Engine thermal management with model predictive control

This item was submitted to Loughborough University's Institutional Repository by the/an author.

Additional Information:

- A Doctoral Thesis. Submitted in partial fulfilment of the requirements for the award of Doctor of Philosophy of Loughborough University.

Metadata Record: [https://dspace.lboro.ac.uk/2134/24274](https://dspace.lboro.ac.uk/2134/24274)

Publisher: © R.I. Abdul Jalal

Rights: This work is made available according to the conditions of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International (CC BY-NC-ND 4.0) licence. Full details of this licence are available at: [https://creativecommons.org/licenses/by-nc-nd/4.0/](https://creativecommons.org/licenses/by-nc-nd/4.0/)

Please cite the published version.
ENGINE THERMAL MANAGEMENT WITH MODEL PREDICTIVE CONTROL

By

RIFQI IRZUAN BIN ABDUL JALAL

A doctoral thesis submitted in partial fulfilment of the requirements for the award of Doctor of Philosophy of Loughborough University

December 2016

Supervisor;

Dr. THOMAS STEFFEN
Dr. ANDREW WILLIAMS

Department of Aeronautical and Automotive Engineering

© by R.I. Abdul Jalal 2016
Abstract

The global greenhouse gas CO\textsubscript{2} emission from the transportation sector is very significant. To reduce this gas emission, EU has set an average target of not more than 95 CO\textsubscript{2}/km for new passenger cars by the year 2020. A great reduction is still required to achieve the CO\textsubscript{2} emission target in 2020, and many different approaches are being considered. This thesis focuses on the thermal management of the engine as an area that promises significant improvement of fuel efficiency with relatively small changes.

The review of the literature shows that thermal management can improve engine efficiency through the friction reduction, improved air-fuel mixing, reduced heat loss, increased engine volumetric efficiency, suppressed knock, reduce radiator fan speed and reduction of other toxic emissions such as CO, HC and NO\textsubscript{x}. Like heat loss and friction, most emissions can be reduced in high temperature condition, but this may lead to poor volumetric efficiency and make the engine more prone to knock. The temperature trade-off study is conducted in simulation using a GT-SUITE engine model coupled with the FE in-cylinder wall structure and cooling system. The result is a map of the best operating temperature over engine speed and load. To quantify the benefit of this map, eight driving styles from the legislative and research test cycles are being compared using an immediate application of the optimal temperature, and significant improvements are found for urban style driving, while operation at higher load (motorway style driving) shows only small efficiency gains. The fuel consumption saving predicted in the urban style of driving is more than 4%.

This assess the chance of following the temperature set point over a cycle, the temperature reference is analysed for all eight types of drive cycles using autocorrelation, lag plot and power spectral density. The analysis consistently shows that the highest volatility is recorded in the Artemis Urban Drive Cycle: the autocorrelation disappears after only 5.4 seconds, while the power spectral density shows a drop off around 0.09Hz. This means fast control action is required to implement the optimal temperature before it changes again.
Model Predictive Control (MPC) is an optimal controller with a receding horizon, and it is well known for its ability to handle multivariable control problems for linear systems with input and state limits. The MPC controller can anticipate future events and can take control actions accordingly, especially if disturbances are known in advance. The main difficulty when applying MPC to thermal management is the non-linearity caused by changes in flow rate. Manipulating both the water pump and valve improves the control authority, but it also amplifies the nonlinearity of the system.

Common linearization approaches like Jacobian Linearization around one or several operating points are tested, by found to be only moderately successful. Instead, a novel approach is pursued using feedback linearization of the plant model. This uses an algebraic transformation of the plant inputs to turn the nonlinear systems dynamics into a fully or predominantly linear system. The MPC controller can work with the linear model, while the actual control inputs are found using an inverse transformation.

The Feedback Linearization MPC of the cooling system model is implemented and testing using MathWork™ Simulink®. The process includes the model transformation approach, model fitting, the transformation of the constraints and the tuning of the MPC controller. The simulation shows good temperature tracking performance, and this demonstrates that a MPC controller with feedback linearization is a suitable approach to thermal management. The controller strategy is then validated in a test rig replicating an actual engine cooling system.

The new MPC controller is again evaluated over the eight driving cycles. The average water pump speed is reduced by 9.1% compared to the conventional cooling system, while maintaining good temperature tracking. The controller performance further improves with future disturbance anticipation by 20.5% for the temperature tracking (calculated by RMSE), 6.8% reduction of the average water pump speed, 47.3% reduction of the average valve movement and 34.0% reduction of the average radiator fan speed.

**Keyword**: Thermal management, Engine cooling system, Fuel Economy, Model Predictive Control, Feedback Linearization, Drive cycles, Simulation and modelling
Acknowledgement

I would like to express my sincerest gratitude and appreciation to my supervisors Dr. Thomas Steffen for his supervision, endless support, invaluable suggestions and dedication throughout the writing of this thesis and the research that preceded it, and to my second supervisor Dr. Andrew Williams, who has given me guidance in doing research, being in helpful discussions, bring in professionalism and providing valuable input during this work. It is impossible to imagine how this thesis would have been written up without their guidance and assistance.

Special thanks also to the members of the Jury who accepted to assess this report and devoted their time to this.

I would like to acknowledge Iain Harber, Paul Reeves, Govinder Sharman, and Harshad Purohit technicians in the Aeronautic and Automotive Department for their assistance in the construction of the test rig. It is a great pleasure to be able to work with them, and to share expertise and ideas.

The dream of doing my PhD would not have been possible without the financial assistance I have received from University Kuala Lumpur and Majlis Amanah Rakyat.

I would also like to extend my gratitude to my friends Farraen Mohd Azmin, Ezhan Johanif and Tengku Nordayana Akma, who are always willing to discuss my work and give interesting comments and suggestions, altogether making the working environment enjoyable. Good luck on their future.

Lastly but of course not the least, I would like to express my utmost gratitude to my wife, Madihah Zakaria, my parents, Abdul Jalal Jafaar and Musalmah Baharudin, and my parents in law, Zakaria Mohd Noor and Fadmah for their understanding, endless patience, support and love throughout the years of my completing this work. To my sons Zaim Rusayidi and Amir Ruzaini, thank you for understanding and giving me the motivation to pursue my dream and ambition, and which is now becoming a reality.
# Table of Content

Abstract...................................................................................................................................................................................... i

Acknowledgement ........................................................................................................................................................................ iii

Table of Content ............................................................................................................................................................................ iv

List of Figures .................................................................................................................................................................................. viii

List of Tables .................................................................................................................................................................................. xx

Notation .......................................................................................................................................................................................... xxi

**CHAPTER 1**  
Introduction................................................................................................................................................................................. 1

1.1. Research Background and Motivation ................................................................................................................................. 1

1.2. Research Aims and Contributions ........................................................................................................................................... 9

1.3. Organization of the Thesis .......................................................................................................................................................... 11

**CHAPTER 2**  
Case of Thermal Management ....................................................................................................................................................... 14

2.1. Friction loss...................................................................................................................................................................................... 15

2.2. Combustion Quality .................................................................................................................................................................... 20

2.3. Emission........................................................................................................................................................................................ 24

2.4. Radiator ........................................................................................................................................................................................ 27

2.5. Summary....................................................................................................................................................................................... 27

**CHAPTER 3**  
Engine Thermal Management Potential ........................................................................................................................................ 30

3.1. Engine Model................................................................................................................................................................................. 31

3.2. Engine Calibration ...................................................................................................................................................................... 42

3.3. Comparison to the Conventional Cooling System .................................................................................................................. 51

3.4. Engine Thermal Management in Drive Cycles ...................................................................................................................... 53

3.5. Summary....................................................................................................................................................................................... 58
<table>
<thead>
<tr>
<th>CHAPTER 4</th>
<th>Drive Cycle Volatility Compared to Controller Response</th>
<th>59</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.1.</td>
<td>Evaluation of Drive Cycle Volatility</td>
<td>60</td>
</tr>
<tr>
<td>4.2.</td>
<td>Transient Response</td>
<td>70</td>
</tr>
<tr>
<td>4.3.</td>
<td>Current Mechanism reliability</td>
<td>79</td>
</tr>
<tr>
<td>4.4.</td>
<td>Summary</td>
<td>81</td>
</tr>
<tr>
<td>CHAPTER 5</td>
<td>Model Predictive Control</td>
<td>83</td>
</tr>
<tr>
<td>5.1.</td>
<td>Background &amp; Concept</td>
<td>84</td>
</tr>
<tr>
<td>5.2.</td>
<td>MPC Advantages for Engine Thermal Management</td>
<td>88</td>
</tr>
<tr>
<td>5.3.</td>
<td>MPC Challenges in Thermal Management</td>
<td>90</td>
</tr>
<tr>
<td>5.4.</td>
<td>Previous MPC Engine Thermal Management</td>
<td>91</td>
</tr>
<tr>
<td>5.5.</td>
<td>Summary</td>
<td>92</td>
</tr>
<tr>
<td>CHAPTER 6</td>
<td>Engine Modelling for MPC</td>
<td>94</td>
</tr>
<tr>
<td>6.1.</td>
<td>Model Objective</td>
<td>94</td>
</tr>
<tr>
<td>6.2.</td>
<td>Cooling System Model</td>
<td>95</td>
</tr>
<tr>
<td>6.3.</td>
<td>Determining the wall temperature target</td>
<td>102</td>
</tr>
<tr>
<td>6.4.</td>
<td>Validation</td>
<td>105</td>
</tr>
<tr>
<td>6.5.</td>
<td>Summary</td>
<td>107</td>
</tr>
<tr>
<td>CHAPTER 7</td>
<td>Linear MPC on Thermal Management</td>
<td>108</td>
</tr>
<tr>
<td>7.1.</td>
<td>Linear MPC Implementations</td>
<td>109</td>
</tr>
<tr>
<td>7.2.</td>
<td>Modelling: MATLAB® System Identification</td>
<td>109</td>
</tr>
<tr>
<td>7.3.</td>
<td>Modelling: Jacobian Linearization</td>
<td>118</td>
</tr>
<tr>
<td>7.4.</td>
<td>Controller Comparison</td>
<td>127</td>
</tr>
<tr>
<td>7.5.</td>
<td>Conclusion</td>
<td>130</td>
</tr>
<tr>
<td>CHAPTER 8</td>
<td>The New Engine Thermal Management Strategy</td>
<td>131</td>
</tr>
<tr>
<td>8.1.</td>
<td>The New Control Concept</td>
<td>131</td>
</tr>
<tr>
<td>8.2.</td>
<td>Feedback Linearization Strategy Overview</td>
<td>133</td>
</tr>
<tr>
<td>8.3.</td>
<td>Feedback Linearization MPC Implementation</td>
<td>139</td>
</tr>
<tr>
<td>8.4.</td>
<td>MPC setup</td>
<td>148</td>
</tr>
<tr>
<td>8.5.</td>
<td>MPC Future Prediction</td>
<td>166</td>
</tr>
<tr>
<td>8.6.</td>
<td>Controller Comparison</td>
<td>172</td>
</tr>
</tbody>
</table>
8.7. Conclusion ........................................................................................................... 174

CHAPTER 9  Experimental Validation of the Controller Dynamics ......................... 175
  9.1. Experiment Setup .................................................................................................. 176
  9.2. Controller implementation .................................................................................... 183
  9.3. MPC Performance Analysis ................................................................................ 202
  9.4. Conclusion ........................................................................................................... 209

CHAPTER 10  Benefits over a Drive Cycle ................................................................. 211
  10.1. Implementations .................................................................................................. 211
  10.2. MPC Setup ........................................................................................................ 221
  10.3. Drive Cycle Performance .................................................................................... 226
  10.4. Conclusion ........................................................................................................... 239

CHAPTER 11  Conclusion and Future Work ............................................................... 240
  11.1. Summary and Conclusion .................................................................................. 240
  11.2. Future Work ....................................................................................................... 242

Reference .................................................................................................................... 244

Appendix A  GT-SUITE Engine Model and Result .................................................. 252
  A.1. Engine Gasoline Natural Aspirated 2.0L engine list ........................................... 252
  A.2. Conventional Cooling system ............................................................................. 255
  A.3. Cylinder Wall Temperature Throughout Drive Cycle ........................................ 255
  A.4. Fuel Consumption Comparison Calculation in Test Cycle ............................... 260
  A.5. Engine Conditions Time Spend ......................................................................... 262

Appendix B  Simulink Model ....................................................................................... 267
  B.1. Engine Cooling System Model in MathWork™ Simulink® ................................... 267
  B.2. Linear Model Estimation Results ....................................................................... 268
  B.3. The System Identification Model Behaviour ..................................................... 270
  B.4. Jacobian Linearization ......................................................................................... 271

Appendix C  Experiment ............................................................................................. 276
  C.1. Water Block Groove Design .............................................................................. 276
C.2. Steady State Model for Experiment ................................................................. 278
C.3. LabVIEW Software Block Diagram .................................................................. 282
C.4. Data for Model Fitting ...................................................................................... 284
C.5. Coolant Capacity Ratio .................................................................................... 285

Appendix D  MPC Performance in Drive Cycles ...................................................... 287
D.1. Results ............................................................................................................ 287
D.2. Fuel Consumption throughout Drive Cycle ...................................................... 295
List of Figures

Figure 1.1: CO₂ emission from various sectors in EU [7]..........................................................2

Figure 1.2: History of CO₂ emission from average new passenger car and future CO₂ emission target by EU [9]..............................................................................................................2

Figure 1.3: Additional manufacturer costs of technology options to reduce CO₂ emissions for an average passenger car [20]............................................................................................................5

Figure 1.4: Coolant temperature set point varies according to the engine speed and load in Audi 18 TSFI 3rd generation engine cooling system [22]............................................................................................6

Figure 1.5: The engine speed and load with the coolant target temperature behaviour based on the 3rd generation engine cooling system from Audi throughout the US06 drive cycle........7

Figure 1.6: Coolant temperature response example from a simulation........................................7

Figure 1.7: Three main factors that have influence in the engine thermal management system. ......9

Figure 1.8: The organisation of the thesis.........................................................................................13

Figure 2.1: Illustration of the boundary, mixed and hydrodynamic friction regimes.....................16

Figure 2.2: Overview of the friction force in Stribeck Curve..........................................................16

Figure 2.3: Overview of engine components fiction behaviour across Stribeck Curve [31–35]........17

Figure 2.4: Typical component fiction in an engine [41]..................................................................18

Figure 2.5: Upper, middle and lower sections of the cylinder liner friction..................................19

Figure 3.1: 2.0L naturally aspirated engine model in GT-SUITE......................................................32

Figure 3.2: Engine brake torque and power from the GT-SUITE model.........................................33

Figure 3.3: Sample of post simulation FE in-cylinder structure temperature results.......................40

Figure 3.4: GT-SUITE model of the engine cooling system............................................................41

Figure 3.5: Model-Based Calibration Process Flow..........................................................................42

Figure 3.6: Input parameters and output response for thermal management model based Calibration.........................................................................................................................45

Figure 3.7: Spark timing boundary limit..........................................................................................46
Figure 3.8: Engine sparks timing from calibration result..............................................................47
Figure 3.9: Engine lambda from calibration result .................................................................48
Figure 3.10: Coolant out temperature from calibration result ..................................................48
Figure 3.11: BSFC map throughout the engine speed and load ..............................................49
Figure 3.12: Exhaust gas temperature throughout the engine speed and load ......................50
Figure 3.13: Gas side cylinder head surface temperature at the valve bride zone throughout the engine speed and load .......................................................................................................50
Figure 3.14: Engine BSFC improvement by optimized steady state thermal management the throughout engine speed and load .................................................................51
Figure 3.15: Cylinder head temperature difference throughout the engine speed and load .......53
Figure 3.16: Calibrated cylinder wall temperature (red), actual cylinder wall temperature (blue) and vehicle speed in NEDC ..........................................................................................................................55
Figure 3.17: Summary results of the fuel consumption reduction by optimizing the cylinder wall temperature in various drive cycles ..........................................................................................................................56
Figure 3.18: Engine speed and load distribution during the Artemis Urban Drive Cycle ..........57
Figure 3.19: Engine speed and load distribution during the Artemis Motorway drive cycle ........57
Figure 4.1: Autocorrelation plot of temperature set point in drive cycle .................................62
Figure 4.2: Autocorrelation plot of temperature set point in Artemis Urban Drive Cycle ........63
Figure 4.3: Autocorrelation plot of temperature set point in Artemis Motorway Drive Cycle ....63
Figure 4.4: Lag of autocorrelation coefficient entering the 95% confidence band for all drive cycles.64
Figure 4.5: Lag plot for Artemis Urban Test Cycle at 0.1 seconds, 1 second, 5 seconds and 10 seconds. ..............................................................................................................................65
Figure 4.6: Lag plot for Artemis Motorway Test Cycle at 0.1 seconds, 1 second, 5 seconds and 10 seconds. ..............................................................................................................................65
Figure 4.7: HWY Drive Cycle with variable width of Hann Window filter ...............................67
Figure 4.8: Power spectral density of cylinder wall temperature set point in the Artemis Urban Test Cycle ..............................................................................................................................68
Figure 4.9: Power spectral density of cylinder wall temperature set point in the Artemis Motorway Test Cycle ..............................................................................................................................69
Figure 4.10: Power spectral density of cylinder wall temperature set point in all drive cycles. ..........69
Figure 4.11: Corner frequency of each drive cycle. .................................................................70
Figure 4.12: Input and output of a system for magnitude and phase response. .........................71
Figure 4.13: Bode plot of cylinder wall temperature response at 3000rpm @ 5bar. ...................72
Figure 4.14: Bode plot of cylinder wall temperature response at 5000rpm @ 10bar. ...............74
Figure 4.15: Bode plot of cylinder wall temperature response at 1000rpm @ 1bar. .................74
Figure 4.16: Typical step response for a second order system....................................................75
Figure 4.17: Cylinder wall temperature response at 1000rpm @ 1bar, 3000rpm @ 5bar and 5000rpm @ 10bar..........................................................76
Figure 4.18: Coolant dynamic viscosity, thermal conductivity and density trend throughout coolant temperature..........................................................................................................................78
Figure 4.19: Comparison of cylinder wall temperature response between the close loop flow rate, close and open loop coolant temperature control of cooling system at 1000rpm @ 1bar. ..........................................................................................................................80
Figure 5.1: Basic structure of Model Based Predictive Control (MPC). .....................................84
Figure 5.2: MPC control strategy scheme. .....................................................................................85
Figure 5.3: Illustration of output response comparison of the conventional PID and MPC with look ahead. ....................................................................................................................89
Figure 6.1: Heat transfer from gas combustion to coolant through the cylinder wall model. ..........95
Figure 6.2: Convection heat transfer coefficient throughout coolant temperature and coolant mass flow rate.................................................................................................................97
Figure 6.3: Heat transfer from coolant to environment in the radiator .......................................98
Figure 6.4: Heat transfer rate from radiator to air at 120°C, 100°C and 80°C coolant radiator inlet temperature..........................................................................................................................100
Figure 6.5: Air mass flow model.....................................................................................................100
Figure 6.6: Transport delays in cooling system.............................................................................102
Figure 6.7: Temperature wall target for mathematical model......................................................103
Figure 6.8: Coolant engine out temperature and water pump signal throughout cylinder wall temperature target ................................................................................................................105
Figure 6.9: Simulink® model output temperatures compared to the GT-Suite model output temperatures.

Figure 6.10: Nonlinear proportionate relationship between the heat transfer at the cylinder head and the overall heat transfer.

Figure 7.1: System Identification flow

Figure 7.2: System identification estimation data and validation data.

Figure 7.3: The best model fitting for linear MPC in MathWork™ System Identification toolbox™.

Figure 7.4: MPC controller with control horizon at 5, 10, 15 and 20.

Figure 7.5: MPC controller performance with lower target temperature and combustion heat at 0 to 200 seconds and 400 seconds to 600 seconds.

Figure 7.6: MPC controller with prediction horizon at 15, 25, 50, 100, 150 and 200.

Figure 7.7: Comparison of wall temperature increase on step response at different equilibrium point.

Figure 7.8: Region partitions for local linearization.

Figure 7.9: Wall temperature Bode plot.

Figure 7.10: Coolant temperature Bode plot.

Figure 7.11: Scheme of Multiple Model Predictive Control (MMPC).

Figure 7.12: Multiple Model Predictive Control performance in random square signal disturbance.

Figure 7.13: MMPC with input rate weight at 0.00001, 0.1, 1 and 5.

Figure 7.14: Illustration of comparison number sub regions defined in MMPC. (a) shows less and (b) shows more number of sub regions being defined.

Figure 7.15: Comparing MMPC with 27 and 8 sub regions.

Figure 7.16: Comparing controller performance between MPC from system identification, MMPC 8 sub regions and MMPC 27 sub regions.

Figure 7.17: Computational cost between MPC with system identification, MMPC 8 and MMPC 27.

Figure 8.1: Engine thermal management with linear MPC and feedback linearization concept.

Figure 8.2: Engine thermal management with linear MPC and feedback linearization.
Figure 8.3: Coolant mass flow rate model by the function of coolant temperature and convection heat transfer coefficient. ................................................................. 136

Figure 8.4: Air mass flow rate output $m_{air}$ from required heat transfer rate $Q_{cool}$, current radiator mass flow rate $m_{rad}$ and current coolant out temperature $T_{out}$. ............................................................. 138

Figure 8.5: Fan signal output $N_{fan}$ determined from required air mass flow rate $m_{air}$ and current eternal air ram speed ................................................................. 138

Figure 8.6: The random input signal ($Q_{comb}$, $Q_{conv}$, $Q_{cool}$) and the measured out signals for identification and validation. ................................................................. 139

Figure 8.7: Measured and simulated output in validation data. ................................................................. 140

Figure 8.8: $Q_{conv}$ constraint throughout coolant engine out temperature and cylinder wall temperature........................................................................................................ 143

Figure 8.9: $Q_{cool}$ constraint throughout $Q_{conv}$ and coolant engine out temperature ($T_{out}$)........ 145

Figure 8.10: Linear constraint surfaces for $Q_{cool}$ ......................................................................................... 146

Figure 8.11: Cross section of the linear constraints and actual limits .......................................................... 148

Figure 8.12: Control outputs and manipulated variables of Feedback Linearization MPC with different control horizons ........................................................................................................................................... 151

Figure 8.13: Comparing RMSE across number of control horizons steps .................................................. 152

Figure 8.14: Overshoot and Undershoot for shorter control horizon during steps response up and down. ........................................................................................................................................... 153

Figure 8.15: MPC prediction trajectory and optimized control movement trajectory at 1180 seconds that create the overshoot in wall temperature ($n_p = 50, n_c = 10$)................................. 154

Figure 8.16: Feedback Linearization MPC with variable control horizon sequences.............................. 156

Figure 8.17: MPC prediction trajectory and optimized control movement trajectory at 580 seconds that create the interval undershoot (Seq 1). ............................................................................................................................. 157

Figure 8.18: Comparing RMSE across number of prediction horizons steps ............................................. 158

Figure 8.19: Control outputs and manipulated variables of Feedback Linearization MPC with different prediction horizons ........................................................................................................................................... 159

Figure 8.20: Larger prediction horizon trigger quicker coolant temperature increases thus improve wall temperature accuracy ........................................................................................................................................... 160
Figure 8.21: Fan and pump signal at various coolant out target temperatures and the optimum target temperature........................................................................................................162

Figure 8.22: Calibrated coolant out temperature target under random square signals disturbance.164

Figure 8.23: RMSE results comparing different wall temperature and coolant out temperature weight setting. .........................................................................................................................165

Figure 8.24: Changes of pump and fan usage compare to without calibrated coolant out temperature target........................................................................................................................................165

Figure 8.25: Comparison between MPC with and without known future inputs.........................................................167

Figure 8.26: Actual reference compared to incorrect prediction 1 and incorrect prediction 2.170

Figure 8.27: Comparing MPC with false reference prediction.................................................................171

Figure 8.28: Comparing controller performance between MPC from feedback linearization, system identification, MMPC 8 sub regions and MMPC 27 sub regions. ..............................................172

Figure 8.29: Computational cost between MPC with feedback linearization, system identification, MMPC 8 and MMPC 27.................................................................173

Figure 9.1: The test rig representing engine cooling system.................................................................................................176

Figure 9.2: The test rig cooling system schematic diagram.................................................................................................................177

Figure 9.3: Water jacket inside the aluminium block design in 3D. .................................................................178

Figure 9.4: Aluminium nitride ceramic heater from Watlow ® ULTRAMIC ® ..........................................................179

Figure 9.5: Surface heater generated heat and temperature for every 10% of power demand. .................................180

Figure 9.6: LabVIEW FPGA block diagram for the test rig consists of thermocouple readings (bottom right), main PWM (top right), secondary PWM (bottom left) and other controller (top left)..................................................................................................................182

Figure 9.7: Overall experiment configuration of PC host, data logger and test rig interconnections. ........................................................................................................................................183

Figure 9.8: MPC feedback linearization implementation in engine thermal management. ........................183

Figure 9.9: Experimental result of heat transfer coefficient throughout the coolant flow rate and temperature..............................................................................................................185

Figure 9.10: Pump signal throughout required coolant flow rate and current valve position. ........................186
Figure 9.11: Required coolant flow rate throughout convection heat transfer coefficient and current coolant out temperature. ................................................................. 186

Figure 9.12: Valve signal throughout required radiator flow rate and current pump flow rate. .......................... 187

Figure 9.13: Rise time for the wall and coolant temperature. ................................................................. 188

Figure 9.14: Wall temperature model output compared to the validation data. ........................................ 189

Figure 9.15: Coolant temperature model output compared to the validation data.............................. 190

Figure 9.16: Radiator flow rate percentage throughout the valve position ............................................. 190

Figure 9.17: Coolant flow rate throughout the valve position and pump signal ..................................... 192

Figure 9.18: Coolant flow rate at the maximum and minimum pump signal throughout valve position. ........................................................................................................ 193

Figure 9.19: Upper constraint and lower constraint for $Q_{conv}$ ............................................................. 194

Figure 9.20: Radiator flow rate throughout valve position and pump coolant flow rate. ....................... 195

Figure 9.21: Radiator out temperature model throughout coolant out temperature and radiator flow rate ........................................................................................................ 196

Figure 9.22: $Q_{cool}$ constraints with throughout wall target temperature $T_{wall}$ (solid surface) and $T_{wall}-10^\circ C$ (faded surface) ........................................................................ 197

Figure 9.23: $Q_{cool}$ constraints throughout wall temperature, coolant out temperature and $Q_{conv}$. ........................................................................................................ 198

Figure 9.24: $Q_{cool}$ upper constraint ........................................................................................................ 199

Figure 9.25: $Q_{cool}$ lower constraints ...................................................................................................... 200

Figure 9.26: Coolant temperature target throughout wall temperature ................................................ 201

Figure 9.27: Rise time comparison between experiment and simulation at max flow rate, half valve position and step heat input from 20% to 80% ................................................................. 201

Figure 9.28: Random $Q_{comb}$ disturbance and for MPC performance analysis ..................................... 202

Figure 9.29: MPC controller performance tracking target temperature in experiment ...................... 203

Figure 9.30: MPC controller constraints for manipulated variables ..................................................... 204

Figure 9.31: Comparison of MPC controller input rate weight of 0.001, 0.01, 0.1 and 1. ..................... 205

