calibration</td>
<td>39</td>
</tr>
<tr>
<td>Figure 7</td>
<td>Double Wiebe Function Correlation to a Two-Stage Burn</td>
<td>43</td>
</tr>
<tr>
<td>Figure 8</td>
<td>Variation in Mass Fraction Burned Curves</td>
<td>45</td>
</tr>
<tr>
<td>Figure 9</td>
<td>Representative Mass Fraction Burned Curve Comparison</td>
<td>45</td>
</tr>
<tr>
<td>Figure 10</td>
<td>In-cylinder Pressure Validation of Combustion Fitting from Calibrated GT-Power Model</td>
<td>48</td>
</tr>
<tr>
<td>Figure 11</td>
<td>Histogram of MAP errors for Calibrated GT-Power Model</td>
<td>49</td>
</tr>
<tr>
<td>Figure 12</td>
<td>Histogram of MAF errors for Calibrated GT-Power Model</td>
<td>49</td>
</tr>
<tr>
<td>Figure 13</td>
<td>Histogram of VE errors for Calibrated GT-Power Model</td>
<td>49</td>
</tr>
<tr>
<td>Figure 14</td>
<td>MAP, Engine Speed Space Reached in GT-Power Data</td>
<td>51</td>
</tr>
<tr>
<td>Figure 15</td>
<td>MAP, Engine Speed Space Reached in Vehicle Over an FTP Cycle</td>
<td>51</td>
</tr>
<tr>
<td>Figure 16</td>
<td>Percentage Burned Mass Variation in GT-Power Data</td>
<td>53</td>
</tr>
<tr>
<td>Figure 17</td>
<td>Burn Duration Variation in Experimental Data</td>
<td>53</td>
</tr>
<tr>
<td>Figure 18</td>
<td>Burn Duration Variation in GT-Power Data</td>
<td>53</td>
</tr>
<tr>
<td>Figure 19</td>
<td>VE Surface for Parked Cams</td>
<td>54</td>
</tr>
<tr>
<td>Figure 20</td>
<td>VE Surface for Cams at (25 cam degrees, 25 cam degrees)</td>
<td>54</td>
</tr>
<tr>
<td>Figure 21</td>
<td>Residual Gas Fraction Surface for Parked Cams</td>
<td>54</td>
</tr>
<tr>
<td>Figure 22</td>
<td>Residual Gas Fraction Surface for Cams at (25 cam degrees, 25 cam degrees)</td>
<td>54</td>
</tr>
<tr>
<td>Figure 23</td>
<td>Illustration of Wave Dynamics in a Helmholtz Resonator</td>
<td>59</td>
</tr>
<tr>
<td>Figure 24</td>
<td>Normalized VE versus Engine Speed Curves for all Manifold Pressures</td>
<td>60</td>
</tr>
<tr>
<td>Figure 25</td>
<td>Effect of Cam Timing on VE (MAP = 85 kPa)</td>
<td>62</td>
</tr>
<tr>
<td>Figure 26</td>
<td>Effect of Cam Timing on Location of VE Peaks</td>
<td>62</td>
</tr>
<tr>
<td>Figure 27</td>
<td>Regression on the Location of the Peak VE</td>
<td>66</td>
</tr>
<tr>
<td>Figure 28</td>
<td>VE Error Surface (RPM 2400 and a MAP of 0.85 bar)</td>
<td>68</td>
</tr>
<tr>
<td>Figure 29</td>
<td>Error in VE as a Function of Cam Positions for All Engine Speeds (MAP = 0.49 bar)</td>
<td>68</td>
</tr>
</tbody>
</table>
Figure 30: VE Error for Full Quadratic Model after Optimization .................................. 73
Figure 31: VE Error at Parked for Full Quadratic Model after Optimization .............. 73
Figure 32: VE Error for Reduced Quadratic Model after Optimization ...................... 75
Figure 33: VE Error at Parked for Reduced Quadratic Model after Optimization ....... 75
Figure 34: Steady-state Operating Points for the Validation Dataset ............................ 79
Figure 35: Comparison between the VE Measured On-Vehicle and the Predicted VE from the VE Function (Drive Trace 2) .............................................................. 79
Figure 36: VE Error in Validating the VE Function with Vehicle Data (20 Coefficients) .......................................................... 80
Figure 37: VE Surface at Parked Generated from Measured Vehicle Data ............... 83
Figure 38: VE Distribution for Measured Vehicle VE .................................................. 85
Figure 39: Difference between Measured VE and Reference Parked VE .................. 85
Figure 40: VE Error Distribution for the Reduced Quadratic VE Function ............... 85
Figure 41: Residual Gas Fraction Error for Optimized MIT Model ............................ 89
Figure 42: Residual Gas Fraction Error for Optimized Senecal Model ....................... 90
Figure 43: Projection of Engine Speed’s Influence on Residual Gas Fraction onto the MAP, Residual Gas Fraction Plane .......................................................... 93
Figure 44: Error Distribution for Inverse MAP Residual Gas Function ..................... 96
Figure 45: Residual Gas Fraction Error Distribution for the Calibration Region (99 kPa) .......................................................... 97
Figure 46: Residual Gas Fraction Error Distribution without Barometric Correction Factor (80 kPa) .......................................................... 98
Figure 47: Residual Gas Fraction Error Distribution with Barometric Correction Factor (80 kPa) .......................................................... 98
Figure 48: Residual Gas Fraction Error Distribution without Barometric Correction Factor (60 kPa) .......................................................... 99
Figure 49: Residual Gas Fraction Error Distribution with Barometric Correction Factor (60 kPa) .......................................................... 99
Figure 50: Residual Gas Fraction Predicted by GT-Power (73 kPa, 4600 RPM) ........ 101
Figure 51: Residual Gas Fraction Predicted by MIT Model (73 kPa, 4600 RPM) ....... 102
Figure 52: Residual Gas Fraction Predicted by Inverse MAP Model (73 kPa, 4600 RPM) .......................................................... 102
Figure 53: Residual Gas Fraction Error for the MIT Model ....................................... 103
Figure 54: Residual Gas Fraction Error for the Inverse MAP Model ......................... 103
<table>
<thead>
<tr>
<th>Symbol</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\beta$</td>
<td>Constant for Senecal, Xin, and Reitz Equation</td>
</tr>
<tr>
<td>$\omega$</td>
<td>Constant for Senecal, Xin, and Reitz Equation</td>
</tr>
<tr>
<td>$\theta$</td>
<td>Crank Angle</td>
</tr>
<tr>
<td>$\gamma$</td>
<td>Specific Heat Ratio</td>
</tr>
<tr>
<td>$\rho_a$</td>
<td>Ambient Density</td>
</tr>
<tr>
<td>$\rho_m$</td>
<td>Manifold Density</td>
</tr>
<tr>
<td>$\eta_v$</td>
<td>Volumetric Efficiency</td>
</tr>
<tr>
<td>$\Phi$</td>
<td>Equivalence Ratio</td>
</tr>
<tr>
<td>$a$</td>
<td>Mass of Fuel Burned at Constant Volume</td>
</tr>
<tr>
<td>AFR</td>
<td>Air-to-Fuel Ratio</td>
</tr>
<tr>
<td>CAD</td>
<td>Computer Aided Drafting</td>
</tr>
<tr>
<td>CFD</td>
<td>Computational Fluid Dynamics</td>
</tr>
<tr>
<td>COV</td>
<td>Coefficient of Variation</td>
</tr>
<tr>
<td>$C_p$</td>
<td>Constant Pressure Specific Heat</td>
</tr>
<tr>
<td>$C_v$</td>
<td>Constant Volume Specific Heat</td>
</tr>
<tr>
<td>$D_e$</td>
<td>Exhaust Valve Diameter</td>
</tr>
<tr>
<td>$D_i$</td>
<td>Intake Valve Diameter</td>
</tr>
<tr>
<td>Acronym</td>
<td>Definition</td>
</tr>
<tr>
<td>---------</td>
<td>----------------------------------------------</td>
</tr>
<tr>
<td>ECAM</td>
<td>Exhaust Cam Timing</td>
</tr>
<tr>
<td>ECU</td>
<td>Engine Control Unit</td>
</tr>
<tr>
<td>EGR</td>
<td>Exhaust Gas Recirculation</td>
</tr>
<tr>
<td>EVC</td>
<td>Crank Angle at Exhaust Valve Closing</td>
</tr>
<tr>
<td>FEA</td>
<td>Finite Element Analysis</td>
</tr>
<tr>
<td>HEGO</td>
<td>Heated Exhaust Gas Sensor</td>
</tr>
<tr>
<td>ICAM</td>
<td>Intake Cam Timing</td>
</tr>
<tr>
<td>IMEP</td>
<td>Indicated Mean Effective Pressure</td>
</tr>
<tr>
<td>IV=EV</td>
<td>Crank Angle when Intake and Exhaust Valve Lifts are Equal</td>
</tr>
<tr>
<td>IVO</td>
<td>Crank Angle at Intake Valve Opening</td>
</tr>
<tr>
<td>Le</td>
<td>Exhaust Valve Lift</td>
</tr>
<tr>
<td>Li</td>
<td>Intake Valve Lift</td>
</tr>
<tr>
<td>m</td>
<td>Total Mass</td>
</tr>
<tr>
<td>ma</td>
<td>Mass of Trapped Air</td>
</tr>
<tr>
<td>(\dot{m}_a)</td>
<td>Mass Flow Rate of Air</td>
</tr>
<tr>
<td>MAF</td>
<td>Mass Air Flow</td>
</tr>
<tr>
<td>MAP</td>
<td>Manifold Pressure</td>
</tr>
<tr>
<td>mf</td>
<td>Mass of Fuel</td>
</tr>
<tr>
<td>MVM</td>
<td>Mean Value Model</td>
</tr>
<tr>
<td>N</td>
<td>Engine Speed</td>
</tr>
<tr>
<td>NOx</td>
<td>Nitrogen Oxides</td>
</tr>
<tr>
<td>OF</td>
<td>Overlap Factor</td>
</tr>
<tr>
<td>Symbol</td>
<td>Description</td>
</tr>
<tr>
<td>--------</td>
<td>----------------------------------</td>
</tr>
<tr>
<td>( p_e )</td>
<td>Exhaust Pressure</td>
</tr>
<tr>
<td>( p_i )</td>
<td>Intake Pressure</td>
</tr>
<tr>
<td>( \rho_m )</td>
<td>Manifold Pressure</td>
</tr>
<tr>
<td>( Q_{hv} )</td>
<td>Lower Heating Value of the Fuel</td>
</tr>
<tr>
<td>( R )</td>
<td>Universal Gas Constant</td>
</tr>
<tr>
<td>( r_c )</td>
<td>Compression Ratio</td>
</tr>
<tr>
<td>RPM</td>
<td>Engine Speed</td>
</tr>
<tr>
<td>SI</td>
<td>Spark Ignition</td>
</tr>
<tr>
<td>( T_m )</td>
<td>Manifold Temperature</td>
</tr>
<tr>
<td>UEGO</td>
<td>Universal Exhaust Gas Sensor</td>
</tr>
<tr>
<td>( V_d )</td>
<td>Displacement Volume</td>
</tr>
<tr>
<td>VE</td>
<td>Volumetric Efficiency</td>
</tr>
<tr>
<td>( V_{Em} )</td>
<td>Manifold Volumetric Efficiency</td>
</tr>
<tr>
<td>WOT</td>
<td>Wide Open Throttle</td>
</tr>
<tr>
<td>( x_r )</td>
<td>Residual Gas Fraction</td>
</tr>
<tr>
<td>( Z_e )</td>
<td>Number of Exhaust Valves</td>
</tr>
<tr>
<td>( Z_i )</td>
<td>Number of Intake Valves</td>
</tr>
</tbody>
</table>
CHAPTER 1

INTRODUCTION

1.1 Introduction

To operate an engine efficiently and produce minimal exhaust pollutants, the air-to-fuel ratio (AFR) has to be very near stoichiometry, 14.6 for a gasoline-fueled engine. Accurately controlling the AFR is very difficult because it involves the interactions of several subsystems. This process centers on estimating the trapped air mass per cylinder. Without an accurate estimate of the air mass, the injected fuel mass based on this estimate could produce an AFR considerably different than the desired ratio. When the intake valves open, the pressure difference between the intake manifold and engine cylinder causes air to flow into the cylinder. As air fills the cylinder and the cylinder volume changes so does the mass flow rate of air. Before an equilibrium air mass can be reached, the intake valves close. The air trapping process is therefore inherently a dynamic process. In addition to these air dynamics, the conditions of the intake manifold can change. When the throttle position is changed, the manifold pressure also changes which adds additional complexity to the air trapping dynamics.

Although the dynamics of air induction are more recognizable, the fuel delivery system also has dynamics. For a port fuel injected engine, only part of the fuel is directly
vaporized. The other portion of fuel collects and forms a puddle on the intake walls. The rate of evaporation of this fuel puddle depends on several factors. It is even possible during strong transients for some liquid fuel to enter the cylinder. Estimating the fuel mass that enters the cylinders requires an estimate of the fuel puddle mass, the evaporation rate and the injected fuel vapor fraction. The leading fuel dynamics model [1] or a similar model which accounts for these phenomena is used in most AFR control models [2, 3, 4].

Clearly both the air and fuel mass estimate problems have uncertainties. To account for and eliminate the effect of these uncertainties, the AFR controller receives feedback from the exhaust gases. Most production vehicles use an oxygen sensor to provide information about the air-to-fuel ratio. Historically heated exhaust gas oxygen (HEGO) sensors which can only output a binary signal were used. Recently universal exhaust gas oxygen (UEGO) sensors which have a linear output have been implemented in some production vehicles. As the name implies, an EGO sensor (heated or universal) responds to the net concentration of oxygen in the exhaust stream. If the concentration is high, then during combustion there was an excess of air (lean AFR). Conversely, a low concentration means that all of the air was used during combustion and the combustion mixture was rich. A HEGO sensor’s output switches near an oxygen concentration corresponding to stoichiometric combustion.

A UEGO sensor is actually a collection of binary HEGO sensors which are modified to providing a piecewise linear response region. Instead of comparing the oxygen concentration of the exhaust stream to ambient air as in a HEGO, a UEGO compare the oxygen concentration of the exhaust stream to a closed chamber. Oxygen is
pumped into the chamber by the UEGO to maintain equilibrium. The amount of oxygen pumped into the chamber is proportional to the AFR.

Regardless of whether a HEGO or UEGO is used, these sensors have limitations. A HEGO sensor is usually used only to indicate if the AFR is rich or lean but not by how much. Therefore, large excursions from stoichiometry are difficult to correct. The exhaust system also has an inherent transport delay. Multiple engine events pass between combustion and when the exhaust gases reach the oxygen sensor. An oxygen sensor also has finite response time. During transient engine operation this transport/sensor delay reduces the effectiveness of the oxygen sensor feedback. Furthermore, this combined delay is not fixed but depends on the operating conditions. Another major issue is the operating temperature range of an oxygen sensor. An oxygen sensor must reach about 600 degrees Fahrenheit before it can be used. During a cold start the ECU can only use an open-loop estimate of the air and fuel masses until the oxygen sensor reaches its operating temperature through its internal heater and heat transfer from the exhaust gases. Most of the tailpipe emissions generated by an engine occur during this period.

For these reasons having an accurate trapped air mass estimate is important. Modern AFR control is a combination of three components. The most important and baseline component is the feed-forward trapped air mass estimate. This is a static value which corresponds to the air mass that would be trapped under cycle-to-cycle steady-state. This open loop estimated is modified by a compensator which accounts for air dynamics. This air estimated is then dynamically modified by feedback from an oxygen sensor.
To predict the trapped air mass, the standard automotive industry practice is to use the non-dimensional quantity volumetric efficiency. Volumetric efficiency is the ratio of the actual mass of air inducted to theoretical maximum mass of air that could be inducted at the same conditions at steady-state. Because an engine’s intake valves are not open long enough to reach equilibrium, the actual trapped mass of air is normally less than the equilibrium value. Through wave dynamics tuning, the volumetric efficiency of an operating condition can be increased. With this type of tuning it is possible to have a volumetric efficiency greater than one hundred percent. Having a high volumetric efficiency is obviously beneficial because more air is inducted. Varying an engine’s operating conditions such as engine speed, throttle position and cam timing changes the volumetric efficiency. To map out the relationship between these parameters and volumetric efficiency, experimental data must be taken.

Advances in actuation technologies for modern SI engines have dramatically increased the complexity of mapping engines due to the large and growing number of degrees of freedom afforded by these new technologies. These technologies include cam phasing, intake manifold tuning devices and multi-staged cam profiles. This thesis describes a method that addresses these issues by relying less on experiments and more on numerical simulations. Several engine simulation software packages are commercially available. These programs are capable of modeling the complete powertrain of a vehicle. This is done with heat transfer models, combustion models, friction models and one-dimension fluid dynamics models. The one-dimensional fluid dynamics models are able to capture the complex wave dynamics occurring in the main flow direction. Fluid motion in the other two auxiliary directions is not modeled but
represented as lumped losses. The first focus of this thesis is to demonstrate the
 calibration procedures required to accurately calibrate a model in one such program, GT-
 Power. The calibration process is tailored for accurate estimation of the VE and residual
gas fraction. As a prototypical example of the method, this work focused on a dual-
independent cam phased engine. To learn more about the benefits and actuator control
strategies of this technology consult [5, 6].

The modeling scheme involves the calibration of an engine model using a sparse
(several hundred points) but space-filling experimental data set in the four-dimensional
operating plane (MAP, RPM, intake and exhaust cam phasing). The procedure focuses
on a rapid and robust calibration of a few key elements of the model using such an
experimental data set to yield accurate air estimation. In particular the procedure is
capable of adequately capturing (without explicitly modeling) the significant changes in
temperature profiles of the intake manifold and the combustion process due to trapped
residual changes resulting from aggressive cam phasing. The model is then extensively
used to simulate a fine map of volumetric efficiency and residual gas fraction over the
engine’s four-dimensional operating range (nearly 200,000 operating points). Instead of
using experimental data, this simulated data can be used to produce the volumetric
efficiency and residual gas fraction maps that are commonly employed in fuel and spark
control. The methodology was implemented and validated on a current production 2.4L
engine, and the results are presented in the thesis. This methodology is applicable to all
engine platforms and tremendously reduces the experimental burden of mapping engines
in a high dimensional actuator space.
Similar procedures have been studied and implemented for SI engines with fixed valvetrains [7]. Modeling with GT-Power has also been applied to more advanced engines but with different calibration procedures and target outputs. A less detailed and systematic application of the process described in this thesis can be found in [8], although the purpose of using such an approach was to calibrate mean value engine models (MVM). In another example, Delphi developed an iterative procedure that uses a GT-Power model to optimize the scheduling of cam lift and timing based on fuel economy and emissions [9].

Following the process descriptions of calibrating and validating a GT-Power model, new representations for both VE and residual gas fraction are presented. Because GT-Power is able to replicate any operating condition quickly and accurately, the complete operating space was simulated. Using the data predicted by GT-Power, new VE and residual gas fraction representations were generated and calibrated. These new models are simple enough to be implemented in a production vehicle but have accuracy better than current practice. An extensive simultaneous validation of the GT-Power model and VE representation was performed using both engine dynamometer data and vehicle data. These models are also robust enough to be applied to any engine platform.

By following the procedures outlined in this thesis, an engine model can be calibrated using a relatively small set of data collected on an engine dynamometer. This data is used to calibrate only the most sensitive components (with respect to air estimation) of a GT-Power model. The model for each component is validated individually and then the complete engine model is validated using a small subset of the original dynamometer data. Once validated, the model can be used to simulate the
complete operating space. The predicted VE and residual gas fraction from the model can then be used to calibrate the new VE and residual gas representations. Even with respect to the engine in a vehicle, the resulting models have high accuracy. As a result quality representations of VE and residual gas fraction can be generated with a significantly shortened calibration process. Compared to current industry practice, this process requires two orders of magnitude less experimental data and most of the process can be automated.
CHAPTER 2

BACKGROUND

2.1 Cylinder Air Mass Predicting and Volumetric Efficiency

As previously stated estimating the trapped cylinder air mass per cycle is essential for precise AFR control. However, there are no sensors capable of directly measuring the cylinder air mass, so it must be estimated. During steady-state the same mass of air is trapped every engine cycle. Obviously an engine is never truly in steady-state especially not when the intake or exhaust valves are moving. Having the exact same fluid dynamics response from cycle-to-cycle is defined as steady-state for an engine. In steady-state the measurement from a mass air flow sensor can be divided by the engine speed to get the trapped air mass per cycle. In transient conditions the manifold conditions change. The dynamics of the filling and emptying of the intake manifold resulting from air entering the engine’s cylinders also changes. These additional dynamics can interfere with a MAF sensor. Thus under transient conditions, the air per cylinder estimate from a MAF sensor is relatively poor. Therefore other sensors or methods must be used to provide an estimate. The most common method is based on the non-dimensional value called volumetric efficiency. Whether the focus is mean-value-models (MVM) [10], AFR
control [11, 2, 3] or exhaust speciation prediction [12], VE is used throughout the automotive industry.

Volumetric efficiency is the ratio of the mass of the air inducted by the engine to the maximum possible mass of air inducted with respect to conditions at some reference location. Using the ideal gas law and a mass air flow sensor, the maximum cylinder mass can be calculated in steady-state. Because volumetric efficiency is non-dimensional, it has the same value for any air temperature. Clearly temperature has an effect on the maximum mass of air that can be inducted into an engine. However, the ratio of the actual air mass to the maximum air mass is constant. The effect of manifold pressure is not as direct as temperature. Manifold pressure not only influences the maximum possible trapped air mass but also the fluid dynamics. When the manifold pressure is increased, the pressure difference between the combustion chamber and the manifold also increases. The mass flow rate of air into the cylinders increases and the flow patterns changes producing a different volumetric efficiency. Therefore volumetric efficiency must be found as a function of the operating conditions (including manifold pressure). Once the volumetric efficiency is known for a given operating condition, the cylinder air mass can then be estimated. Again the ideal gas law can be used to calculate the maximum possible cylinder air mass. The product of this maximum value and the volumetric efficiency is the predicted cylinder air mass.

Even though volumetric efficiency is a steady-state concept, it can be used to predict the air mass during transients. Although the cam positions, manifold pressure or engine speed may be changing, they are nearly constant with respect to a single engine cycle. The propagation of waves in the intake system is considerable faster that any of
the operating conditions can change. The trapped mass predicted for an operating condition under transient operating is therefore close to the actual mass.

2.1.1 Steady-state Detection

Because a mass air flow sensor is only accurate under steady conditions, detecting steady-state is very important both in data collection and in the final engine controller. Incorrectly interpreting a data point as steady-state can cause large errors. A mass air flow sensor is only accurate in steady-state so the volumetric efficiency calculated from this measurement could have a relatively large error during transients. In steady-state the measurement from a mass air flow sensor provides an accurate air estimate. It may be necessary to compare the air estimate from the mass air flow sensor to the prediction from the volumetric efficiency calculation. This comparison could be used for diagnostics or to obtain a dynamic correction factor. Hence, having an algorithm that accurately detects steady-state is very important. It is impossible to measure the flow patterns from cycle-to-cycle and compare them. Instead, steady-state detection is based on the assumption that if the operating conditions do not change then the fluid dynamics should remain constant from cycle-to-cycle. When the operating conditions vary less than some threshold, the engine is assumed to be in steady-state.

2.1.2 Hardware Limitation and Robustness

The biggest constraint in implementing a volumetric efficiency based air estimate comes from the ECU. Both the memory and processing power of the ECU are limited. Volumetric efficiency depends on the operating conditions. If the volumetric efficiency
was stored directly using a table, then a massive amount of storage space would be required. For example, an engine with four degrees of freedom such as the one used in this study would require a table of $33 \times 30 \times 14 \times 14$ dimensions. This size of table may not be necessary, but it is difficult to systematically reduce the table’s size. Reducing the size of the table would decrease the accuracy and would still exceed the memory limitation of an ECU. A solution which requires fewer variables to be stored is necessary, yet the VE must be represented as a function of the operating conditions. However, the calculations involved in function must not require a lot processing power. Addition and multiplication requires much less calculation resources than division and exponents do.

The VE representation must also be robust enough to be applied to any engine. This includes turbocharged and supercharged applications as well as cam phasing, multistaged cam platforms and powertrains with intake tuning valves. Each one of these systems adds complexity and presents unique modeling difficulties. Instead of an operation condition defined by four variables, it may take 5, 6 or even more in future applications. Representing a system defined by this many independent variables directly with a table is unfeasible even if an ECU’s memory storage capacity was expanded.

2.2 Trapped Residual Gas Mass

One of the hardest measurements to take and to predict is the residual gas mass. Residual gases come from a combination of three sources. The most direct source is exhaust gas that is recirculated externally (external EGR). A small pipe with a control valve (EGR valve) connects the exhaust to the intake. By adjusting the position of the
valve, the amount of external EGR entering the intake can be controlled. Modern engines rely less and less on this method of controlling EGR. Another source of residual gases is a result of clearance volume and the exhaust valve timing. Once the exhaust valves are closed, a portion of the combustion gases will always remain. Depending on the exhaust timing, the cylinder volume at exhaust valve closing will change. This volume is always greater than the clearance volume. The mass of residual gases that cannot be expelled is a function of exhaust timing, exhaust temperature and exhaust pressure.

The final source of residual gases is due to reverse flow of exhaust gases through the intake ports. Intake and exhaust valves do not open and close instantaneously. Because they are opened through kinematic linkages to a cam shaft, they open and close gradually. The average valve is at least partially open for about 240 crank angle degrees. For most production engines, both the exhaust and intake valves will be open for a short time between the exhaust and intake strokes. Exhaust gases are very hot and at or above ambient pressure, whereas intake gases are normally cool and below ambient pressure. When both the intake and exhaust valves are open, a pressure gradient will cause exhaust gases to flow into the intake. This phenomenon is more pronounced as the valve overlap is increased. Consequently, engines with independent cam phasing mainly rely on this technique for controlling residual gas concentrations. Both the residual gases resulting from backflow and from external EGR are dependent on the intake manifold pressure and the exhaust manifold pressure. The external EGR is obviously also dependent on the EGR valve position.

Measuring the combined residual gases from each source is nearly impossible to measure, but they are very important for several reasons. Residual gases have already
burned and therefore will not react during combustion. When residual gases are combined with fresh or unburned air, the residuals gases act as a diluent which affects both the gas temperature and composition (most notably CO₂). Depending on operating conditions, this effect could be desirable or detrimental. The gas temperature dictates the fuel evaporation rate and chemical kinetics. While the dilution of the gas composition affects the heat rise during combustion. A more diluted mixture will have produce less NOₓ but will also decrease the speed of combustion. It is therefore very important to know and to be able to control the residual fraction.

During startup and at idle, combustion stability is paramount. Idle naturally has very high cycle-to-cycle variation due to low speed and weak turbulent mixing, so adding a diluent would increase that variation even more. If the residual gas fraction is too high, then the engine may produce incomplete combustion or even misfire. Compounding the problem, an engine is very susceptible to high residual gas fractions at idle. Because the engine speed at idle is so low, the gases in the engine have more time to reach equilibrium. The manifold pressure is also very low at idle. Therefore more exhaust gases flow back into the intake during valve overlap.

Under normal part load conditions, residual gases can raise the overall efficiency of an engine. As long as the residual gas fraction is not too large, complete combustion can still occur. The efficiency is increased because residual gases reduce the need to throttle. Instead of exclusively using the throttle to meet the demanded torque, a combination of the throttle and residual fraction can be used. Opening the throttle increases the intake manifold pressure and the amount of gas that the cylinder will induct. When residual gases are added, the throttle can be opened more. The manifold pressure
is higher (hence less pumping loses), but not all of the gas inside the manifold is air. Some of the gas is residual gas, so the total amount of unburned air inducted into the cylinder remains the same. By opening the throttle more, less fluid frictional losses are encountered.

Another major benefit of using residual gases is the reduction of harmful NO\textsubscript{x} emissions. The formation of NO\textsubscript{x} is strongly temperature dependent. Decreasing the peak combustion temperature a few hundred degrees can reduce NO\textsubscript{x} emission several times. After combustion the major chemical species are nitrogen gas, carbon dioxide and water. Carbon dioxide and water to a lesser extent have a larger heat capacity than air. For a given energy input of fuel, the combustion temperature is decreases as the residual content is increased. At very high temperatures, nitrogen gas reacts with the available oxides and form nitrogen oxides, NO\textsubscript{x}. Decreasing the combustion temperature strongly reduces the amount of thermal NO\textsubscript{x} that can form. Once NO\textsubscript{x} has formed it does not naturally decompose even after the combustion temperature decreases. Using a catalyst such as a three-way catalytic converter can reduce most of the NO\textsubscript{x} that has formed. However, a catalytic converter must reach an activation temperature before it works efficiency, and its efficiency may degrade over time. Therefore it is still important to reduce the engine out NO\textsubscript{x} emissions.

Finally, an estimate of the residual gas fraction is needed for proper spark timing. As mentioned above the residual gas fraction dilutes the fresh charge. For combustion to occur, the local air-to-fuel ratio must be near stoichiometric. When a diluent is present, the likelihood of having a sufficient air-to-fuel ratio around a given fuel molecule decreases. Therefore, increasing the residual gas fraction decreases the oxygen
concentration which reduces the burn rate and increases the burn duration. To produce the most work, half of the fuel mass should combust before top dead center and half should combust after top dead center. When combustion is symmetric to top dead center, the highest peak pressure will be achieved. Accordingly as the residual gas fraction is increased, the spark timing must be advanced to ensure that combustion starts early enough to generate a peak pressure at the right location. Maximum break torque is generally seen when the peak pressure occurs around 6-10 degrees after top dead center. When a more accurate residual gas fraction estimation is available to the spark controller, the engine will have better drivability and run more efficiently.

The largest difficulty of measuring the residual gas fraction comes from the reverse flow of exhaust gases. As previously stated, the amount of valve overlap strongly influences the residual gas fraction. For modern engines with cam phasing, the amount of valve overlap dynamically changes. Additionally because cam profiles are curved, the relationship between cam timing and valve overlap is nonlinear. This relationship has been studied but it is difficult to describe physically. Another problem is that the valves are moving so fast that equilibrium is never reached. In the crank angle domain, the intake and exhaust valves are simultaneously open for the same number of crank angle degrees for fixed cam timing regardless of the engine speed. However the motion of a fluid is governed by time domain relationships, so engine speed has a major effect. As engine speed is increased, there is less time for combustion gases to flow back through intake valves and therefore less combustion gas reaches the intake system.

Even if the exact mass of combustion gases that travel back through the intake valves could be predicted, the mass of this reverse flow that makes it back into the engine
must still be predicted. Although most of the residuals will reenter the cylinders, it is possible for a portion to remain in the intake system. To find the exact residual gas fraction, the concentration of the gases in the engine’s cylinders after intake valve closing must be measured. The gases will include oxygen, carbon dioxide, carbon monoxide, hydrocarbons, nitrogen, nitrogen oxides and others. If no residual gases were present, then the gas composition would almost completely be oxygen, nitrogen and fuel vapor. Therefore by comparing the cylinder’s composition to the composition of the intake gases and the exhaust gases, the residual fraction can be determined.

The implementation of this measurement system is certainly not trivial. An exhaust gas analyzer does not output the gas composition in real time. Compared to the period of an engine (tens of milliseconds), an analyzer has a very large time constant on the order of tens of seconds. Therefore synchronizing the intake, cylinder and exhaust compositions could be a problem. The largest problem comes from sampling the gases in the cylinder without altering the engine’s dynamics. Obviously physical modification of the combustion chamber will be required. Only the composition after the intake valve closes and before combustion is of interest. Therefore, a method for separating out this data is necessary. In a study performed by Albert [13], the head of an engine was modified to include a miniature valve. Approximately every 10 to 15 engine cycles the engine was intentionally misfired and the valve was opened. To collect enough data for the emission analyzer, many cycles of normal combustion followed by misfire were needed. This process is very complicated and costly. It is therefore impractical for an automobile manufacturer to perform.
2.2.1 Existing Physics-based Residual Gas Models

It would be beneficial to find a method of predicting the residual gas fraction without experimental data. Using physics, several predictive relationships have been generated. The first major study was performed at MIT by Fox, Cheng and Heywood [14]. It introduced a measure for valve overlap called the “overlap factor.” This factor was developed to quantify the volume flow rate of combustion gases that could reenter the intake system. The effective area of the intake and exhaust valves is very important. Because the valves are moving, the effective area varies as a function of the crank angle. Another consideration is that the exhaust valve area is decreasing and the intake valve area is increasing. Therefore the intake valve area is initially the largest flow restriction but it switches to the exhaust valve area as the exhaust valve is closing. All of these factors were combined to generate (1) shown below.

\[
OF = \frac{Z_i D_i \int_{IVO}^{WEV} L_i d\theta + Z_e D_e \int_{IVO}^{EVC} L_e d\theta}{V_d \int_{WEV}^{WEV} L_e d\theta}
\]

(1)

Figure 1 shows the variation in overlap factor as a function of cam timing for the engine studied in this thesis.
In this equation the opening of the intake valves is defined as the crank angle when the valve lift increases above 0.15 mm. Conversely, the exhaust valve closing is defined as the crank angle when the valve lift decreases below 0.15 mm. The MIT paper proposed the residual gas fraction relationship shown in (2).

\[
x_r = C_1 \left( \frac{p_e}{p_i} \right)^{2\gamma} \left( \frac{OF}{N} \right) \left( \frac{p_e - p_i}{p_a} \right) + C_2 \left( \frac{p_e}{p_i} \right) \left( \frac{1}{r_e} \right)^\gamma \Phi
\]

This equation was developed and validated over a relatively sparse operating range. Only two engine speeds, 900 and 1500 RPM, were tested and only two cam profiles (which both had relatively low valve overlap). The total number of data points was only 36. Even though the equation was largely based on physics concepts, it incorporated two calibration constants, C1 and C2. After tuning these constants the MIT paper was able to find a good correlation. However because the number of data points was so small, the
equation may not have had that high of accuracy when used outside of its calibration region. This same handful of points could probably have been fit to many other equations. Because the fluid dynamics of reverse flow is so complicated, it would be very difficult to completely describe directly using physics. Hence the MIT equation had two coefficients to account for the unmodeled physics such as wave dynamics.