Figure 9.32: Comparing coolant out temperature target weight ......................................................... 207
Figure 9.33: Comparing MPC controller performance between with and without known future disturbance. ................................................................. 208

Figure 9.34: Comparing MPC controller performance between with and without known future disturbance during wall temperature warm-up. ........................................ 209

Figure 10.1: Engine and cooling system model build in GT-SUITE ............................................. 212

Figure 10.2: Local combustion heat transfer rate $Q_{comb}$ at optimized coolant out temperature. ........................................ 213

Figure 10.3: Nonlinearly proportional relation between $Q_{conv}$ and $Q_{cool}$. ........................................ 214

Figure 10.4: $Q_{other}$ throughout the engine speed and load .................................................................................. 215

Figure 10.5: Wall temperature model fitting for GT-SUITE engine thermal management. .................... 216

Figure 10.6: Coolant temperature model fitting for the GT-SUITE engine thermal management. ......... 217

Figure 10.7: $Q_{conv}$ upper and lower linear constraints in the GT-SUITE model simulation. ................ 218

Figure 10.8: The $Q_{cool}$ upper and lower limits in GT-SUITE simulation. .............................................. 219

Figure 10.9: Three linear upper constraints representing the complex surface of the $Q_{cool}$ upper limits .......................................................................................................................... 220

Figure 10.10: Two linear lower constraints representing the complex surface of the $Q_{cool}$ lower limits ........................................................................................................................ 221

Figure 10.11: Coolant reference temperature throughout engine speed and load ........................................ 222

Figure 10.12: MPC result with the initial setup in GT-SUITE simulation .............................................. 223

Figure 10.13: Illustration of MPC inputs and outputs before the temperature undershoot. ..................... 224

Figure 10.14: Comparison of the $Q_{conv}$ input rate weight to the pump signal response. ..................... 225

Figure 10.15: MPC controller performance compared to the conventional cooling system in the New European Drive Cycle .............................................................................. 227

Figure 10.16: Fuel consumed throughout the NEDC by the conventional cooling system compared to MPC and a perfect controller of ideal head temperature. .......................... 228

Figure 10.17: MPC controller performance compared to the conventional cooling system in the US Supplemental Federal Test Procedure ........................................................................ 229

Figure 10.18: Fuel consumed throughout the US06 by the conventional cooling system compared to MPC and a perfect controller of ideal head temperature ..................................... 230
Figure 10.19: MPC controller performance with, and without the future knowledge of disturbance in NEDC. ................................................................. 231

Figure 10.20: Fuel consumed throughout the NEDC by the MPC without the known future disturbance compared to MPC with the known future disturbance and a perfect controller of ideal head temperature. ................................................................. 232

Figure 10.21: MPC controller performance with, and without the future knowledge of disturbance in US Supplemental Federal Test Procedure. ................................................................. 233

Figure 10.22: Fuel consumed throughout the US06 by the MPC without the known future disturbance compared to MPC with the known future disturbance and a perfect controller of ideal head temperature. ................................................................. 234

Figure 10.23: Wall temperature RMSE comparison between the conventional cooling system, MPC with, and without the future knowledge of disturbance. ................................................................. 235

Figure 10.24: Fuel consumption reduction full potential reduction, MPC with and without the future knowledge of disturbance of compared to conventional cooling system. ................................................................. 236

Figure 10.25: Water pump speed average comparison between the conventional cooling system, MPC with, and without the future knowledge of disturbance. ................................................................. 237

Figure 10.26: Average valve movement comparison between the conventional cooling system, MPC with and without the future knowledge of disturbance. ................................................................. 238

Figure 10.27: Fan average speed comparison between the conventional cooling system, MPC with and without the future knowledge of disturbance. ................................................................. 238

Figure A.1: Engine bore and stroke frequency based on Table A.1 .............................................................................. 254

Figure A.2: Fan strategy in the GT-SUITE cooling system model. .............................................................................. 255

Figure A.3: Calibrated head temperature (red), actual head temperature (blue) and vehicle speed (green) throughout the NEDC. .............................................................................. 256

Figure A.4: Calibrated head temperature (red), actual head temperature (blue) and vehicle speed (green) throughout the WLTC .............................................................................. 256

Figure A.5: Calibrated head temperature (red), actual head temperature (blue) and vehicle speed (green) throughout the Artemis Urban Cycle. .............................................................................. 257

Figure A.6: Calibrated head temperature (red), actual head temperature (blue) and vehicle speed (green) throughout the Artemis Rural Road Cycle. .............................................................................. 257
Figure A.7: Calibrated head temperature (red), actual head temperature (blue) and vehicle speed (green) throughout the Artemis Motorway Cycle. .......................... 258

Figure A.8: Calibrated head temperature (red), actual head temperature (blue) and vehicle speed (green) throughout the FTP-75kph. .......................................................... 258

Figure A.9: Calibrated head temperature (red), actual head temperature (blue) and vehicle speed (green) throughout the US06. .......................................................... 259

Figure A.10: Calibrated head temperature (red), actual head temperature (blue) and vehicle speed (green) throughout the HWY. .......................................................... 259

Figure A.11: Total fuel consumed throughout the test cycle calculation. ......................................... 260

Figure A.12: BSFC throughout engine speed and load at 170°C. .................................................. 261

Figure A.13: BSFC throughout engine speed and load at 200°C. .................................................. 261

Figure A.14: BSFC throughout engine speed and load at 230°C. .................................................. 262

Figure A.15: Time spend of engine speed and load in Federal Test Procedure 75kph Test Cycle. .... 263

Figure A.16: Time spend of engine speed and load in US06 Supplemental Federal Test Procedure. 263

Figure A.17: Time spend of engine speed and load in The Highway Fuel Economy Test Cycle. ...... 264

Figure A.18: Time spend of engine speed and load in New European Driving Cycle. ................. 264

Figure A.19: Time spend of engine speed and load in Worldwide harmonized Light vehicles Test Cycle. ........................................................................................................... 265

Figure A.20: Time spend of engine speed and load in Artemis Urban Test Cycle. ......................... 265

Figure A.21: Time spend of engine speed and load in Artemis Rural Road Test Cycle. ................ 266

Figure A.22: Time spend of engine speed and load in Artemis Motorway Test Cycle. ................. 266

Figure B.1: Engine cooling system model in the MathWork™ Simulink® ...................................... 267

Figure B.2: State Space models fitting result .................................................................................. 269

Figure B.3: Transfer function models fitting result ......................................................................... 269

Figure B.4: Step response result from ......................................................................................... 270

Figure B.5: Bode Plot result of wall temperature. ........................................................................ 270

Figure B.6: Bode Plot result of coolant out temperature. ............................................................... 271

Figure B.7: Wall temperature Bode plot for Config B (MMPC27). ................................................. 274
Figure B.8: Coolant temperature Bode plot for Config B (MMPC27) .......................................................... 274
Figure B.9: MMPC27 with input rate weight at 0.00001, 0.1, 1 and 5. ............................................................. 275
Figure C.1: 10 types of water block groove design for the experiment. .......................................................... 276
Figure C.2: The water block temperature result from Autodesk® CFD during maximum heat and coolant flow rate being applied.................................................................................................................. 277
Figure C.3: The coolant flow result from Autodesk® CFD during maximum heat and coolant flow rate being applied................................................................................................................................. 277
Figure C.4: LabVIEW Software block diagram.................................................................................................... 282
Figure C.5: LabVIEW FPGA block diagram........................................................................................................ 283
Figure C.6: Estimation and validation data for experiment wall model fitting...................................................... 284
Figure C.7: Estimation and validation data for experiment coolant model fitting................................................ 285
Figure D.1: MPC controller performance with the future knowledge of disturbance, without the future knowledge of disturbance and conventional cooling system in Artemis Urban Cycle.................................................................................................................................. 288
Figure D.2: MPC controller performance with the future knowledge of disturbance, without the future knowledge of disturbance and conventional cooling system in Artemis Rural Road Cycle.................................................................................................................................. 289
Figure D.3: MPC controller performance with the future knowledge of disturbance, without the future knowledge of disturbance and conventional cooling system in Artemis Motorway Cycle.................................................................................................................................. 290
Figure D.4: MPC controller performance with the future knowledge of disturbance, without the future knowledge of disturbance and conventional cooling system in NEDC.................................. 291
Figure D.5: MPC controller performance with the future knowledge of disturbance, without the future knowledge of disturbance and conventional cooling system in WLTC. ....................... 292
Figure D.6: MPC controller performance with the future knowledge of disturbance, without the future knowledge of disturbance and conventional cooling system in FTP-75kph........... 293
Figure D.7: MPC controller performance with the future knowledge of disturbance, without the future knowledge of disturbance and conventional cooling system in HWY......................... 294
Figure D.8: MPC controller performance with the future knowledge of disturbance, without the future knowledge of disturbance and conventional cooling system in US06. ....................... 295
Figure D.9: Fuel consumed throughout the NEDC by a perfect controller of ideal head temperature compared to the conventional cooling system, MPC with and without the known future disturbance. ................................................................. 296

Figure D.10: Fuel consumed throughout the WLTC by a perfect controller of ideal head temperature compared to the conventional cooling system, MPC with and without the known future disturbance. ........................................................................................................ 296

Figure D.11: Fuel consumed throughout the Artemis Urban Cycle by a perfect controller of ideal head temperature compared to the conventional cooling system, MPC with and without the known future disturbance. ........................................................................................................ 297

Figure D.12: Fuel consumed throughout the Artemis Rural Road Cycle by a perfect controller of ideal head temperature compared to the conventional cooling system, MPC with and without the known future disturbance. ........................................................................................................ 297

Figure D.13: Fuel consumed throughout the Artemis Motorway Cycle by a perfect controller of ideal head temperature compared to the conventional cooling system, MPC with and without the known future disturbance. ........................................................................................................ 298

Figure D.14: Fuel consumed throughout the FTP 75kph by a perfect controller of ideal head temperature compared to the conventional cooling system, MPC with and without the known future disturbance. ........................................................................................................ 298

Figure D.15: Fuel consumed throughout the US06 by a perfect controller of ideal head temperature compared to the conventional cooling system, MPC with and without the known future disturbance. ........................................................................................................ 299

Figure D.16: Fuel consumed throughout the HWY by a perfect controller of ideal head temperature compared to the conventional cooling system, MPC with and without the known future disturbance. ........................................................................................................ 299
List of Tables

Table 2.1: Summary of the effect of toxic emission changes from thermal management. .................. 26
Table 2.2: Summary of the thermal management influence in the engine output............................ 29
Table 3.1: Engine model specification for temperature optimization........................................... 31
Table 3.2: Geometry attributes for the FE in-cylinder structure in GT-SUITE. .................................. 40
Table 3.3: Calibration parameter limits ..................................................................................... 43
Table 3.4: Vehicle model specification in GT-Suite....................................................................... 54
Table 7.1: Linear MPC setup for simulation.................................................................................. 113
Table 7.2: Equilibrium points setting.......................................................................................... 120
Table 8.1: Constraints for actuator limits and working operation limits in engine thermal management.............................................................................................................. 141
Table 8.2: Feedback Linearization MPC setup for simulation...................................................... 149
Table 8.3: List of variable control horizon sequences................................................................. 155
Table 8.4: Weight setting between wall temperature and coolant out temperature with calibrated coolant out temperature target................................................................................. 162
Table 8.5 : MPC Coolant out weight setting. .............................................................................. 163
Table 10.1: Feedback Linearization MPC setup for GT-SUITE simulation.................................. 221
Table A.1: List of engine specification......................................................................................... 252
Table B.1: Specification of Simulink® engine cooling system model parameter ....................... 268
Table B.2: Equilibrium points list on 8 sub-regions configuration............................................... 272
Table B.3: Equilibrium points list on 27 sub-regions configuration............................................. 273
Table C.1: The groove designs result from Autodesk CFD simulation......................................... 277
Table C.2: List of engine power rating and coolant capacity........................................................ 286
Notation

Abbreviations

AlN  Aluminium Nitride
Artemis  Assessment and Reliability of Transport Emission Models and Inventory Systems
ArtemisM  Artemis Motorway Cycle
ArtemisR  Artemis Rural Road Cycle
ArtemisU  Artemis Urban Cycle
BMEP  Brake Mean Effective Pressure
BSFC  Brake Specific Fuel Consumption
CAGE  Calibration Generation
CFD  Computational Fluid Dynamic
CO  Carbon Monoxide
CO₂  Carbon dioxide
CPU  Central Processing Unit
DoE  Design of Experiment
EU  European Union
ECU  Engine Control Unit
EGR  Exhaust gas recirculation
FEA  Finite Element Analysis
FFT  Fast Fourier Transform
FMEP  Friction Mean Effective Pressure
FPGA  Field-Programmable Gate Array
FTP75kph  United State Federal Test Procedure 75kph
GPS  Global Positioning System
GT  Gamma Technologies
HC  Hydrocarbon
HDL  Hardware Description Languages
HWY  United State Highway Fuel Economy Driving Schedule
IMEP  Indicated Mean Effective Pressure
LTI  Linear Time Invariant
LQR  Linear Quadratic Regulator
MBC  Model-based Calibration
MBT  Maximum Brake Torque
MIMO  Multi-Input Multi-Output
MMPC  Multiple Model Predictive Control
MPC  Model Predictive Control
MPCA  Model Predictive Control Allocation
NEDC  New European Drive Cycle
NI  National Instrument™
NMPC  Nonlinear Model Predictive Control
NOₓ  Nitrogen Oxide
PC  Personal Computer
PI  Proportional–Integral
PID Proportional–Integral–Derivative
PRESS Predicted Residual Sum of Squares
PSD Power Spectral Density
PWM Pulse Width Modulator
QP Quadratic Optimization Problem
RBF Radial Basis Function
RMSE Root-Mean-Square Error
SI Park Ignition
SISO Single Input Single Output
TCP Transmission Control Protocol
TFSI Turbocharged Fuel Stratified Injection
US06 United State Supplemental Federal Test Procedure
VI Visual Editor
VHDL Very High Speed Integrated Circuit Hardware Description Languages
WLTC Worldwide Harmonized Light vehicles Test Procedures
WOT Wide open throttle

Symbols

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
<th>Unit</th>
</tr>
</thead>
<tbody>
<tr>
<td>$m$</td>
<td>Mass flow</td>
<td>kg/s</td>
</tr>
<tr>
<td>$m_{air}$</td>
<td>Air mass flow through radiator</td>
<td>kg/s</td>
</tr>
<tr>
<td>$m_{by}$</td>
<td>Coolant mass flow throughout bypass pipe</td>
<td>kg/s</td>
</tr>
<tr>
<td>$m_c$</td>
<td>Coolant mass flow throughout engine</td>
<td>kg/s</td>
</tr>
<tr>
<td>$m_{rad}$</td>
<td>Coolant mass flow throughout radiator</td>
<td>kg/s</td>
</tr>
<tr>
<td>$\Delta u_j$</td>
<td>Manipulated variables movement $j$</td>
<td></td>
</tr>
<tr>
<td>$h_{gas}$</td>
<td>Heat transfer coefficient from combustion gas to cylinder wall</td>
<td>W/mK</td>
</tr>
<tr>
<td>$h_r$</td>
<td>Heat transfer coefficient from coolant to radiator wall</td>
<td>W/mK</td>
</tr>
<tr>
<td>$h_w$</td>
<td>Heat transfer coefficient from cylinder wall to coolant</td>
<td>W/mK</td>
</tr>
<tr>
<td>$A_e$</td>
<td>Flame surface area</td>
<td>m$^2$</td>
</tr>
<tr>
<td>$C_c$</td>
<td>Coolant heat capacity</td>
<td>J/K</td>
</tr>
<tr>
<td>$C_w$</td>
<td>Cylinder wall heat capacity</td>
<td>J/K</td>
</tr>
<tr>
<td>$D$</td>
<td>Cylinder bore</td>
<td>m</td>
</tr>
<tr>
<td>$J$</td>
<td>Cost function</td>
<td></td>
</tr>
<tr>
<td>$L$</td>
<td>Hydraulic diameter</td>
<td>m</td>
</tr>
<tr>
<td>$N_e$</td>
<td>Engine speed</td>
<td>rpm</td>
</tr>
<tr>
<td>$N_{fan}$</td>
<td>Radiator fan speed</td>
<td>rpm</td>
</tr>
<tr>
<td>$N_{pump}$</td>
<td>Water pump speed</td>
<td>rpm</td>
</tr>
<tr>
<td>$S_L$</td>
<td>Laminar flame speed</td>
<td>m/s</td>
</tr>
<tr>
<td>$S_T$</td>
<td>Turbulence flame speed</td>
<td>m/s</td>
</tr>
<tr>
<td>$T_{air}$</td>
<td>Environment temperature</td>
<td>°C</td>
</tr>
<tr>
<td>$T_{gas}$</td>
<td>Combustion gas temperature</td>
<td>°C</td>
</tr>
<tr>
<td>$T_{in}$</td>
<td>Coolant engine in temperature</td>
<td>°C</td>
</tr>
<tr>
<td>$T_{out}$</td>
<td>Coolant engine out temperature</td>
<td>°C</td>
</tr>
<tr>
<td>$T_{rad}$</td>
<td>Coolant radiator out temperature</td>
<td>°C</td>
</tr>
<tr>
<td>$T_w$</td>
<td>Cylinder wall temperature</td>
<td>°C</td>
</tr>
<tr>
<td>$bmep$</td>
<td>Brake mean effective pressure</td>
<td>bar</td>
</tr>
<tr>
<td>Symbol</td>
<td>Definition</td>
<td>Unit</td>
</tr>
<tr>
<td>--------</td>
<td>------------</td>
<td>------</td>
</tr>
<tr>
<td>$c_c$</td>
<td>Coolant specific heat capacity</td>
<td>J/kgK</td>
</tr>
<tr>
<td>$d$</td>
<td>Disturbance</td>
<td></td>
</tr>
<tr>
<td>$f_{mep}$</td>
<td>Friction mean effective pressure</td>
<td>bar</td>
</tr>
<tr>
<td>$f_{mep_{\text{ref}}}$</td>
<td>Friction mean effective pressure reference</td>
<td>bar</td>
</tr>
<tr>
<td>$i$</td>
<td>Series of number</td>
<td></td>
</tr>
<tr>
<td>$j$</td>
<td>Number of variables</td>
<td></td>
</tr>
<tr>
<td>$k$</td>
<td>Time step</td>
<td></td>
</tr>
<tr>
<td>$l_r$</td>
<td>Thermostat valve lift position (0: fully close; 1: fully open)</td>
<td></td>
</tr>
<tr>
<td>$m_b$</td>
<td>Burned mass</td>
<td>kg</td>
</tr>
<tr>
<td>$m_e$</td>
<td>Entrained mass of unburned mixture</td>
<td>kg</td>
</tr>
<tr>
<td>$n_c$</td>
<td>Control horizon length</td>
<td></td>
</tr>
<tr>
<td>$n_p$</td>
<td>Prediction horizon length</td>
<td></td>
</tr>
<tr>
<td>$n_{\text{pump}}$</td>
<td>Water pump signal (0: no flow; 1: max flow)</td>
<td></td>
</tr>
<tr>
<td>$n_u$</td>
<td>Number of input variables</td>
<td></td>
</tr>
<tr>
<td>$n_y$</td>
<td>Number of output variables</td>
<td></td>
</tr>
<tr>
<td>$p_{\text{gas}}$</td>
<td>In-cylinder pressure</td>
<td>N/m$^2$</td>
</tr>
<tr>
<td>$r$</td>
<td>Set point</td>
<td></td>
</tr>
<tr>
<td>$r_j$</td>
<td>Set point $j$</td>
<td></td>
</tr>
<tr>
<td>$t$</td>
<td>Time</td>
<td>s</td>
</tr>
<tr>
<td>$t_s$</td>
<td>Sampling time</td>
<td>s</td>
</tr>
<tr>
<td>$u$</td>
<td>Manipulated variable</td>
<td></td>
</tr>
<tr>
<td>$u_j$</td>
<td>Manipulated variables $j$</td>
<td></td>
</tr>
<tr>
<td>$u_{\text{target}}$</td>
<td>Input target value</td>
<td></td>
</tr>
<tr>
<td>$v_{\text{cool}}$</td>
<td>Coolant velocity</td>
<td>m/s</td>
</tr>
<tr>
<td>$v_{\text{car}}$</td>
<td>Vehicle speed</td>
<td>Km/hr</td>
</tr>
<tr>
<td>$v_{\text{gas}}$</td>
<td>In-cylinder gas speed</td>
<td>m/s</td>
</tr>
<tr>
<td>$w_{\Delta u}$</td>
<td>Input movement cost function weight</td>
<td></td>
</tr>
<tr>
<td>$w_u$</td>
<td>Input cost function weight</td>
<td></td>
</tr>
<tr>
<td>$w_y$</td>
<td>Output cost function weight</td>
<td></td>
</tr>
<tr>
<td>$x$</td>
<td>States</td>
<td></td>
</tr>
<tr>
<td>$x_0$</td>
<td>Initial state condition</td>
<td></td>
</tr>
<tr>
<td>$y$</td>
<td>Outputs</td>
<td></td>
</tr>
<tr>
<td>$y_j$</td>
<td>Output variable $j$</td>
<td></td>
</tr>
<tr>
<td>$\eta_{\text{VE}}$</td>
<td>Volumetric efficiency</td>
<td></td>
</tr>
<tr>
<td>$\mu$</td>
<td>Dynamic viscosity</td>
<td>kg/ms</td>
</tr>
<tr>
<td>$\mu_{\text{cool}}$</td>
<td>Coolant dynamic viscosity</td>
<td>kg/ms</td>
</tr>
<tr>
<td>$\mu_{\text{ref}}$</td>
<td>Dynamic viscosity reference</td>
<td>kg/ms</td>
</tr>
<tr>
<td>$\rho_{\text{cool}}$</td>
<td>Coolant density</td>
<td>kg/m$^3$</td>
</tr>
<tr>
<td>$\rho_u$</td>
<td>Unburned density</td>
<td>kg/m$^3$</td>
</tr>
</tbody>
</table>
CHAPTER 1  Introduction

The aim of the thesis is to present a new approach to the thermal management of internal combustion engines. It is based on the determination of an optimal temperature for each operating condition, and on the control of the cylinder wall temperature using multivariable Model Predictive Control (MPC) with both variable coolant temperature and flow rate. Due to the highly non-linear nature of automotive system, the use of optimised maps is common. The use of MPC on the other hand is rare in the automotive industry [1], while it is popular in other engineering sectors (process control, chemical industries and building thermal management) due to its ability to handle multivariable constrained control problems [2–6]. Instead, the automotive industry is still dominated by single variable controllers, which are often used in combination or in cascade to form complex control schemes. This thesis demonstrates the advantages of using a predictive multivariable controller for thermal management.

1.1. Research Background and Motivation

CO₂ Legislation

Transportation is one of the sectors that contribute a significant share of global greenhouse gas emission CO₂. This sector contributes 20% of EU countries overall CO₂ emission, which ranks just after the energy industry, and similar to production (Figure 1.1). Figure 1.1 illustrates CO₂ emission from various sectors in EU countries. The majority (about 60%) of the transportation sector emissions are generated by passenger cars [7]. Consequently, the EU is aggressively fighting climate change by applying a stringent legislation in the passenger car to reduce the greenhouse gas emission. The EU proposed to add emission labels to new cars, and the EU also encourages state members to vary the tax pricing based on the vehicle CO₂ emission level [8].
The EU sets an average target of not more than 130g CO$_2$/km for a new passenger car by the year 2015, which will reduce it to 95g CO$_2$/km by the year 2020. Figure 1.2 shows the historic average CO$_2$ emissions from passenger cars and the EU target. Actual emissions in 2013 were already recorded lower than the target for 2015 [9]. However, a significant reduction is still required to achieve CO$_2$ emission target in 2020.

Improving the engine fuel efficiency is the most effective way to reduce CO$_2$ emissions, because the majority of primary energy is lost in the engine. Y. Gao and M. Checkel (2007)
stated that a constant average ratio of 3.1 kg CO$_2$ is produced by every kilogram of fuel [10]. In other words; reducing the engine fuel consumption means that the vehicle CO$_2$ emission is reduced as well. Therefore, many new technologies are rapidly developed and introduced in the market such as hybrid vehicle, electric vehicle and alternative fuel to improve engine fuel efficiency.

Despite the increasing availability of alternative powertrains in the market, they are not popular with mainstream customers. For an example, electric vehicles adoption is hindered by a lack of a sufficiently developed charging infrastructure. Other barriers such as the high cost of motors, batteries and power converters, lack of after sale supports and safety concerns can hinder the adoption [11]. Consequently, the overall impact of electric vehicles remains limited, and current fuels such as gasoline are expected to dominate the passenger car powertrain for many years to come.

Consequently, most automotive companies and research institutions invest significant research effort into further improving gasoline engine efficiency, as this is necessary to reach the EU emission targets. Recent developments include engine downsizing, continuous variable transmission, light weight components, gasoline direct injection, variable valve timing, variable valve lift, friction reduction materials, dual clutch transmission, starts-stop system and increased electrification.

Engine thermal management is one of the areas that promise improved engine fuel efficiency by reducing engine warm-up duration and optimising the operating temperature. The spark ignition engine produces the worst fuel efficiency and large amount of pollution emission especially CO and HC during engine warm-up [12]. Some studies showed that 50% to 80% of total HC and CO emission is emitted in the case of New European Driving Cycle and US FTP 75 cycles that come from the first 300 seconds of the warm-up phase [13–15]. During the warm-up or cold-start period, fuel efficiency can be as low as 10% due to friction, thermal loss and insufficient air-fuel mixing.
Emerging of Electro-Mechanical Components

A key trend in automotive systems is the increased electrification of components. This includes the introduction of automatic transmission, electric power steering, electronic fuel injection, active torsion bars, electro-mechanical brakes, automatic transmission and motorized throttle. Together with advanced control mechanisms, these mechatronic components enhance the driving experience, safety and efficiency. This also includes the cooling system, specifically electric water pump, motorized thermostat valve and variable speed electric radiator fan, which replace the conventional wax thermostat valve, crank driven water pump and radiator fan [16]. In addition to the packaging and efficiency advantages, these systems also offer new control opportunities compared to the passive type conventional cooling system that has remained virtually unchanged for decades [17,18].

Because only minimal hardware changes are required, advanced engine cooling systems are considered highly cost effective to improve engine fuel efficiency (see Figure 1.3). Figure 1.3 below shows the comparison of various technologies and the implementation costs and its price for every one percent of CO₂ reduction. The additional cost for an optimized cooling system and an advanced cooling system is only about €20 and €40 for every one percent of CO₂ reduction, which is close to the bottom end of the cost spectrum. In particular, these methods are much more cost-effective than a hybrid power train, a start-stop system and even direct injection. Moreover, the implementation price is also low compared to most of other technologies. Therefore, the implementation of active thermal management should be considered before implementing more costly technologies like downsizing and direct injection [16,19,20].
Figure 1.3: Additional manufacturer costs of technology options to reduce CO\textsubscript{2} emissions for an average passenger car [20].

Current State of the Art

Recently, some OEMs have started to adopt an active type of thermal management system due to the availability of reliable electric water pumps and valve. For example, the Audi 18 TFSI 3\textsuperscript{rd} generation engine uses a state of the art cooling system with variable coolant temperature set points [21]. The temperature set point varies according to the engine speed and load, from 105°C at low engine speed and load to 85°C at high engine speed and load. Figure 1.4 below shows a map of the coolant temperature set point over the engine speed and load.
Figure 1.4: Coolant temperature set point varies according to the engine speed and load in Audi 18 TSFI 3rd generation engine cooling system [22].