Several other similar equations for the residual gas fraction have also been composed. Equation (3), proposed by Senecal, Xin, and Reitz [15], was intended to be an extension of the MIT equation to diesel engines.

\[
x_r = C_1 \left( \frac{RT_i}{p_e} \right)^{\frac{1}{2}} \left( \frac{p_e}{p_i} \right)^{\frac{\gamma - 1}{2\gamma}} \frac{r_c - 1}{r_c} \left( \frac{OF}{N} \right) \left( 1 + \beta \right)^{\frac{\gamma - 1}{\gamma}} \frac{p_e - p_i}{(1 + \beta + \omega)^{\frac{1}{2}}} + C_2 \left( \frac{p_e}{p_i} \right)^{\frac{1}{\gamma}} \left( \Phi^2 - .5295\Phi + .5295 \right) \frac{\left( 1 + \beta \right)^{\frac{\gamma - 1}{\gamma}}}{(1 + \beta + \omega)}
\]

(3)

\[
\beta = \frac{\alpha m_f Q_{hv}}{m C_p r_c^{\gamma - 1} T_i}
\]

(4)

\[
\omega = \frac{(1 - \alpha) m_f Q_{hv}}{m C_p r_c^{\gamma - 1} T_i}
\]

(5)

The backflow of combustion gases through the exhaust and intake valves does not depend on how the fuel is burned. However, the composition of the exhaust gases is different for diesels engines. Most of the time diesel engines are operated very lean, so the combustion gases can still contain a lot of unburned air. Another reason the equation was developed is because the MIT equation did not fit well enough. This is another indication that not all of the physics are captured, and the MIT equation cannot be blindly applied outside its calibration region. The Senecal, Xin, and Reitz equation added many
new terms and two new constants, (4) and (5). In the MIT equation all of the values were measurable or constant. For the Senecal, Xin, and Reitz equation, a new non-dynamically measurable quantity was introduced. This value designated by alpha represents the mass of fuel burned at constant volume. The implementation of this equation is therefore more difficult. It would require finding a relationship between the mass of fuel burned at constant volume and the operating conditions. Additionally because this equation has so many dynamically changing terms, the calculation time would be very long.

To implement the MIT or Senecal equations, multiple coefficients have to be calibrated. Therefore experimental residual gas fraction data is needed. Unless the coefficients for every type engine were exactly the same, experimental data would be needed for every engine. The cost of modifying an engine to capture residual gas fraction data is substantial. Another equation developed based on the MIT equation was designed so that experimental data would not be necessary [16]. In the MIT model not all of the physical principles were captured, so two calibration constants were used. Equation (6) developed by Siviero, Suglia and Cavina has essentially the same form as the previous models but an explicit relationship between engine characteristics and the calibration constant, (7), was constructed.
\[ x_r = \sqrt{\frac{1}{C} \pi \sqrt{\frac{2}{3}} \frac{r_c - 1}{r_c}} \left( \frac{OF}{N} \right) \left( \frac{RT_i}{p_e} \right) \left( \frac{p_c}{p_f} \right)^{\frac{k+1}{2k}} + \frac{1}{C} \frac{r_c - 1}{r_c} \phi V_{W0} \left( \frac{p_e}{p_i} \right)^{\frac{1}{k}} \] (6)

\[ C = \left[ 1 + \frac{LHV}{c_s \epsilon T \left( \frac{m_{cyl}}{m_f} \right) r_c^{k-1}} \right]^{\frac{1}{k}} \] (7)

Although this model requires no calibration, it is not a complete first principle explanation of the process. Many assumptions are made to arrive at the final form.

2.3 Engine Modeling Software

Computer-aided engine modeling is becoming very widely used throughout the automotive industry and research community. Computers are so powerful and inexpensive that engine modeling can provide a substantial saving of both time and money. Although an engine model is a simplified approximation of a real engine, they can be calibrated to achieve the desired accuracy. Most commercially available engine modeling software is based on one-dimensional fluid dynamics equations. They also contain heat transfer and friction models which often required calibration. In general the purpose of engine modeling can be grouped into two categories: engine design or virtual engine representations.

2.3.1 Design using Engine Models

Building and testing a prototype engine is very expensive. Most designs involve at least one of the following: computer aided drafting (CAD), computational fluid
dynamics (CFD), finite element analysis (FEA) and/or engine simulation software. These tools help to keep the number of prototype designs to a minimum. Engine modeling software is also very important in this process. For example an engine model may be used to design an intake or exhaust system. Tuning the length of the intake and exhaust piping is most important for these applications. An engine model may also be used to select cam profiles or design a turbocharger. For most design applications, the engine model does not require extensive calibration. Because the engines being modeled have yet to be built (and may never be), the models cannot be calibrated with experimental data. Fortunately, components like wall temperatures can be approximated as constants with respect to the operating conditions without affecting the results. At the design stage the trends are much more important than the actual magnitudes. Inaccuracies in the wall temperature, for example, will affect the magnitude of the outputs, but the trends will remain the same.

2.3.2 Virtual Representations using Engine Models

The other classifications of engine models are those created to exactly represent a specific physical engine. In this case experimental data is needed to ensure the model produces the exact same outputs as the engine. Because engine simulation software cannot account for every phenomenon occurring in a real engine (like combustion and complex chemical kinetics), it is impossible to create an engine model that has every output perfectly match its experimental counterpart. Instead, a model is normally adjusted until a single variable very nearly agrees between simulation and experiment. That is not to say that all of the other variables will be grossly inaccurate. A careful
calibration will produce a model that predicts the trends of most quantities very well. The magnitudes and offsets, however, may deviate slightly from experimental data.

A model could be tuned to match the torque output or trapped air mass per cylinder or exhaust gas composition. For any variable to agree, several components in the model need to be calibrated including a combustion and heat transfer model. Once a model is tuned to represent a specific engine, an operating condition can be accurately simulated instead of running it experimentally. Additionally, the effect of small changes in the design can be seen without having to build a new engine. Another important use of this type of model is control design. A model could be used for optimizing the spark timing, cam timing, residual gases or many other control variables. Because a model can be run on multiple computers, hundreds of thousands of cases can be simulated within weeks. Testing these many points on an engine dynamometer would take months or years and many expensive prototype engines.

2.4 GT-Power

When used properly computer-based, object-oriented modeling can be an invaluable tool. One program commonly used in the automotive industry is GT-Power by Gamma Technologies. GT-Power explicitly resolves the one-dimensional wave dynamics throughout the entire engine model. The code resolves explicitly the unsteady wave dynamics in the entire engine. While GT-Power is mainly designed for steady-state (in terms of cycle-to-cycle variation) computations, the solution methodology is iterative in nature and estimates all of the relevant physical variables at each time-step until the solutions converge. There are several important keys to building an accurate engine
model using a one-dimensional simulator like GT-Power. First, the geometry must be
scrupulously interpreted and defined for accurate predictions of the wave dynamics.

Furthermore and perhaps more importantly, the geometry is projected onto its
one-dimensional equivalent. Hence, the simplification of complex three-dimensional
fluid dynamics onto a one-dimensional framework does not allow the simple translation
of geometric data into the dimensions of equivalent one-dimensional flow element.
Secondly, the simplifying assumptions are determined by the desired model outputs (VE
and residual gas fraction in this case). Similarly the factors that will be used to calibrate
the model need to be determined.

Figure 2 shows a layout view of a GT Power model for a 4 cylinder engine with 4
valves per cylinder. Each cylinder has 2 on the intake side and 2 on the exhaust side.
Each square represents a volume with a specified length, diameter, curvature, surface
roughness, and wall temperature as well as some additional specifications. Each round
object represents a vertical plane between volumes meant to define an orifice defined by
a flow loss coefficient. The layout features sub-models for each of the cylinders, which
require geometric information together with several user-defined models for calculating
or estimating the combustion process and heat transfer among others. These sub-models
allow the software to estimate quantities like in-cylinder pressure and exhaust gas
composition. The layout also shows a sub-model for the crank-train attached to all of the
cylinders, which models the inertia and kinematical crank characteristics.
The detail present in Figure 2 makes it clear that calibrating such a model can be challenging. Each of the objects contains many tunable variables, some of which are functions of the operating conditions. Fortunately, GT-Power estimates many necessary parameters including the majority of the flow loss coefficients due to geometry changes between connected volumes. Many of these lumped losses are meant to account for the net effects of the actual three-dimensional flow phenomena. Nonetheless, there are a few critically important variables that are specifically designed to be user inputs including: wall temperatures (or a heat transfer model), surface roughness, flow loss coefficients for the throttle and valves, geometric values, ambient conditions, and models for the combustion and heat release inside of the cylinders.

2.4.1 How GT-Power Generates a Solution

For GT-Power to account for the one-dimensional gas dynamics, each component in the model is discretized or separated into many smaller components. These
subcomponents have very small volumes and the fluid’s scalar properties in these volumes are assumed to be constant. The scalar properties of a fluid include pressure, temperature, density and internal energy. Each volume also has vector properties that can be transferred across its boundaries. These properties include mass flux and fluid velocity. Figure 3 illustrates the difference between vector and scalar properties. GT-Power determines the change in the scalar properties by solving simultaneous one-dimensional equations.

![Staggered Grid](image)

Figure 3: Illustration of Component Discritization [17]

The first equation ensures the conservation of mass as seen in (8).

\[
\frac{dm}{dt} = \sum_{\text{boundaries}} m_{\text{flux}}
\]  

(8)

Related to the mass equation is the conservation of momentum relationship shown in (9).

\[
\frac{d(m_{\text{flux}})}{dt} = \frac{dpA + \sum_{\text{boundaries}} (m_{\text{flux}}u) - 4C_f \rho u^2 \frac{dx}{D} - C_p \left( \frac{1}{2} \rho u^2 \right) A}{dx}
\]  

(9)

These two equations are predominately used to determine the trapped air and residual gas masses. They are also used to determine the fuel dynamics and mass flow rates. GT-Power also uses (10) which ensures that energy is conserved.
\[
\frac{d(me)}{dt} = p \frac{dV}{dt} + \sum_{\text{boundaries}} m_{\text{flux}} H - h_g A(T_{\text{gas}} - T_{\text{wall}}) \tag{10}
\]

The final equation is an exergy balance as shown in (11) which is implicitly solved.

\[
\frac{d(\rho HV)}{dt} = \sum_{\text{boundaries}} (\rho u_{\text{eff}} H) + V \frac{dp}{dt} - h_g A(T_{\text{gas}} - T_{\text{wall}}) \tag{11}
\]

Using these equations, the heat transfer from a volume to the walls or another volume can be determined. These final two equations determine the gas temperatures which also influences the trapped air and residual masses.

2.5 Calibration Techniques

Throughout this thesis, several empirical relationships are formulated. The phenomena being approximated are so complicated that explicit relationships cannot be determined from first principles. Empirical relationships are often better than those based on first principle approximations. An empirical relationship can be significantly less complex and still be accurate enough for many applications. Depending on the desired accuracy and form of the equation, different calibration techniques are used. The following is a brief description of the algorithms used in this thesis.

2.5.1 Least Squares

Least squares is a technique that minimizes the average of the square errors (L^2-norm). Least squares is a linear algebra concept where a vector is projected onto a vector space. The equation being solved is the matrix equation shown in (12).
Both the $A$ matrix and the $y$ vector are known, but the $x$ vector is not. If $y$ is an element of the column space of the $A$ matrix, then an exact solution exists. When least squares is applied to experimental data, this is never the case. Fortunately, the projection of the dependent vector $y$ onto the column space of $A$ can be found. The solution to (12) is the vector $x$ which when multiplied by matrix $A$ most closely approximates $y$ (the square error is minimized).

Even though least squares is based on linear systems theory, it can be applied to a subclass of nonlinear models. Least squares can be used to find the coefficients of any model where the coefficients enter the equation linearly. If each term of a model that is multiplied by a calibration coefficient is arranged into the columns of a matrix, then the optimal coefficients are the least squares solution to the above equation. With today’s computing power the optimal coefficients can be found within a few seconds. Therefore, least squares is one of the best techniques for finding the coefficients of a model that can be put into a linear form. For more information on least squares consult [18].

2.5.2 Stepwise Linear Regression

A stepwise linear regression algorithm is a least squares and probabilistic based optimization tool [19]. When dealing with a process that cannot be explained directly using first principles, an experimental correlation is used to model that process. Although the variables that may influence the process might be known, the importance of each term is normally unknown. A technique like least squares will generate the best
model using every term regardless of their significance. In most practical applications excluding insignificant terms from a model is as important as the accuracy of the model. Because only insignificant terms are removed, the overall accuracy of the model will remain the same while significantly reducing the storage and calculation burden for evaluation.

The stepwise linear regression algorithm works by iteratively calculating the statistically importance of each independent variable. Initially the algorithm starts with just a constant model. Then for each step, the statistically most important independent variable given the current model is added to the model. Once a new term is added to the model, least squares is applied to the new model to find the optimal coefficients. For this newly calibrated model the importance of each remaining independent variable is recalculated, and the most important term is added to the model. Again least squares is used to calibrate this new model. This process continues until there are no remaining independent variables that have a statistical correlation above some predefined threshold. The final model, therefore, only contains the most important independent variables, and the model’s fit has the smallest square error because least squares optimization was used.

2.5.3 Gradient-based Nonlinear Optimization

When the calibration coefficients of a model do not enter a model linearly, then a nonlinear optimization tool must be used. Most such optimization techniques for nonlinear problems are gradient-based methods. Gradient-based optimization algorithms search for minimum solutions by following the partial derivatives of a function. In practice gradient based optimization is an iterative procedure that is based on numerical
approximations of derivatives such as Newton’s method. To begin the optimization, an initial guess must be provided. Next, the partial derivatives with respect to the model’s error are approximated in every direction. The coefficients of the model move in the direction where the model’s error is reduced the most. At this new point the partial derivatives are reevaluated and the process repeats until there is no improvement possibility in any direction. Although gradient-based optimization algorithms are relatively quick if the initial guess is in the basis of attraction of the solution, their solutions are not necessarily the globally optimal solutions. If an equation has many local minima, then it is possible for a gradient based optimization to get “stuck” on a local minimum solution. Reference [20] provides a good overview of different gradient based methods.

2.5.4 Genetic Algorithms

In contrast to gradient-based optimizations, a genetic algorithm is a global nonlinear optimization technique. The complete theory behind genetic algorithms can be found in [21]. Using the same concepts as evolution in biology, a genetic algorithm simulates mating between individuals in a large population. Each individual is a set of coefficients corresponding to the model being optimized. The equations are evaluated using every set of coefficients (individuals) and the population is ranked. For this project the ranking was based on the summation of the square error between the equation’s prediction and the experimental data. However, very complicated nonlinear cost functions can be used. The very best individuals are directly carried over to the next generation but the majority of the new population is created by mating. Mating in a
genetic algorithm is a method of splicing the coefficients between two individuals. The resulting new population will have coefficients similar to the previous population, but their fitness relative to the equations will be different. Again, the equations are evaluated using every set of coefficients. Because the best individuals were kept from the previous generation, the new best individual is at least as good as before. To ensure that the population doesn’t become stagnant, random mutations are introduced into the population during mating. After simulating enough generations, this process will find the best individual which is the set of coefficients that have the best fit (smallest square error). Unlike mathematical (gradient-based) techniques which have formal proofs for convergence (or lack thereof), genetic algorithms do not formally guarantee convergence. However when applied to complex high dimensional problems, their performance is often far superior.

2.6 Model Sensitivity

As stated above, many of the components in GT-Power require calibration with experimental data. Depending on the output variable of interest, the importance of each component’s calibration accuracy will change. If the focus of the model is on emissions production or torque output, then the combustion parameterization is the most important and sensitive component. When the major concern is air flow (like this thesis), the thermal environment is considerably more important than the combustion model. To ensure that each of the parameters that require calibration is sufficiently accurate, an extensive sensitivity study was performed. Sensitivities can be interpreted as discrete approximations of partial derivatives of the outputs with respect to the inputs. These are
inherently local to a point in the parameter space. Furthermore, they can be in absolute or relative terms. The relative sensitivity of each parameter was calculated using (13) and this formulation is used throughout this thesis.

\[
\text{Sensitivity} = \frac{\frac{y(x_0) - y(x_0 + \Delta x)}{y(x_0)}}{\frac{x_0 - (x_0 + \Delta x)}{x_0}}
\]

\( y = \text{Dependent Variable} \)

\( x_0 = \text{Independent Variable} \)

\( \Delta x = \text{Change in Input} \)  \hspace{1cm} (13)

The independent parameter is perturbed slightly from a set point, and the outputs are recorded. By taking a ratio between the normalized difference in the output and the normalized difference in the input, the sensitivity can be found. This is effectively a non-dimensional partial derivative of the output with respect to the input. Figure 4 shows an example of the approximation being performed. The sensitivity of each parameter was calculated for multiple set points. Because airflow is the focus of the model, the sensitivity was calculated with respect to volumetric efficiency.
It is important to have an understanding of what type of sensors will have a large effect on the model’s accuracy before attempting to acquire experimental data to calibrate and tune this model. Understanding the ultimate goal of this procedure, which is to accurately estimate the VE and residuals, is also critical. Equation (14) shows the formal definition of VE when referenced to intake manifold conditions.

\[
VE = \sum_{i=1}^{N_{cyl}} \frac{\int \dot{m}_{valve,i}}{V_D \rho_a \text{man}}
\]

\(VE = \text{Volumetric Efficiency}\)
\(N = \text{Number of Cylinders}\)
\(\dot{m}_{valve,i} = \text{Mass Air Flow Into Cylinders}\)
\(V_D = \text{Displacement Volume}\)
\(\rho_a \text{man} = \text{Air Density In Manifold}\)

It is clear that there are many influences in the intake that can alter the VE by manipulating either the flow itself or the manifold density.

As expected there are several parameters which have a large effect on the model’s estimation of VE. The measurement of these parameters when collecting experimental
data is critically important for the calibration and validation of the model. The actual method for calculating this sensitivity using GT-Power is simple; run the model at a set of nominal operating points, then vary the input parameters as desired and observe the change in the output parameters. This evaluation is done strictly in simulation. As a representative example, the results from a sensitivity study of intake wall temperature are shown in Figure 5. Although the VE curve doesn’t change shape significantly, the locations of the peaks change. If the wall temperatures are incorrectly modeled, then the VE errors on the flanks of the VE curve can be immense. The results from the sensitivity study showed that the most important parameters relative to VE are the intake runner wall temperatures and the exhaust backpressure. The effects of the pre-throttle piping and combustion parameters were found to be minimal.
Figure 5: Example Sensitivity Study Showing the Effect of Wall Temperature on VE
CHAPTER 3

EXPERIMENTAL SETUP AND GT-POWER CALIBRATION PROCEDURE

3.1 Experimental Setup

All testing was performed at The Ohio State University’s Center for Automotive Research in an engine dynamometer test cell or chassis dynamometer. The testing was performed on a modern 2.4 liter four cylinder gasoline preproduction engine. The fuel is port injected and spark ignited. Both the intake and exhaust cam timing can be controlled independently with a 25 cam degree range. Because it is a preproduction engine, its prototype calibration varied slightly from the production one. In addition several physical modifications were made to the engine. The radiator and cooling system were removed and replaced with a water cooling tower system. To prevent wave dynamics from interfering with the LFE, the intake system was attached to a 50 gallon drum that draws room temperature air.

In addition to these modifications, many other components were instrumented so that data could be collected. A laminar flow sensor was attached to the entrance of the intake manifold to measure the mass flow rate of the air entering the intake. Two pressure sensors were placed in the intake system. One measured the ambient pressure, and the other measured the manifold air pressure. The intake system was also outfitted
with several thermocouples to measure the temperature distribution for the intake. Each runner in the intake manifold had two thermocouples to measure the wall temperature and air charge temperature. Another set of thermocouples was placed to measure the pre- and post-throttle temperatures. The exhaust system was also outfitted with several thermocouples. Another important modification to the exhaust was the addition of multiple heated exhaust gas oxygen (HEGO) sensors and universal exhaust gas oxygen (UEGO) sensors.

Additionally, a high precision angle displacement sensor with a one degree resolution was added to the crankshaft. The dynamometer measures the engine speed, torque output and power output. To measure the in-cylinder pressure, a special spark plug with a built-in pressure sensor was used. The data acquisition for the pressure sensor was triggered by the crank angle sensor such that a pressure measurement was taken at every crank angle degree. On the exhaust side, a Horiba MEXA 7500 gas analyzer was used to analyze the emissions and determine the composition of the gases.

Excluding the in-cylinder pressure sensors, the data was acquired at 100 Hz through a 12 bit data acquisition system. The interface used for this was Labview version 7. On the controller side, an ETAS system was used. ETAS is a sophisticated control computer which directly allows the monitoring and adjustment of many control parameters inside the ECU of the engine. With ETAS both the intake and exhaust cam timings could be controlled. Although ETAS can control other parameters such as spark timing, all other parameters were left unchanged. This uniformity provided the basis for determining the effects of cam phasing on volumetric efficiency and the variables needed for the GT-Power model.
3.2 Experimental Procedure

As previously stated, the goal of the testing was to determine the impact of intake cam timing, exhaust cam timing, engine speed and manifold air pressure on VE. Because time is a major constraint, the testing points were chosen in a semi-random manner such that the whole range of engine conditions was covered with a relatively low number of points. The data set collected is sparse in the four-dimensional parameter space yet it is space filling. Each operation point, which consisted of distinct values for the four parameters discussed earlier, was allowed to reach steady state after which thirty seconds of data was taken. The data was then averaged for each of the operating points. For the in-cylinder pressure data, 200 cycles of data was collected and the data was averaged for each crank angle degree.

Approximately 600 operating points were chosen to best represent the entire engine conditions. The engine speed ranged from 1000 to 5000 RPM and the manifold air pressure ranged from 0.2 to 1 bar (controlled via the throttle). Both the intake and the exhaust cams had a range of 0 to 25 cam angle degrees. The step sizes for the operating conditions were 200 RPM, 2 cam angle degrees and 2 degrees of throttle angle. Because the manifold air pressure was controlled with the throttle angle, the manifold air pressure was not limited to a discrete set of values. Figure 6 shows the distribution of the operation points and a two-dimensional projection of the manifold pressure, engine speed plane.
Figure 6: Randomly Sampled, Space-filling Dataset to be used for the Model Calibration

3.3 GT-Power Calibration Overview

Once the experimental data was collected, the GT-Power model was tuned to exactly (within the limits of the sensor’s accuracy) represent the physical engine. Calibrating a GT-Power model required several steps. In each of these steps, critical GT-Power parameters or sub-models were tuned so that their outputs matched experimental data. In the preliminary phase the model was exercised to conduct numerical sensitivity experiments to identify which parameters and sub-models have the strongest impact on the prediction accuracy of the air mass estimation. This led to calibrating three categories of objects: the major flow loss elements (throttle and valves), wall temperatures for key locations and the combustion heat release. Unrepeatable operating points, such as those resulting in poor combustion (coefficient of variation (COV) of indicated mean effective pressure (IMEP) greater than 4%, off-idle), were removed for the experimental dataset before calibration. This filtered dataset was then separated into a calibration dataset and a validation dataset to be used later.
3.4 Throttle and Valve Calibration

Though the properties of most of the orifices are calculated by GT-Power from geometry, the discharge coefficient of the throttle varies with throttle angle. Manifold pressure and throttle angle are very strongly correlated. By convention manifold pressure is normally used in engine control. GT-Power, however, does not allow direct specification of manifold pressure. Instead a throttle angle is specified, and GT-Power solves for the manifold pressure. GT-Power requires a model relating the discharge coefficient of the throttle to the throttle angle. Typically the discharge coefficient of the throttle would be measured on a flow bench. Then a polynomial or spline approximation is used to capture this dependence. Although minor, changing the cam timing has an influence on the resulting manifold pressure even for constant throttle positions. Therefore a manifold pressure based throttle controller was implemented in the GT-Power. Using manifold pressure as feedback, the throttle angle was adjusted until the desired manifold pressure was achieved. Because manifold pressure is used in engine control instead of throttle angle, small adjustments in throttle angle made by the throttle controller do not affect the results.

Calibrating the intake and exhaust valves is a straightforward procedure, although it requires a high-quality flow bench and accompanying equipment and sensors. When conducting flow measurements on the cylinder head, it is best to create boundary flow conditions that are as similar as possible to those that would occur on the complete running engine. To that end, parts of the intake as well as the exhaust runner were attached to the head during flow measurements. Flow measurements for the intake and exhaust valves were obtained independently. By stepping through the valve lift at small
increments while measuring flow through the head, a discharge coefficient versus valve
lift was measured. During high valve overlap conditions flow reversal occurs, so the
discharge coefficients for the valves were tabulated in both the forward and reverse flow
directions. A simple one-dimensional table lookup for discharge coefficient versus lift
was generated for each valve in both the forward and reverse directions.

3.5 Steady-State Wall Temperature

As described before, model sensitivity studies were used to simplify the model by
specifying only a few appropriately chosen wall temperatures. This is considerably
easier to calibrate (based on a few experimentally measure temperatures), than to
calibrate a complete heat transfer model for the whole engine like in [22]. These steady-
state temperatures corresponded to the intake manifold wall, the intake runner walls, and
the catalyst wall. In order to account for changes in the ambient temperature that might
confound the regression models, only the temperature departures from ambient were used
in the regression models. The relationships between the wall temperature and the
operating conditions are very complicated and cannot be described using first principles.
Instead, empirical correlations were used. The equation for each wall temperature model
is a quadratic regression of the independent variables. The general form of a model is
shown in (15).

\[ F = a_1 \text{RPM} + a_2 \text{MAP} + a_3 \text{ICAM} + a_4 \text{ECAM} + a_5 \text{RPM} \times \text{MAP} + a_6 \text{RPM} \times \text{ICAM} \\
+ a_7 \text{RPM} \times \text{ECAM} + a_8 \text{MAP} \times \text{ICAM} + a_9 \text{MAP} \times \text{ECAM} + a_{10} \text{ICAM} \times \text{ECAM} \\
+ a_{11} \text{RPM}^2 + a_{12} \text{MAP}^2 + a_{13} \text{ICAM}^2 + a_{14} \text{ECAM}^2 \]  

(15)
To determine these correlations, least squares optimization was used. Even though the accuracy of the wall temperature is paramount for VE accuracy, a global fit of data provided sufficient accuracy.

3.6 Combustion Model

Combustion modeling is clearly the most challenging aspect of all engine simulations. From the sensitivity study it was found that specifying a heat release profile is accurate enough for a model focused on airflow. However the heat release profiles varied dramatically as a function of operating conditions. Likewise, under high valve overlap conditions burning occurred in two stages. To model this heat release, a cascade of two Wiebe functions was used. As shown in (16), several coefficients were needed to properly fit the data.

\[
x_b = \chi \left( 1 - \exp \left[ -a_1 \left( \frac{\theta - \theta_0}{\theta_D} \right)^{m_z+1} \right] \right) + (1 - \chi) \left( 1 - \exp \left[ -a_2 \left( \frac{\theta - \theta_0}{\theta_D} \right)^{m_z+1} \right] \right)
\]  

(16)

An illustration of the double Wiebe approximation fit to the experimental data is shown in Figure 7. Even though a Wiebe approximation is a gross simplification of the combustion process, the sensitivity of the combustion model relative to VE was very low. Hence, a simple Wiebe function parameterization is sufficient for the problem of air estimation.
3.6.1 Heat Release Calculations

The heat release rate cannot be directly measured, but instead must be inferred from the in-cylinder pressure measurements. Using (17), the differential changes in volume and pressure can be related to the heat release rate.

\[
\frac{dQ}{d\theta} = \frac{\gamma}{\gamma - 1} p \frac{dV}{d\theta} + \frac{1}{\gamma - 1} \frac{V}{d\theta} \frac{dp}{d\theta}
\]  

(17)

In this equation \( \gamma \) is the specific heat ratio which was assumed to have a constant value of 1.4. The pressure, \( p \), is the in-cylinder pressure and \( V \) is the volume. Because all of the data was triggered at every crank angle degree, the differential volume and pressure values were found using the finite difference approximation. The volume in the cylinders was never measured, but the dimensions of the engine are known. Using the connecting rod length, the radius of the cylinder and the compression ratio, the volume inside the combustion chamber can be related to the crank angle using (18).
The heat release rate generated from these equations is not a direct measure of the heat released by the fuel. Instead, this equation generates the apparent or net heat release. Because the gases inside the cylinders have different temperatures than the cylinders’ walls, there is heat transfer between the gases and the walls. Before combustion, heat flows from the walls to the gases; after combustion, heat flows from the gases to the walls. The measured heat release rate is therefore the heat released by the fuel minus the heat transfer losses. Although the net heat release has a slightly different shape than the true heat release curve, the heat release of the fuel was approximated by the net heat release.

3.6.2 Mass Fraction Burned Curves

Once these calculations were made for each cylinder, the values of the four cylinders were averaged. The resulting average mass fraction burn values were graphed against crank angle degrees and analyzed. Across the operation conditions, the mass fraction burned curves varied dramatically as shown in Figure 8. Figure 9 shows a comparison between mass fraction burned curves for a low valve overlap condition and one with high valve overlap. The mass burn fraction curve on the left (low valve overlap) is very smooth and has only one inflection point. Conversely, the mass fraction burned curve corresponding to high valve overlap (right graph) has a very abrupt change in its curvature and three inflection points. For the first part of the curve, the heat release is very rapid, but after only a short burn duration the heat release rate slows substantially. Because the intake and exhaust valves are simultaneously open for so long, the ratio of
fresh charge to recirculated combustion products is much lower. This mixture is able to burn quickly initially but as the fuel is burned the oxygen becomes scarce. Without an adequate amount of oxygen, the burn rate is reduced.

![Variations in MFB Curves](image1)

Figure 8: Variation in Mass Fraction Burned Curves

![Normal vs High Valve Overlap](image2)

Figure 9: Representative Mass Fraction Burned Curve Comparison
3.6.3 Predicting the Mass Fraction Burned Curves

Once all of the mass fraction burned curves were created, a gradient based nonlinear optimization was performed to find the optimal coefficients for the double Wiebe function (16). GT-Power does not have this same functional form, but instead uses 10-90% burn duration in crank angle degrees and the crank angle corresponding to 50% burned (anchor angle). Therefore, the burn durations and anchor angles were extracted from both Wiebe functions for each operating point. Although the variation between operating conditions was smaller in this representation, it was still too large for a global model. Instead the operating space was separated into 9 regions in the MAP, RPM plane. The same technique used to calibrate the wall temperature models was applied to each zone of the combustion model for a total of 63 models (9 zones x 7 coefficients). For each region and each coefficient a quadratic regression of the operating conditions (MAP, RPM, ICAM and ECAM) was performed using least squares. By specifying the coefficients of the mass fraction burned model, GT-Power calculates the heat release automatically.

3.7 Model Validation

The inputs to the calibrated GT-Power simulation are MAP, RPM, ICAM, ECAM, an initial guess at throttle angle and ambient temperature and pressure. From these GT-Power calculates engine air flow rate (MAF), MAP (throttle adjusted until desired MAP achieved), manifold air temperature, exhaust temperature, cylinder pressure, volumetric efficiency (from trapped fresh charge mass), and residual gas mass. An engine on a dynamometer when set to the same “inputs” as the GT-Power model also
calculates these same physical quantities. Therefore, the best model validation possible is to compare GT-Power’s calculated outputs with the outputs that the engine “calculates.” Moreover, since all but the residual gas fraction are readily measured, this is easy to do. A subset of the experimental dataset not used for calibration was used for validation.