However, this map is based on steady state operating points. Because thermal effects can be comparatively slow, it is an open question whether or not the control system is able to cope with the fast engine speed and load change behaviour. This is an important argument because engine outputs will degrade such as engine power, fuel consumption and emission if the actual coolant temperature is not accurate. This is important because a mismatch between the desired and optimal temperature will conflict with the calibration and degrade engine power, fuel consumption and emissions. Engine speed and load in the actual driving condition can change faster than the coolant temperature can respond, especially on heavy acceleration and sudden deceleration. Figure 1.5 illustrates the engine speed (red line) and load (blue line) behaviour throughout one of the legislation drive cycles. Coolant temperature low response is due to the thermal inertia present in the engine wall structure and also in the coolant. Therefore, the coolant temperature experiences longer time during warm-up and cool-down compared to the engine speed and load change (Figure 1.6).
Figure 1.5: The engine speed and load with the coolant target temperature behaviour based on the 3\textsuperscript{rd} generation engine cooling system from Audi throughout the US06 drive cycle.

Figure 1.6: Coolant temperature response example from a simulation.

There are two principles ways of speeding up a control loop. The first one is to use fast actuators with large control authority. For a cooling system, this would require a wide
possible temperature range and very high or very low flow rates. This option is usually
difficult and costly to implement, because major design changes need to be done, which
would increase the development cost and it is impractical for mass production. In a cooling
system, further issues arise, such as physical limits, packaging and high energy consumption.
For instance, a lot of energy is needed for the radiator fan to change the coolant
temperature quickly, because the fan would be operating outside its ideal operating region.
This high energy demand in return would cause poor fuel efficiency, defeating the original
objective.

The second option is to use a more sophisticated control system, which makes the best use
of the existing components. Since this option does not require major design changes, it does
not lead to the same cost increase and packaging issues. However, the downside of this
option is that it requires advanced control algorithms and much more detailed information
to handle disturbances quickly without suffering too much from uncertainties and non-
linear effects.

Disturbances in the engine thermal management come from three main areas:

- the heat generated by the combustion
- driver action (engine speed and torque, which in turn determine the temperature
  reference)
- Environmental factors including the air temperature, wind and road slope.

Figure 1.7 below illustrates the interaction between the three elements. In the near future,
with the availability of equipment such as the GPS navigation system and driver style
identification, vehicle parameter monitoring, on-board camera and radar sensors, it will be
possible to anticipate most of these factors with a reasonable degree of certainty [23].
Advanced knowledge of future disturbances is a distinct advantage in control, and
therefore, future engine thermal management systems should consider a predictive type
controller that can integrate this knowledge.
1.2. **Research Aims and Contributions**

The research aim of this thesis is the development of a novel multivariable thermal management system that manipulates both available variables (the coolant temperature and flow rate) to best effect.

The system uses the Multi-Input Multi-Output (MIMO) Model Predictive Control (MPC) to deal with the freedom and the constraints presented by having two control inputs. Engine thermal management is known to be nonlinear system where the direct linear MPC implementation is not straightforward. Furthermore, manipulating both the coolant temperature and flow rate can amplify the nonlinearity.

No similar attempt can be found in the literature, and the closest work by M. Bruckner et al. [24] used linear MPC with only one control variable for the thermal management. While it does include an electric water pump, the coolant temperature was maintained at 90°C using a conventional wax thermostat.

In order to implement the MPC controller, a number of further contributions are required:
• The feedback linearization method is used in this work to solve the nonlinearity problem for the linear MPC. Feedback linearization is one of the methods to linearize the nonlinear system by algebraically transform the nonlinear systems dynamics into fully or partly linear system.

• It is obvious that two control inputs provide more control authority than one. However, there is no systematic study on how the response time of the thermal management can be reduced, and how the fast response to the flow rate and the comparatively slow response to the temperature can be used to obtain the best effect. This work quantitatively compares the temperature response between using a single actuator and using both controller actuators. Then, the response is compared with the engine speed and load changes in the drive cycles. Based on this analysis, the gap between the required and the actual response time can be reduced significantly.

This work is focused on the control problem itself, and therefore two related problems have been excluded from consideration. The first related issue is the warm-up process from the cold-start to optimal engine operating temperature. It will not be considered here, although it is an important part of the engine thermal management. The warm-up process also requires extensive consideration of the after-treatment system, which is beyond the scope of this work.

Secondly, the prediction of driver behaviour and outside environment is not pursued here, instead it is assumed to be addressed separately. The reason is that it is a completely unrelated research area, dealing with human and open systems, requiring a very different field of expertise. Numerous studies on predicting the driver behaviour have already been published [25,26]. The proposed engine thermal management would also work very well with autonomous vehicles, where prediction is an integral part of the autonomous control strategy [27–29]. In this case, all the knowledge could be available for this engine thermal management as future prediction.
1.3. **Organization of the Thesis**

A brief description of the thesis organization is given below.

**CHAPTER 2: Case of Thermal Management**

This chapter focuses on the literature review and discusses the state of the art based on relevant presentations. It explains how thermal management can reduce the fuel consumption and improve the engine power output. Several details are discussed, including the engine friction, heat transfer, and exhaust emissions, which form the building blocks for a novel thermal management system.

**CHAPTER 3: Engine Thermal Management Potential**

This chapter covers the modelling of the engine and the combustion process, to establish the influence of the engine thermal management on fuel economy. The model is created in GT-SUITE, and includes established models for parts of the engine as discussed in the literature review. Based on this model, an engine calibration is performed for different temperatures. The result demonstrates the potential for improved efficiency.

**CHAPTER 4: Drive Cycle Volatility**

The engine calibration is a static process, but in reality the operating point of optimum temperature is constantly changing. Therefore, it is impossible to implement a calibration accurately, because the dynamics of the system can cause a delay. This is especially true for thermal management, because the thermal time constant is comparatively large. This chapter discusses two important questions: (1) the water pump and electric servo valve effect on the temperature response speed, and (2) the thermal management expected environment during real driving from eight legislative and research drive cycles. This gives an indication on the importance of having both the water pump and electric servo valve rather than using one of them in the engine thermal management to confront the high volatility of temperature set points in the actual driving.
CHAPTER 5: Model Predictive Control

Engine thermal management with the water pump and electric servo valve as its control input requires a controller that is capable of handling the Multi-Input Multi-Output (MIMO) system, where the Model Predictive Control (MPC) is known to be good at. This chapter gives an introduction of the MPC background, concept, and challenges when applied on the engine thermal management. Previous attempts on using MPC on the engine thermal management are also discussed.

CHAPTER 6: Engine Modelling for MPC

This chapter explains the mathematical model developed in the MathWork™ Simulink®. The model is a less complex model compared to the model made in the GT-SUITE (CHAPTER 3) purely for controller development. The Simulink® model will be used for the controller study when comparing the conventional method MPC and the newly developed Feedback Linearization MPC.

CHAPTER 7: Linear MPC on Thermal Management

This chapter demonstrates the difficulty and problem encountered from the implementation of the conventional linear MPC for the nonlinear engine thermal management. Two linearization approaches are discussed; the linearization from the MathWork™ System Identification and the Jacobian Linearization. The MPC controller step-by-step implementation as well as the controller performance and its influence to computational burden are explained. The result will be an important indicator to a new linearization approach for the MPC with respect to the engine thermal management problems.

CHAPTER 8: The New Engine Thermal Management Strategy

This chapter presents the new proposed thermal management strategy in detail and its implementation method. Feedback linearization is applied to the model in order to address
the strong non-linearity in the engine cooling system. This allows reformulating the problem in a near linear form that is then solvable via the MPC.

CHAPTER 9: Experimental Validation of the Controller Dynamic

The validation of the new MPC controller is done on a test rig constructed, similar to the engine cooling layout with a water jacket, an aluminium block, an electric pump, an electric servo valve, a fan and a radiator. This chapter experimentally demonstrates the new MPC controller dynamic behaviour in handling engine thermal management as well as a validation of the new controller concept.

CHAPTER 10: Benefits over a Drive Cycle

This chapter introduces the real world environment to the new Feedback Linearization MPC controller. The new controller performance is evaluated in eight legislative and research drive cycles as discussed in CHAPTER 4 and coupled with the engine, cooling system and vehicle model developed in the GT-SUITE (CHAPTER 3). The performance is also compared to the conventional cooling system in terms of the temperature set points and the actuator power consumption.

The overall thesis organisation and its interconnection between chapters can be seen in Figure 1.8 below.

Figure 1.8: The organisation of the thesis.
This chapter explains all the benefits of thermal management based on the literature review. Four main aspects on how thermal management can benefit the engine outputs are being discussed in this chapter:

- Reduced friction loss;
- Improved combustion;
- Reduced emissions; and
- Better radiator use.

An engine model in GT-SUITE (CHAPTER 3) is built to produce the similar engine output characteristics mentioned above. This literature information also gives the perspective in engine calibration to optimize engine outputs.
2.1. **Friction loss**

**Engine Friction**

One of the basic approaches to improve engine efficiency is by reducing the engine’s mechanical friction loss. Many changes have been studied to reduce this mechanical friction loss, such as; journal bearing coating, variable flow rate oil pump, reduced valve spring tension and roller tappet valve train.

Mechanical friction losses typically consume around 11% of the engine indicated output at full load. Because friction force is nearly constant to an engine speed, the relative loss increases significantly at the part load conditions [30]. Since engines are dimensioned for the peak power demand, most of the normal driving conditions are in part load. Therefore any reduction in the mechanical friction can significantly improve the fuel economy.

Lubricants are used to reduce friction and to prevent mechanical wear from metal to metal surfaces acting on each other. Both functions are highly dependent on the lubrication characteristics. The conditions can be divided into three regimes:

- boundary friction regime,
- hydrodynamic friction regime and
- mixed friction regime

As illustrated in Figure 2.1 below, the boundary friction regime has the highest fiction force and wear. It occurs when the oil film between two rubbing surfaces is so thin that it is not completely formed, causing surface contact and direct load transfer. In contrast, the hydrodynamic friction is when the two surfaces are fully separated by the oil film and the load is supported by the lubrication viscous pressure force. Mixed friction is a combination of both the boundary friction and hydrodynamic friction.
The Stribeck Curve in Figure 2.2 shows these three regimes and the resulting friction coefficient over the lubrication parameter. The variation of the friction regimes is a function of dimensionless lubrication parameter $\mu \times N/P$ where $\mu$ is the dynamic viscosity, $N$ is the relative speed between two rubbing surfaces and $P$ is load to the surface.

Engine friction can be considered as mostly hydrodynamic friction. Of course, there are variations across the components and operating points, but more than 70% of the total engine friction is in the hydrodynamic friction regime. Figure 2.3 below illustrates the engine components across the friction regimes. Differences can occur as a result of variations in the surface relative speed, oil viscosity and load across engine components [31–35].
Figure 2.3: Overview of engine components friction behaviour across Strubeck Curve [31–35].

The minimal friction is at the boundary between the mixed friction and the hydrodynamic friction regime. Since engine friction is caused mostly in the hydrodynamic friction regime, it could be reduced by lowering the lubrication parameter. This can be achieved by either lowering the surface relative speed $N$, oil viscosity $\mu$, or increasing the load $P$. Speed and load cannot be changed easily in an existing engine; therefore the oil viscosity is the variable to manipulate.

The most common methods to reduce lubrication viscosity are either using a lower oil grade [36,37] or increasing the oil working temperature [32,33,35,38]. The temperature has the advantage that it can be changed quickly to adapt to new engine conditions, while the oil grade is a choice that applies for a long time. The main danger of using low viscosity oil is that more engine components may enter mixed and boundary frictions during extreme conditions.

The oil temperature can be controlled using the oil cooling system, which can help to improve engine efficiency as well as protect engine component from wear when required.
The oil temperature can be increased by the controller for a lower friction, hence better engine efficiency. On the other hand, the temperature can be reduced when the engine enters high torque conditions to protect from wear in boundary and mixed friction regime.

An alternative of having a direct oil temperature controller (e.g. via oil temperature thermostat) is by having coolant thermal management. This can be done as the engine lubrication system removed the heat by transferring it to the coolant in the oil heat exchanger before the heat can be released to the environment.

**Piston Friction**

The piston assembly contributes the largest share of mechanical frictions. The piston skirt and piston rings contacts with cylinder liner contribute 30% to 40% of the total engine friction [39,40], as shown in Figure 2.4 [41]. Piston friction is a fundamental effect of any reciprocating engine: it cannot be avoided completely, but it can be minimised.

![Figure 2.4: Typical component friction in an engine](image)

Many approaches have been done to reduce the piston skirt and piston ring contact friction with cylinder liner. Initial attempts concentrated on the mechanical design. Some of the options include reducing the piston skirt area, reducing the piston mass, changing from the plateau honed to dined-honed liner, having an additional friction pad at the piston skirt, and reducing ring tension and width. A combination of piston rings pack modifications and an additional friction pad at piston skirt shows the best effect: a study by Leong et al. (2007)
recorded a friction reduction of 38% [42]. Reducing the piston mass alone by 25% helps to lower the friction by only 0.07 bars of FMEP at 1500rpm [30].

Engine thermal management is a complementary approach to reduce the piston rings pack and skirt friction with or without the design modification. The oil film viscosity between the piston and cylinder liner has a major influence in the piston friction. The oil sump temperature only has a negligible influence in the piston friction, because the oil film between the piston and cylinder liner is extremely thin (only a few tens of μm). As a result, the oil film heat capacity is very low, and the lubrication temperature follows the cylinder liner temperature nearly completely [33].

Piston friction happens over a range that can be divided into three sections; upper section, middle section (where this is the highest piston speed point) and the lower section of cylinder liner as in Figure 2.5. The high temperature and low speed cause the boundary friction regime and mixed friction regime to dominate at the upper section, while the hydrodynamic friction is found in the middle and lower sections. Due to the larger distance covered, the middle section contributes the highest friction load [33,43].

![Figure 2.5: Upper, middle and lower sections of the cylinder liner friction.](image)

The most effective way of reducing piston friction is by reducing the middle section friction. The oil temperature at the middle section liner surface is dependent on the coolant temperature variation around the cylinder liner. An increased coolant temperature will reduce the thin oil film viscosity and therefore lower the middle section friction.
2.2. **Combustion Quality**

Combustion quality is a key factor in achieving better fuel efficiency, higher power and lower emission. The combustion quality can be improved by manipulating the engine temperature in four ways (although not at the same time):

- reduce heat loss,
- reduce knock tendency,
- increase volumetric efficiency, and
- improve air-fuel mixture quality.

**Reduce Heat Loss**

Almost 70% of fuel energy is being rejected as heat loss from the total fuel energy without serving as a useful power to run the vehicle [10]. The two main shares of the energy loss go into the exhaust gas and into the cooling system. The movement of the coolant is the main element that extracts from the engine into the environment to prevent the engine from overheating. To achieve this, the combustion chamber is surrounded by a complex design of the water jacket which contains the coolant.

Heat loss by convection to the coolant is typically roughly comparable to the engine power output at full load conditions. When using a mechanical water pump, the proportion of heat going into the coolant can increase at part load conditions [44]. This is partly due to the fixed flow rate of the mechanical water pump to the engine speed, which has to be designed to protect the engine from excessive heat for the worst cases. Because of the fixed ratio, the pump flow rate is proportionate to the engine speed, but not to the engine load. This leads to excessive coolant flow rate at the part load condition due to the fact that the water pump speed could not respond to the load changes, which in turn means that the combustion temperature is much lower than the ideal.

One way to reduce the heat loss during part load is by reducing the heat convection to the coolant. The heat convection can be reduced by either lowering the coolant flow rate or increasing the temperature. It was experimentally proven by Willumeit et al. (1984) [44]...
that fuel consumption reduction can be achieved up to 20% by reducing the convection heat at the part load. This reduces the heat loss into the cylinder wall, increases the combustion temperature and pressure, thus improving the combustion quality and increasing the work per cycle.

**Reduce Knock Tendency**

It is a common understanding that, higher compression ratio and higher intake boost pressure can increase the engine efficiency and performance, but this is limited by the knock in the spark ignition engine. Furthermore, engine knock is the main reason that the engines could not reach the maximum brake torque ignition timing (MBT). Running in the knock condition could lead to catastrophic engine failure.

Knock is an abnormal combustion phenomenon characterised by the spontaneous ignition of an unburnt portion of the air-fuel mixture before it is reached by the flame front. The occurrence depends both on the ignition delay of low temperature oxidation reaction (which is a function of composition, temperature and pressure) and the speed of the flame front. This phenomenon usually happens at high engine load and low speed, which is a serious problem especially for modern downsized engines.

Two common methods in the engine calibration to prevent knock are delaying the ignition timing and enrichment of the air-fuel mixture. Delaying ignition timing may seem counter-intuitive, because it means that an even longer ignition delay is required. Although it is an effective method at high load, the delayed ignition reduces the cylinder pressure. The enrichment of air-the fuel mixture reduces the combustion speed through dilution. However, both changes have an adverse impact on fuel consumption and high emission.

Design changes such as intercooler rating, direct injection and a squish combustion chamber are typically used to reduce the knock tendency. Studies have shown that knocks can also be reduced by lowering the cylinder head temperature [45–48]. For an example; Kobayashi et al. (1984) [49] used a separate block and cylinder head cooling circuit to gain a better control of knock behaviour to improve the fuel consumption and performance. The results show that lowering the cylinder head temperature suppressed the knock even in higher
compression ratio configuration. The experiment demonstrates that the engine performance output improved by 10% at high speed and fuel economy improving by 5% at part load and 7% during idling.

The main effect is that a lower cylinder wall temperature helps cooling the air-fuel mixture during the inlet and compression stroke prior to ignition. A low cylinder head temperature in the intake port also reduces the air-charge temperature before entering the combustion chamber. In addition, the temperature of the residual gas in the combustion chamber is also reduced by a lower cylinder head and piston top temperature. All these effects contribute to a lower air-fuel mixture temperature, which help to avoid knock.

Overall, thermal management allows spark ignition engines to run with more advanced spark timing, higher compression ratio and potentially higher boost pressure without encountering any knock. Of course, excessive cooling of the cylinder wall surface will create an unnecessary increase in the heat loss and friction, which can cause poor performance and fuel consumption, so a compromise needs to be reached.

**Increase Volumetric Efficiency**

Volumetric efficiency is an indication of the engine breathing capability. It is defined as the ratio of the amount of air that enters the cylinder to the cylinder volume displacement at reference pressure and temperature as in equation (1) below. Improving the engine performance by improving the volumetric efficiency is a very popular technique in motorsport. This is due to the fact that higher volumetric efficiency means more O2 can be used for the fuel to create more energy from the combustion. This includes valvetrain modification, intake system modification, exhaust system modification and even turbocharging.

\[
\eta_{VE} = \frac{[\text{volume of air taken into cylinder at reference point}]}{[\text{cylinder swapt volume}]} \tag{1}
\]

Temperature and pressure are the two main factors influencing air density and therefore charge efficiency. Air heating from the contact with the cylinder head and the cylinder wall can reduce the volumetric efficiency. A lower surface temperature will cause the air charge
temperature to reduce, which creates higher air charge density. Trapped air residual pressure in the cylinder can also be reduced via a lower cylinder temperature. As a result, improved cooling can contribute to a better volumetric efficiency because more air charge can enter the cylinder.

Thomas et al. (2011) [50] experimentally demonstrate that lower inlet port and cylinder wall temperature can improve the engine volumetric efficiency and power output. It was also reported that the engine friction will increase as the cylinder wall temperature is reduced. The increment of the friction load is still low compared to the increment of the indicated combustion pressure from higher volumetric efficiency at certain engine operations especially at full load, resulting in higher engine torque at the expense of fuel economy. The high volumetric efficiency from the lower cylinder wall temperature is not effective at lower engine load, because the engine friction load increase rate is greater than the indicated combustion pressure.

**Improve Air-Fuel Mixture Quality**

Air-fuel mixing is another factor that influences the combustion quality and therefore influences the engine output performance and emission. The best results are typically achieved with a more homogenous air-fuel mixture for spark ignition engines.

The fuel vaporization rate is one of the main factors in determining the air-fuel mixture quality. A low vaporization rate causes the build-up of a liquid fuel film on the wall, especially in the intake ports of multi-ports fuel injector engine. Excessive quantities of liquid fuel film can cause large combustion variations, vehicle driveability problems, bad emission and high fuel consumption [51,52]. Unfortunately, only 20% of fuel is directly vaporized from the fuel injector spray under most conditions [51]. The remaining 80% of fuel injected forms a liquid fuel film on the wall that has to evaporate.

Using a higher wall temperature can help to improve the fuel vaporization rate, as do higher air and fuel temperature. Chen et al. (1996) [51] demonstrated that higher wall temperature improves the evaporation rate. Gasoline fuel is a mixture of a large number of individual hydrocarbons that have different boiling temperature points. Too low wall temperature may
cause only some of these to evaporate, which can then affect the fuel composition and cause problems in the combustion. The hydrocarbons will completely evaporate at temperatures higher than 180°C, so this is a suitable temperature for the inlet port wall.

2.3. Emission

As stated earlier, friction reduction and better combustion quality from the thermal management can improve engine efficiency. Indirectly this also reduces CO₂ gas emission [10].

Toxic emission gases such as Nitrogen Oxide (NOₓ), Carbon monoxide (CO) and Hydrocarbon (HC) are affected even stronger by the thermal management practice. These toxic gases can have a very harmful impact on the environment and the health of the people affected. To address this challenge, road vehicles have permitted target limits for the toxic emissions under European legislation, with EURO 6 being the current application version.

Hydrocarbon (HC)

HC emission is mainly caused by the unburned hydrocarbon from the combustion process. This may be caused by an insufficient amount of air in the cylinder, bad mixing between fuel and air, or insufficient time or temperature to burn the fuel completely, especially in small crevices around the combustion chamber wall. The unburned hydrocarbon emitted via the exhaust valve is then considered as toxic emissions.

It is very well known that temperature has a large impact on the combustion process, and therefore on the amount of HC emission. A high coolant temperature increases the in-cylinder wall temperature, which again encourages the hydrocarbon to evaporate and to burn completely.

Chanfreau et al. (2001) [53] recorded a 17% HC gas emission reduction by increasing the coolant out temperature to 115°C in the transient drive cycle. Meanwhile, Couetouse and Gentile (1992) [54] recorded a 10% to 20% HC gas emission reduction in steady state
conditions. Russ et al. (1995) [55] and Guillemot et al. (1994) [56] stated that by increasing either the cylinder block or cylinder head temperature, it leads to significant reduction in HC emission. Moreover, Willumeit et al. (1984) [46] identified that HC reduction up to 50% are possible at low speed if the cylinder wall top side temperature is maintained at 200°C.

**Nitrogen Oxide (NO\textsubscript{x})**

NO\textsubscript{x} gas emission shows the opposite behaviour with temperature: NO\textsubscript{x} tends to form at high temperatures, especially with excess oxygen available. NO\textsubscript{x} gas emission therefore increases with the coolant temperature. The key reaction is that at high temperatures and pressures, nitrogen and oxygen react with each other to produce NO\textsubscript{x}. NO\textsubscript{x} emissions are therefore a problem especially at high load and high speed conditions. High coolant temperature increases the combustion temperature and therefore the rate of NO\textsubscript{x} production.

Couëtouse and Gentile (1992) [54] recorded that NO\textsubscript{x} gas emission increases up to 10% to 20% in steady state conditions with the elevated coolant temperature. Interestingly, Chanfreau et al. (2001) [53] stated that there was no noticeable increase in NO\textsubscript{x} gas during the transient drive cycle. This was possible due to the fact that the friction loss reduction from thermal management causes lower IMEP which leads to lower combustion pressure and temperature.

**Carbon Monoxide (CO)**

CO emission is basically caused by fuel that is only partially oxidized, resulting in CO rather than CO\textsubscript{2}. Although not technically a hydrocarbon gas, the formation of CO follows a similar pattern as HC – the oxidation process is not completed due to the lack of time or temperature. This commonly happens in rich air-fuel mixture that causes oxygen starvation and leads to an increase in CO emission. In other words, CO emission mainly depends on the air-fuel mixture quality rather than the combustion temperature. However, a higher cylinder wall temperature could reduce CO emission by improving the air-fuel mixture.

Couëtouse and Gentile (1992) [54] recorded that CO emission was unchanged when increasing the coolant temperature. Chanfreau et al. (2001) [53] reported about CO
emission reduction in a transient drive cycle with thermal management. Meanwhile, Santhosh et al. (2011) [50] stated that CO emission is slightly increased if the coolant temperature reduces during the wide open throttle (WOT) steady state test.

Table 2.1 below is a summary of thermal management influence on toxic emissions as gathered from the literature. The experimental studies demonstrate that the impact of the cylinder wall temperature on toxic emissions behaviours is considered with the theoretical model of combustion. The HC and to a lesser degree CO emission can be reduced with a higher cylinder wall temperature, meanwhile NO\textsubscript{x} emission shows the opposite effect – that they can be reduced with a lower cylinder wall temperature.

<table>
<thead>
<tr>
<th>Test by</th>
<th>Test type</th>
<th>Thermal Management style</th>
<th>Emission changes</th>
</tr>
</thead>
</table>
| Couëtouse and Gentile (1992) [54] | 3 constant speed test            | Coolant temperature 115°C | • 10% to 20% HC reduction  
• 10% to 20% NO\textsubscript{x} increase  
• CO unchanged |
| Chanfreau et al. (2001) [53]       | FTP75 + HWFET Drive cycle test   | Average Coolant temperature higher 10°C than normal | • 15% CO reduction  
• 17% HC reduction  
• NO\textsubscript{x} unchanged |
| Russ et al. (1995) [55]           | 1500rpm and 3.8 bar IMEP.       | Varying coolant temperature at head and block separately from 70°C to 110°C. | • 8% HC reduction in every 10 increase  
• 4.6% HC reduction in every 10 increase at block only.  
• 3.1% HC reduction in every 10 increase at head only.  
• There are no details on CO and NO\textsubscript{x} |
| Guillemot et al. (1994) [56]      | Steady State                     | Changing coolant temperature at block only and cylinder head only from 90°C to 35°C. | • HC increase from 50% to 75%. With head contributing 70% of the increment.  
• There are no details on CO and NO\textsubscript{x} |
| Santhosh et al. (2011) [50]       | WOT Steady state                 | Maintaining coolant temperature at 60°C and 70°C from 95°C. | • Significant increase in HC especially at low speed  
• Slight increase in CO  
• Slight decrease in NO\textsubscript{x} |
| Willumeit et al. (1984) [44]      | Steady State (carburettor engine) | Maintaining top cylinder wall temperature at 200°C. | • HC reduction 10% to 50%.  
• There are no details on CO and NO\textsubscript{x} |
2.4. **Radiator**

The radiator is a key component of the cooling system and therefore of any thermal management system. Radiators in automotive engineering emerge with the introduction of the water-based cooling systems. The function of the radiator is to remove excessive heat produced by the combustion to the environment – it works as a heat exchanger between the water and air.