To ensure that the GT-Power model can accurately calculate volumetric efficiency, the model of each component that affects VE starting with the intake manifold air temperature was validated. The wall temperature in the model is specified for each point. However, the air temperature is what affects the VE, so this is the quantity that was validated. For every case the air temperature agreed to within a single degree Celsius. The next quantity that was validated was the in-cylinder pressure. A good quantity to compare in terms of the combustion is the in-cylinder pressure, because it has the largest effect on the VE. Small errors in the mass fraction burned curve model could translate to amplitude and/or offset errors in cylinder pressure. As shown in the representative pressure trace of Figure 10, the estimated cylinder pressure is very close to the measured experimental data. The peak pressures of both curves are aligned both in crank-angle location as well as in magnitude.
The next metric that was validated was the pressure in the exhaust system. During the calibration stage, the geometry for the exhaust was built. The errors are very negligible with the exception of high pressure points. Even at these high pressure points, the relative errors are within 2%. The most important validation metrics are the manifold pressure and mass air flow rate. Figure 11 and Figure 12 show the error distributions for MAP and MAF, respectively, for the validation data set. Once again, the agreement between the GT-Power-calculated values and the dynamometer measurements is excellent. The MAP errors are normally distributed with no average error thanks to the throttle controller. The MAF errors also have a normal distribution with negligible average error. Figure 13 shows a histogram of the differences between VE calculated by GT-Power and VE measured on the dynamometer. As with MAP and MAF, the errors are nearly normal with an average error of only 0.11%. Similar comparisons for the other measurable quantities showed similar performance. Because it cannot be physically
measured, the only variable that could not be validated in this manner was residual gas fraction. In view of the accuracy of the other comparisons and that the residual gas fraction is governed by the same gas exchange wave dynamics as the fresh air charge, it is highly unlikely that the calibrated simulation would deviate markedly from the actual trapped residual gas mass. Furthermore, the required level of accuracy for the fresh air charge prediction is considerably more stringent than the residual gas fraction which is primarily used to schedule spark.

Figure 11: Histogram of MAP errors for Calibrated GT-Power Model

Figure 12: Histogram of MAF errors for Calibrated GT-Power Model

Figure 13: Histogram of VE errors for Calibrated GT-Power Model
3.8 Virtual Engine Mapping and Data Filtering

After the GT-Power Model is validated, it can simulate any operating condition with approximately the same accuracy as experimentally measured data. The first step in simulating an operating condition is to calculate the wall temperatures and combustion parameters from the sub-models. These sub-models are based on quadratic regressions of the operating conditions. Although they have been validated, it is important to place bounds on the sub-model outputs. Quadratic functions are very good at predicting data inside the space covered by the calibration data. The data is essentially being intelligently interpolated. Outside of this space the quadratic function must extrapolate the trends. Extrapolating a quadratic function can result in some extreme values. The calibration data was chosen to be space filling so that the nearly all of the operating space was covered. However, it is still possible that small areas in the four-dimensional operating space were not covered. By placing bounds on the outputs of the sub-models, the errors associated with extrapolating a quadratic function are significantly reduced. The bounds for the function were chosen based on the physical data. The largest and smallest measured values for each variable were extended by 25% and the bounds were placed there.

GT-Power can simulate any operating condition, even ones that are non-physical. Before an operating condition is run, it may not be known if that operating condition could be run on an engine. Therefore, it is important to exclude these conditions based on the outputs of GT-Power. For completeness every engine speed from 600 to 6400 in increments of 200, MAP from 0.15 to 1.05 bar in 0.04 increments and every cam position
from 0 to 25 cam degrees in 2 degree steps was tested. Many of these operating conditions would result in poor burn, misfire or high COV and some are not even reachable. As shown in Figure 14 it was not possible to reach very low manifold pressures at low engine speeds. This is consistent with a physical engine. To provide a comparison, the manifold pressures and engine speeds encountered during a federal transportation procedure (FTP) are shown in Figure 15.

Two of the best indications of a physically realizable operating condition are combustion duration and residual gas fraction. When the residual gas fraction is too high, combustion becomes very slow and unrepeatable. Regular combustion has a COV of IMEP less than 4 percent. If the residual gas fraction is too high, then this variation increases dramatically. Although an engine can still run when the residual gas fraction is high, the engine controller tries to avoid these operating conditions. Including these extreme points in a model will not improve the correlation at the extreme points because they have so much variation. An operating condition with a residual gas fraction of 30
percent or more will produce a COV of much higher than 4 percent. For the four cylinder engine studied in this thesis, the filter cutoff for the residual gas fraction was 30 percent.

The second physically based filtering metric is burn duration. The combustion model used in GT-Power is a double Wiebe function. This functional form can represent stable combustion very well. Unstable combustion, however, cannot be captured with this model. Even if it could be modeled, these points should be excluded, because they have high variation. For an engine to completely burn all of the fuel in less than 25 crank angle degrees, there would have to be multiple flame fronts. Therefore any operating condition that GT-Power predicts to have a burn duration of less than 25 crank angle degrees should be removed.

The residual gas fraction condition filters out a substantial amount of data. As shown in Figure 16 many simulated cases resulted in more residuals than fresh air. These cases directly correspond to small throttle openings and high valve overlap. Figure 17 and Figure 18 compare the variation in the burn duration of the experimental data to the simulated data. Obviously the simulated cases with burn durations near zero need to be excluded, because they represent complete misfire. This only removes a handful of cases. Finally to ensure that each GT-Power case is run in a timely manner, a limit should be placed on the number of iterations before convergence. If an iterations limit is used, then not every case may converge. The case that did not converge must be filtered out too. Again these cases correspond to high valve overlap and small throttle openings which could not be run on an engine. After all of the filtering, about a quarter of the data was removed.
After the complete range of operating conditions is simulated and filtered, the relationship between volumetric efficiency (or residual gas fraction) and the operating conditions can be determined. To illustrate the effect of the operating conditions on volumetric efficiency, the VE surfaces for parked cams and fully overlapped cams (ICAM = 25 cam degrees, ECAM = 25 cam degrees) are shown in Figure 19 and Figure 20. The engine speed affects the wave dynamics of the intake system and therefore strongly affects the volumetric efficiency. Conversely, the manifold pressure affects the
volumetric efficiency in a nearly linear manner. The volumetric efficiency surface changes with cam position, but the local peaks occur in approximately the same position. Compared to the volumetric efficiency, the residual gas fraction has a much more regular variation with respect to cam position. Figure 21 and Figure 22 compare the residual gas fraction surfaces for parked cams and fully overlapped cams. In both cases the manifold pressure has a strong inverse effect. The effect of engine speed, however, is only important under high overlap operating conditions.

![Figure 19: VE Surface for Parked Cams](image)

![Figure 20: VE Surface for Cams at (25 cam degrees, 25 cam degrees)](image)

![Figure 21: Residual Gas Fraction Surface for Parked Cams](image)

![Figure 22: Residual Gas Fraction Surface for Cams at (25 cam degrees, 25 cam degrees)](image)
CHAPTER 4

FORMULATION OF THE NEW VOLUMETRIC EFFICIENCY REPRESENTATION

4.1 VE Modeling Accuracy Requirements

Maintaining a high accuracy in the representation of VE is a very difficult problem. To provide a quality air estimate for AFR control, the standard deviation of VE errors should be less than 3%. When an engine is in a vehicle, the variation is much higher than in laboratory conditions. This added variability must be accounted for in the accuracy metric of a VE representation. Therefore the target standard deviation of the VE error for the VE representation relative to experimental data is 2%.

For conventional engines with fixed cam timing, representing VE is simple. A two-dimensional table indexed by manifold pressure and engine speed can be used. Complex engines with additional degrees of freedom require higher dimensional tables to represent VE. An engine with cam phasing would require a four-dimensional table. About a hundred times more table entries would have to be stored and many more calculations would be required to lookup the VE. These additional storage and calculation requirements make VE table representations undesirable for engines with many degrees of freedom.
Several hybrid techniques involving tables have been attempted. In one such approach developed by Ford Motor Company [23], the operating space represented in the table is reduced to only a few key points. To calculate the cylinder mass, they use an affine equation based on MAP. The two volumetric efficiency coefficients for this equation are stored in a three-dimensional table indexed by engine speed, intake cam position and exhaust cam position. Instead of using a rectangular grid, only three entries are stored per engine speed. These three points correspond to the default (parked) cam pair, the cam pair which produces the best fuel economy and the maximum torque cam pair. The spacing of this table is irregular and requires a sophisticated method for interrogating the table. The method purposed in this paper was to calculate the Euclidian distance between the current cam pair and the table’s three cam pairs at the current engine speed. Then, use an inverse distance interpolation to calculate the predicted volumetric efficiency coefficients.

Although the amount of calibration space required for this method is small, the calculation algorithms are complex. The accuracy of this method is very good near the three modes stored in the table (parked, peak fuel economy and max power). As the distance away from these points increases, the accuracy decreases. Operating conditions inside the triangle defined by the table points should have reasonable accuracy because the VE does not change very much. Points outside this triangle probably have poor volumetric efficiency. Because the predicted VE is a weighted average of the three table entries, the prediction is bounded by the smallest table entry. Therefore, the operating conditions that produce poor VE would be strongly over predicted. Unfortunately during transients (when VE is most important), the cam pairs will almost never follow any of the
three modes and could easily follow a path that lies outside the table’s triangle. Additionally, this method is based heavily on the operating conditions which produce peak power and peak fuel economy. The operating conditions that produce these two conditions are very different from engine to engine. Therefore the accuracy of this method can also vary dramatically. Significant problems can arise if points become collinear (for example the default cam condition corresponds to the best fuel economy). Although this approach has good accuracy in certain regions, many regions would have unacceptable errors.

4.2 Data Trends between Operating Conditions and VE

VE is dependent on the operating condition, so a VE representation must also be a function of the operating conditions. A complete physics description is not available, so the form of the function must be designed. To find a function that captures the relationship between the operating conditions and VE, it is important to understand each variables’ impact on VE. For this engine four variables define the operating condition. Dimensions greater than three become hard to visualize. Each parameter was studied by fixing two of the variables and sweeping the remaining two.

4.2.1 Nearly Linear Relationship between MAP and VE

The manifold pressure has the most direct effect on volumetric efficiency. As the manifold pressure increases, the volumetric efficiency increases. Reconsider Figure 19 of Section 3.8 which showed the relationship between MAP and engine speed on VE when the cams were parked. Engine speed does not alter the relationship between MAP
and VE. Regardless of the engine speed, VE and MAP have a nearly linear relationship. At different cam positions the shape of the VE surface changes. Figure 20 also in Section 3.8 shows the dependence of VE on MAP and engine speed when the cams are fully actuated. The curvature in the engine speed direction noticeably changes but not in the MAP direction. Only at very low and very high manifold pressures does the dependence of VE on MAP become noticeably nonlinear.

4.2.2 Resonance Structure in Engine Speed

From the previous two surface plots, it is clear that the engine speed has a complex impact on VE. The sharp peaks and valleys in VE are caused by resonances in the intake system. While the engine is running, the intake system acts like a Helmholtz resonator. A Helmholtz resonator is a tube with one end open and the other end connected to a large closed volume. Figure 23 illustrates the gas flow in a Helmholtz resonator. When the air inside a Helmholtz resonator is excited, the large gas volume at the closed end acts like a spring. As air moving in the tube region reaches the closed end, momentum is transferred to the gas volume. The gas inside the volume is momentarily compressed as the momentum is absorbed. Once velocity of the gas at the entrance to the gas volume has decreased to zero, the gas volume transfers momentum back to the gas. Then the gas inside the tube travels out the open end and the vacancy of gas creates a low pressure zone. Some gas will flow back into the tube and the process will repeat. The dimensions of a Helmholtz resonator’s tube such as length and cross sectional area determine the oscillation frequencies or resonance frequencies of the gas.
In an engine the cylinders act as a large closed gas volume. The intake runner length and cross sectional area are constant so the resonance frequencies of the intake system are relatively constant. While the concept of Helmholz resonators is good to explain the presence of resonances and anti-resonances in VE, the interactions are far more complex because of the multiple branching systems of pipes and progressive boundary conditions imposed by the opening and closing of valves. When an intake valve opens, air begins to flow into the cylinder. While the cylinder is filling, resonance waves are propagating inside the intake system. The location of the resonance wave when the intake valves closes affects the mass of air inducted by the engine. If the intake valve closes when the resonance wave reaches the cylinder, then an extra mass of air will be trapped in the cylinder. Conversely if the intake valve closes when the resonance wave leaves the cylinder, then less air will be trapped. As the engine speed changes so
does the position of the resonance wave with respect to intake valve closing. This causes
the VE to have peaks and valleys. Regardless of MAP, the locations of the VE peaks and
valleys relative to engine speed are constant. To illustrate this claim, the VE for each
manifold pressure was normalized by dividing by the VE at 600 RPM. Figure 24 shows
this normalized VE at parked for all manifold pressures.

![Figure 24: Normalized VE versus Engine Speed Curves for all Manifold Pressures](image)

4.2.3 Effect of ICAM and ECAM

Although the cam timing can be changed, the duration in crank angles that the
intake valves are open stays constant. The time in crank angles between the excitation of
the intake system (intake valve opening) and intake valve closing is fixed. Therefore the
location of the resonance wave with respect to intake valve closing is also constant.
Independent of cam timing, the peaks and valleys of VE occur at nearly the same engine
speeds. However, cam timing can alter the magnitude of the peaks and valleys slightly.
Even though the duration of the intake event is fixed, the state of the cylinder when the
intake valve opens is a function of the intake and exhaust cam position. Like most outputs, the dependence of VE on cam timing is mainly due to valve overlap. When the valve overlap is high, the cylinder at intake valve opening contains a large mass of hot exhaust gas at a pressure above ambient. Conversely when there is no valve overlap, the cylinder contains a small mass of exhaust gas near or below ambient pressure. Although the pressure waves propagate similarly, the difference between the intake manifold pressure and the cylinder pressure varies with cam position. It is this pressure difference that causes the excitation of the intake system. Therefore the peaks occur at nearly the same engine speeds, but the amplitudes change. Figure 25 shows the effect of cam timing on VE. VE is graphed as a function of engine speed for all possible cam pairs under a constant manifold pressure of 85 kPa. The peaks and valleys align but the magnitude changes. Consider the prominent VE peak around 2400 RPM. Figure 26 shows the location of the local VE peak as a function of the cam positions. Except under the most extreme overlap condition (ICAM = -25 cam degrees, ECAM = 25 cam degrees), the locations of VE peaks are nearly constant.
4.3 Possible VE Model Forms

Although a polynomial regression of the independent variables (engine speed, manifold pressure, intake cam position and exhaust cam position) could be performed, it would not have sufficient accuracy. The nearly linear effect of manifold pressure could
be captured, but the other three variables are much harder to capture. Because the VE has about five local minima and maxima, at least a sixth order polynomial in engine speed would be needed. Even with this high order polynomial, the curve fit does not represent the VE very well. More importantly, calculating six exponentials of different orders would be very taxing for the ECU. One of the current automotive practices is to represent variables like VE using a table. For an engine with cam phasing, this table would be four-dimensional. In between table entries the VE is linearly interpolated. To accommodate the memory storage limits, the table resolution must be relatively coarse. Because the table must be so small, the errors are relatively high.

4.4 Comparison between VE at Parked and VE Away from Parked

Using a four-dimensional table requires too much memory and using a polynomial regression lacks accuracy. However, combining these two techniques can produce good results. A two-dimensional table can provide a baseline approximation which can then be modified by quadratic terms. The hardest variable to represent is engine speed, so it needs to be one of the table inputs. Manifold pressure was chosen to be the other input to make the procedure universal. Not all engines have the same actuators and independent variables like intake and exhaust cam phasing. However manifold pressure and engine are variables on every engine. The cam positions chosen for the baseline table were the parked or unactuated position. Before the engine reaches its target operating temperature, the cams stay parked and when warmed-up parked is frequently encountered. It is the default cam position and where the cams go if something malfunctions. Lastly combustion is stable with parked cams for all manifold
pressure and engine speed combinations. Therefore, the best reference cam position to use for a VE table is parked cams.

To provide a baseline comparison metric, the volumetric efficiency distribution statistics are provided in Table 1. The variation of volumetric efficiency across the operating space is very large and difficult to capture. However, much of the variation can be reduced by using a parked VE table reference. Using a table size of 17 by 17 and GT-Power data, a parked table was generated. If the effect of the cam positions was completely ignored and the volumetric efficiency was only estimated via a parked table, then the volumetric efficiency errors would be those shown in the right column of Table 1. The standard deviation in volumetric efficiency is reduced from 14.50% to 4.30% by simply using a parked table reference. Although this is more than twice as large as desired (2%), it does represent a dramatic reduction.