Radiator design can have an important impact on vehicle fuel consumption and performance. Until the late 1990s, almost all vehicle radiators were made from aluminium. Aluminium, compared to non-ferrous metal (brass and copper), is lightweight, cheap, and combines a number of positive attributes such as high heat transfer coefficient, strength, corrosion resistance, convenient processing and process quality. An aluminium radiator can be smaller and lighter than other non-ferrous radiators. A smaller radiator reduces the aerodynamic resistance and the weight of the vehicle, where both improves the economy [57]. In fact, the aerodynamic impact of the radiator has been studied in great detail, and blocking the air flow through the radiator when it is not required has been found to reduce the drag of a vehicle [58].

High coolant temperature enables the use of a smaller radiator, because the higher temperature difference leads to better radiator efficiency. Charyulu et al. (1998) [59] showed that even a 1°C coolant temperature increment significantly improves the radiator heat removal. Meanwhile Krüger et al. [60] stated that by increasing the coolant temperature by 5°C, the air mass flow can be reduced by 10%, which can reduce the fan drive power requirement by as much as 30%.

2.5. **Summary**

The work established in the literature reviews shows that the engine outputs can be improved by manipulating the in-cylinder wall temperature and coolant temperature. The relevant effects can be seen in the friction reduction, improved air-fuel mixing, reduced heat
loss, increased engine volumetric efficiency, advanced spark timing, suppressed knock and reduced radiator fan speed. In addition, toxic emission such as CO, HC and NO\textsubscript{x} can also be reduced. Table 2.2 summarizes the overall thermal management effects for each element under consideration. Some elements show a clear trend with temperature, but for some the trade-offs are still unclear. The amount of improvements depends on the engine type and design.
Table 2.2: Summary of the thermal management influence in the engine output.

<table>
<thead>
<tr>
<th>Items</th>
<th>Low temperature</th>
<th>High temperature</th>
<th>Remark</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Engine Component Frictions</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Piston assembly (without upper cylinder liner section)</td>
<td>x</td>
<td>O</td>
<td>The highest friction component in the engine Effective before entering the mixed friction regime.</td>
</tr>
<tr>
<td>Piston upper cylinder liner section</td>
<td>O</td>
<td>x</td>
<td>Effective before entering the hydrodynamic friction regime.</td>
</tr>
<tr>
<td>Valvetrain</td>
<td>O</td>
<td>x</td>
<td>Effective before entering the hydrodynamic friction regime.</td>
</tr>
<tr>
<td>Cranktrain</td>
<td>x</td>
<td>O</td>
<td>Effective before entering the mixed friction regime.</td>
</tr>
<tr>
<td>Oil pump</td>
<td>x</td>
<td>O</td>
<td>Effective before entering mixed friction regime.</td>
</tr>
<tr>
<td><strong>TOTAL FRICTION</strong></td>
<td>x</td>
<td>O</td>
<td>Effective before friction increases back at some point.</td>
</tr>
<tr>
<td><strong>Combustions</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Heat loss reduction</td>
<td>x</td>
<td>O</td>
<td>High wall temperature reduces the combustion heat transfer to wall.</td>
</tr>
<tr>
<td>Anti-knock</td>
<td>O</td>
<td>x</td>
<td>Only effective at the knock region.</td>
</tr>
<tr>
<td>Volumetric efficiency</td>
<td>O</td>
<td>x</td>
<td>Significant at full load.</td>
</tr>
<tr>
<td>Air-fuel mixture quality</td>
<td>x</td>
<td>O</td>
<td>Not so effective if the wall temperature is above 180°C.</td>
</tr>
<tr>
<td><strong>TOTAL COMBUSTION</strong></td>
<td>Δ</td>
<td>Δ</td>
<td>Could not be specifically defined. Depending on the engine speed and load.</td>
</tr>
<tr>
<td><strong>Emissions</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NO&lt;sub&gt;x&lt;/sub&gt;</td>
<td>O</td>
<td>x</td>
<td>Reduced coolant temperature generally reduces NO&lt;sub&gt;x&lt;/sub&gt;.</td>
</tr>
<tr>
<td>CO&lt;sub&gt;2&lt;/sub&gt;</td>
<td>Δ</td>
<td>Δ</td>
<td>CO&lt;sub&gt;2&lt;/sub&gt; highly dependent on engine efficiency.</td>
</tr>
<tr>
<td>CO</td>
<td>Δ</td>
<td>Δ</td>
<td>CO is highly dependent on air fuel mixture quality.</td>
</tr>
<tr>
<td>HC</td>
<td>x</td>
<td>O</td>
<td>High component temperature significantly vaporizes HC.</td>
</tr>
<tr>
<td><strong>TOTAL EMISSIONS</strong></td>
<td>Δ</td>
<td>Δ</td>
<td>Could not be specifically defined. Each emission component reacts differently.</td>
</tr>
<tr>
<td><strong>Radiator</strong></td>
<td>x</td>
<td>O</td>
<td>Higher coolant temperature improves radiator efficiency.</td>
</tr>
<tr>
<td><strong>Components durability</strong></td>
<td>O</td>
<td>x</td>
<td>Generally lower temperature extends component life. Thermal shock will also increase thermal fatigue.</td>
</tr>
</tbody>
</table>

Remark:
- **O**: Better
- **Δ**: Uncertain
- **x**: Worst
This chapter will elaborate on the details of the engine model setup and the resulting thermal management potential. The engine model is based on one of the reference models in GT-SUITE and it is modified to suit the thermal management and engine output study as described in CHAPTER 2. The model variables like the spark timing, air-fuel ratio and cylinder wall temperature are calibrated to optimize the engine thermal efficiency. The engine thermal management potential is described by applying the optimized engine model in the driving cycles and comparing it with the conventional engine cooling system. The driving cycles are based on academic and legislative test cycles, including the Artemis cycles and US EPA Federal test cycles.

The GT-SUITE engine model will be later used to create a plant model in MATLAB™ Simulink® for the thermal management control strategy study. The Simulink® model is a model reduction of the GT-SUITE model to reduce the complexity but it still has similar dynamic behaviour.
3.1. **Engine Model**

The analysis of thermal management is performed in numerical simulations using the GT-SUITE software from Gamma Technologies. This software is designed for the automotive field work simulation and analysis. It can be linked with several third-party software such as MathWork™ Simulink®, if additional functionality is required. The ability to link GT-SUITE and MathWork™ Simulink® is advantageous for control analysis purposes, and this will be used at the later stage of this research.

Due to budget constraints and the difficulty of measuring small BSFC changes, the results could not be quantitatively verified on a test engine. The advantage of the mode is that it can provide data over a wide range of conditions without risk of damage, small difference can be easily quantified, and it reduces the time and effort compared to an experiment.

**Engine specification**

A 4 cylinder 2.0L naturally aspirated engine is chosen for this purpose, mainly due to the fact that it is one of the common types of passenger vehicle engine categories [61]. Furthermore, a generic, non-manufacture specification sample model of this engine type is also available in the GT-SUITE. This work is based on the sample engine model, which is later modified to incorporate the thermal management model and the MathWork™ Simulink® controller. The engine model specification is as in Table 3.1 below.

<table>
<thead>
<tr>
<th>Displaced volume (cc)</th>
<th>1999</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bore (mm)</td>
<td>86</td>
</tr>
<tr>
<td>Stroke (mm)</td>
<td>86.07</td>
</tr>
<tr>
<td>Number of cylinder</td>
<td>4</td>
</tr>
<tr>
<td>Compression ratio</td>
<td>10</td>
</tr>
<tr>
<td>Fuel Injection Type</td>
<td>Port injection</td>
</tr>
<tr>
<td>Number of valve per cylinder</td>
<td>4</td>
</tr>
<tr>
<td>Engine type</td>
<td>Gasoline Naturally Aspirated</td>
</tr>
</tbody>
</table>
Figure 3.1: 2.0L naturally aspirated engine model in GT-SUITE.

The engine power and torque in the sample model is rated at 99.2kW@6500rpm and 180.4Nm@4000rpm. The performance of the engine model is considered to be slightly lower than the average real world natural aspirated engines (Figure 3.2). This is due to the fact that the engine model specification is not being fine-tuned such as its cam timing, intake manifold dimension and exhaust manifold layout. However, the fine tuning is not required as the effect is not related to the engine thermal management.
The most common concept is by using the friction model dependent to oil viscosity as the temperature determines the oil viscosity property. The friction model using oil viscosity is as shown in equation (2) below [42,66,67]:

\[
fmep = fmep_{ref} \left( \frac{\mu}{\mu_{ref}} \right)^n
\]

\(fmep\) = Friction Mean Effective Pressure (FMEP) [bar]
\(fmep_{ref}\) = FMEP at reference oil temperature [bar]
\(\mu\) = Oil dynamic viscosity [kg/ms]
\(\mu_{ref}\) = Oil dynamic viscosity at reference oil temperature [kg/ms]

The index \(n\) is based on the engine. Typically, it is 0.19 to 0.24 for gasoline engine and 0.25 to 0.32 for diesel engine [42]. The oil dynamic viscosity can be calculated using the temperature-dependent Vogel equation as shown in equation (3) below [37,42]:

\[\mu = \frac{A}{T - B}\]

\(A\) and \(B\) are constants specific to each oil type.

The index \(n\) is based on the engine. Typically, it is 0.19 to 0.24 for gasoline engine and 0.25 to 0.32 for diesel engine [42]. The oil dynamic viscosity can be calculated using the temperature-dependent Vogel equation as shown in equation (3) below [37,42]:

\[\mu = \frac{A}{T - B}\]

\(A\) and \(B\) are constants specific to each oil type.
\[ \mu = k_v \cdot \exp \left( \frac{\theta_1}{T + \theta_2} \right) \]  

(3)

\[ \frac{\mu}{T} = \text{Oil dynamic viscosity [kg/ms]} \]

\[ T = \text{Current oil temperature [°C]} \]

Where, \( k_v, \theta_1 \) and \( \theta_2 \) are constants determined for specific oil. However, this is true, as long as the sum of the overall engine friction is in the hydrodynamic regime. Unfortunately, the engine will eventually enter the mixed friction regime when the coolant or oil temperature is increased to a certain level [65,68]. The actual friction curve starts to deviate from the theoretical friction model line over the engine speed when it enters the mixed friction regime. As in Strubeck Curve, the effect is much stronger at lower engine speed than at high speed.

A friction model that can simulate the mixed friction behaviour is therefore required for a thorough thermal management study. As explained in CHAPTER 2, most of the thermal management advantages are during hot conditions. The hot conditions are near to the material limit working temperature, high knock tendency, low volumetric efficiency, high NOx as well as the fact that the overall engine friction starts to enter the mixed friction region where all these will give an opposite effect to the engine output. This will create a trade-off in determining the engine temperature to achieve higher engine efficiency. An accurate friction model will create an accurate trade-off in the thermal management strategy.

Fischer’s engine friction model (developed by G. Fischer (1999) [65]) captures the engine mixed friction regime behaviour. This engine friction model is a function of the lubricant temperature, coolant temperature, brake specific mean pressure (BMEP) and engine speed. The friction model was validated using a few spark ignition engines. The friction model requires two engine friction mean effective pressure (FMEP) points as the input reference points. Both inputs should be measured at 0bar BMEP with 90°C coolant and lubricant temperature. Preferably, the input reference points should be, one at lower engine speed and another one at higher engine speed. The equation is shown below:

\[ f_{mep} = C_0 + C_1\left( A_0 + A_1 \cdot N_e + A_2 \cdot N_e^2 \right) + \Delta P_{BMEP} \]  

(4)
\( f_{mep} \) = Friction Mean Effective Pressure (FMEP) [bar]

\( N_e \) = Current engine speed [rpm]

The coefficient \( A_0, A_1 \) and \( A_2 \) are the function of the conditional temperature as in equations (5), (6) and (7) below:

\[
A_0 = 1.0895 - 1.079 \cdot 10^{-2} T_{cond} + 5.525 \cdot 10^{-5} T_{cond}^2
\]

\[
A_1 = 4.68 \cdot 10^{-4} - 5.904 \cdot 10^{-6} T_{cond} + 1.88 \cdot 10^{-8} T_{cond}^2
\]

\[
A_2 = -4.35 \cdot 10^{-8} + 1.12 \cdot 10^{-9} T_{cond} - 4.79 \cdot 10^{-12} T_{cond}^2
\]

Where, \( T_{cond} \) is the engine’s current conditional temperature. The conditional temperature represents both the oil and coolant temperatures. The oil temperature affects the friction of the engine cranktrain, valvetrain and oil pump. Meanwhile, the frictional losses of the piston assembly are influenced by the coolant temperature. Overall, both temperatures act in approximately equal measure to the friction losses of the overall engine friction. The average of both temperatures is used when different coolant and oil temperature are fed. The oil temperature is taken at the oil sump, while the coolant temperature is taken at the coolant engine out temperature.

The \( C_0 \) and \( C_1 \) are determined by the two reference points at the reference conditional temperature. The \( C_0 \) and \( C_1 \) are expressed by the equation below:

\[
C_0 = f_{mep\_ref \_1} - C_1 \cdot (A_{0 \_ref} + A_{1 \_ref} \cdot N_{e \_1} + A_{2 \_ref} \cdot N_{e \_1}^2)
\]

\[
C_1 = \frac{f_{mep\_ref \_1} - f_{mep\_ref \_2}}{A_{1 \_ref} (N_{e \_1} - N_{e \_2}) + A_{2 \_ref} (N_{e \_1}^2 - N_{e \_2}^2)}
\]

\( f_{mep\_ref \_1} \) = FMEP at first reference point [bar]

\( f_{mep\_ref \_2} \) = FMEP at second reference point [bar]

\( N_{e \_1} \) = Engine speed at first reference point [rpm]

\( N_{e \_2} \) = Engine speed at second reference point [rpm]

The \( A_{0 \_ref}, A_{1 \_ref} \) and \( A_{2 \_ref} \) are the \( A_0, A_1 \) and \( A_2 \) as in equations (5), (6) and (7) at the reference conditional temperature.
The combustion pressure changes influence to the engine friction by friction force in piston assembly and cranktrain; as well as changes made by the oil film viscosity by the cylinder wall temperature are expressed in $\Delta P_{\text{BMEP}}$ part as equation (10) below:

$$
\Delta P_{\text{BMEP}} = B_0 + B_1 b\text{mep}
$$

$b\text{mep} = \text{Brake mean effective pressure [bar]}$

The coefficient $B_0$ and $B_1$ are both the function of the engine speed and conditional temperature as equations (11) and (12) below:

$$
B_0 = -2.625 \cdot 10^{-3} + 3.75 \cdot 10^{-7} N_e + 1.75 \cdot 10^{-5} T_{\text{cond}} + 2.5 \cdot 10^{-9} N_e T_{\text{cond}}
$$

$$
B_1 = 8.95 \cdot 10^{-3} + 1.5 \cdot 10^{-7} N_e + 7.0 \cdot 10^{-5} T_{\text{cond}} - 1.0 \cdot 10^{-9} N_e T_{\text{cond}}
$$

$N_e = \text{Current engine speed [rpm]}

T_{\text{cond}} = \text{Current conditional temperature [°C]}

**Combustion model**

The core of any internal combustion engine is the combustion process. Besides its provision of power to the engine, it also provides a heat energy source from the chemical reaction in the combustion chamber. Therefore, the combustion model is important in thermal management simulations. The model calculates fuel burn rate per unit of time or crank angle. Burn rate is the rate at which air and fuel molecules are transferred from the unburned zone to the burned zone.

Two types of combustion model template available in GT-SUITE [69]. The first is the non-predictive combustion model. This type of combustion model simply imposes a burn rate from experiment data as a function of the crank angle, for an example the pre-set Spark Ignition Wiebe Combustion Model. However, the burn rate will not change regardless of the condition in the cylinder, which means that the effect for varying spark-timing and air motion cannot be studied.
The second type is called the predictive combustion model. Opposite to the non-predictive combustion model, the predictive combustion model reacts to changes to the condition in the cylinder. Therefore, this model requires information on the spark timing, start of spark location, in-cylinder composition, in-cylinder flow, cylinder dimension, fuel properties and interaction between the flame and wall. The advantage of the predictive type of combustion model requires no measurement or test except for the initial model correlation. Furthermore, it is self-adjusting for transient conditions due to the fact that its burn rate is dependent on the mentioned input information. Therefore, the predictive type combustion model, the “EngCylCombSITurb” (SI Turbulent Flame Combustion Model) is suitable in this work to get the advantage of the thermal management to engine spark timing and knock suppression. However, one significant drawback of the predictive type combustion model is that the model will cause slower simulation time than the non-predictive combustion model following the added complexity of the model calculations. The combustion model calculation is based on the equation below [69]:

\[ \frac{dM_b}{dt} = \frac{(M_e - M_b)}{\tau} \]  
\[ \frac{dM_e}{dt} = \rho_u A_e (S_T + S_L) \]  
\[ \tau = \frac{\lambda}{S_L} \]

- \( m_b \) = Burned mass [kg]
- \( m_e \) = Converted unburned gas by combustion [kg]
- \( t \) = Time [s]
- \( \rho_u \) = Unburned gas density [kgm\(^{-3}\)]
- \( A_e \) = Flame surface area [m\(^2\)]
- \( S_T \) = Turbulence flame speed [ms\(^{-1}\)]
- \( S_L \) = Laminar flame speed [ms\(^{-1}\)]
- \( \tau \) = Time constant [s]
- \( \lambda \) = Taylor microscale length [m]

This combustion model is based on a two-zone predictive combustion model. It is assumed that the flame front propagates in a spherical manner by the function of the laminar and turbulent flame speed. The unburned mixture of air and fuel, converted into the combustion, \( m_e \) is proportional to the turbulent flame speed \( S_T \) and laminar flame speed \( S_L \). Meanwhile, the burn rate is proportional to the unburned gas behind the flame.
front, \((m_e - m_p)\) divided by a time constant \(\tau\). This time constant is the ratio of Taylor microscale, \(\lambda\) and the laminar flame speed. The turbulence intensity and length scale are provided by the embedded in-cylinder flow model.

Another important element that affects the spark timing is the knock limit as explained in the previous chapter. The knock provides constraints to the engine optimization problem. Therefore, the knock model is required to predict knock occurrences and remove them from the viable solutions of the optimization process. The knock model is used based on Douaud and Eyzat’s knock model [70]. The inputs for the knock model are the combustion chamber geometry, octane number and spark plug location. For the current study, a simple combustion chamber with a flat piston and head surface is specified. Spark plug location is positioned at the centre and 1mm from the head surface. Meanwhile, the octane number is set at 95.

The knock model defines knock occurrence based on the average induction time integral calculated at every individual surface contact with the bulk of unburned gas. The induction time integral is evaluated from the intake valve closing angle to the knock initiation. The combustion cycle is considered knocking when the induction time integral equation (16) below is one.

\[
I(t) = \int_{t_0}^{t_{\text{ign}}} \frac{dt}{\tau} = 1
\]

\(I\) = Induction integral time  \\
\(t\) = Time [s]  \\
\(t_{\text{ign}}\) = Time of autoignition [s]  \\
\(t_0\) = Time at intake valve close [s]  \\
\(\tau\) = Induction time [s]

Meanwhile equation (17) below is the induction integral time in the function of the crank angle.

\[
I(\theta) = \frac{1}{6 \cdot N_{\text{engine}}} \times \int_{\theta_0}^{\theta_{\text{ign}}} \frac{1}{\tau} d\theta = 1
\]

\(I\) = Induction integral time  \\
\(\theta\) = Crank angle [°]  \\
\(N_{\text{engine}}\) = Engine speed [rpm]  \\
\(\tau\) = Induction time [s]
\[ \theta_0 = \text{Crank angle at intake valve close [°]} \]
\[ \theta_{\text{ign}} = \text{Crank angle at autoignition [°]} \]

The induction time, \( \tau \), is a function of the instantaneous cylinder pressure and unburned gas temperature as in equation (18) below:

\[ \tau = M_1 \cdot 5.72 \cdot 10^6 \left( \frac{ON}{100} \right)^{3.402} \cdot p^{-1.7} \cdot \exp \left( \frac{3800}{M_2 T} \right) \quad (18) \]

- \( \tau \) = Induction time [s]
- \( M_1 \) = Knock induction time multiplier
- \( M_2 \) = Activation energy multiplier
- \( ON \) = Fuel octane number
- \( p \) = Instantaneous cylinder pressure [Pa]
- \( T \) = Instantaneous unburned gas temperature [K]

**Heat transfer model**

The heat transfer in the cylinder is modelled using “WoschniGT” correlation. It closely emulates the classical Woschni correlation without swirl, but the treatment of heat transfer coefficient during the opening of the valve is different. This model is recommended by Gamma Technology when the swirl data is not available [70].

GT-SUITE also has the Finite Element Method (FEM) in-cylinder structure for the thermal management analysis. The FE in-cylinder is as an interconnection to the heat transfer between the gas combustion and engine cooling system. The Combustion model fed the FE cylinder wall structure with instantaneous gas-side boundary data as the heat source of the engine. This creates the opportunities to include a detailed component thermal analysis in combustion chamber design studies. The in-cylinder structure is discretized into regions of the cylinder head, cylinder liner, piston, valves and ports as in Figure 3.3. This will create a better temperature distribution around the in-cylinder structure as compared to a lump of mass method. The in-cylinder structure dimension is given in Table 3.2.
Figure 3.3: Sample of post simulation FE in-cylinder structure temperature results.

Table 3.2: Geometry attributes for the FE in-cylinder structure in GT-SUITE.

<table>
<thead>
<tr>
<th>Attribute</th>
<th>Object value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Head</strong></td>
<td></td>
</tr>
<tr>
<td>Material</td>
<td>Aluminium</td>
</tr>
<tr>
<td>Head deck thickness</td>
<td>8 mm</td>
</tr>
<tr>
<td>Dome height</td>
<td>10 mm</td>
</tr>
<tr>
<td><strong>Piston</strong></td>
<td></td>
</tr>
<tr>
<td>Material</td>
<td>Aluminium</td>
</tr>
<tr>
<td>Top deck thickness</td>
<td>12 mm</td>
</tr>
<tr>
<td>Piston height</td>
<td>60 mm</td>
</tr>
<tr>
<td>Skirt thickness</td>
<td>4.5mm</td>
</tr>
<tr>
<td>Piston ring thickness</td>
<td>2 mm</td>
</tr>
<tr>
<td><strong>Cylinder liner</strong></td>
<td></td>
</tr>
<tr>
<td>Material</td>
<td>Iron</td>
</tr>
<tr>
<td>Wall thickness</td>
<td>9 mm</td>
</tr>
<tr>
<td>Cylinder length</td>
<td>150 mm</td>
</tr>
<tr>
<td>Head-water jacket top distance</td>
<td>0 mm</td>
</tr>
<tr>
<td>Head-water jacket bottom distance</td>
<td>130 mm</td>
</tr>
</tbody>
</table>

**Cooling System Model**

The cooling system model is built following a typical engine cooling system configuration. It consists of the engine block water jacket, cylinder head water jacket, wax thermostat, radiator, mechanical water pump and others as shown in Figure 3.4 below.
Figure 3.4: GT-SUITE model of the engine cooling system.

The model is based on the reference close loop cooling system configuration in GT-SUITE, with modification to suit the thermal optimisation work. For example, the original model has a separate head and block coolant flow. A modification is made to change the flow path from a parallel flow to a series flow as found in most production engines. The flow enters into the cylinder block first, and exits out of the cylinder head. The characteristic of water pump flow rate over water pump speed is modified to achieve a coolant out and coolant in temperature difference of 8°C in the full load condition. Furthermore, the wax thermostat position is relocated at the engine coolant out from the radiator’s coolant out location, because this is the standard position chosen for best thermal control. The thermostat controls the coolant out temperature at 90°C.

The heat produced by the combustion model and piston friction is first transferred to the FE in-cylinder wall structure, from which it enters the water jacket and engine oil before being released to the atmosphere by the radiator. Lump masses represent the outer engine wall structure and transmit heat from the water jacket to the atmosphere. However, the engine oil is a simple circuit model which only maintains the temperature at 120°C. The fixed oil temperature is chosen to remove any influence on engine friction and combustion. This allows to focus on the effect of changing the conditions of the heat transfer to the water jacket.
3.2. **Engine Calibration**

The non-predictive combustion model requires the value of the engine spark timing and lambda to produce combustion outputs. Therefore, the calibration tradeoff of spark timing and lambda is done to get the optimized value based engine outputs. The coolant out temperature is also calibrated to demonstrate the optimized engine efficiency as a result of thermal management.

The MathWork™ Model-Based Calibration (MBC) Toolbox™ is used to reduce the engine calibration’s time and effort. The MBC toolbox creates statistical models of engine output responses and generates optimal calibrations. Typical calibration methods are complex, because the problem has many degrees of freedom and can be challenging even for experience engineers with computer support. For example, there are too many variable combinations to test individually, and the trade-offs between performance, efficiency, emissions, and reliability are complex. Figure 3.5 below shows the MBC process flow used to address these challenges.

![Model-Based Calibration Process Flow](image)

**Figure 3.5:** Model-Based Calibration Process Flow.

**Calibration objective**

The calibration objective is simplified to focus on the lowest engine BSFC for all engine speeds and loads. This is different from a production calibration, where additional factors
such as durability, drivability, comfort and performance and included, which lead for example to air-fuel enrichment at full load to gain higher performance. Three parameters are calibrated: spark timing, lambda and coolant out temperature. The calibration limits are also considered, as shown in Table 3.3 below:

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Limit</th>
</tr>
</thead>
<tbody>
<tr>
<td>Knock probability</td>
<td>0</td>
</tr>
<tr>
<td>Maximum exhaust temperature</td>
<td>850°C</td>
</tr>
<tr>
<td>Maximum cylinder head temperature at gas side surface</td>
<td>250°C</td>
</tr>
</tbody>
</table>

Also noteworthy is the fact that, besides the coolant temperature, the coolant flow rate can also influence the engine performance output. The coolant flow rate can alter the heat transfer rate between the cylinder wall and water jacket. However, the coolant flow rate is not included in this calibration, since the empirical Fisher engine friction model cannot predict friction changes from the cylinder wall temperature changes that result from the coolant flow.

**Design of Experiment (DoE)**

Design of Experiment (DoE) is a method to generate a reduced test plan by specifying the range of the engine calibration input variable over the range of engine operating conditions. The DoE generates the test plan based on the calibration model setup. Three types of model setups are available in the MathWork™ MBC Toolbox;

- One-stage;
- Two-stage; and
- Point-by-point.

**Model Setup: One-stage**

One-stage model setup uses the given parameter to generate a global model response. It is used to generate test plans that vary all variables simultaneously and identify and model the relationship among the variables. This model setup requires less test data for the model fitting and is suitable for less complex engine response model.
**Model Setup: Two-stage**

The two-stage test strategies have two separate variations; local and global. The local variation which is usually a single control variable is sweep while the global variables held as constant. For example, it is by sweeping the spark timing at a given engine speed, load, valve timing and air-fuel ratio. Here, the spark timing is the local parameter, while the other parameters are the global parameters. The sweep is then repeated in the other variation of the global set points. The MBC toolbox uses the local model to calculate the global model of the engine response. Each local model has certain coefficients such as the max and knot. Several global models are fitted to the different coefficients of the local model. These coefficients are referred to as the response features of the local models. The local model is usually a low order linear model to limit the number of coefficients. One-stage and two-stage model setups are not suitable for high complexity engine response.

**Model Setup: Point-by-Point**

The point-by-point model setup also has local and global variation test strategies. However, the point-by-point model setup does not create global models as does the two-stage model setup. It builds a localized model at each global set point. This enables the fitting of the model with higher accuracy. Point-by-point command-line functionality seeks to handle the complexity of developing designs for each operating point. However, the downside is that the models created do not provide estimated response between the operating points.

The point-by-point model setup is being used for this work. This is due to the fact that the model complexity is increased by adding the coolant out as an additional parameter. One-stage and two-stage model setups could not produce satisfactory engine output response accuracy. Coolant temperature has less influence on the engine response at certain engine operating points; therefore, the inaccurate model could cause the coolant temperature reading to become unreasonable. Accurate engine responses such as the exhaust temperature and cylinder wall temperature require high accuracy due to the fact that the strategy runs near the temperature limits and it could damage the engine. Figure 3.6 below shows the local parameters, global parameter and engine response for this work.
The DoE of the point-by-point model setup is separated by local and global test plans. The global test plan is distributed randomly with higher concentration at high engine load. The local test plan is distributed randomly with the parameter range depending on the engine operating condition. Additional test points will be added for better model accuracy. The DoE created is then transferred to the GT-SUITE for data collections.

**Data Modelling**

Data collected from GT-Suite according to the designed DoE is imported back into the MBC Model Fitting apps in the MBC toolbox. The MBC Model Fitting apps will generate the statistical models of the engine output response, which will then be used by the MathWork™ Calibration Generation (CAGE) toolbox during the calibration optimization. The apps can also perform a variety of pre-processing operations, including filtering to remove unwanted data, transforming or scaling raw data, grouping test data, and matching test data to experimental designs.

Three models are made; BSFC, cylinder head temperature and exhaust temperature model. The cylinder head and exhaust temperature models are made to ensure that the calibration optimization solution is within the cylinder head and exhaust working temperature.
Knock limit is predicted by applying the boundary model. The boundary is to make the calibration optimization solution within the boundary area. Figure 3.7 below shows the boundary model (blue colour) and the knock limit (red surface) at 2000rpm@10bar. The boundary is set as Hull Convex to ensure that everything outside the recorded data is considered outside the boundary. Additional data are taken when the knock limit surface is insufficient.

Figure 3.7: Spark timing boundary limit.

A few variation of the Radial Based Functions (RBF) are used in the model generation as it has better accuracy than any polynomial model for the complex model. The best output responses are selected by comparing the root-mean-square error (RMSE) and predicted residual sum of square RMSE (PRESS RMSE). The RMSE is the indication of the model error compared to each data point. PRESS RMSE indicates that the model is not overly sensitive to any single data point or a measure of the predictive power of the models. Overfitting problem or the model’s unnecessarily complexity can be verified by having PRESS RMSE to be relatively bigger than RMSE. Below are the RMSE and PRESS RMSE equations:

$$RMSE = \sqrt{\frac{\sum_{t=1}^{n}(\hat{y}_t - y_t)^2}{n}}$$  \hspace{1cm} (19)
\[ PRESS RMSE = \sqrt{\frac{\sum_{t=1}^{n}(\hat{y}_t - y_t)^2}{n}} \]  

\[ y_t = \text{Test point data value at } t \]
\[ \hat{y}_t = \text{Test point of fitted model value at } t \]
\[ n = \text{Number of total test point} \]
\[ \hat{y}_{(t)} = \text{Test point of model (fitted without test data } t \text{) at } t \]

**Optimized Calibration**

*Calibration process*

The Calibration Generation (CAGE) apps under the MathWork™ MBC Toolbox applied the models created during the model fittings to generate the engine spark timing, lambda and coolant out temperature look-up tables. The lookup tables are filled by optimizing the BSFC. The optimization problem is solved within the given constraints and input boundaries. Figure 3.8, Figure 3.9 and Figure 3.10 are the post optimization result of the engine spark timing, lambda and coolant out temperature throughout the engine speed and load:

![Figure 3.8: Engine sparks timing from calibration result.](image-url)
The coolant out temperature table shows that the coolant out temperature was 120°C at low engine load throughout the engine speed and gradually it became cooler at higher load.
The coolant temperature dropped as low as 50°C at 3000rpm@max load, although it is questionable whether this is achievable under practical conditions.

The coolant out temperature distribution matches the literature. High coolant temperature at low engine load improves the fuel consumption by reducing the friction and heat loss. Lower coolant out temperature improves the fuel consumption by reducing the knock tendency thus allowing more advanced spark timing.

**Validation**

Finally, the engine spark timing table, lambda table and coolant out temperature table are embedded back into the GT-SUITE engine model. The engine runs throughout the engine operation to validate the calibration. The engine BSFC result pattern agrees with the common spark ignition engine BSFC where the lowest BSFC is at low engine speed and high load region. The overall exhaust gas temperature and cylinder head surface temperature is lower than its temperature limits. Figure 3.11, Figure 3.12 and Figure 3.13 show the engine BSFC and the constraint temperature.

![BSFC map throughout the engine speed and load.](image)

Figure 3.11: BSFC map throughout the engine speed and load.
Figure 3.12: Exhaust gas temperature throughout the engine speed and load.

Figure 3.13: Gas side cylinder head surface temperature at the valve bridge zone throughout the engine speed and load.
3.3. **Comparison to the Conventional Cooling System**

Additional spark timing and lambda calibration tables are made based on 90°C coolant out temperature to represent a conventional cooling system. This forms a baseline for comparison of the optimized coolant out temperature. In general, the result shows that the engine fuel consumption reduction is better at low engine speed and load, where the optimised temperature is higher than the standard temperature. The fuel consumption reduction is up to 25.49 g/kW.hr (or 3.52%). This value is comparable to experimental result in the literatures, although some higher improvements have been reported [30,44–46,48].

As the engine load increased, the fuel consumption reduction became smaller. At around 80% of the maximum torque, the difference becomes negligible, because the optimised temperature is identical to the conventional coolant temperature. At maximum torque, again a small improvement can be achieved, this time with a lower coolant temperature (which provides better knock protection) Figure 3.14 shows the optimized cooling system BSFC improvement throughout the engine load and speed range.

![Figure 3.14: Engine BSFC improvement by optimized steady state thermal management throughout engine speed and load.](image-url)
It should also be noted that further improvement could be recorded if the thermal management approach is also to include the engine lubrication system. The lubricant temperature can further reduce the total engine friction [34,35].

Another important piece of information for the new thermal management strategy is the distribution of the cylinder wall temperature throughout the engine speed and load. This is due to the fact that the temperature will be used for the new thermal management strategy as the reference set point. Unlike the spark timing and lambda, the cylinder wall temperature could not be changed from one set point to another set point instantaneously for every engine cycle. The cylinder wall temperature has high dynamic behaviour from the thermal inertia, heat source, radiator performance and actuator (water pump and thermostat) delay. Therefore, a wide distribution of the cylinder wall temperature will give an indication that the coolant out temperature will have problems when the engine runs at a wide operating range and in quick succession.

The temperature difference between the highest and lowest points of the optimized combustion wall temperature range is significantly less than the difference for a conventional cooling system. The optimized wall temperature increases up to 23.5°C at low engine speeds and load while it reduces by 12.4°C at high load. The smaller temperature gap between the highest and the lowest temperature leads to potentially longer engine life. This is due to the reduced low cycle fatigue during engine transients. Figure 3.15 below shows the cylinder head temperature difference between the optimized coolant out temperature and the conventional cooling system.
3.4. **Engine Thermal Management in Drive Cycles**

The BSFC improvements of the calibration work so far apply to steady state conditions, and it depends on the engine speed and load. In reality, engine speed and load are constantly changing, and steady state is never reached. Seven drive cycles are taken to represent the driving behaviour on the actual road. The eight drive cycles are:

- US Federal Test Procedure 75kph (FTP75kph),
- US Highway Fuel Economy Driving test Schedule (HWY),
- US Supplemental Federal Test Procedure (US06),
- New European Drive Cycle (NEDC),
- Artemis Urban Cycle,
- Artemis Rural Road Cycle,
- Artemis Motorway Cycle, and
- Worldwide harmonized Light vehicles Test Procedures (WLTC).

![Figure 3.15: Cylinder head temperature difference throughout the engine speed and load.](image)
The vehicle model from a generic vehicle specification available in the GT-Suite is taken to generate the engine speed and load changes throughout the drive cycles. The vehicle model is linked with the engine and cooling system model to simulate the actual cylinder wall temperature. The vehicle’s detailed specification is shown in Table 3.4.

Table 3.4: Vehicle model specification in GT-Suite.

<table>
<thead>
<tr>
<th>Specification</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vehicle mass (kg)</td>
<td>1426</td>
</tr>
<tr>
<td>Drag coefficient</td>
<td>0.31</td>
</tr>
<tr>
<td>Vehicle frontal area (m²)</td>
<td>1.82</td>
</tr>
<tr>
<td>Transmission type</td>
<td>4-speed Automatic</td>
</tr>
<tr>
<td>1st gear ratio</td>
<td>2</td>
</tr>
<tr>
<td>2nd gear ratio</td>
<td>1.45</td>
</tr>
<tr>
<td>3rd gear ratio</td>
<td>1</td>
</tr>
<tr>
<td>4th gear ratio</td>
<td>0.667</td>
</tr>
<tr>
<td>Final drive ratio</td>
<td>3.55</td>
</tr>
</tbody>
</table>

Both the calibrated cylinder wall temperature set point and simulated cylinder wall temperature are measured throughout the test. It should also be noted that the engine is a hot engine where the engine does not require running a warm-up period during the test.

**Results**

Figure 3.16 shows the sample of result of the calibrated and actual cylinder wall temperature during the NEDC. From the result, the actual cylinder wall temperature runs lower than the calibrated cylinder wall temperature set point. In addition, the optimized cylinder wall temperature changes are considered to be highly volatile due to the fact that the changes are directly connected to the rapid engine speed and load changes.
Figure 3.16: Calibrated cylinder wall temperature (red), actual cylinder wall temperature (blue) and vehicle speed in NEDC.

From here, the fuel consumption estimation can be calculated by comparing the current engine BSFC based on its current head temperature. The overall result of the fuel consumption saving is shown in Figure 3.17 below. It demonstrates that the optimized combustion wall temperature improves the fuel consumption for all drive cycles. The fuel consumption reduction of more than 2% is predicted in Artemis Urban Cycle, FTP75kph, NEDC and Artemis Rural Road Cycle. The Artemis Urban shows the highest fuel consumption improvement which is up to 2.36%. Meanwhile, the US06, WLTC, HWY, and Artemis Motorway Cycle are smaller but still make relevant fuel consumption reduction. The lowest fuel consumption reduction is found for the Artemis Motorway cycle at 1.71%.
Figure 3.17: Summary results of the fuel consumption reduction by optimizing the cylinder wall temperature in various drive cycles.

Discussion

The highest fuel consumption reduction of more than 2% is found in urban drive cycles with an average speed below 35km/hr. Fuel economy benefits are also found in high speed drive cycles, but these are significantly smaller. This shows that the optimized combustion wall temperature control has the highest potential for fuel economy improvement during urban driving.

The reason is that the engine mostly runs at low engine speed and low engine load in urban driving style. Figure 3.18 shows the distribution of the engine speed and engine load during the Artemis Urban drive cycle. It had a high share of engine loads below 3bar and engine speeds below 3000rpm. This area is where most of the fuel economy benefits can be realized, as shown in Figure 3.14. Meanwhile, the engine ran at higher speed and higher load for the motorway driving style. Figure 3.19 shows the distribution of the engine speed and engine load during Artemis Motorway drive cycles. It had a higher share of loads at 4bar and above. The engine speed would exceed 3000rpm, in contrast to the urban style of driving.
Figure 3.18: Engine speed and load distribution during the Artemis Urban Drive Cycle.

Figure 3.19: Engine speed and load distribution during the Artemis Motorway drive cycle.

The results also show that the coolant temperature needs to run near the temperature limits to achieve optimized combustion wall temperature. Furthermore, the wall
temperature also needs to respond quickly, because the reference temperature is subject to rapid changes. This is a tough requirement, because changes in temperature require significant thermal flows due to the high thermal masses involve in the engine, and fast changes of coolant temperature are typically neither possible nor desirable. Developing an engine thermal management using an advanced predictive controller will be the main challenge for this thesis.

3.5. **Summary**

The results show that controlling the combustion wall temperature improves the engine efficiency in several ways. Increasing the temperature reduces the piston assembly friction, reduces the heat loss and improves the air-fuel mixture. Lowering the temperature can improve the engine volumetric efficiency and reduce the knock tendency. Based on these conflicting effects, an optimized combustion wall temperature is found for different engine speeds and loads in steady state conditions. The result shows that the fuel consumption saving is higher especially at lower engine speeds and loads.

The fuel consumption saving predicted from the optimized combustion wall temperature in the urban style of driving is more than 4%. Meanwhile, the saving in the motorway style of driving is lower, yet still meaningful. The highest fuel saving is 4.38% during the FTP-75kph and the lowest fuel consumption saving is 1.93% on the Artemis Motorway cycle.
On a real engine, it is impossible to consistently and immediately maintain the ideal thermal management set point temperature as explained in the previous chapter. The reason is that the set point temperature changes faster than the actual cylinder wall temperature can follow. This is a direct consequence of the dynamic of the system, which includes the thermal inertia, transport delays and the actuator response.

The cylinder wall temperature can be controlled via the coolant temperature or the flow rate. The coolant temperature in turn depends on the balance of heat absorbed from the engine and the heat removed in the radiator. The two control inputs are the water pump speed and thermostat valve position, which affect the coolant flow rate and temperature.

This chapter focuses on the dynamics of the temperature set point changes and its volatility in typical drive cycles. The results will be compared with the actual cylinder wall temperature response. The same legislated and research drive cycles are used as in the previous chapter.
4.1. **Evaluation of Drive Cycle Volatility**

Understanding the changes of the temperature set point during actual driving conditions is important for the design of the temperature controller. If the set point volatility is too high (changes are too fast), it indicates that the system will not be able to follow in time, and the controller may be more difficult to design and that it would still be less effective. This is because a slow controller response together with high frequency set point changes leads to large average control errors.

Statistical methods are used to determine quantitatively the volatility of the set point. These methods are:

- autocorrelation,
- lag plot and
- power spectral density (PSD).

**Autocorrelation**

The autocorrelation function is a common tool for characterising the volatility of randomness in a dataset. It is determined by computing the autocorrelation of the data values over varying time differences. A fixed time displacement between two points in time is also called lag; for example, the lag \( x(t) \) versus \( x(t - 1) \) is 1 second. In an independent random data, the autocorrelation coefficient would be near zero for all time lag separations other than 0, but most practical data show a correlation over certain periods of time. The autocorrelation coefficient is defined as follows:

\[
R_h = \frac{C_h}{C_0}\tag{21}
\]

Where;
\[ C_h = \frac{1}{N} \sum_{t=1}^{N-h} (Y_t - \bar{Y})(Y_{t+h} - \bar{Y}) \]

\[ C_0 = \frac{1}{N} \sum_{t=1}^{N-h} (Y_t - \bar{Y})^2 \]

and;

\[
R_h = \text{Autocorrelation coefficient at lag } h \\
C_h = \text{Autocovariance function at lag } h \\
C_0 = \text{Variance} \\
Y_t = \text{Sample at time } t \\
\bar{Y} = \text{Sample mean} \\
h = \text{Time lag [s]} \\
N = \text{Sample size}
\]

Applied to a finite data set, the autocorrelation is a stochastic measure that approximates the true behaviour of the generating system. It is therefore important to establish a confidence band, and the standard 95% confidence is used here. The autocorrelation coefficient values that are in the confidence band represent 95% of the correlation between the current data and its lags hypothetically have no correlation. The confidence band is based on the Bartlett formula as shown in equation (22) below [71]:

\[
z_{1-\alpha/2} \left( \frac{1}{N} \left( 1 + 2 \sum_{i=1}^{k} R_i^2 \right) \right)
\]

(22)

Where;

\[
z = \text{Quantile function of the standard distribution} \\
\alpha = \text{Significant level (95%)} \\
N = \text{Sample size} \\
k = \text{lag} \\
R = \text{Autocorrelation coefficient at lag } h
\]

Figure 4.1 below shows the autocorrelation of the combustion wall temperature set point for a number of drive cycles. All cycles have high correlation until 1 second lag. For higher lags, the autocorrelation coefficient starts to drop gradually. The NEDC cycle shows a significantly strong autocorrelation into 10 second lag, after which the autocorrelation
coefficient drops very steeply. This is due to the specific NEDC pattern, which is based on regular segments of constant vehicle speed and steady speed change. NEDC is known not to replicate any of the real drive cycles, so this behaviour is not representative of real world driving.

Figure 4.1: Autocorrelation plot of temperature set point in drive cycle.

In general, urban type drive cycles show a faster drop of the autocorrelation than the motorway type of drive cycles. As expected, the Artemis Urban Drive Cycle has the lowest autocorrelation coefficient. The autocorrelation deteriorates even below 1s, enters the 95% confidence band at 5.4 seconds, which means that the correlation is no longer significant. Figure 4.2 shows an autocorrelation plot of Artemis Urban Drive Cycle and its confidence band throughout the lags. Meanwhile, the Artemis Motorway Drive Cycle has the highest autocorrelation at long lags. Remains mostly unaffected for several seconds, and it enters the 95% confidence band at 66.67 seconds. Figure 4.3 shows an autocorrelation plot of the Artemis Motorway Drive Cycle and its confidence band throughout the lags.
Figure 4.2: Autocorrelation plot of temperature set point in Artemis Urban Drive Cycle.

Figure 4.3: Autocorrelation plot of temperature set point in Artemis Motorway Drive Cycle.

Figure 4.4 below shows the lag in seconds of each drive cycle where the autocorrelation coefficient enters the 95% confidence band. The figure quantifies that the urban types of drive cycles have much higher temperature set point volatility compared to the motorway type of drive cycle.
Lag Plot

The autocorrelation is a linear approach, which is not entirely justified for drive cycles that are subject to strict limits. The lag plot is another way of visualising the correlation through time. The plots display observations for a time series against a later set of observations. It can be used to check if a data set or time series is random or not – visible patterns are a sign that the data set is non-random.

Figure 4.5 and Figure 4.6 illustrate the temperature set point lag plot at 0.1, 1, 5 and 10 seconds for both Artemis Urban Drive Cycle and Artemis Motorway Drive Cycle. The lag plot shows that both drive cycles show a mostly linear pattern at 0.1 second lag, which signifies a strong positive correlation. However, Artemis Urban Drive Cycle already shows a significant number of data points outside the linear pattern at 1 second, unlike the Artemis Motorway Drive Cycle, which is still very linear. Artemis Urban shows that the data is far away from the linear pattern at 5 and 10 seconds, while the Artemis Motorway still has a high density of data at the linear line throughout the lag plots.
Figure 4.5: Lag plot for Artemis Urban Test Cycle at 0.1 seconds, 1 second, 5 seconds and 10 seconds.

Figure 4.6: Lag plot for Artemis Motorway Test Cycle at 0.1 seconds, 1 second, 5 seconds and 10 seconds.
Power Spectral Density

Power spectral density (PSD) is another way of looking at the volatility of a signal. It reveals the same information as the autocorrelation, but the representation is in the frequency domain, not in the time domain, which can be an advantage for controller design. As a general rule, a controller should have a frequency response that includes the relevant frequencies of the set point signal, as the controller performance would suffer if dominant frequencies are higher than the controller bandwidth.

Power Spectral Density (PSD) is again a statistical method, used to reveal typical frequencies in time series data. It is a common method in finance and environment trend study [72–76], but it is rarely used in automotive areas. There are automotive applications focusing on the vibration analysis, such as the vibration of road loads, diagnosis of bearing conditions and engine vibrations [77–80], because these effects are easy to see in the frequency domain.

Power Spectral Density (PSD) describes how the signal power content of a transient signal is distributed over different frequencies. It is calculated using the Fourier Transform, and in practice the Fast Fourier Transform (FFT) is being used. The power spectral density $S_{AA}(f)$ can be computed from the FFT as in the equation below:

$$S_{AA}(f) = \frac{FFT(A) \cdot FFT^*(A)}{N}$$

(23)

Where $FFT(A)$ is the Fast Fourier Transform of signal $A$ and $FFT^*(A)$ denotes the complex conjugate of $FFT(A)$. $N$ is the number of points in the acquired time-domain signal.

The PSD is applied to the target cylinder wall temperature over a number of drive cycles. Because the result is very noisy, signal filtering is applied in MathWork™ MATLAB®. A Hann Window is used as the filter to remove any noise and expose the set point power spectral density characteristics across the frequency. A wider window filter width is applied towards higher frequency to counter the decreasing signal levels that can be lost in the noise. The narrow window width effective to highlight the PSD signal characteristic at low frequency is not found to be adequate at higher frequency, and vice versa. Figure 4.7 below
demonstrates the PSD result without filtering compared to the various filter window widths. The equation below is the Hann Window Filter equation:

\[
Y(f) = \sum_{n=0}^{N-1} X\left(f - \frac{N - 1}{2} + n\right) \cdot W(n)
\]

\[
W(n) = 0.5 \left(1 - \cos\left(\frac{2\pi n}{N - 1}\right)\right)
\]

\[0 < n < N - 1\]

Where;

- \(Y(f)\) = Post filter PSD at frequency \(f\) [C°/Hz dB]
- \(X(f)\) = Actual PSD results at frequency \(f\) [C°/Hz dB]
- \(W(n)\) = Hann Window
- \(n\) = Sample number
- \(N\) = Window width in discrete-time

Figure 4.7: HWY Drive Cycle with variable width of Hann Window filter.

The PSD characteristics of all drive cycles are broadly similar. The frequency range can be divided into two regions: a flat signal power is observed in the lower frequency region, followed by a drop off (decrease) towards higher frequencies. This is a typical shape for any
signal with temporal correlation resulting from low pass filtering. The slope of the drop is consistently around 25dB per decade for all drive cycles – which is just above the 20db per decade for a first order filter.

The frequency that separates the flat region from the drop is called the corner frequency. It is an indicator for the required frequency bandwidth - the thermal management controller response should be fast enough to allow the corner frequency to pass nearly unattenuated to achieve satisfactory control performance. The corner frequencies are different for every drive cycle, as shown in the two extreme cases in Figure 4.8 (Artemis Urban Drive Cycle) and Figure 4.9 (Artemis Motorway Drive Cycle).

Figure 4.8: Power spectral density of cylinder wall temperature set point in the Artemis Urban Test Cycle.
Figure 4.9: Power spectral density of cylinder wall temperature set point in the Artemis Motorway Test Cycle.

The Artemis Urban Drive Cycle is the most volatile cycle with the highest corner frequency (0.09Hz). It is followed by the Artemis Rural Road Drive Cycle (0.045Hz) and HWY (0.04Hz). The lowest corner frequency is seen in the Artemis Motorway Drive Cycle with 0.015Hz. Figure 4.10 shows the comparison of the PSDs of the cylinder wall temperature set point for all drive cycles, while Figure 4.11 compares the corner frequencies.

Figure 4.10: Power spectral density of cylinder wall temperature set point in all drive cycles.
Summary

The results from both PSDs and autocorrelations are consistent with each other, in qualitative and quantitative terms. The Artemis Urban Drive Cycle is the most volatile drive cycle while the Artemis Motorway Drive Cycle is the least volatile. The Artemis Urban Drive Cycle has shown a random (uncorrelated) behaviour for frequencies below 0.09Hz or lags above 5.4 seconds. The Artemis Motorway Drive Cycle however only reaches randomness after 67 second lag or below a corner frequency of 0.015Hz.

4.2. Transient Response

The autocorrelation, lag plot and PSD results help to understand the dynamic of the cylinder wall temperature set point over common drive cycles. They show that the set point temperature has a frequency bandwidth of up to 0.09Hz, and the signal becomes essentially random (uncorrelated) after about 5 seconds. This indicates that the controller should have a higher bandwidth than 0.09Hz, and a response time significantly faster than 5 seconds, in order to cope with most driving environments.
Coolant flow rate and coolant temperature are the two main control inputs that are available for the control of the cylinder wall temperature. Therefore, both inputs are analysed using the same tools:

- Bode plot (frequency domain) and
- step response (time domain).

The results are then compared to the previous autocorrelation and power spectral density conclusions.

**Bode Plot**

The Bode plot is a tool used to represent a frequency response of a system. This is done by comparing a sine wave input signal to the system output signal. The Bode plot consists of the magnitude part and the phase part. The magnitude plot is the ratio of the input and output signal amplitude and the phase plot represents the output signal lag as shown in Figure 4.12 below.

![Bode Plot Diagram](image)

Figure 4.12: Input and output of a system for magnitude and phase response.

The water pump speed and coolant in temperature are used as the Bode plot input to represent the flow and temperature control. The output for both is the cylinder wall
temperature at the cylinder head. Because the controller design comes later in the design process, only the open loop transfer function is considered here. It is expected that a good controller can improve the response speed by a certain margin.

To simplify the analysis, the engine cooling circuit flow is considered an open circuit flow – this removes the secondary effects caused by the recirculation which could detract from the main results. The coolant warm-up and cool-down performance cannot be analysed this way, but it is mainly determined by the engine heat supply and the radiator performance (and it can be improved using other technologies, such as the coolant heat storage [81]). Delays of the actuators are neglected.

The sinusoidal amplitude inputs of the coolant temperature and water pump speed are set to $90^\circ C \pm 30^\circ C$ and $3000rpm \pm 1500rpm$. The cylinder wall temperature amplitude at an engine speed of $3000rpm @ 5bar$ BMEP is very similar for both input signals. Therefore, the input amplitudes are normalized to make them comparable to each other. The Bode plot in Figure 4.13 confirms that the output amplitude is almost identical for both input signals.

![Bode plot of cylinder wall temperature response at 3000rpm @ 5bar.](image)

Figure 4.13: Bode plot of cylinder wall temperature response at 3000rpm @ 5bar.
The dotted lines in the Bode plot in Figure 4.13 above represent the PSD critical frequencies of the Artemis Motorway (0.015Hz) and the Artemis Urban (0.09Hz) cycle. The grey area includes a phase delay of more than 180° where the control action becomes infeasible, as the response is inverted compared to the low frequencies.

The plot shows that the water pump and the temperature control have almost the same response, but the flow control has a slight advantage at higher frequency. The combined use of both (flow and temperature) gives a stronger response throughout the frequency range when compared to a single control input. The critical phase -180° is reached at around 1Hz, but the gain at this frequency is only 0.01 or 1%, which means that the control is no longer effective before reaching this point.

The response magnitude at the corner frequency of the Artemis Motorway Drive Cycle is above -3dB or 0.707: it is 0.63, 0.70 and 1.3 for the temperature, flow and combined control. However, the magnitude drops significantly at the 0.09Hz mark, the corner frequency of the Artemis Urban Drive Cycle – it is only 0.16, 0.21 and 0.36 for temperature, the flow and combined control. This means that the controller will have to use large input signals to achieve the desired temperature in time, and it may be difficult to achieve acceptable control for the Artemis drive cycle due to the physical limits on the inputs.