<table>
<thead>
<tr>
<th>All Units %</th>
<th>VE</th>
<th>All Units %</th>
<th>Parked Table Reference Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>Standard deviation</td>
<td>14.50</td>
<td>Standard Deviation</td>
<td>4.298</td>
</tr>
<tr>
<td>Mean</td>
<td>74.18</td>
<td>Average Error</td>
<td>2.166</td>
</tr>
<tr>
<td>Average Absolute</td>
<td>11.83</td>
<td>Average Absolute</td>
<td>3.928</td>
</tr>
<tr>
<td>Difference from Mean</td>
<td></td>
<td>Error</td>
<td></td>
</tr>
<tr>
<td>Min, Max</td>
<td>36.69, 103.7</td>
<td>Max Error</td>
<td>21.65</td>
</tr>
</tbody>
</table>

Table 1: Distribution of VE and Distribution of VE Relative to Parked VE Table

4.5 Shifting the VE Peaks via Engine Speed Warping

The cam positions mainly affect the magnitude of the volumetric efficiency relative to parked. However, some cam combinations, generally corresponding to high
valve overlap, can shift the location of the VE peaks and valleys. Even though the location of these features stays relatively constant under most cam positions, any differences from the baseline table value can cause major errors. Remember the results shown in Section 2.6 for the sensitivity study of intake wall temperature. Small changes in the peak VE location can cause dramatic errors. Using a simple quadratic regression of the intake and exhaust cam positions, the shifts in the peak VE locations were modeled. Equation (19) was optimized using data from the VE peak at 2400 RPM.

$$\frac{RPM_{VE_{peak}}}{RPM_{parked}} = C + a_1ICAM + a_2ECAM + a_3ICAM \times ECAM + a_4ICAM^2 + a_5ECAM^2 \quad (19)$$

As shown in Figure 27, the location of this VE peak varies regularly with respect to cam position. Although the other peaks have different trends, they can all be captured with the same technique. The deviations in peak locations are smooth functions of cam position.
4.5.1 Engine Speed Warping VE Function

The first model tried to capture the small variations in peak locations by modifying the engine speed used in looking up the reference parked table. Equation (20) shows the form of the model and the terms in each function.

\[
VE(\text{ICAM}, \text{ECAM}, \text{MAP}, \text{RPM}) = f_{\text{add}}(\text{ICAM}, \text{ECAM}) + \\
VE_{\text{parked}}(\text{MAP}, \text{RPM} \ast f_{\text{warp-mult}}(\text{ICAM}, \text{ECAM}))f_{\text{mult}}(\text{ICAM}, \text{ECAM})
\]

\[
f_{\text{warp-mult}}(\text{ICAM}, \text{ECAM}) = 1 + a_1 \text{ICAM} + a_2 \text{ECAM} + a_3 \text{ICAM} \ast \text{ECAM} + a_4 \text{ICAM}^2 + a_5 \text{ECAM}^2
\]

\[
f_{\text{add}}(\text{ICAM}, \text{ECAM}) = b_1 \text{ICAM} + b_2 \text{ECAM} + b_3 \text{ICAM} \ast \text{ECAM} + b_4 \text{ICAM}^2 + b_5 \text{ECAM}^2
\]

\[
f_{\text{mult}}(\text{ICAM}, \text{ECAM}) = 1 + c_1 \text{ICAM} + c_2 \text{ECAM} + c_3 \text{ICAM} \ast \text{ECAM} + c_4 \text{ICAM}^2 + c_5 \text{ECAM}^2
\]

The VE table is indexed by engine speed and manifold pressure. However instead of looking up VE using the actual measured engine speed, a slightly modified value is used.
The true engine speed is modified by a function of the cam position such that the peaks and valleys align with the parked VE. By design this term reduces to one when the cams are parked. An alternate form was also tried. Instead of a multiplicative engine speed corrector, an additive corrector was explored. Equation (21) represents this additive version of this VE representation.

\[
VE(ICAM, ECAM, MAP, RPM) = f_{add}(ICAM, ECAM) + \\
VE_{parked}(MAP, RPM + f_{warp_{add}}(ICAM, ECAM))f_{mult}(ICAM, ECAM)
\]

\[
f_{warp_{add}}(ICAM, ECAM) = a_1ICAM + a_2ECAM + a_3ICAM \ast ECAM + a_4ICAM^2 + a_5ECAM^2
\]

\[
f_{add}(ICAM, ECAM) = b_1ICAM + b_2ECAM + b_3ICAM \ast ECAM + b_4ICAM^2 + b_5ECAM^2
\]

\[
f_{mult}(ICAM, ECAM) = 1 + c_1ICAM + c_2ECAM + c_3ICAM \ast ECAM + c_4ICAM^2 + c_5ECAM^2
\]

These two forms will be referred to as the engine speed warping VE functions.

For a single manifold pressure (85 kPa) and engine speed (2400 RPM), the coefficients for both models were optimized. Because these equations are so complex and nonlinear, they cannot be optimized using least squares. Instead a genetic algorithm was used. The errors for both models were similar and almost negligible. A surface plot of the VE error for the additive engine speed warping model is presented in Figure 28. For every manifold pressure and engine speed combination, the optimum coefficients were found. The VE errors were almost always less than 2 percent. All of the VE error surfaces as a function of cam position are graphed together in Figure 29. The magnitude of the engine speed warping was commensurate with the differences in the peak VE location. This indicates that the optimum engine speed warping terms were indeed the ones that made the VE curves align. Although the model had a good fit when it was
optimized for single manifold pressure and engine speed pairs, the coefficients varied for different operating pairs. The variation in the optimized model coefficients was not regular enough to be modeled. To capture this variation, a table could be used. However, each coefficient would need a table. Given the memory constants of the ECU, this is unfeasible to implement.

For this model to be practical, a single set of coefficients must hold for the entire operating range. To find a global set of coefficients, all of the data was used in the optimization. The optimal coefficients for the engine speed warping part of the model were nearly zero. Although in local operating regions the engine speed warping provided great benefit, this benefit disappeared when the entire region was considered. To evaluate the effectiveness of the engine speed warping term on the global dataset, the term was removed and the equation was re-optimized. After optimized the model without the engine speed warping was actually better than the model which included it.
If the genetic algorithm was allowed to run long, the model with the warping term would have eventually reached a solution which produced results at least as good as the model without the warping. However, the improvement would be almost negligible and the computation time would be very long.

Table 2 compares the VE errors for the model with and without the engine speed warping term. The errors for both equations are unacceptable (standard deviation both greater than 3%), but an important conclusion can be drawn. As a global solution, the possible improvement that could be gained from including the engine speed warping term is outweighed by the increased computation time and increased complexity.

<table>
<thead>
<tr>
<th></th>
<th>With Engine Speed Warping</th>
<th>Without Engine Speed Warping</th>
</tr>
</thead>
<tbody>
<tr>
<td>Standard Deviation (%)</td>
<td>3.10</td>
<td>3.05</td>
</tr>
<tr>
<td>Average Absolute (%)</td>
<td>1.92</td>
<td>1.91</td>
</tr>
</tbody>
</table>

Table 2: Comparison between Error of Engine Speed Warping Model with and without the Warping Term

4.6 Full Quadratic VE Function

Although the engine speed warping did not work as a global model, the parked table reference showed promise. By multiplying the parked table by a function of cam positions and adding to it another function of only cam positions, the volumetric efficiency error was decreased. The second model is an extension of the first model without the engine speed warping. By removing the engine speed warping function,
more terms can be added to the function that multiply and add to the parked table. The
form of the second volumetric efficiency function is shown in (22).

\[ f_{\text{mult}}(ICAM, ECAM, MAP, RPM) = a_1 ICAM + a_2 ECAM + a_3 ICAM \times ECAM + a_4 ICAM^2 + a_5 ECAM^2 + a_6 MAP + a_7 ICAM \times MAP + a_8 ECAM \times MAP + a_9 MAP^2 + a_{10} RPM + a_{11} RPM^2 + a_{12} RPM \times ICAM + a_{13} RPM \times ECAM + a_{14} RPM \times MAP + 1 \]

\[ f_{\text{add}}(ICAM, ECAM, MAP, RPM) = b_1 ICAM + b_2 ECAM + b_3 ICAM \times ECAM + b_4 ICAM^2 + b_5 ECAM^2 + b_6 MAP + b_7 ICAM \times MAP + b_8 ECAM \times MAP + b_9 MAP^2 + b_{10} RPM + b_{11} RPM^2 + b_{12} RPM \times ICAM + b_{13} RPM \times ECAM + b_{14} RPM \times MAP + b_{15} \]

\[ VE(ICAM, ECAM, MAP, RPM) = VE_{\text{parked}}(MAP, RPM) f_{\text{mult}}(ICAM, ECAM, MAP, RPM) + f_{\text{add}}(ICAM, ECAM, MAP, RPM) \]

(22)

Instead of making the additive and multiplicative function only a function of the cam
positions, they are now full quadratics of the cam positions, manifold pressure and engine
speed. The number of coefficients was nearly doubled in the process, though. Again the
multiplicative term contained a constant term equal to one. Unlike the previous model,
the additive function contains a constant offset term \(b_{15}\). This term will ensure that the
model has zero mean error. From now on this model will be referred to as the full
quadratic VE function.

When the cams were parked in the engine speed warping VE function, this
multiplicative term was equal to one. For the full quadratic VE function the
multiplicative term does not reduce to one. However, including the offset provides a
good initial guess. An initial condition of all zeros corresponds exactly to the parked
table reference VE. This reduces the optimization time when a genetic algorithm is used.
In the current form, however, the inputs can be arranged such that least squares can be
used. First the reference parked VE for every experimental data point is found. Next the
product of each term in the multiplicative function with the parked VE is calculated. Each individual term in the additive function is also found. Finally all of the terms in the multiplicative and additive functions are combined together to form a matrix. The columns of the matrix correspond to different products of the operating conditions and the parked VE. For example the parked VE multiplied by the square of engine speed is a term and so is the product of ICAM and ECAM. Because least squares can be used, the optimization coefficient can be found nearly instantaneously. Additionally, having a form which has a good guess for optimization via a genetic algorithm provides flexibility to optimize using any criteria. For certain applications this may be very important to reduce the maximum errors to within a range. This criteria cannot necessary be achieved with the least squares solution. However by adjusting the cost function to include the maximum error, it can be found using a genetic algorithm.

The optimal coefficients were found using both a genetic algorithm and least squares. When the cost function for the genetic algorithm was the summation of the square errors, the two optimization techniques provided the same solution. The least squares solution provided great results that were within the desired error range. Table 3 compares the VE errors between just the parked reference table and the full quadratic VE function. Modifying the parked VE reference table with the additive and multiplicative quadratic functions reduced the VE errors significantly. Because the VE function contained a constant term ($b_{15}$), the error also had a zero mean. The standard deviation of the VE error was only 1.882% which is better than the desired 2%.
Although the overall accuracy of the VE function was very good, it did have errors when the cams were parked. Obviously the parked table VE could be directed used instead of the complete VE function, however this would introduce discontinuities. Every time the cams entered or left parked, the volumetric efficiency prediction would jump. Figure 30 and Figure 31 show the VE error distribution for the complete operating space and just for parked respectively. In general the VE function overestimated the VE at parked by about 1 – 2 %. A sudden change of 1 – 2 % in volumetric efficiency could be felt by the vehicle’s operator. Therefore, it is desirable to remove the error at parked without introducing any discontinuities. Another minor issue is the number of terms in the VE function. If some of the terms could be removed, then the ECU would be taxed less. Since the volumetric efficiency is calculated every engine cycle for every cylinder, reducing the computational complexity even slightly can have a significant benefit.
4.7 Reduced Quadratic VE Function

To solve these problems many modified forms of the full quadratic VE function were tested. The biggest issue was removing the VE error at parked. When the cams are parked nearly all of the terms in the full quadratic VE function reduce to zero. The only remaining terms are MAP, MAP^2, RPM, RPM^2, MAP*RPM and the constant offset term. If all of these terms are removed, then the accuracy is reduced significantly. The manifold pressure and engine speed terms are significant in parts of the quadratic VE function. However most of the effects of manifold pressure and engines speed are already captured by the parked VE table. Only the effects of the intake and exhaust cam timing are not accounted for in the parked table. The statistical importance of manifold pressure and engine speed in the full quadratic VE function means that the impact of cam timing is not uniform with respect to manifold pressure and engine speed. By multiplying MAP, MAP^2, RPM, RPM^2, MAP*RPM with both ICAM and ECAM, this non-uniformity can be captured and the terms reduced to zero at parked. The only
remaining nonzero term at parked is the constant offset. Without the offset term, the VE function could have a nonzero mean error. However after calibration using least squares, the mean is nearly zero, so the offset term was also removed.

The other concern with the VE function was the number of terms. Using a stepwise linear regression for optimization instead of just least squares, the number of terms was reduced from 28 to 20. Only the most important terms were included in the model generated using a stepwise linear regression. Therefore the accuracy of the VE function was not adversely affected by removing these terms. After making these modifications, the VE function takes the form shown in (23).

\[
f_{\text{mult}}(ICAM, ECAM, MAP, RPM) = 1 + a_1 ICAM + a_2 ECAM + a_3 ICAM \ast ECAM + a_4 ICAM^2 + a_5 ECAM^2 + a_6 ICAM \ast MAP + a_7 ECAM \ast MAP + a_8 MAP^2 \ast ICAM \ast ECAM + a_9 RPM \ast ICAM \ast ECAM + a_{10} RPM^2 \ast ICAM \ast ECAM
\]

\[
f_{\text{add}}(ICAM, ECAM, MAP, RPM) = b_1 ICAM + b_2 ECAM + b_3 ICAM \ast ECAM + b_4 ICAM^2 + b_5 ECAM^2 + b_6 ICAM \ast MAP + b_7 ECAM \ast MAP + b_8 RPM \ast ICAM + b_9 RPM \ast ECAM + b_{10} RPM \ast MAP \ast ICAM \ast ECAM
\]

\[
\text{VE}(ICAM, ECAM, MAP, RPM) = \text{VE}_{\text{parked}}(MAP, RPM) f_{\text{mult}}(ICAM, ECAM, MAP, RPM) + f_{\text{add}}(ICAM, ECAM, MAP, RPM)
\]

From now on this equation will be referred to as the reduced quadratic VE function.

Even though the reduced quadratic VE function has fewer terms than the full quadratic VE function, the accuracy was actually better. The terms removed were statistically insignificant and the error at parked was removed by adjusting some of the terms. As a result the overall errors were small although not exactly zero mean. Figure 32 shows the error distribution over the entire operating space. The error for only the operating conditions with parked cams is shown in Figure 33. The parked reference table only has a finite number of entries (17 by 17) so the error is not exactly zero. Linearly
interpolating between table entries introduces a small amount of error. Table 4 compares the errors between the full quadratic and reduced quadratic VE functions. In nearly every respect the reduced quadratic VE function performs better. Not only are the overall standard deviation and the average absolute error smaller, but also the error at parked is almost zero. Moreover, this is accomplished with fewer terms. The only downside is that the reduced quadratic VE function did have a slightly larger maximum error.

![Figure 32: VE Error for Reduced Quadratic Model after Optimization](image1)

![Figure 33: VE Error at Parked for Reduced Quadratic Model after Optimization](image2)

<table>
<thead>
<tr>
<th>All Units %</th>
<th>Full Quadratic VE Function</th>
<th>Full Quadratic VE Function (Parked Only)</th>
<th>Reduced Quadratic VE Function</th>
<th>Reduced Quadratic VE Function (Parked Only)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Standard Deviation</td>
<td>1.882</td>
<td>1.812</td>
<td>1.733</td>
<td>0.207</td>
</tr>
<tr>
<td>Average Error</td>
<td>0.00</td>
<td>-0.8426</td>
<td>-0.0152</td>
<td>-0.0314</td>
</tr>
<tr>
<td>Average Absolute Error</td>
<td>1.31</td>
<td>1.633</td>
<td>1.12</td>
<td>0.0464</td>
</tr>
<tr>
<td>Max Error</td>
<td>15.9</td>
<td>9.245</td>
<td>16.38</td>
<td>1.842</td>
</tr>
</tbody>
</table>

Table 4: VE Error Comparison between Full Quadratic and Reduced Quadratic VE Functions
4.8 Vehicle Validation

After the final form of the VE function was chosen and calibrated with GT-Power data, it was compared against vehicle data. The main use of volumetric efficiency is dynamically predicting the trapped air mass per cylinder. To completely validate the VE function, it needs to be validated in the environment where it will be used (in a vehicle). The original calibration of the GT-Power model was done using experimental data taken on an engine dynamometer. This same fully instrumented engine was transplanted to a vehicle. To keep costs down, the quality of sensors used in a production vehicle are slightly less accurate than the ones used in the engine dynamometer. Trying to reproduce a production vehicle as close as possible, production quality sensors (instead of lab grade) were used to calculate volumetric efficiency. Most production vehicles do not have a thermocouple in the intake system to measure the intake air temperature; they use a model of the air temperature. Designing and calibrating an intake air temperature model was outside of the scope of the thesis, so a direct thermocouple measurement was used. Excluding this difference, the setup used to validate the VE function on-vehicle is production equivalent.