Further simulation is performed with higher and lower engine loads, as shown in Figure 4.14 (5000rpm @ 10bar) and Figure 4.15 (1000rpm @ 1bar) below. Both results confirm that the combined control has a better frequency response than the single variable control, although the difference is less pronounced.
It can be seen that the flow control becomes more effective at high engine load, but it deteriorated at low engine load. This can be explained by the fact that the coolant flow at
constant temperature does not carry any additional energy to warm up the cylinder wall structure. It can only cool down the cylinder by increasing the heat transfer coefficient from running at higher flow rate. Therefore, the flow control is not significant at low load where the heat flow is very low. The flow control is more effective at high engine load due to the availability of the high amount of heat. The results for the coolant temperature control are the opposite: it is more effective at low engine load. This indicates that the flow control gives advantage in controlling the cylinder wall temperature at high engine load and coolant temperature control is more effective at low load. The combination of the flow and temperature control is effective throughout all engine conditions.

**Step Response**

Step response is a time domain characteristic that measures the response of a system to sudden input change from one steady-state to another steady-state. It contains important information such as overshoot, rise time, settling time, dead time, as well as the indication for stability and the order of the system. Figure 4.16 below illustrates a typical step response of a second order system and its characteristics.

![Step Response Diagram](image)

Figure 4.16: Typical step response for a second order system.

The step response is based on the same model with 90°C±30°C and 3000rpm±1500rpm of temperature and flow control inputs. It is taken at three engine operating points: 1bar @ 1000rpm, 5bar @ 3000rpm and 10bar @ 5000rpm (Figure 4.17). The result shows that it
confirms the Bode plot results: the coolant temperature control has a better response at low load, and the flow control is better at high load.

Figure 4.17: Cylinder wall temperature response at 1000rpm @ 1bar, 3000rpm @ 5bar and 5000rpm @ 10bar.

In all cases, the cylinder wall temperature shows a dominant first order system step response, but characteristically for any non-linear system, the rise and fall behaviours differ noticeably. The rise magnitude is greater than the fall magnitude, but the time delay is longer. This effect is more pronounced for the flow control, and it can be explained by the heat transfer equation below:

\[
\dot{Q}_{wc} = \alpha \cdot A \cdot (T_w - T_c)
\]

(25)

Where;
\[
\alpha = \frac{Nu \cdot k_c}{L}
\]  

(26)

\[
Nu = f(Re, Pr)
\]

(27)

\[
Re = \frac{\rho_c v_c L}{\mu_c}, Pr = \frac{c_p \mu_c}{k_c}
\]

(28)

And;

\[
\dot{Q}_{wc} = \text{Heat transfer rate between wall and coolant [W]}
\]

\[
\alpha = \text{Heat transfer coefficient [W/m}^2\text{K]}
\]

\[
A = \text{Heat transfer area [m}^2\text{]}
\]

\[
Nu = \text{Nusselt Number}
\]

\[
k_c = \text{Coolant thermal conductivity [W/mK]}
\]

\[
L = \text{Characteristic length [m]}
\]

\[
Re = \text{Reynolds Number}
\]

\[
Pr = \text{Prandtl Number}
\]

\[
\rho_c = \text{Coolant density [kg/m}^3\text{]}
\]

\[
v_c = \text{Coolant velocity [m/s]}
\]

\[
\mu_c = \text{Dynamic viscosity [kg/ms]}
\]

\[
c_p = \text{Coolant heat specific [J/kgK]}
\]

The geometric parameters, \(A\) and \(L\) are the only constants, given by the water jacket design. The coolant velocity \(v_c\) is proportional to the coolant mass flow rate. Thermal conductivity \(k_c\), density \(\rho_c\), and viscosity \(\mu_c\) depend to some degree on the coolant temperature, although the impact may be small in typical operating conditions. Figure 4.18 below shows the coolant thermal conductivity, density and viscosity relationship to the coolant temperature.
The coolant heat transfer coefficient can also be rewritten as a function of the coolant mass flow rate $\dot{m}_c$ and temperature $T_c$:

$$\alpha = f(\dot{m}_c, T_c) \quad (29)$$

It follows from equations (25) and (29) that the coolant temperature affects the cylinder wall temperature not only because of the temperature difference, but also via changes in the heat transfer coefficient. The flow rate is easier to analyse, because it only affects the heat transfer coefficient. The flow rate has a much bigger effect on the heat transfers coefficient than the coolant temperature. The heat transfer coefficient affects the response time of the cylinder wall temperature: a low heat transfer coefficient causes a slow and limited response.

The effect can clearly be seen when comparing the rise and fall of the response time of the cylinder wall temperature. The cylinder wall temperature increases by the lower heat transfer coefficient from lower mass flow rate that causes longer response time compared to the coolant temperature changes. The heat transfer coefficient for coolant the temperature control is almost constant during this process while the heat transfer coefficient in the flow control is reduced.
The difference between the rise and fall magnitude is an indication for the nonlinear characteristic of the cylinder wall temperature behaviour. A linear time invariant model is required by most controllers, such as the Model Predictive Control or PID control. The non-linear effects can cause difficulties in the controller design, degraded responses and even instabilities.

4.3. **Current Mechanism reliability**

From the Bode plot and step response results, reviewing the current electric water pump and motorized valve (or thermostat valve) capability in the actual thermal management is also important. Both the electric water pumps and motorized valve will be discussed and its advantages and disadvantages compared in terms of:

- running in close loop cooling circuit and
- actuator power consumption.

**Run in Close Loop Cooling Circuit**

Changing the coolant temperature using a motorized valve is limited by the radiator performance and engine heat supply. The cylinder wall temperature requires high coolant temperature at lower engine speed and load, and low coolant temperature at higher engine speed and load, where this can be referred to in Figure 1.4. This situation would create more problems in the motorized valve control, due to the fact that the high coolant temperature is required when the engine heat supply is low, and vice versa. Thus, a very high radiator performance is essential to remove the high amount of heat created by the engine faster, whereas the technology like the coolant heat storage [15] can provide additional heat to increase the coolant temperature when the engine produces less heat. However, the design changes may require high cost, weight and packaging problem.

The electric water pump performance is not entirely limited by the heat from the engine and radiator performance. The comparison between the flow control, temperature control
with and without the radiator can be seen clearly in Figure 4.19. The flow control has a slight slower rise response from 90°C to 120°C, compared to the open loop cooling circuit temperature control. However, the flow control has far better response than the close loop cooling circuit temperature control.

![Graph showing cylinder wall temperature response and control input variable between close loop flow rate and close and open loop coolant temperature control of cooling system at 1000rpm @ 1bar.]

**Figure 4.19**: Comparison of cylinder wall temperature response between the close loop flow rate, close and open loop coolant temperature control of cooling system at 1000rpm @ 1bar.

**Power Consumption**

In general, the electrification of the water pump and valve power efficiency is lower compared to the mechanical water pump. This is due to the power loss in converting energy from the engine rotation mechanical energy to the electric energy by alternator then converting the electric energy back to the mechanical movement energy by the electric water pump and motorized valve. In addition, more power loss has been observed from the alternator efficiency less than 55% [82,83] along with the water pump and valve driver efficiency. In contrast, the mechanical water pump runs directly by a belt or chain attached...
to the engine rotations, while the wax thermostat valve movement is from the expansions and contractions of the thermostat’s wax pallet.

Power loss in the mechanical water pump is higher than the electrical water pump at the same flow rate. This is due to the fact that the mechanical water pump has high side load from the belt tension. This causes the mechanical water pump required to have large bearings and structure of the bearing. On the contrary, the electric water pump can have its impeller to be design in parallel with the drive motor. This eliminates the needs to have large bearing thus, reducing the pump friction [17,84]. Furthermore, H. Wsewolod et al. stated that the electric pump could reduce 90% of power loss at 1500rpm at lower flow rate [84]. This reduction is still far greater, even with the low alternator efficiency.

This works in the opposite way for the motorized valve. The motorized valve requires an additional power source to run compared to the wax thermostat. The wax thermostat movement energy is derived from the extraction of the coolant heat waste energy. However, the motorized valve power consumption is not more than 15W at maximum operation (depending on the type and model). Furthermore, the motorized valve gives the potential to have a variable coolant temperature for better engine efficiency.

4.4. Summary

The cylinder wall temperature set point volatility in eight types of drive cycles is quantitatively determined using autocorrelation, lag plot and power spectral density. All statistical methods show that high volatility is recorded in urban driving. The set point for the Artemis Urban Drive Cycle is considered random after only 5.4 seconds while its power spectral density critical frequency is 0.09Hz.

The results are then compared with the open loop cooling circuit control transient response. The results show that the flow control is effective at high combustion temperature while the temperature control is effective at low combustion temperature. The combined control has higher magnitude response compared to the flow and temperature control throughout the
frequency. However, the output magnitude even for the combined control is less than 40% of the input magnitude at 0.09Hz for critical frequency in the Artemis Urban Driving Cycle. This means that effective control is possible, but challenging, and it will depend on a fast, effective controller and good control authority in the actuators.

Overall, using a variable coolant flow rate with an electric water pump might be superior to controlling the coolant temperature with a motorized mixing valve. This is because the flow control shows a faster response compared to the temperature control in the close loop cooling circuit. Furthermore, the electric water pump consumes less power than the mechanical water pump, while the motorized valve requires an additional power source to run compared to the wax thermostat.
Model Predictive Control (MPC) is an advanced control method which is already gaining popularity in many industries; especially process and chemical industries. This is because of its success in handling the Multi-input and Multi-output (MIMO) system without violating safety constraints [3,85,86]. This makes it possible to operate closer to hard constraints than the conventional controller, and in return gain larger profit [87].

Despite these advantages, MPC is still considered new in the automotive industry; i.e. MPC can only be found in research work but not in production vehicles. The reasons are due to the implementation complexity of MPC and high demand of processing power that a current Engine Control Unit (ECU) is not able to cope with. However, the requirement of better engine efficiency and the rapid increase in the ECU computational performance will turn MPC controllers viable in on-the-road vehicles in near future.

This chapter explains the MPC background, including concepts and previous attempts on applying MPC in engine thermal management.
5.1. **Background & Concept**

**Background**

Model Based Predictive Control (MPC) is one of the computer control algorithms that utilize an explicit process model to predict the future response of a system as shown in Figure 5.1 below. Model Predictive Control can be traced back to the 1960s, and it was starting to become popular from 1980s for the oil and chemical industries. Its main advantage is the ability to handle multivariable constrained control problems [2].

![Figure 5.1: Basic structure of Model Based Predictive Control (MPC).](image)

In the automotive industry, MPC is still considered a new approach (unlike in the process and chemical industries [1], where it is well established). Currently, the Proportional-Integral controller (PI controller) with gain scheduling is the most common control strategy in the automotive industry. The drive for better efficiency in modern engines leads to more control inputs, and with the increasing complexity and stronger interactions, the PI control approach is becoming more limiting. For modern engines, advanced controller can provide better decoupling, more accurate tracking, and more robust stability. Better CPU (Central Processing Unit) performance at low cost means that advanced controllers are becoming more attractive for the automotive industries. For this reason, MPC is studied here as a promising technology for the control of future engines.
Concept

The basic concept of MPC in the Single-input Single-output system (SISO) is illustrated in Figure 5.2 below. The current time $t(k)$ and previous information of states is fed into the build-in model in MPC to predict the dynamic of the plant output over a finite time horizon $t(k + n_p)$. This finite time horizon is called prediction horizon $n_p$. Within this horizon, the controller will compute the optimum future trajectory of the manipulated variable $u$ along the control horizon $t(k + n_c)$ based on minimum MPC move cost function $J$. The first input in the optimal manipulated variable, $u$ sequence is then sent into the plant. These processes are then started again for the new current time $t(k + 1)$. The prediction horizon $n_p$ keeps being shifted forward at every time step; and for this reason MPC is also called the receding horizon control.

![Figure 5.2: MPC control strategy scheme.](image)

In MPC, there are three important components:

- The explicit model for predicting future plant dynamic behaviour,
- The MPC move cost function, and
- The constraints.
**Model for predicting future plant dynamic**

MPC predicts the plant future behaviour by using a time discrete dynamic model. The dynamic model in MPC is commonly a linear time invariant (LTI) and it is in the form of state space model as shown in equation (30) below:

\[
\begin{align*}
x(k + 1) &= Ax(k) + Bu(k) \\
y(k) &= Cx(k)
\end{align*}
\] (30)

Where \(x(k), u(k)\) and \(y(k)\) are the state, input and output vectors at the \(k\)th sampling instant of the plant. The input vectors consist of manipulated variables and measured disturbance. Therefore, it allows multivariable control strategy. The model will predict the dynamic plant behaviour along the prediction horizon at every optimization interval.

**Cost Function**

The optimal sequence of manipulated variables \(u\) is derived from the numerical minimization of the following quadratic cost function at each sampling instant (31):

\[
\min J(k) = \sum_{j=1}^{n_y} J_{y,j} + \sum_{j=1}^{n_u} J_{u,j} + \sum_{j=1}^{n_u} J_{\Delta u,j}
\] (31)

Where \(J_{y,j}, J_{u,j}\) and \(J_{\Delta u,j}\) are the output reference tracking \(j\), manipulated variables tracking \(j\) and manipulated variable movement suppression \(j\). Meanwhile, \(n_y\) and \(n_u\) are the number of the output variables and input variables.

The output reference tracking part \(J_{y,j}\) is to keep selected plant outputs near specified reference values throughout the prediction horizon \(n_p\). The manipulated variables tracking part \(J_{u,j}\) aims to keep selected manipulated variables near specified target values throughout the \(n_p - 1\) horizon. The manipulated variable movement suppression part punishes the rapid changes in the manipulated variables, which can be used to make the input signal smooth and to reduce oscillations. All the cost function parts are tuned based on weight \(w_{y,j}, w_{u,j}\) and \(w_{\Delta u,j}\) as shown in equations (32), (33) and (34). The higher a weight is, the more effort the controller makes to keep the corresponding deviation low.
\[
J_{y,j} = \sum_{i=1}^{n_p} w_{y,j}[y_j(k + i|k) - r_j(k + i|k)]^2
\]  
\[
J_{u,j} = \sum_{i=0}^{n_p-1} w_{u,j}[u_j(k + i|k) - u_{\text{target},j}(k + i - 1|k)]^2
\]  
\[
J_{\Delta u,j} = \sum_{i=0}^{n_p-1} w_{\Delta u,j}[u_j(k + i|k) - u_j(k + i - 1|k)]^2
\]

Where \(y_j, r_j, u_j\) and \(u_{\text{target},j}\) are the output component \(j\), set point component \(j\), manipulated variable component \(j\) and manipulated variable target component \(j\) at \(k\) time interval.

**Constraints**

The MPC solution is subjected to constraints imposed to the quadratic optimization problem (QP). MPC can handle any linear constraint, but typically only upper and lower bound constraints are used; by applying linear constraints as in equation (35) below:

\[
y_{j,\text{min}}(k + i|k) \leq y_j(k + i|k) \leq y_{j,\text{max}}(k + i|k)
\]
\[
u_{j,\text{min}}(k + i|k) \leq u_j(k + i|k) \leq u_{j,\text{max}}(k + i|k)
\]
\[
\Delta u_{j,\text{min}}(k + i|k) \leq \Delta u_j(k + i|k) \leq \Delta u_{j,\text{max}}(k + i|k)
\]

\[
Eu_j(k + i|k) + F y_j(k + i|k) + G d_j(k + i|k) \leq H
\]

Where \(y_j, u_j, \Delta u_j,\) and \(d_j\) are the output variable component \(j\), manipulated variable component \(j\), manipulated variable movement component \(j\) and measured disturbance component \(j\). In equation (36), \(E, F, G\) and \(H\) are constant matrices; they represent a generalised linear constraint for the MPC controller. It would be possible, but not convenient, to use only equation (36). Internally, all constraints are converted in this form, and they are respected throughout the prediction horizon \(n_p\) in the optimal QP solution.
5.2. **MPC Advantages for Engine Thermal Management**

There are several reasons behind the success of MPC in the process and chemical industries. The most important reason is the way in which MPC handles constraints. MPC can handle actuator physical limits as well as output process limits [3,85,86]. This enables MPC to run at set points near to limits, thus creating a more profitable process [87].

A control strategy that can handle constraint is of great interest for thermal management. Besides handling actuator mechanism limits such as water pump speed limits and valve opening limits, thermal management also requires a controller that can prevent coolant and wall temperature from exceeding operating temperature limits. This is very important because over temperature could cause a catastrophic failure to the engine. However, optimal operating may require getting very close to the limit to improve the engine thermal efficiency; i.e. reducing the friction by running at a higher temperature and reducing the heat losses as shown in CHAPTER 3.

Another important advantage for thermal management is that MPC can handle complex multivariable process naturally. This is a significant advantage compared to SISO approaches such as a PID controller [86]. Both the electric water pump and the electric mixing valve influence the coolant out and cylinder wall temperature and should be used as control variables in thermal management. It will be more complex if disturbances such as combustion heat, engine bay temperature and oil temperature are taken into account. Therefore, thermal management is most effective using a Multi-Inputs Multi-Outputs (MIMO) type controller such as MPC.

As mentioned in the previous chapter, the transient response results show that even the combined control with open loop cooling circuit could not effectively cope with the urban type of drive cycles temperature set point volatility. The magnitude responses in the Bode plot are less than 40% of the input magnitude at 0.09Hz and they start to be random after 5.4 seconds in Artemis Urban Drive Cycle. The close loop cooling circuit will only worsen the frequency response. Therefore, a physical redesign of the cylinder head water jacket and material could improve the transient response by reducing thermal mass, but this option would come at a high cost and time requirement for the development.
MPC offers an important advantage for fast tracking, that can compensate the system response to some degree, by anticipating the set point changes to reduce the error [88]. This requires advance knowledge of changes in set point. MPC controller can then react before the reference or disturbance changes (A reactive controller such as a conventional PID control only responds after the controller detects error.) This prediction enables the MPC to reduce the error during the changes as the MPC controller concept can minimize the error throughout the prediction horizon rather than minimizing the error based on the current time measurement as shown in Figure 5.3.

![Controller Output Diagram](image)

**Figure 5.3**: Illustration of output response comparison of the conventional PID and MPC with look ahead.

In the near future, with the availability of equipment such as the GPS navigation system, V2V (vehicle to vehicle) and V2X (vehicle to vehicle, infrastructure etc.) communication, vehicle parameter monitoring, on-board camera and others, it will be possible to anticipate driver actions and outside environment changes with reasonable certainty [23]. Therefore, MPC with known future reference holds great promise for engine thermal management.
5.3. **MPC Challenges in Thermal Management**

Despite the mentioned MPC advantages, the implementation of the MPC controller for engine thermal management is not a straightforward application. MPC depends on a linear model of the plant, and a poor model would deteriorate the MPC performance [85,89,90]. This highlights a key challenge in thermal management: due to the changing flow rates, the system is considered highly nonlinear [89]. The main part of nonlinear is the heat transfer from the combustion gas temperature to the cylinder wall and the convection heat transfer from the engine wall to the coolant jacket, which both depend on the flow rate. Furthermore, engine thermal management also experiences variable transport delay due to the variable water pump flow rate.

Heat transfer from the combustion gas to the cylinder wall model can be presented as the Woschni/Huber combustion (37) model below [91].

\[
h_{gas} = 130D^{-0.2}p_{gas}^{0.8}T_{gas}^{-0.53}(C_1v_{gas})^{0.8}
\]

Where \( h_{gas} \) is the heat transfer coefficient from combustion gas to the cylinder wall, \( D \) is the cylinder bore, \( p_{gas} \) is the in-cylinder pressure, \( T_{gas} \) is the combustion gas temperature and \( v_{gas} \) is the gas velocity. \( C_1 \) is a constant that depends on engine intake, compression, power and exhaust stroke. Meanwhile, the convection heat transfer coefficient from the engine wall structure to coolant is dependent on the coolant temperature [92,93] and flow rate[94].

However, nonlinear dynamic models can also be handled by MPC. The model can be described as nonlinear differential as in equation (38) and nonlinear cost function as in (39) below [95]:

\[
\dot{x} = f(x(t), u(t)), \quad x(0) = x_0
\]

\[
\min_j \int_t^{t+n_p} F(x(t), u(t)) \, dt
\]
The problem is that Nonlinear Model Predictive Control (NMPC) solving a nonlinear dynamic optimization problem with nonlinear constrain is highly computationally demanding. Furthermore, the global optimization cannot be guaranteed to succeed in every optimization cycle due to the fact that it is a non-convex, constrained nonlinear optimization problem. Moreover, the high computational burden also leads to the optimization not being able to be solved by MPC within a limited time and which could cause computational delay [96]. This is not desirable in a fast process such as engine thermal management where the environment changes very rapidly [97].

The nonlinear problem can be converted back into a linear problem by applying a linearization method. Linearizing the model at a certain operating condition is one of the common methods in the nonlinear control problem. However, this method is not suitable in the engine thermal management as the engine uses a wide range of operating conditions, and these would lead to very different linearized models. CHAPTER 7 explains how the conventional linear MPC is not a proper solution for engine thermal management. The proposed approach as the engine thermal management solution is feedback linearization which will be explained in CHAPTER 8.

5.4. Previous MPC Engine Thermal Management

MPC in thermal management is becoming popular in some industries such as in building thermal management [3–6]. It has been proven that MPC can reduce energy consumption without scarifying the thermal comfort. However, MPC in engine thermal management is still considered unpopular. This can be explained by the fact that the environment in the engine thermal management is different compared to building thermal management; the engine runs at a wide operating range and in quick succession in a normal driving condition. Furthermore, it is a fast tracking with relatively slower actuator responses.

M. Bruckner et al. (2006) [24] developed an MPC controller for an electric water pump to maintain the cylinder head temperature at a constant temperature, 95°C. Meanwhile, the coolant temperature is only maintained by the conventional wax thermostat. The model for
MPC was acquired from the model identification method with 78% and 74% accuracy for the cylinder wall and coolant out temperature. The higher accuracy can be understood from its linearization near the cylinder head and coolant temperature. Having the cylinder head and coolant temperature as variables makes the model identification method unsuitable and the fitting is poor. M. Bruckner et al. (2006) also stated that the known future input helps improve the cylinder head temperature accuracy. Besides, the reduction of NO$_x$ was also recorded by having the known future input.

C. Vermillion et al. (2011) [98] proposed a modular control strategy that combines the Model Predictive Control Allocation (MPCA) with an inner loop reference model. The objective is to control the oil temperature. It uses electric valve and heater rather than oil pump. The oil pump speed is proportionate to the engine speed. It consists of three parts; inner loop reference model design, outer loop control and MPCA optimization.

The outer loop was a PI controller for the tracking of the set point and rejection of the disturbance. The inner loop reference model was to create a desirable target for the MPCA optimization from the outer loop input. In the meantime, the MPCA optimization was to manipulate both actuators to minimize the error between the outputs of the reference model. One particular undesirable strategy by the author is that the heater provides constant 2.25kW heat even after the oil temperature has reached the reference temperature.

H. Wu et al. (2014) [89] developed a thermal management with MPC control, but it was mainly to control the heat exchanger temperature for EGR testing rather than engine thermal management. The thermal management circuit layout is very different from the engine thermal management considered here.

### 5.5. **Summary**

MPC is an advanced control strategy, which is still new in the automotive industry especially engine thermal management. MPC can generally provide an excellent control strategy to
engine thermal management except that the thermal management requires a nonlinear solution of MPC. However, Nonlinear MPC is considered as a computational burden when dealing with fast environment such as engine thermal management. Therefore, in CHAPTER 8 will present the proposed strategy for MPC in thermal management shall be presented by using the feedback linearization approach.
CHAPTER 6  Engine Modelling for MPC

This chapter explains the mathematical model developed in MathWork™ Simulink® for the MPC study. The model is a less complex version of the model made in GT-SUITE (CHAPTER 3). It consists of simple wall, water jacket, radiator, valve and water pump, all represented as a lumped parameter model. However, the nonlinear aspects such as the variable transport delay and variable heat transfer coefficient are fully represented. This model will highlight the difficulty in implementing a typical linear MPC approach (CHAPTER 7) compared to the proposed Feedback Linearization MPC (CHAPTER 8).

6.1. Model Objective

A mathematical model of engine thermal management is created in MathWork™ Simulink® as a controls oriented model as compared to the high-fidelity model in GT-SUITE. For example, the GT-SUITE engine model uses a complex cylinder wall structure to analyse the heat flow and temperature distribution. Other elements like water pump model can simulate other phenomena like surging and choking which are not required in the MPC controller development. The controls oriented model captures the key dynamic and non-linear effects of the plant, while reducing complexity and simulation time. It enables more effective simulation work, which helps in controller development and improvement.
The developed Simulink® model focuses on the wall temperature and coolant out temperature behaviour, while still including the important nonlinearities in the cooling system. The sources of nonlinearity include the heat transfer coefficients, radiator heat transfer rate and air flow rate through radiator. The combustion heat produced has been simplified by linear combustion heat based on power demand rather than engine speed and load.

6.2. Cooling System Model

The thermal model is constructed from first principle. It considers two key components of the cooling system with heat exchange:

- the internal cooling system: engine block and pump,
- the external cooling system: radiator, bypass and mixer.

All other components are assumed to be adiabatic.

Internal Cooling System

The internal heat transfer is from combustion gas to coolant as shown in Figure 6.1. It consists of three parts; combustion gas, cylinder wall and coolant.

![Figure 6.1: Heat transfer from gas combustion to coolant through the cylinder wall model.](image-url)
The cylinder wall represents both a thermal capacitance and a thermal resistor. It accumulates or releases energy depending on the overall thermal flux that runs through it. Accordingly, the enthalpy balances can be formulated using a lumped parameter approach. The enthalpy balance for the cylinder wall is:

\[ C_w \frac{dT_{wall}}{dt} = \dot{Q}_{comb} - \dot{Q}_{conv} \] (40)

\[ \dot{Q}_{comb} = \text{Combustion gas heat transfer rate to wall [W]} \]
\[ \dot{Q}_{conv} = \text{Convection heat transfer rate from wall to coolant [W]} \]
\[ T_{wall} = \text{Cylinder wall temperature [K]} \]
\[ C_w = \text{Cylinder wall heat capacity [J/K]} \]

The combustion gas heat transfer rate to wall \( \dot{Q}_{comb} \) is based on predefined steady state data. The heat transfer rate from wall to coolant \( \dot{Q}_{conv} \) is calculated using the following linear relationships:

\[ \dot{Q}_{conv} = A h_w (T_w - T_{out}) \] (41)

where \( h_w \) depends on the geometry;

\[ h_w = \frac{Nu \cdot k}{L} \] (42)

and the Nusselt number \( Nu \) is a non-linear function;

\[ Nu = f(Re,Pr) \] (43)

\[ \dot{Q}_{conv} = \text{Convection heat transfer rate [W]} \]
\[ A = \text{Heat transfer rate area [m}^2\text{]} \]
\[ h_w = \text{Heat transfer coefficient [W/m}^2\text{K]} \]
\[ T_w = \text{Cylinder wall temperature [K]} \]
\[ T_{out} = \text{Coolant out temperature [K]} \]
\[ Nu = \text{Nusselt number [-]} \]
\[ k = \text{Coolant thermal conductivity [W/mK]} \]
\[ L = \text{Characteristic length [m]} \]
\[ Re = \text{Reynolds number [-]} = \frac{vL}{\nu} \]
\[ Pr = \text{Prandtl number [-]} = \frac{c_p \mu}{k} \]
Equation (42) means that the heat transfer coefficient \( h_w \) is not constant, but a function of both temperature and flow rate. A non-linear model is found from steady state of heat transfer coefficient \( h_w \) from the previous GT-SUITE engine model as shown in Figure 6.2.