The correlation between an engine on a dynamometer and on a vehicle is not as straightforward as it might appear. All of the intake and exhaust piping is the same in both cases, but the thermal environment can be significantly different. On an engine dynamometer the engine is in an open room containing near ambient, motionless air. Conversely, an engine in a vehicle is enclosed by the hood and body panels of the vehicle. Through radiation from the engine block, the under-hood environment is much hotter and less uniform than the test cell. The temperature of the engine and therefore the
radiative heat transfer rate is not controlled as tightly in a vehicle as it is in a test cell. In a test cell a cooling tower connected to a cold water source regulates the coolant temperature. To remove the heat from an engine on a vehicle a radiator and fan are used. Although some of the heat is removed by air circulation generated by the radiator fan, the hood on a vehicle can trap a significant amount of heat. When a vehicle is moving, air flows through the front air dam and around parts of the engine. The mass flow rate of this air is dependent on the vehicle speed and the speed of the radiator fan. The convective heat transfer resulting from this airflow is not present in an engine dynamometer test cell.

Even though volumetric efficiency is ideally independent from temperature, this is only true if the temperature profiles of the intake and exhaust systems are similar. Because the under-hood environment affects each section of piping differently, the volumetric efficiency can change from dynamometer to vehicle. The temperature profiles also have more variation on a vehicle due to changes in vehicle speed, fan speed, coolant temperature and loading. Not only must the engine power the vehicle, but also any accessories like the air conditioning system. Because the GT-Power model was calibrated with dynamometer data, these differences have to be overcome. In addition the sensors used on a vehicle are not as accurate. As a result the volumetric efficiency errors of any VE function will be larger. If these errors are unacceptable, the GT-Power model can be tuned based on vehicle data. Any number of parameters can be modified to improve the correlation between the GT-Power model (VE function) and the vehicle. Examples of tunable components are the surface roughness, pipe diameter, pipe length, bend angle, discharge coefficient and discretization length. If this type of tuning is
required, then the operating space will have to be simulated again. Using this data, the VE function will then have to be recalibrated.

To validate the reduced quadratic VE function generated in the thesis, VE data from several city and highway drive cycles was recorded. As described earlier, VE can only be measured in steady-state. Therefore, transient data points were removed from the data. The definition of steady-state on a vehicle is much less stringent than on an engine dynamometer. Steady-state on a vehicle is defined as having the same cycle-to-cycle air dynamics. Many quantities such as vehicle speed, intake wall temperature, coolant temperature and radiator fan speed could be varying during steady-state in this definition. These possible sources of variation contribute to small unavoidable errors in VE.

The steady-state (as define above) data points for the combined city and highway drive cycles covered the majority of the operating space. Figure 34 shows the steady-state manifold pressures, engine speed combinations encountered in one of the datasets. The agreement between the VE function and the vehicle were just outside the range of desired accuracy (within 3% standard deviation). Most of the large errors were at the extreme corners of the operating space. By tuning the GT-Power model slightly, however, very good results were achieved. The VE predicted by the VE function is compared to the true VE measured on the vehicle in Figure 35. The VE function is able to track the true vehicle within a standard deviation of 3%.
Figure 34: Steady-state Operating Points for the Validation Dataset

Figure 35: Comparison between the VE Measured On-Vehicle and the Predicted VE from the VE Function (Drive Trace 2)
Figure 36 shows the distribution of the VE errors between the VE function and the vehicle for this drive trace. The error statistics are shown in Table 5. Although the errors in this dataset had a nonzero mean, the offset was small. Other datasets had offsets of a similar magnitude but in the opposite direction. Depending on the operating conditions of the drive cycle the offset changed. This is an indication of the overall correlation between the VE function and vehicle is nearly zero mean. The magnitude of these errors is comparable with the current practice VE representations. However, the necessary calibration space is considerably less than what a four-dimensional table representation would require.

![Figure 36: VE Error in Validating the VE Function with Vehicle Data (20 Coefficients)](image-url)
### Table 5: VE Error of the VE Function Compared to Vehicle Data

<table>
<thead>
<tr>
<th></th>
<th>Vehicle to VE Function Comparison</th>
</tr>
</thead>
<tbody>
<tr>
<td>Standard Deviation</td>
<td>2.833</td>
</tr>
<tr>
<td>Mean Error</td>
<td>0.656</td>
</tr>
<tr>
<td>Mean Absolute Error</td>
<td>2.375</td>
</tr>
</tbody>
</table>

#### 4.9 Robustness of VE Function and Calibration Process

One of the largest benefits of the VE representation is its ability to smooth noisy vehicle data. To generate the most accurate volumetric efficiency model, the most accurate volumetric efficiency data should be used. Both GT-Power and real vehicle data have some uncertainty. GT-Power data is precisely able to predict the volumetric efficiency of an engine model. However, the engine model cannot exactly represent the physical engine. The predicted volumetric efficiency data is very smooth but may not have the exact curvature of the physical engine. Measuring the volumetric efficiency on a vehicle has the potential to be a more faithful representation of the engine. However the volumetric efficiency measurement on a vehicle is subject to the limitations of the sensors. In many cases a vehicle is not driven on a chassis dynamometer to generate volumetric efficiency. Specifically targeting an operating condition and waiting for steady-state on a chassis dynamometer is expensive and time consuming. In addition, the air flowing through the engine compartment as a car is moving cannot easily be replicated.
On a production vehicle, volumetric efficiency will be used to provide an open loop estimate of the trapped air mass during transient driving conditions. When a vehicle is in steady-state, the mass air flow sensor can be used to provide the air prediction. Therefore most volumetric efficiency data is extracted from measurements recorded while a vehicle is driven in the city or on a highway. To extract volumetric efficiency from a dataset, accurate air per cylinder, air temperature and manifold pressure measurements are needed. All of these sensors are susceptible to noise and sensor dynamics. As previously stated, a mass air flow sensor is only accurate when the engine is operating in steady-state. However the algorithm used to detect steady-state may misinterpret unsteady conditions as steady. Even after filtering the signals, these error sources create a noisy volumetric efficiency measurement. Figure 37 shows the volumetric efficiency surface generated from vehicle data for parked cams. This data came from a 1.6 liter four cylinder engine with cam phasing. The jaggedness of the surface is a direct result of the noisy volumetric efficiency measurement.
The reduced quadratic VE function is inherently smooth. If vehicle data is used to calibrate the volumetric efficiency model, then the artifacts from the noisy data are removed leaving only the underlying volumetric efficiency features. To illustrate this effect, vehicle data taken from a 1.6 liter engine with cam phasing was used to calibrate the reduced quadratic VE function. The data used to create the parked VE surface plot shown above came from the same engine. Table 6 shows the distribution statistics of the raw VE and the VE relative to the parked table. Compared to the 2.4 liter engine, the overall distribution of VE for the 1.6 liter engine is considerably less (8.02% versus 14.5%). However the error distribution when only the parked VE table is used as a model is much higher (6.35% versus 4.3%). Therefore the multiplicative and additive functions have much more error to remove. This wide of a distribution is an indication that the vehicle VE data has a high degree of uncertainty and is not very repeatable.
<table>
<thead>
<tr>
<th>All Units %</th>
<th>Raw Vehicle Data</th>
<th>All Units %</th>
<th>Parked Table Reference Only</th>
<th>Reduced Quadratic VE Function</th>
</tr>
</thead>
<tbody>
<tr>
<td>Standard Deviation</td>
<td>8.02</td>
<td>Standard Deviation</td>
<td>6.35</td>
<td>3.01</td>
</tr>
<tr>
<td>Mean</td>
<td>50.99</td>
<td>Mean Error</td>
<td>3.37</td>
<td>-0.08</td>
</tr>
<tr>
<td>Average Absolute Difference from Mean</td>
<td>50.99</td>
<td>Mean Absolute Error</td>
<td>5.77</td>
<td>2.36</td>
</tr>
<tr>
<td>Min, Max</td>
<td>21.8, 72.32</td>
<td>Max Error</td>
<td>33.19</td>
<td>25.54</td>
</tr>
</tbody>
</table>

Table 6: Volumetric Efficiency Comparison between Vehicle Data and Reduced Quadratic VE Function Calibrated from Vehicle Data

Even with this variability the reduced quadratic VE function was able to represent the data well. The standard deviation as also shown in Table 6 is right at the desired 3% metric and the mean is nearly zero. The error distribution for the raw VE data (Figure 38) and just the parked VE models (Figure 39) were very wide. However the spread was reduced substantially when the reduced quadratic VE function was applied (Figure 40). Some of the “error” in the reduced quadratic VE function is actually do to measurement errors, so the true correlation between the representation and the engine’s VE is even better. Regardless of the origin of the data (dynamometer, vehicle or GT-Power), the reduced quadratic VE function provides an accurate fit.
4.10 Applying the Model to Other Engine Platforms

Although the VE model was designed for an engine with dual independent cam phasing, it can be adapted to work for almost any platform. For forced induction systems, the model can remain in the exact same form. Compared to naturally aspirated engines, forced induction engines can achieve higher manifold pressures. The
relationship between volumetric efficiency and manifold pressure is very well defined. By simply expanding the manifold pressure range and spacing in the parked volumetric efficiency table (and possibly adding more entries), the model can capture the additional effects of forced induction on volumetric efficiency. Even engines with multi-staged cam profiles or intake tuning valves can use the same volumetric efficiency model. An engine with two-staged cams can be thought of as two engines with different cam profiles. Instead of having a single volumetric efficiency model, a model would be needed for each cam profile. Although more calibration space is needed, the function’s memory requirement is not much for an engine with 5 degrees of freedom.
CHAPTER 5

FORMULATION OF NEW RESIDUAL GAS FRACTION REPRESENTATION

5.1 Comparison between Existing Models and GT-Power Predicted Residual Gas Fraction

The residual gas fraction of an operating condition is a non-dimensional quantity representing the ratio of trapped residual gas mass to the total trapped mass. In this thesis the residual gas fraction is represented as a decimal and therefore ranges from 0 to 1 in magnitude. A residual gas fraction of 1 corresponds to the complete absence of fresh air. Before a new model form was explored, the physics models generated by MIT and Senecal were compared to the residual gas fraction predicted by the GT-Power model. As a first comparison the suggested coefficients for the MIT and Senecal equations were used. Using these coefficients produced abysmal results. Both models had near zero mean errors but large standard deviations. The MIT equation provided a particularly poor approximation of the residual gas fraction; the standard deviation was 0.0511. It is important to remember that complete combustion only occurs when the residual gas fraction is less than 0.30. If GT-Power was not available to predict the residuals, then the coefficients of either model could not be optimized. This would result in errors of this
magnitude. Obviously these errors limit the possible uses of the residual gas fraction in engine control.

Using GT-Power to predict the residual gases provides the opportunity to optimize these models. Both models were optimized using least squares. Compared to the suggested coefficients, the optimized coefficients enhanced the model’s fit significantly. For this improvement to occur, the magnitude of the coefficients changed markedly. Table 7 compares the results from the base MIT model to the optimized model. The $C_1$ coefficient increased by an order of magnitude, whereas the $C_2$ coefficient decreased by about 30 percent. As previously stated, the MIT model was only validated with two cam profiles and both of them had low valve overlap factors (0.52 and 1.43). The effect of cam timing via the overlap factor is only captured in the first term of the model. Because of the engine studied in this thesis can have such large overlap factors (>10), it is not that surprising that the optimal $C_1$ coefficient changed so much. This large change, however, is an indication of the inadequacies of the MIT model. The distribution of errors for the model is shown in Figure 41. Although optimization reduced the standard deviation of the error from 0.0511 to 0.0343, the errors are still unacceptable.
The optimization of the Senecal equation produced similar results. The error distribution is shown in Figure 42, and the error statistics are organized in Table 8. Even though the Senecal equation is much more complex and contains an additional calibration factor, the standard deviation of the error (0.0316) was only slightly better than the MIT model. When factoring ease of implementation, the Senecal equation’s slightly improved
accuracy is strongly outweighed by its increased complexity. The Senecal equation has several additional exponential terms and many more multiplications and divisions. Evaluating the Senecal equation in an ECU would be several times more computationally expensive than the MIT equation.

![Residual Gas Fraction Error for Optimized Senecal Model](image)

**Figure 42: Residual Gas Fraction Error for Optimized Senecal Model**

<table>
<thead>
<tr>
<th>Method</th>
<th>Senecal, Xin, Reitz. Method</th>
</tr>
</thead>
<tbody>
<tr>
<td>Standard Deviation</td>
<td>0.0316</td>
</tr>
<tr>
<td>Average Error</td>
<td>-1.50E-04</td>
</tr>
<tr>
<td>Average Absolute Error</td>
<td>0.0239</td>
</tr>
<tr>
<td>Max Error</td>
<td>0.1677</td>
</tr>
<tr>
<td>C1</td>
<td>0.025</td>
</tr>
<tr>
<td>C2</td>
<td>1.7132</td>
</tr>
<tr>
<td>Alpha</td>
<td>0.8697</td>
</tr>
</tbody>
</table>

**Table 8: Residual Gas Fraction Error Statistics for Senecal Function**
5.1.1 Shortcomings of Current Physics-based Models

Even though the current residual gas models are based on physics, they contain several approximations. Truly representing the residual gas fraction with physics would require the independent integrations of the mass flow rates of residuals and fresh air. The current models are basically mean value models (MVM). Besides the calibration coefficients, the valve overlap factor is the least physically based component. Although the major trends are captured by the overlap factor, it does not account for cam timing and wave dynamics. If both the intake and exhaust cam timings are shifted in the same direction by the same amount, then the valve overlap factor would remain constant. However the residual gas fraction could change dramatically. Consider the case when the exhaust valve closes nominally at top dead center and both timings are advanced by the same amount. The cylinder volume when the exhaust valve closes will be larger, so the mass of exhaust gas left in the cylinder after the exhaust valve closes will also be larger. This will translate into a higher mass of residual gases when the intake valve closes and a higher residual gas fraction. The overlap factor also presents another problem because it is not a constant value. A table indexed by the intake and exhaust cam positions must be used to store the overlap factor.

5.2 Data Trend between Operating Conditions and Residual Gas Fraction

Physics-based models are often preferred to empirical models because they perform better outside their original calibration region. Extrapolating the trends captured in an empirical model can be very dangerous. When high order polynomials are used, the accuracy of an empirical model degenerate quickly the further outside the calibration
range it is evaluated. In the case of the MIT model where only a few data points could be taken, a physics-based model is appropriate. GT-Power can rapidly simulate every possible operating condition, so the problem of extrapolation is removed. Therefore, an empirical model was explored. To create an accurate empirical model, the model form must be able to capture the influence of each variable. Before any model forms were proposed, the relationship between the data and the operating conditions was explored. Exhaust pressure impacts the trapped residual mass, but the variation is minimal compared to the variation in intake manifold pressure. For a non-turbocharged engine, the normal range of exhaust pressures is from about 1 to 1.2 bar. Most vehicles do not directly measure exhaust pressure either. For these reasons, the explicit effect of exhaust pressure was left out of the model.

Unlike volumetric efficiency which has a complex dependence on engine speed, the residual gas fraction of an engine is relatively insensitive to engine speed. Wave dynamics influences the residual gas mass but it affects the fresh air mass. For most cam timing pairs, the intake valve is open when the exhaust valve closes. Therefore, both the fresh air and the combustion gases are affected by the wave dynamics of the exhaust system. As the engine speed changes the effect of the wave dynamics of the exhaust system changes, but the residual gas fraction remains relatively constant.