![Figure 6.2: Convection heat transfer coefficient throughout coolant temperature and coolant mass flow rate.](image)

The enthalpy balance for the coolant in engine block is based on coolant transport and convection:

\[
C_c \frac{dT_{out}}{dt} = \dot{m}_c \cdot c_p \cdot T_{in} - \dot{m}_c \cdot c_p \cdot T_{out} + \dot{Q}_{conv}
\]

- \( C_c \) = Coolant heat capacity [J/K]
- \( T_{out} \) = Coolant engine out temperature [K]
- \( T_{in} \) = Coolant engine in temperature [K]
- \( \dot{m}_c \) = Coolant mass flow rate [kg/s]
- \( c_p \) = Coolant specific heat capacity [J/kgK]
- \( \dot{Q}_{conv} \) = Convection heat transfer rate from wall to coolant [W]
The coolant mass flow rate $\dot{m}_c$ is determined by the water pump. It is assumed that coolant mass flow rate has a linear relationship with water pump signal; a reasonable approximation for centrifugal pumps typically found in engine cooling system. The normalized water pump signal $n_{pump}$ is in the range of 0 to 1 (no flow to maximum flow). The coolant mass flow rate equation is a simple linear scaling operation as follows:

$$\dot{m}_c = n_{pump} \cdot \dot{m}_{c\ max}$$

(45)

$m_c$ = Coolant mass flow through engine [kg/s]  
$n_{pump}$ = Water pump signal []  
$\dot{m}_{c\ max}$ = Maximum coolant mass flow through engine [kg/s]

Thermal change across the pump is neglected, which means the coolant temperature does not change across the pump.

**External Cooling System**

The external cooling system consists of the valve, the bypass pipe, the radiator and fluid mix junction as shown in Figure 6.3.

![Figure 6.3: Heat transfer from coolant to environment in the radiator.](image-url)
The function of valve is to regulate the flow split between bypass pipe and radiator, which ultimately determines the coolant temperature. The dynamic response time of the valve is neglected; and the flow split is considered a linear relationship to the valve position. The valve lift $l_r$ is normalized to the range of 0 to 1; where 0 means all flow through the bypass, and 1 through the radiator. The flow equation is as follows:

$$
\dot{m}_c = l_r \cdot \dot{m}_{rad} = (1 - l_r) \cdot \dot{m}_{by}
$$

$m_c$ = Coolant mass flow through engine [kg/s]

$l_r$ = Valve lift [ ]

$m_{rad}$ = Coolant mass flow through radiator [kg/s]

$m_{by}$ = Coolant mass flow through bypass pipe [kg/s]

The radiator is modelled as a lump thermal capacitance with convective heat loss. Capacitance that leads to dynamic of the radiator from the enthalpy balance is described as:

$$
C_{c\ rad} \frac{dT_{rad\ out}}{dt} = \dot{m}_{rad} \cdot c_p \cdot T_{rad\ in} - \dot{m}_{rad\ out} \cdot c_p \cdot T_{rad} + \dot{Q}_{rem}
$$

$C_{c\ rad}$ = Coolant heat capacity at radiator [J/K]

$\dot{m}_{rad}$ = Coolant mass flow rate through radiator [kg/s]

$c_p$ = Coolant specific heat capacity [J/kgK]

$T_{rad\ out}$ = Coolant radiator out temperature [K]

$T_{rad\ in}$ = Coolant radiator in temperature [K]

$\dot{Q}_{rem}$ = Heat transfer rate to air [W]

The heat loss is modelled from the steady state behaviour of heat transfer rate from coolant to air ($\dot{Q}_{rem}$). The heat transfer rate is a non-linear function of air mass flow rate, coolant radiator in temperature and coolant mass flow rate through radiator. It is determined from the previous GT-SUITE model as shown in Figure 6.4.
Figure 6.4: Heat transfer rate from radiator to air at 120°C, 100°C and 80°C coolant radiator inlet temperature.

The air mass flow rate through the radiator is also defined using a steady state lookup table from previous GT-SUITE data. It is a function of external air ram to radiator and radiator fan speed as shown in Figure 6.5. This relationship is less important for the controller, because the radiator temperature is not considered a variable of the controller plant model.

Figure 6.5: Air mass flow model.
Coolant from both the bypass pipe and the radiator outlet then joins and mixes at the fluid mix junction before entering the engine. Physically, the coolant mix junction is a simple Tjunction connector of both pipes. Assuming the coolant specific heat is constant, the coolant temperature relation is defined as a simple mixing relationship:

\[
T_{in} = \frac{\dot{m}_{by} \cdot T_{by} + \dot{m}_{rad} \cdot T_{rad}}{\dot{m}_c}
\]

(48)

- \(T_{in}\) = Coolant engine in temperature [K]
- \(T_{by}\) = Coolant bypass temperature [K]
- \(T_{rad}\) = Coolant radiator temperature [K]
- \(\dot{m}_{by}\) = Coolant mass flow rate through bypass pipe [kg/s]
- \(\dot{m}_{rad}\) = Coolant mass flow rate through radiator [kg/s]
- \(\dot{m}_c\) = Coolant mass flow rate through engine [kg/s]

**Coolant Transport delay**

The delay in the coolant circuit is caused by two factors: thermal mass of the components, and the thermal mass and transport delay of the coolant fluid in the pipes and components. Both influence the transient behaviour of the system, and while the thermal mass is mostly constant, the transport delay is strongly dependant on the coolant flow rate. At a low flow rate, the coolant takes a longer time to move through the pipe, and the delay is higher. At a high flow rate, the opposite effect applies, and the delay is reduced. These variable delays are a significant non-linearity of the dynamic behaviour, and they are a challenge to be included in any control model. The transport delay is calculated using this equation:

\[
\tau = \frac{V \cdot \rho}{\dot{m}}
\]

(49)

- \(\tau\) = Time delay in pipe from engine [s]
- \(V\) = Volume of pipe [m³]
- \(\rho\) = Coolant density [kg/m³]
- \(\dot{m}\) = Coolant mass flow rate through pipe [kg/s]

There are three transport delay terms as shown in Figure 6.6:

- the pipe from valve to radiator (\(\tau_1\)),
- the pipe from radiator to fluid mixture point (\(\tau_2\)) and
- the bypass pipe (\(\tau_3\)).
Assumptions are being made that there are no coolant heat loss to ambient along the pipes and pressure drop across component is neglected.

6.3. Determining the wall temperature target

The wall temperature target of an engine is calibrated based on the lowest BSFC throughout engine speed and load as was explained in CHAPTER 3 (Figure 3.11 and Figure 3.13). The combustion heat produced by an engine is closely correlated to the fuel consumption, but it is not directly linear to engine speed and load. In this mathematical model, the target cylinder wall temperature is therefore based on the combustion heat input $\dot{Q}_{comb}$ and not on the engine speed and load. The heat input $\dot{Q}_{comb}$ determines the target wall temperature $T''_w$ (Figure 6.7).
Figure 6.7: Temperature wall target for mathematical model.

The required heat to remove the wall ($\dot{Q}_{\text{conv}}$) is also calculated to examine the working environment for the pump and valve on the temperature target setting. This is to ensure the temperature target is a feasible target for both actuators. This can be done by setting the coolant out temperature $T'_{\text{out}}$ to be reminded at 120°C as long as the water pump speed did not reach the maximum speed. The coolant out temperature should be reduced to remove more heat after the water pump speed has reached the maximum speed.

From equation (41) and Figure 6.2, the coolant out temperature range can be calculated as the equations below:

$$h_w = \frac{\dot{Q}_{\text{conv}}}{A(T'_{\text{wall}} - T'_{\text{out}})}$$

$$h_w = f(T_{\text{out}}, \dot{m}_c)$$

$h_w$ = Convection heat transfer coefficient [W/m$^2$K]
$\dot{Q}_{\text{conv}}$ = Convection heat transfer rate [W]
$A$ = Convection heat transfer area [m$^2$]
$T'_{\text{wall}}$ = Target cylinder wall temperature [K]
$T'_{\text{out}}$ = Coolant engine out temperature at 120°C [K]
$\dot{m}_c$ = Coolant mass flow rate through engine [m/s]

as long as:

$$\dot{m}_c < \dot{m}_{\text{v max}}$$
\[
\dot{m}_c = \text{Coolant mass flow rate [m/s]}
\]
\[
\dot{m}_{c\text{ max}} = \text{Maximum coolant mass flow rate [m/s]}
\]

If

\[
\dot{m}_c = \dot{m}_{c\text{ max}}
\] (52)

The engine coolant out temperature \(T_{\text{out}}\) will drop to achieve the desired \(\dot{Q}_{\text{conv}}\) as in the equation below:

\[
T_{\text{out}} = T'_{\text{wall}} - \frac{\dot{Q}_{\text{conv}}}{Ah'_{w}}
\] (53)

\(T_{\text{out}}\) = Coolant engine out temperature [K]
\(T'_{\text{wall}}\) = Target cylinder wall temperature [K]
\(\dot{Q}_{\text{conv}}\) = Convection heat transfer rate [W]
\(A\) = Convection heat transfer area [m\(^2\)]
\(h'_{w}\) = Convection heat transfer coefficient at maximum coolant mass flow rate [W/m\(^2\)K]

The result shows that coolant temperature only drops until 80°C (Figure 6.8 first graph). This is a reasonable range. If the temperature drops significantly further, for example to 50°C; it is considered too cold for any engine, and it would also require a larger radiator to achieve the required heat rejection.

The results show that a reasonable coolant out temperature working range is from 80°C to 120°C. The water pump also shows a good range of control (Figure 6.8 second graph). The water pump did not reach the minimum pump speed to reach the lower wall temperature target. This gives some reserve to the pump to control the wall temperature.
Figure 6.8: Coolant engine out temperature and water pump signal throughout cylinder wall temperature target.

6.4. Validation

The Simulink® model is fed with the same input signals (combustion heat $\dot{Q}_{comb}$ to the wall, the water pump speed, the valve position and the fan speed) as the GT-SUITE model to compare the wall and the coolant temperature responses. The wall temperature output achieves very high precision, while the coolant temperature output shows a slightly larger deviation (Figure 6.9). This is to be expected, because the wall model is reasonably concise and linear apart from the convection, while the coolant temperature is subject to many non-linear effects.
The main difference between the models is that the water jacket in the GT-SUITE is not represented as a single homogenous volume. The water jacket in the GT-SUITE model absorbs the heat from the engine wall structure locally, and it is not distributed evenly (Figure 6.10). This is different in the Simulink model as all the heat absorbed from the combustion to the wall is then being uniformly transferred directly to the coolant. This creates difference between the models, which is amplified by the closed cooling circuit and potential differences in the radiator model.

However, in terms of the dynamic response, the Simulink model can still be considered as a representative model of the engine cooling system. This is because except for the coolant temperature deviation in certain situations, the Simulink model is representing the main characteristic of the engine cooling system. The model reduction in the Simulink® gives faster a result which is suitable for this work of controller development and analysis.
6.5. **Summary**

A control oriented model of the full cooling system including cylinder wall and radiator has been created in MathWork™ Simulink®. It is a simpler model compared to GT-SUITE model made in CHAPTER 3, in that it uses only a few lumped states, plus a transport delay for the coolant. The wall temperature and coolant temperature transient behaviour in Simulink® model are considered close to the GT-SUITE model even through there is a steady state gap in the coolant temperature model. The Simulink® model will be used in CHAPTER 7 and CHAPTER 8 for the new MPC development and study.
This chapter explains two conventional approaches for applying a linear MPC controller to a nonlinear system. The two approaches are:

- Linearization of the model using simulation data (using the MathWork™ System Identification toolbox™); and
- Jacobian Linearization leading to a number of models around different operating points.

Linearization via the MathWork™ System Identification toolbox™ creates a linear model for MPC directly from the recorded input and output data. The Jacobian Linearization method, on the other hand, linearizes the model around several operating points, creating a set of linear models. This again leads to a several different MPC controllers design around the different operating points, which collectively are able to handle the system nonlinearity throughout the operating region. This chapter try to address the difficulty of finding linear models for thermal management as deliberated in CHAPTER 6, as well as the idea of input linearization. The controllers are being evaluated for their performance and computational burden. The result will be an important indicator which is the most suitable MPC approach in engine thermal management.
7.1. **Linear MPC Implementations**

All real systems are nonlinear, but linear controllers including MPC are used successfully in systems that have minimal nonlinearities or where the nonlinearities can be neglected. Any significant nonlinearity behaviour in the system will reduce the applicability of a linear controller and may lead to a reduction in performance. The critical step for finding a linear control is the linearization of the model. Linearization is a linear approximation of the nonlinear system that is performed at an equilibrium point, and generally expected to be valid within a certain region around this point. So a linear model approximation is good when working near the equilibrium point, but it may not be successful far away from this point. The use of system identification and Jacobian Linearization are two typical linearization methods, and their implementation for engine thermal management will be explained in this chapter.

7.2. **Modelling: MATLAB® System Identification**

System identification is a technique in building a mathematical model of dynamic system from measured input and output data. Linearization via system identification does not require a non-linear model or an operating point, but it will generate an optimal linear model for the given data instead. A black box system identification method can determine the mathematical relation of inputs and outputs without going into the details of what is actually happening inside the system. This helps to build a reasonably simple model of potentially complex systems, which may be hard to model based on the first principle (using physical laws and component behaviours) due to its reasonable time.

The system identification procedure consists of three steps (Figure 7.1):

1. designing suitable input signals,
2. measuring the system responses using the predetermined input signals, and
3. finally using the collected data to estimate the mathematical model typically followed by a validation step.
The approach is called black-box modelling. It is based on the experiment data only, without any knowledge of the system itself.

Data Preparation

System Identification Inputs Setup

Three input signals are selected; combustion heat $\dot{Q}_{\text{comb}}$, pump speed $N_{\text{pump}}$ and valve position $l_r$ (Figure 7.2). The fan speed is fixed at 0.1 to reduce complexity during model fitting. The input signals are randomly generated and some steady state sections are added to create a 5000-seconds time line. The first 3000 seconds are used for model estimation and the remaining 2000 seconds are used as a validation data. The system response (wall and coolant temperature) is monitored to ensure the working temperature limits of the experimental setup. The highest reference for the wall temperature has to be below 250°C whilst the maximal coolant temperature should be less than 120°C.
Identification and Validation

A number of state space and transfer function models with different configurations have been created and compared in the MathWork™ System Identification toolbox™. The best model is a second order state space model, but even this only achieves a moderate fit of 73.5% for the wall temperature and a very poor fit of 32.8% for the coolant temperature in the validation data set (Figure 7.3). The difference can be seen to be both in steady state values and in the dynamics of the model. These results are disappointing, especially for the coolant temperature, which was already found to be more complex. This indicates that there are strong nonlinear effects on the coolant temperature caused by transport delay in the pipe, nonlinear heat transfer from the wall and nonlinear heat transfer at radiator.
The fitted wall and coolant temperature model is able to capture the transient dynamic of the actual data, but it struggles with the correct gain. The highest error of the wall temperature is 10°C (at 3328 seconds and 4480 seconds), but most of the time the linear model stays only a few degrees from the actual wall temperature data. The coolant temperature model is not able to replicate the actual output data with any degree of certainty. It has an error of more than 5°C most of the time, and exceeding 10°C a few times, which is disappointing given the small range of temperatures in the experiment. It also shows a much slower response. Using this model for control could cause significant problems because of the large model uncertainty.

Figure 7.3: The best model fitting for linear MPC in MathWork™ System Identification toolbox™.

**MPC Setup for the System Identification Model**

The MPC controller is calibrated to achieve the optimum performance, which means that it is quite aggressive. The calibration starts with the choice of calibration of prediction and control horizon, which has to be sufficiently long for the stability of the system and to capture the main dynamics. The input and output weights are set for an aggressive pursue of good control performance, which means that the controller is typically constrained by system limits.
Initial Setup

The MPC initial setup for this calibration purpose is used as in Table 7.1 below:

Table 7.1: Linear MPC setup for simulation.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sample time, $t_s$</td>
<td>0.5 seconds</td>
</tr>
<tr>
<td>Prediction horizon, $n_p$</td>
<td>50</td>
</tr>
<tr>
<td>Control horizon, $n_c$</td>
<td>15</td>
</tr>
<tr>
<td>Output weight – $T_w$</td>
<td>10</td>
</tr>
<tr>
<td>Output weight – $T_c$</td>
<td>0.1</td>
</tr>
<tr>
<td>Input rate weight – $\dot{Q}_{conv}$</td>
<td>0.00001</td>
</tr>
<tr>
<td>Input rate weight – $\dot{Q}_{cool}$</td>
<td>0.00001</td>
</tr>
</tbody>
</table>

The sample time duration or time step is the most critical parameter, because it also affects the feasible length of the control horizon. It is chosen to be significantly faster than the relevant process in the system. Based on M. Morari et al. [99] the sample time can be determined using the equation below:

$$\text{sample time} = 0.03 \times \text{settling time} \tag{54}$$

In this case, settling time is not constant but varies from 15 seconds to 50 seconds as shown in CHAPTER 4 under step response. The minimum settling time is selected to ensure the sample time can be used for all transient responses. Based on equation (54) above, the sampling time should be 0.45 seconds which can be rounded up to 0.5 seconds. This is a typical value for a thermal system, although control schemes with less complex controllers and without a prediction horizon may use shorter sampling times.

The wall temperature output weight is set very high to ensure that the controller focuses on close and fast tracking of the reference signal, since the wall temperature has a direct influence on the engine performance. In contrast, the coolant target weight is low, so that the controller has a wide freedom in choosing appropriate coolant temperatures. This also
applies to the manipulated variables rates weight, which is set very low to achieve a fast response.

For testing, the target reference for the wall temperature is based on the combustion heat input as shown in Figure 6.7, while coolant reference signal is set constant at 80°C.

Control Horizon

It is well known that a very short control horizon can lead to an unstable controller. This can be seen here (Figure 7.4): the wall temperature shoots higher than the working temperature limit (250°C). A very long control horizon can also cause issues, because limits may be projected incorrectly, and the computation time may increase unreasonably. For this application, the control horizon is set at 15, since this shows the best wall temperature tracking.
It is noticeable that from 0 to 200 seconds and from 400 seconds to 600 seconds, the MPC could not track the target temperature properly; the actual temperature is significantly higher than the target temperature. Testing has shown that this is not due to physical limits of the system (e.g. the cooling fan), but caused by model mismatch between the experiment and the linear model. This can be demonstrated by reducing the target temperature and the combustion heat, which still causes the controller to fail in tracking the target temperature.
The water pump and the valve are not at maximum, which means the physical system is capable of stronger cooling.

**Figure 7.5:** MPC controller performance with lower target temperature and combustion heat at 0 to 200 seconds and 400 seconds to 600 seconds.

**Prediction Horizon**

A number of different values are tested for the prediction horizon: 15, 25, 50, 100, 150 and 200. The long values of 150 and 200 creates unwanted over and undershoot (Figure 7.6) since the over and undershoot are so significant which cause the wall temperature to run over the temperature limit (250°C). On top of that, the manipulated variables are fluctuating...
aggressively (especially the valve position before 200 seconds). Prediction horizons from 15 to 100 show a more stable control performance and a better temperature tracking. Overall the prediction horizon of 15 gives the best result in temperature tracking. The prediction horizon is usually longer than the control horizon, often at least twice. A shorter prediction horizon may lead to a more aggressive behaviour of the MPC controller.

The MPC using the system identification model delivers reasonable tracking of the wall target temperature, which is expected since the model captures this variable well. The calibration only requires choosing the control and prediction horizon. Further calibration of
MPC input and output weights are not required, since the controller performance and stability of the manipulated variable are considered good enough.

7.3. **Modelling: Jacobian Linearization**

Another approach to finding a linear model is the linearization of the system using Jacobian Linearization. This uses a linear approximation of the non-linear system dynamics at a given equilibrium point. The Jacobian Linearization method typically works well as long as the system stays close to the equilibrium point, since the model accuracy will decrease as the distance from the equilibrium point increases. Selecting the equilibrium point is crucial to exploit the nonlinear behaviour using Jacobian Linearization, and if no one point is found to be sufficient, several equilibrium points and therefore several different linear models may be necessary to achieve acceptable performance across the operating range.

**Equilibrium Points Selection**

The engine coolant system has many equilibrium points, and it is not obvious which one to choose. The factors affecting the vehicle engine are its speed and load, but the cooling system can also run at different cooling temperatures and flow rates. This leads to nonlinearities like the variable transport delay in the cooling system, which depends on the flow rate, as well as nonlinear heat transfer between coolant and solid parts. A single equilibrium point is not able to cover the engine operating range. Therefore, a number of equilibrium points will be used here.

The selection of the variables and states for the equilibrium points also influence the control behaviour and performance. The number of equilibrium points will also increase sharply with the number of variables and partitions per variable. The number of equilibrium points determines the number of MPC controller that have to run in parallel, which increases the computational complexity. Even though, the inactive controllers’ optimisation calculation is turned off, all controllers update its state estimate in parallel to minimise bumps during controller transitions.
Local Linearization Region

The variables are selected based on the nonlinear influence on the system outputs. Coolant flow $m_c$ is the major factor of nonlinearity in the system model (CHAPTER 6). The coolant flow rate $\dot{m}_c$ influences the heat transfer from wall to coolant and causes variable transport delay. Coolant out temperature $T_{out}$ also has a nonlinear heat transfer rate that influences the wall besides the radiator, although this effect is easier to manage. The radiator flow rate $\dot{m}_{rad}$ is another variable that has a nonlinear influence to the heat transfer rate and to the transport delay, but its influence on the wall temperature is quite indirect.

Selecting radiator flow rate as one of the variables to be partitioned can cause conflict with partitioned coolant out temperature $T_{out}$ variable, and an equilibrium point may not be achievable when both $T_{out}$ and $\dot{m}_{rad}$ are fixed. For example, coolant out temperature could not be set at a very high degree with high radiator flow rate when the heat from combustion is not high. Furthermore, $\dot{m}_{rad}$ is directly dependent to pump flow rate $\dot{m}_c$. The combustion heat rate $\dot{Q}_{comb}$ is considered linear to the output wall temperature $T_w$ based on equation (40). However, it could not be considered as linear in this situation when it is coupled with water pump speed and valve position as controller manipulated variables. The $\dot{Q}_{comb}$ would not produce a linear wall temperature dynamic at different equilibrium point of $m_c$ and $T_c$. For example, wall temperature increases by two different equilibrium points with the same initial and different final value of $\dot{Q}_{comb}$ (Figure 7.7).

![Figure 7.7: Comparison of wall temperature increase on step response at different equilibrium point.](image)
Equilibrium points are placed at the centre of each partition of the operating range, and each partition can reflect a different dynamic behaviour. For many systems, the point of transition of the behaviour is obvious, but that is not the case here, since the change is gradual. There is no specific method to determine appropriate regions for thermal management system, and therefore a grid system is used, which is common in automotive engineering. Therefore, all three variables (Combustion heat transfer \( \dot{Q}_{comb} \), pump signal \( n_{pump} \) and coolant out temperature \( T_{out} \)) are divided evenly into two partitions in this work which creates eight evenly place equilibrium points (Figure 7.8).

![Figure 7.8: Region partitions for local linearization.](image)

Table 7.2: Equilibrium points setting.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Equilibrium point setting</th>
</tr>
</thead>
<tbody>
<tr>
<td>Combustion heat transfer ( \dot{Q}_{comb} ) signal</td>
<td>0.25</td>
</tr>
<tr>
<td>Coolant out temperature ( T_{out} )</td>
<td>80°C</td>
</tr>
<tr>
<td>Pump signal ( n_{pump} )</td>
<td>0.25</td>
</tr>
</tbody>
</table>

*Coolant Out Temperature Equilibrium*

The combustion heat transfer \( \dot{Q}_{comb} \) and pump signal \( n_{pump} \) can be easily set to the equilibrium point since both are the inputs of the system. However, coolant temperature
$T_{\text{out}}$ is a measurement, and it requires the choice of suitable system inputs to achieve the desired value. The coolant temperature is not linearly proportional to the valve opening, so no direct mapping occurs. The correct inputs to achieve the desired coolant temperature are found by system inversion using the MATLAB® optimization function with constraints ‘fmincon’. The constraints are set for maximum and minimum allowable valve position, whilst other input readings are set at equilibrium required point during the optimization.

There are a few points that do not reach equilibrium conditions after the valve position optimization. These are caused by physical limits, for example where the combustion heat $Q_{\text{comb}}$ is too high for the given low coolant flow rate $\dot{m}_c$ and low coolant temperature $T_{\text{out}}$. An approximation can be achieved by setting the radiator fan speed to maximum, which does lead to the desired equilibrium.

**Local Linearization Model**

The models are linearized at each equilibrium points in MATLAB® via ‘linearize’ function. This function performs a linearization on the individual components of the model, and then combines them into a linear model of the system. The linearization is based on snapshot of the system state at 5,000 seconds for the system, which is used as the steady state. The Bode plots of the linearized models are shown in Figure 7.9 and Figure 7.10. The wall and coolant temperature responses from combustion heat signal are almost consistent for all equilibrium points. On the other hand, the pump and valve responses show strong dynamic variation over the different equilibrium points.
The linearization of the models at all equilibrium points is fast, certainly much faster than an experimental identification. The experimental identification using small perturbation around the operating point would be preferable, since it avoids errors of the non-linear model.
MPC Setup for Jacobian Linearization

Multiple Model Predictive Control

A linear MPC controller that uses different linear models representing the different sub regions is called a Multiple Model Predictive Control (MMPC). MMPC is a common method to solve control problem in nonlinear system [100–102]. In addition to a standard MPC, the MMPC has a switch criterion that determines which sub region and therefore which model is active. The MMPC receives the switching signal and selects the corresponding controller from the controller bank. The active controller then solves the MPC optimization problem to determine the optimal plant manipulated variables for the current input signals. The MMPC switches from one controller to another as operating conditions change. Here the switch is selected based on created sub regions as shown in Figure 7.8. Typically, the combination of more models results in higher accuracy of the nonlinear approximation, because smaller regions can be used. The local linearization models are used to design MMPC scheme to assess its performance.

![Figure 7.11: Scheme of Multiple Model Predictive Control (MMPC).](image)

Result of MMPC

The MMPC performance is evaluated using the same random square signal disturbance defined previously. The MMPC setup also follows the same setup as in the previous MPC with system identification model. Unfortunately, the MMPC experiences high input fluctuation, which means that it fails to track wall temperature at several points (Figure
Strong fluctuations are noticeable for the valve signal almost throughout the test duration. The fluctuation may be caused by model inaccuracies, or by switching between different controllers.

![Graphs showing temperature and signal changes](image)

Figure 7.12: Multiple Model Predictive Control performance in random square signal disturbance.

The obvious countermeasure is to make the controller less aggressive by increasing the input rate weight. The values 0.00001, 0.1, 1 and 5 are tested, and the stability improves slightly as the weight increases especially from 900 seconds to 1100 seconds (Figure 7.13). However, there is no significant improvement in the temperature tracking.
Number of Sub Regions

It is expected that a higher number of sub regions can improve the controller stability and tracking performance. This is because the transition between models will be smoother, and the linear models are more accurate. Figure 7.14 illustrates the principle of MMPC comparing between low and high number of sub regions in a two-dimensional state-space system.
Figure 7.14: Illustration of comparison number sub regions defined in MMPC. (a) shows less and (b) shows more number of sub regions being defined.