The operating condition with the strongest and most regular influence on residuals is manifold pressure. If the manifold pressure is very low, then the pressure differential between the cylinder and the intake manifold is large. The flow rate of combustion gases back into the intake is proportional to this pressure difference, so the concentration of combustion gases in the intake is higher. Moreover, at low manifold pressures an engine
is not able to efficiently induct the air/residual gas mixture very well. The combustion chamber already contains some combustion gases that were not forced out during the exhaust stroke and only a small mass of the air/residual gas mixture fills the combustion chamber during the intake stroke. Therefore a large portion of the gas in the cylinders at the end of the intake stroke is residual gas. This combination of effects causes the residual gas fraction to increase dramatically at low manifold pressures. Figure 43 highlights the relationship between manifold pressure and residual gas fraction at parked. Figure 43 is a projection of this effect for every engine speed. Across the entire operating range, engine speed only causes about 1 to 2 percent variation in residual gas fraction.

![Figure 43: Projection of Engine Speed’s Influence on Residual Gas Fraction onto the MAP, Residual Gas Fraction Plane](image)

Cam timing obviously has a major impact on the residual gas fraction. To capture this effect, the previous models used an overlap factor. This shape could be reproduced using quadratics although the change would not be as sharp. The raw data
also showed that the cam timing has a translational effect on the dependence between MAP and the residual gas fraction. The residual gas fraction is still inversely dependent on MAP regardless of the cam timing. However, the curve is shifted to the right and shifted up relative to parked depending on the cam timing.

5.3 New Residual Gas Fraction Model

Based on the conclusions drawn from the residual gas fraction data generated by GT-Power, a new residual gas fraction model was designed. The most important effect to capture was the dependence of residuals on manifold pressure. Clearly they have some type of inverse relationship. To determine what type of functional form approximated the relationship best, several curve fits of the data shown in Figure 43 of the previous section were performed. The forms tested included an inverse square relationship, an inverse square root relationship and a simple inverse relationship. The inverse square root and simple inverse approximations performed the best. Because the simple inverse is less complex, it was chosen over the inverse square root approximation. In addition to testing these forms, an optimization of the exponent of the manifold pressure was also performed. The best exponent was found to be 0.73, however it did not provide any significant improvement over the simple inverse. Therefore the simple inverse manifold pressure relationship was chosen to be the basis of the residual gas fraction representation.

As stated in the previous section, the cam timing essentially shifts the manifold pressure relationship up and right relative to parked. By including a multiplicative and additive function of cam position, this effect can be captured. Although the engine speed
did not show a significant effect on the residual gas fraction, it was included in the model.

The two terms multiplying and adding to the inverse manifold pressure function are full quadratic regressions of cam timing and engine speed. Equation (24) shows the form of the new residual gas fraction model.

\[
RGF = \frac{A(ICAM, ECAM, RPM)P_{baro}}{MAP} + B(ICAM, ECAM, RPM)
\]

\[
A(ICAM, ECAM, RPM) = a_0 + a_1 ICAM + a_2 ECAM + a_3 RPM + a_4 ICAM * ECAM + a_5 ICAM * RPM + a_6 ECAM * RPM + a_7 ICAM^2 + a_8 ECAM^2 + a_9 RPM^2
\]

\[
B(ICAM, ECAM, RPM) = b_0 + b_1 ICAM + b_2 ECAM + b_3 RPM + b_4 ICAM * ECAM + b_5 ICAM * RPM + b_6 ECAM * RPM + b_7 ICAM^2 + b_8 ECAM^2 + b_9 RPM^2
\]

Barometric pressure is included in the multiplicative term to provide an altitude correction. The effect of this term will be detailed in the next section. From now on this model will be referred to as the inverse MAP residual gas fraction model.

Using the residual gas fraction predicted by GT-Power, the inverse MAP residual gas fraction model was globally calibrated with least squares. A step wise regression was not used because the number of coefficients was already low, and the calculation is less taxing than the physics models. Even so, the fit was considerably better than the physics models. The error distribution shown in Figure 44 had a zero mean and a standard deviation of 0.0193. The error statistics can be referenced in Table 9. Compared to the MIT model, the average absolute error was reduced by nearly half (0.026 compared to 0.0136). Even though the inverse MAP residual gas fraction model was not explicitly derived from physical laws, the representation captured each variable’s impact on residuals remarkably.
5.4 Barometric Pressure Compensation

Although the accuracy at standard atmospheric pressure was very good, excluding exhaust pressure from the residual gas fraction model could cause a problem at high altitudes. At high altitudes the ambient pressure could be significantly lower than at sea level. In the United States, it is possible to encounter barometric pressures ranging from 60 to 100 kPa. The exhaust pressure will be similarly impacted. The residual gas
fraction is dependent on the difference between the intake manifold pressure and the exhaust pressure. Without an altitude correction factor the residual gas fraction model would inherently assume an exhaust pressure of about 110 kPa, since the model was calibrated with data at an atmospheric pressure of 100 kPa. This would result in gross over prediction of the residuals. The physics-based models account for altitude effects naturally with an exhaust pressure term. For the inverse MAP residual gas fraction function, the barometric pressure term corrects for altitude effects. To verify its effectiveness, a separate validation exercise was performed. Simulating different barometric pressures in GT-Power is as simple as changing the ambient pressure. For this secondary validation, the residual gas fraction model was calibrated at a barometric pressure of 99 kPa. For means of comparison, a model without the barometric pressure term was also calibrated. Figure 45 shows the distribution of error for the calibration dataset which is the same for both models.

Figure 45: Residual Gas Fraction Error Distribution for the Calibration Region (99 kPa)
Next, the barometric pressure in GT-Power was changed to 80 kPa and several thousand operating conditions were simulated. The residual gas fraction data was then compared to the two models calibrated at 99 kPa. The error distributions for the model without barometric compensation and the model with barometric pressure are shown in Figure 46 and Figure 47. As expected the model without any altitude correction overpredicted the residual gas fraction resulting in a mean error of -0.0294. The mean error for the model with an altitude correction was nearly zero. The standard deviation of the error was also much larger for the uncorrected model (0.0284 compared to 0.0176).

![Figure 46: Residual Gas Fraction Error Distribution without Barometric Correction Factor (80 kPa)](image1)

![Figure 47: Residual Gas Fraction Error Distribution with Barometric Correction Factor (80 kPa)](image2)

Another dataset at 60 kPa was also simulated in GT-Power. Again the residual gas fraction predicted in GT-Power was compared to ones predicted by the two models. At a barometric pressure nearly 40% lower than the one used to calibrate the model, the model without an altitude correction factor performed abysmally. The model overpredicted the residual gas fraction by 0.0778 on average. Not only did the model have a major error bias, but it also had a very wide error distribution (Figure 48).
corresponding to a standard deviation of 0.0434. Approximating the residual gas fraction as a constant would be a better representation in this case. The results for the residual gas fraction model with the barometric pressure correction term were still great. As seen in Figure 49 the error distribution is almost identical to the calibration dataset and the one at 80 kPa. The mean of the distribution was less than 0.005 and the standard deviation was 0.0176. Table 10 compares all of the error statistics for the each barometric pressure and each model. Clearly including the barometric compensation factor is necessary. The error statistics are actually better at 80 and 60 kPa than the calibration dataset at 99 kPa, thereby showing the model is invariant to barometric pressure.

Figure 48: Residual Gas Fraction Error Distribution without Barometric Correction Factor (60 kPa)

Figure 49: Residual Gas Fraction Error Distribution with Barometric Correction Factor (60 kPa)
<table>
<thead>
<tr>
<th>Barometric Pressure (kPa)</th>
<th>Without Correction</th>
<th>With Correction</th>
<th>Without Correction</th>
<th>With Correction</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>99</td>
<td>80</td>
<td>80</td>
<td>60</td>
</tr>
<tr>
<td>Standard Deviation</td>
<td>0.0197</td>
<td>0.0284</td>
<td>0.0176</td>
<td>0.0434</td>
</tr>
<tr>
<td>Average Error</td>
<td>3.06E-16</td>
<td>-0.0294</td>
<td>9.04E-04</td>
<td>-0.0778</td>
</tr>
<tr>
<td>Average Absolute Error</td>
<td>0.0145</td>
<td>0.0308</td>
<td>0.0134</td>
<td>0.0778</td>
</tr>
<tr>
<td>Max Error</td>
<td>0.1317</td>
<td>0.1025</td>
<td>0.0945</td>
<td>0.199</td>
</tr>
</tbody>
</table>

Table 10: Effect of Barometric Pressure on the Inverse MAP Residual Gas Fraction Model

5.5 Comparison between Inverse MAP Residual Gas Fraction Model and Physics-based Models

Not only does the inverse MAP residual gas model have a smaller error distribution than the physics-based models, but it is also invariant to manifold pressure. To illustrate the effectiveness of the inverse MAP representation, the effect of cam timing was explored at a representative operating point. Because the residual gas fraction can easily exceed 0.30 at low manifold pressures, a manifold pressure of 73 kPa and an engine speed of 4600 RPM were chosen. The globally optimized versions of each model were used for this comparison. The true impact of cam timing on residual gas fraction as predicted by GT-Power is shown in Figure 50. As expected the highest residual gas fraction occurs when the valves have the most overlap (ICAM = 25 cam degrees, ECAM = 25 cam degrees). However it may be slightly surprising to see that the minimum value did not occur at parked. Although the residual gas fraction at parked was small, the
The smallest value was at (ICAM = 0 cam degrees, ECAM = 12 cam degrees). The valve overlap for this cam pair is still almost zero, but the exhaust valve closes at about top dead center. Therefore there is almost no reverse flow and the cylinder volume at exhaust valve closing is at a minimum.

Because the MIT model is based on the valve overlap factor which is independent of valve position relative to cylinder position, the predicted residual gas fractions for low valve overlap cam pairs is poor. As shown in Figure 51, the minimum is incorrectly predicted to be at parked. The effect of intake and exhaust cam timing is also predicted to be symmetric by the MIT model. Clearly, the true residual gas fraction is not symmetric. Another major difference is the magnitude of the residual gas fraction at

Figure 50: Residual Gas Fraction Predicted by GT-Power (73 kPa, 4600 RPM)
fully actuated cams. The MIT model predicts a residual gas fraction of 0.16 compared to the true value of 0.10.

Starkly contrasting the MIT model, the inverse MAP residual gas fraction model represents the true fraction very well. The predicted residual gas fraction surface presented above in Figure 52 has the same curvature as the true surface. Firstly, the minimum value is correctly predicted at ICAM = 0 cam degrees and ECAM = 12 cam degrees. Secondly, the local curvature at the four extreme corners is very similar to the actual residual gas fraction surface. Not only are the trends accurately predicted, but also the magnitudes are very accurate too.

The residual gas fraction error surfaces also demonstrate the superior accuracy of the inverse MAP residual gas fraction model. The errors for the MIT model (Figure 53) are as high as 0.06 and not zero mean. The regularity of the error surface indicates that simple trends were missed. Unlike the MIT model, the inverse MAP model (Figure 54) has a zero mean error and all errors are within 0.015. Although the errors are smooth,
they are not nearly as easy to represent as the MIT model’s errors. This signifies that the inverse MAP model captures the trends of the residual gas fraction as well as quadratic terms could.

Figure 53: Residual Gas Fraction Error for the MIT Model

Figure 54: Residual Gas Fraction Error for the Inverse MAP Model
CHAPTER 6

CONCLUSIONS AND RECOMMENDATIONS

6.1 Conclusions and Recommendations

This thesis focused on the extensive use of suitably calibrated engine simulation software to estimate the gas charge (fresh air and residual) in modern spark ignition (SI) engines. Modern engines with multiple additional degrees of freedom (due to an ever richer actuator set) are a challenge to efficiently calibrate by traditional direct engine mapping, as this approach suffers from the “curse of dimensionality”. On the other hand, one-dimensional wave dynamics engine simulation software is a very robust and powerful tool especially for complex engines. The methodology developed in this thesis was investigated and validated for a modern production SI engine with dual independent cam phasing. The goal is to replace most of the experimental data collection procedures with engine simulations so that the overall calibration time and costs can be considerably reduced. However, the benefits of this approach can only be realized if the engine simulations are a good proxy for reality. This can only achieved if the engine model can be suitably calibrated with a minimal experimental data set yet yield models with the required level of accuracy over the entire range of operating conditions of interest. Even relatively simple one-dimensional wave dynamics engine models (such as the GT-Power
software used in this study) have a very large numbers of parameters to adjust or tune to match real engine behavior. The challenge is to focus on a limited numbers of parameters which have a high level of influence on the output quantities of interest (volumetric efficiency and residual gas fraction in this case). By restricting the accuracy requirements to only these gas charge variables and systematically investigating the sensitivities of the model outputs to its various parameters, it is possible to “suitably” calibrate such a model with a very limited experimental data set. Using less than 1,000 experimental data points to suitably calibrate select elements of the model, it is possible to generate a globally valid model to predict gas charge parameters and use this model to generate a virtual engine map consisting of several hundred thousand operating points. Engine simulations can target any operating condition in the high-dimensional actuator space (4-D in this study), even non-physical ones, to provide a complete mapping of the output charge variables with the desired grid spacing. All possible operating conditions encountered by an engine can be used to calibrate a volumetric efficiency and residual gas fraction model. Therefore, these empirical regressions resulting from such virtual mapping can be free of any extrapolation errors.

In this thesis, both the VE and residual gas fraction maps were evaluated using the predicted values obtained by an appropriately calibrated GT-Power model. Very importantly, it should be noted that while fresh charge can be measured experimentally at steady-state with conventional instrumentation, residual gas fractions which are internal to the cylinder are nearly impossible to measure directly, lest some very expensive research grade instrumentation and very complex experimental procedures completely ill-suited for a calibration-production environment. Hence, the GT-Power model-based
approach used in this thesis intrinsically yields a richer data set than could be achieved in a calibration test cell.

However, this type of detailed mapping in multiple dimensions (regardless of whether it is done physically or obtained through a virtual mapping by extensive simulations) becomes a challenge to store in on-board controllers (ECUs, ECMs …) due to balancing storage space and accuracy requirements with computational burden for evaluation of the charge at each engine event. To address these storage, accuracy and computationally efficient representation issues, new representations were developed. For the fresh charge, the VE data uses a reduced quadratic VE representation by which the VE can be represented with the same level of high accuracy of the current methodologies, requires minimal calibration storage space (tables) and has a relatively low computational burden which can be handled by today ECUs. For the residual gas data, an inverse MAP residual gas fraction representation was used which also has an accuracy which is much better than the current models and is invariant to barometric pressure changes. This increased accuracy could expand the possible control uses of residual gas fraction estimates.

Both new representations (fresh charge and residual gas fraction) match or exceed current accuracy requirements. Furthermore, a significant benefit of this process is its systematic and robust approach. A properly calibrated GT-Power model can be used to represent any type of engine. Similarly, the VE and residual gas fraction representations are flexible enough to be applied to any engine platform including upcoming engine platforms with additional degrees of freedom (variable valve lift mechanism, etc.), where the number of independent variables prohibits or significantly impedes traditional “brute
force” mapping approaches. Hence, all of these benefits come with a significantly reduced experimental burden and a much reduced calibration time.

For the engine studied in this thesis, for example, the results depend on 4 dimensions (MAP, RPM, ICAM, ECAM). Collecting enough experimental data to fully map the volumetric efficiency of an engine with four degrees of freedom takes a significant amount of calibration effort as these mapping tables may represent 100,000 to 200,000 operating conditions. Instead of spending months in a test cell to collect data to extensively populate a four-dimensional space, a few days in a test cell are sufficient to gather the required model calibration data (less than 1000 operating conditions), automatically process this limited experimental data to calibrate the model and using a cluster of computers generate the virtual mapping results in a few days. As long as the engine model is properly calibrated, the errors associated with the target outputs (namely volumetric efficiency and residual gas fraction) are about the same or better than data collected experimentally.

The process outlined in this thesis considerably reduces the calibration effort (about 100 times less experimental data and an overall shortened process timeline). Aside from collecting and processing a small experimental dataset, this process could almost be completely automated. Because least squares or stepwise regressions are used to generate the sub-models for GT-Power, all of this can be done automatically. Using these models to generate the inputs to the GT-Power can also be done automatically. Even extracting output data from GT-Power and filtering it based on residual gas fraction, burn duration and convergence can be automated. Finally the coefficients of the
VE and residual gas fraction representations can be calibrated systematically using least squares.

In summary, the new proposed process (virtual mapping and enhanced representation of the both the fresh charge and residual gas fraction) which was investigated and validated in this thesis represents 1-2 orders of magnitude decrease in calibration time, while maintaining or exceeding overall accuracy and is readily extendable to other upcoming engine technologies with additional degrees of freedom.
LIST OF REFERENCES