Unfortunately, this improvement cannot be replicated on the experiment. To test the idea, all three variables (combustion heat transfer rate $\dot{Q}_{comb}$, pump signal $n_{pump}$ and coolant out temperature $T_{out}$) are divided into three partitions which created 27 sub regions. The MMPC with 27 sub regions (MMPC27) shows no significant improvement compared to MMPC with 8 sub regions (MMPC8) (Figure 7.15). On the contrary, the MMPC27 controller actually shows a higher input controller fluctuation. This could indicate that the switching is causing the problems since the switching occurs more often with more regions.
7.4. **Controller Comparison**

**Controller Performance**

Linear MPC from system identification model provides a stable controller compared to both 8 and 27 sub regions MMPC (Figure 7.16). The MPC from system identification model does not have input controller fluctuation throughout the random steps response. The wall...
temperature tracking performance is also better than both MMPCs. However, it resulted to big tracking offset errors like 200 seconds to 400 seconds and 1,150 seconds to 1,500 seconds. The errors show that the MPC with system identification model could not capture the nonlinear system dynamic precisely. The MPC with system identification model and MMPC only have a short prediction horizon (15 steps or 7.5 seconds) since the occurrences of both over and undershoot wall temperature happen at a longer prediction horizon setting. The short prediction horizon is not preferred if known future disturbance is to be exploited for improving controller performance.

Figure 7.16: Comparing controller performance between MPC from system identification, MMPC 8 sub regions and MMPC 27 sub regions.
Implementation Difficulty

Implementation of MPC controller via system identification is direct and simple. It only requires data collection for the model fitting in MathWork™ System Identification toolbox™ and the model is ready to be used in MPC. No requirement for detail mathematical model to be built up prior to the modelling fitting. However, cautious tuning is required to work around the poor model fit. Using an aggressive controller can cause unwanted over and undershoot, which could lead to an engine failure.

Using Jacobian linearization implementation or MMPC for engine thermal management in real world would be very difficult. First of all, the controller performance in this case no better than the previous approach, and therefore not worth the additional effort, time and cost of developing and linearizing the model. In addition, the equilibrium point is not easy to achieve in some situations. The local model has to be linearized by using system identification with small input perturbation and this process is repeated in other sub regions. A higher number of regions will add additional effort, both on the design and the implementation of the controller. Therefore, MPC with a model from system identification is found to be a more robust and practical method than the Jacobian linearization (or MMPC).

Computational burden

The computational cost is one of the important criteria in implanting a controller in real world hardware system. Computational cost measures the execution time required per time step during simulation. A High computational cost will cause overruns on the real-time processor, which means that the result arrives later than required. Depending on the scheduling system, a task overrun can cause abrupt performance degradation, and it can even affect other functions.

The computational burden comparison is performed by timing the average simulation time duration (using “tic” and “toc” Matlab® function) , and dividing it by the simulated time. This is pessimistic, because it includes the simulation of the model, but the computation time is dominated by the controller, so the error is reasonably small. It is obvious that MPC with system identification model runs at a much lower computational cost compared to MMPC.
All models run faster than real time, but that is expected given the computing power of a desktop PC (Intel® Core™ i5-2400 CPU @ 3.10Ghz and 8GB RAM).

Figure 7.17: Computational cost between MPC with system identification, MMPC 8 and MMPC 27.

7.5. Conclusion

This chapter compares MPC with system identification model, Multiple MPC with 8 sub regions and 27 sub regions. The comparison shows clearly that MPC with model system identification is better in terms of controller performance, implementation process and computational burden. However, the controller performance should be improved further due to the requirement of controller for engine thermal management system that can track target wall temperature as close as possible. High temperature error could lead to unwanted engine failure such as engine knocking and overheating. The next chapter will explain a new MPC controller to mitigate the concerns.
CHAPTER 8 The New Engine Thermal Management Strategy

As explored in the preceding chapter, thermal management deals with a strongly non-linear system, and linear MPC cannot be applied directly without significant compromises (as explained in CHAPTER 7).

A new engine thermal management strategy is introduced to solve the problem. The strategy is to achieve the lowest brake specific fuel consumption (BSFC) by controlling cylinder wall temperature using linear MPC with feedback linearization. The feedback linearization approach is applied to turn the non-linear problem into a linear one by finding new input signals to the system that compensate the non-linear dynamics. Therefore, this chapter presents the details of addressing the thermal management challenge using linear MPC in combination with feedback linearization.

8.1. The New Control Concept

Controller Objective

As previously mentioned, the control objective is to achieve the lowest brake specific fuel consumption (BSFC) by controlling cylinder wall temperature. As analysed in CHAPTER 3, the
cylinder wall temperature has a direct influence to the engine output, while the coolant temperature only has an indirect effect.

Both electric water pump and valve lift are used as manipulated variables by the controller here. This is the first attempt to use linear MPC to regulate cylinder wall temperature using both electric water pump and valve lift. However, linear MPC is not a straight forward solution to control this non-linear problem, and nonlinear MPC would require excessive computational power. Therefore, the preferred approach is to adapt linear MPC for the thermal management problem with the help of feedback linearization. Figure 8.1 illustrates the concept of MPC with feedback linearization in engine thermal management. The inputs, outputs and known disturbance are linearized by the input transformation. This creates a linear plant model that forms the basis of the MPC design.

Figure 8.1: Engine thermal management with linear MPC and feedback linearization concept.

Feedback Linearization

Linearization is a common method for dealing with a nonlinear control problem. Feedback linearization is one way of linearizing the behaviour of a system. Feedback linearization uses an algebraic transformation to turn the nonlinear systems dynamics into a fully or predominantly linear system. This is achieved by applying inputs to the system that compensates the non-linear dynamics. The feedback can be combined with inputs and states transformation. It is a method of transforming the initial dynamic system model into a comparable dynamic system model of a simpler form [103].
The approach of feedback linearization is very different from conventional Jacobian linearization used before, which is an approximation that works only in close proximity of the equilibrium point.

### 8.2. Feedback Linearization Strategy Overview

A mathematical model is developed to capture the thermal behaviour of the engine cooling system. Feedback linearization is applied to the resulting model.

**Overall Control Strategy**

The overall proposed control strategy is illustrated in Figure 8.2. The coolant out temperature and the wall temperature are measured variables of the nonlinear plant. The control variables are water pump speed, valve lift and radiator fan, and the main disturbance signals are engine speed, engine load, air velocity (or vehicle speed) and the environment temperature.

- **Figure 8.2:** Engine thermal management with linear MPC and feedback linearization.
The nonlinear effect of the control variables (water pump speed, valve lift and fan speed) on the system is linearized by replacing them with two transformed input variables: the wall convective heat flow $\dot{Q}_{\text{conv}}$ and the coolant transport heat flow $\dot{Q}_{\text{cool}}$. The disturbance is also changed into a more manageable variable: $\dot{Q}_{\text{comb}}$. All three linearized inputs are heat transfer rates; the disturbance is the amount of heat transfer rate from combustion to wall $\dot{Q}_{\text{comb}}$, while the two control variables are the heat transfer rate from wall to coolant $\dot{Q}_{\text{conv}}$ and heat transfer rate from coolant in the engine to the outside environment $\dot{Q}_{\text{cool}}$. In terms of these inputs, the plant is nearly linear, because the non-linearity leading to the heat flows have been excluded from the model. The feedback linearization can be seen in the light orange colour area in Figure 8.2.

**Feedback Linearization Methods**

The feedback linearization method is implemented based on thermodynamic first principle as explained in CHAPTER 6. It is reminded that the cooling system used here is representing the basic mathematical engine cooling system heat transfer. Other heat source and heat sink should be added to have a better model representing real world system dynamic behaviour.

The linearization is being divided into two parts: thermodynamic system interacting in the combustion wall and coolant. The first part is as follows:

$$C_w \frac{dT_{\text{wall}}}{dt} = \dot{Q}_{\text{comb}} - \dot{Q}_{\text{conv}}$$

(55)

$$\dot{Q}_{\text{comb}} = \text{Combustion transfer rate [W]}$$

$$\dot{Q}_{\text{conv}} = \text{Convection heat transfer rate [W]}$$

$$T_{\text{wall}} = \text{Cylinder wall temperature [K]}$$

$$C_w = \text{Cylinder wall heat capacity [J/K]}$$

and the second part is the coolant temperature equation:

$$C_c \frac{dT_{\text{out}}}{dt} = \dot{Q}_{\text{conv}} - \dot{Q}_{\text{cool}}$$

(56)

$$\dot{Q}_{\text{conv}} = \text{Convection heat transfer rate [W]}$$

$$\dot{Q}_{\text{cool}} = \text{Radiator heat transfer rate [W]}$$

$$T_{\text{out}} = \text{Coolant out temperature [K]}$$
\[ C_c = \text{Coolant heat capacity [J/K]} \]

**Pump Speed**

Equation (55) represents the cylinder wall temperature model in terms of combustion heat disturbance \( \dot{Q}_{\text{comb}} \) and convection heat transfer \( \dot{Q}_{\text{conv}} \). The convection heat transfer controller variable is based on equation:

\[
\dot{Q}_{\text{conv}} = Ah_w(T_w - T_{out}) \tag{57}
\]

\[
\begin{align*}
\dot{Q}_{\text{conv}} &= \text{Convection heat transfer rate [W]} \\
A &= \text{Heat transfer rate area [m}^2]\text{]} \\
h_w &= \text{Heat transfer coefficient [W/m}^2\text{K]} \\
T_w &= \text{Cylinder wall temperature [K]} \\
T_{out} &= \text{Coolant out temperature [K]}
\end{align*}
\]

The heat transfer coefficient \( h_w \) is determined as a steady state data throughout coolant mass flow rate and temperature.

\[
h_w = f(\dot{m}_c, T_{out}) \tag{58}
\]

\[
\begin{align*}
h_w &= \text{Heat transfer coefficient [W/m}^2\text{K]} \\
\dot{m}_c &= \text{Coolant out temperature [°C]} \\
T_{out} &= \text{Coolant out temperature [K]}
\end{align*}
\]

Here, the coolant mass flow rate \( \dot{m}_c \) is representing the required water pump speed to achieve the controller demand of convection heat transfer rate \( \dot{Q}_{\text{conv}} \). Inverting both equations (57) and (58) can determine the required coolant mass flow rate \( \dot{m}_c \) as below.

\[
\begin{align*}
h_w &= \frac{A(T_w - T_{out})}{\dot{Q}_{\text{conv}}} \\
\dot{m}_c &= f(h_w, T_{out}) \tag{59}
\end{align*}
\]

\[
\begin{align*}
\dot{Q}_{\text{conv}} &= \text{Convection heat transfer rate [W]} \\
A &= \text{Heat transfer rate area [m}^2]\text{]} \\
h_w &= \text{Heat transfer coefficient [W/m}^2\text{K]} \\
T_w &= \text{Cylinder wall temperature [K]} \\
T_{out} &= \text{Coolant out temperature [K]}
\end{align*}
\]

The coolant mass flow is modelled by the function of convection heat transfer coefficient and coolant out temperature and the result is as Figure 8.3 below:
Figure 8.3: Coolant mass flow rate model by the function of coolant temperature and convection heat transfer coefficient.

Valve Position

Equation (56) represents the coolant outlet temperature model based on convection heat $\dot{Q}_{conv}$ and heat transport $\dot{Q}_{cool}$. The $\dot{Q}_{cool}$ is the heat absorb by the coolant in and it can be represent by equation below:

$$\dot{Q}_{cool} = \dot{m}_c \cdot c_p \cdot (T_{out} - T_{in})$$

- $\dot{Q}_{cool}$ = Radiator heat transfer rate [W]
- $\dot{m}_c$ = Coolant mass flow rate through engine [kg/s]
- $c_p$ = Coolant specific heat [J/kgK]
- $T_{out}$ = Coolant engine out temperature [K]
- $T_{in}$ = Coolant engine in temperature [K]

The coolant inlet temperature $T_{in}$ is the result of mixing the coolant from the bypass and the radiator in the coolant mix junction. The bypass coolant temperature is considered
equal to coolant outlet (ignoring the transport delay). Therefore, coolant inlet temperature can be modelled as:

\[
T_{in} = \frac{\dot{m}_{by} \cdot T_{out} + \dot{m}_{rad} \cdot T_{rad}}{\dot{m}_{c}}
\]

(61)

- \( \dot{m}_{by} \) = Coolant mass flow rate through bypass pipe [kg/s]
- \( \dot{m}_{rad} \) = Coolant mass flow rate through radiator [kg/s]
- \( T_{rad} \) = Coolant radiator out temperature [K]

From equation (60) and (61), removed heat transfer rate (\( \dot{Q}_{cool} \)) can also be defined as:

\[
\dot{Q}_{cool} = \dot{m}_{rad} \cdot c_p \cdot (T_{out} - T_{rad})
\]

(62)

- \( \dot{m}_{rad} \) = Coolant mass flow rate through radiator [kg/s]
- \( c_p \) = Coolant specific heat capacity [J/kgK]
- \( T_{out} \) = Coolant engine out temperature [K]
- \( T_{rad} \) = Coolant radiator out temperature [K]

The coolant mass flow rate is considered proportional to the valve opening \( l_r \); therefore, the valve lift can be determined from the equation below:

\[
\dot{m}_{c} = l_r \cdot \dot{m}_{rad}
\]

(63)

**Radiator Fan**

A required heat transfer rate at radiator \( \dot{Q}_{cool} \) gives the amount of air mass flow rate through radiator \( \dot{m}_{air} \). The relationship also depends on the current coolant mass flow rate at radiator \( \dot{m}_{rad} \) and coolant out temperature \( T_{out} \) (Figure 8.4). The air mass flow rate \( \dot{m}_{air} \) then leads to the actual system input, the fan speed \( N_{fan} \), using the steady state data as in Figure 8.5.
Figure 8.4: Air mass flow rate output $\dot{m}_{\text{air}}$ from required heat transfer rate $\dot{Q}_{\text{cool}}$, current radiator mass flow rate $\dot{m}_{\text{rad}}$ and current coolant out temperature $T_{\text{out}}$.

Figure 8.5: Fan signal output $N_{\text{fan}}$ determined from required air mass flow rate $\dot{m}_{\text{air}}$ and current external air ram speed.
8.3. Feedback Linearization MPC Implementation

Modelling of the Thermal Behaviour

Figure 8.6 shows input and output signals for identification and validation of the model. The first 3000 seconds are used for identification and the remaining 2000 seconds for validation. It is the same data being used in System Identification for linear MPC in CHAPTER 7.

![Figure 8.6: The random input signal ($\dot{Q}_{\text{comb}}$, $Q_{\text{conv}}$, $Q_{\text{cool}}$) and the measured out signals for identification and validation.](image)

The model is modelled as separated models; wall temperature model and coolant out temperature model. It tends to make the identification process more critical when the model is a single model since the parameter fitting tends to be ill posed. The identification of the model is using a state space approach, resulting in the following state space model:
\[
\begin{bmatrix}
\dot{x}_1 \\
\dot{x}_2 \\
\dot{x}_3 \\
\dot{x}_4
\end{bmatrix} =
\begin{bmatrix}
0.0016 & -0.0013 & 0 & 0 \\
0.0030 & -0.0025 & 0 & 0 \\
0 & 0 & -0.2364 & -0.0064 \\
0 & 0 & 0.0233 & 5.1303e-04
\end{bmatrix}
\begin{bmatrix}
x_1 \\
x_2 \\
x_3 \\
x_4
\end{bmatrix}
+ \begin{bmatrix}
-1.3283e-04 & -1.3243e-04 & 0 \\
-2.0173e-04 & 2.0104e-04 & 0 \\
0 & -1.5309e-05 & -1.6188e-06 \\
0 & 1.3127e-06 & 1.4990e-07
\end{bmatrix}
\begin{bmatrix}
Q_{\text{comb}} \\
Q_{\text{conv}} \\
Q_{\text{cool}}
\end{bmatrix}
\] 

\[
\begin{bmatrix}
T_w \\
T_{\text{out}}
\end{bmatrix} = \begin{bmatrix}
-1.1064e-03 & 653.5372 & 0 & 0 \\
0 & 0 & -493.4215 & -4.7616e-03
\end{bmatrix}
\begin{bmatrix}
x_1 \\
x_2 \\
x_3 \\
x_4
\end{bmatrix}
\]

The output that the model reproduces from inputs is measured in percentage using the equation:

\[
\text{FIT} = \left(1 - \frac{|y - \hat{y}|}{|y - \bar{y}|}\right) \times 100
\]

Where \(y\) is the measured output, \(\hat{y}\) is the predicted model output and \(\bar{y}\) is the mean of \(y\). The FIT value for cylinder wall model is 97.40% and for the coolant out temperature model is 81.52%. Figure 8.7 shows the comparison of measured output data and model response to the validation data.

![Wall temperature](image1)

![Coolant out temperature](image2)

Figure 8.7: Measured and simulated output in validation data.
The lower fit of coolant out temperature model can be explained by nonlinearity of variable transport delays in the system. However, the model fit is still considered rescannable for the MPC to handle. The temperature error throughout the validation data is not more than ±5°C.

**MPC Constraints**

*Physical limits*

One of the main reasons for using an MPC approach is that it can deal with state and input limits. Implementing suitable constraints is critical for any thermal management system addressing different operating conditions. Constraints can also be used to ensure that the plant stays close to the linearization point, since for further away states, the linearization may no longer be applications. Operating outside the desired conditions could lead to poor control system performance and engine overheating. Table 8.1 below states the physical and actuator constraint used in the engine thermal management system.

<table>
<thead>
<tr>
<th>Variables</th>
<th>Max limits</th>
<th>Min limit</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coolant out temperature, $T_{out}$</td>
<td>120°C</td>
<td>80°C</td>
</tr>
<tr>
<td>Cylinder wall temperature, $T_w$</td>
<td>245°C</td>
<td>170°C</td>
</tr>
<tr>
<td>Water pump speed, $N_{pump}$</td>
<td>5000 rpm</td>
<td>500 rpm</td>
</tr>
<tr>
<td>Valve lift, $l_r$</td>
<td>0.95</td>
<td>0.05</td>
</tr>
</tbody>
</table>

It is assumed that the engine has already reached the operating temperature; therefore, the minimum coolant out temperature is set at 80°C. Small but non-zero minimum values are chosen for water pump speed and valve lift as well as maximum valve lift in order to limit the time delay caused by the coolant transport. The variable time delay caused by varying coolant speeds is a key factor of nonlinearity of the system, and limiting this to a moderate range allows the application of a linear controller. In addition, a minimal pump speed helps to avoid local overheating in the water jacket. Although short excursion below this minimum
may be acceptable, typically the limit is enforced at all times, and the MPC design follows this convention.

**Water Pump Constraint**

The water pump speed and valve lift constraint cannot be applied directly to the MPC, because the variables are no longer part of the linearized plant after feedback linearization. Rather, both constraints have to be transformed and imposed on the new inputs: \( \dot{Q}_{conv} \) and \( \dot{Q}_{cool} \). This approach turns the linear constraint into a nonlinear limit, and a reasonable linear approximation has to be found.

From equations (57) and (58), there are three variables that determined the \( \dot{Q}_{conv} \) physical limits: coolant mass flow rate \( \dot{m}_c \), coolant out temperature \( T_{out} \) and cylinder wall temperature \( T_w \). The \( \dot{Q}_{conv} \) physical limits are calculated throughout coolant out temperature and cylinder liner temperature at maximum coolant flow rate (for maximum limit) and minimum coolant flow rate (for minimum limit).

The result for the \( \dot{Q}_{conv} \) physical limits throughout the coolant out temperature and cylinder wall temperature is shown in Figure 8.8 below. The upper surface is the \( \dot{Q}_{conv} \) at maximum coolant mass flow rate and vice versa. Both maximum and minimum limits create almost a flat surface except for the maximum limit where a slight convex surface is presented. This enables linear \( \dot{Q}_{conv} \) constraints for MPC to be imposed approximately to actual water pump limits.
Figure 8.8: $\dot{Q}_{conv}$ constraint throughout coolant engine out temperature and cylinder wall temperature.

The linear surface constraints are made based on model fitting in MATLAB®. The results are:

Maximum $\dot{Q}_{conv}$ constraint equation:

$$\dot{Q}_{conv} \leq -104.5 \cdot T_{out} + 356 \cdot T_{wall} - 2.014 + 04$$  \hspace{1cm} (66)

$\dot{Q}_{conv}$ = Convection heat transfer rate [W]
$T_{out}$ = Coolant engine out temperature [K]
$T_{wall}$ = Cylinder wall temperature [K]

Minimum $\dot{Q}_{conv}$ constraint equation:

$$\dot{Q}_{conv} \geq -8.25 \cdot T_{out} + 37.53 \cdot T_{wall} - 3509$$  \hspace{1cm} (67)

$\dot{Q}_{conv}$ = Convection heat transfer rate [W]
$T_{out}$ = Coolant engine out temperature [K]
\[ T_{\text{wall}} = \text{Cylinder wall temperature [K]} \]

**Valve Constraints**

The maximum constraint for removed heat transfer rate \( \dot{Q}_{\text{cool}} \) is when valve opening as well as radiator fan is at its maximum and via versa. From equation (62), three variables determine the heat transfer rate \( \dot{Q}_{\text{cool}} \): coolant radiator mass flow rate \( \dot{m}_{\text{rad}} \), engine coolant out temperature \( T_{\text{out}} \) and coolant radiator out temperature \( T_{\text{rad}} \). Two out of the three stated variables are not in the MPC controller variables; radiator coolant mass flow rate \( \dot{m}_{\text{rad}} \) and coolant radiator out temperature \( T_{\text{rad}} \). Therefore, a different approach is considered to replace the two variables.

The radiator mass flow rate depends on engine coolant mass flow rate at maximum valve lift for maximum constraint and vice versa. The engine coolant out mass flow rate \( \dot{m}_{c} \) can be replaced by \( \dot{Q}_{\text{conv}} \) as shown in equations (62) and (63). The cylinder wall temperature and coolant engine out temperature parts of the equation are based on the results in Figure 8.9.

\( \dot{Q}_{\text{cool}} \) is then plotted throughout \( \dot{Q}_{\text{conv}} \) and coolant out temperature. The result is as shown in Figure 8.9. The upper surface represents \( \dot{Q}_{\text{cool}} \) limit at maximum valve lift and the lower surface represents \( \dot{Q}_{\text{cool}} \) limit at minimum valve lift. Although it is hard to observe from the figure, the surfaces dominated by a strong convex curvature, with a small lightly concave area near maximum valve lift.
A linear constraint could not represent this complex limit surface properly. Therefore, three linear constraints are created to approximate the limit surface. The minimum limit is sufficiently flat to be presented with a single linear constraint (Figure 8.10). The linear constraints are conservative to ensure the main operating points are within the constraint area, but this will lead to a slight loss of control authority and performance. The controller performance will drop significantly, or it could even become unstable, if the constraints are too relaxed, allowing invalid inputs. Constraint A (red surface in Figure 8.10) is fitted from the lowest point at 120°C, lowest point at 80°C and middle point at 100°C. Constraint B (blue surface in Figure 8.10) is fitted from the middle point at 100°C, highest point at 100°C and middle point at 80°C. Constraint C is fitted from the middle point at 80°C, highest point at 80°C and highest point at 100°C.
Figure 8.10: Linear constraint surfaces for $\dot{Q}_{cool}$.

The results for the surface model fitting equations are:

Upper $\dot{Q}_{conv}$ constraint equation:

\[
\dot{Q}_{cool} \leq -2.607 \cdot \dot{Q}_{conv} - 271.3 \cdot T_{out} + 3.071 \times 10^4
\]

\[
\dot{Q}_{cool} \leq -1.603 \cdot \dot{Q}_{conv} - 618.9 \cdot T_{out} + 5.419 \times 10^4
\]

\[
\dot{Q}_{cool} \leq -0.7985 \cdot \dot{Q}_{conv} - 1445 \cdot T_{out} + 1.013 \times 10^5
\]

\[
\dot{Q}_{cool} = \text{Removed heat transfer rate at radiator [W]}
\]

\[
\dot{Q}_{conv} = \text{Convection heat transfer rate [W]}
\]

\[
T_{out} = \text{Coolant engine out temperature [K]}
\]

Lower $\dot{Q}_{conv}$ constraint equation:

\[
\dot{Q}_{cool} \geq -0.4266 \cdot \dot{Q}_{conv} - 31.65 \cdot T_{out} + 3914
\]
\[ \dot{Q}_{\text{cool}} = \text{Removed heat transfer rate at radiator [W]} \]
\[ \dot{Q}_{\text{conv}} = \text{Convection heat transfer rate [W]} \]
\[ T_{\text{out}} = \text{Coolant engine out temperature [K]} \]

The linear constraints do not achieve a perfect approximation of the actual constraint (Figure 8.11). The surfaces cross section at 120°C shows that the linear constraints limit the radiator short of its full heat rejection potential, while the lower constraint is very close to the actual minimum constraint (bottom of Figure 8.11). The deviation of the upper limit is not important, because a coolant temperature of 120°C is only used at low load, and the valve would not typically be fully open. In contrast, the valve minimum constraint is much more important because it covers a typical operating point of low combustion heat and low convection heat.

The cross section at 80°C is different (top of Figure 8.11). The typical operating point is a high load and high heat transfer. The linear surfaces constraints here are a much better fit of the actual upper constraint to ensure the valve is able to operate at fully capacity when needed. The lower constraint has a more significant deviation, which is again not usually relevant, because the valve is not usually fully closed when the load is high.
MPC parameter tuning is guided by the desire to achieve good reference tracking, sufficient disturbance rejection and robustness against model mismatch. The dynamic behaviour of the plant strongly depends on MPC parameters selection besides the MPC model accuracy.
In general, unconstraint MPC with a good linear model will have better performance with a long horizon. However, the computational effort increases with the horizon length, because a larger problem has to be solved in each MPC iteration. In contrast, MPC with a short horizon can be unstable or show poor performance. Plant nonlinearity and dead times cause a model mismatch, which adds to the complexity of the parameter selection. Therefore, the MPC tuning selection is being explained in this part.

The tuning of the MPC parameters is only focusing on the controller performance to track the wall temperature target. This is to achieve the fuel consumption reduction and engine performance output as explained in CHAPTER 2.

**Control and Prediction Horizon**

The influences of MPC control and prediction horizon are also investigated on a random square signal disturbance with the same initial MPC parameter setup as previous. The initial MPC setup is as Table 8.2 below:

<table>
<thead>
<tr>
<th>Variable</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sample time, $t_s$</td>
<td>0.5 seconds</td>
</tr>
<tr>
<td>Prediction horizon, $n_p$</td>
<td>50</td>
</tr>
<tr>
<td>Control horizon, $n_c$</td>
<td>25</td>
</tr>
<tr>
<td>Output weight – $T_w$</td>
<td>10</td>
</tr>
<tr>
<td>Output weight – $T_c$</td>
<td>0.1</td>
</tr>
<tr>
<td>Input rate weight – $\dot{Q}_{conv}$</td>
<td>0.00001</td>
</tr>
<tr>
<td>Input rate weight – $\dot{Q}_{cool}$</td>
<td>0.00001</td>
</tr>
</tbody>
</table>

Here, the manipulated variables rate weights are set low for two reasons: one is to achieve fast responses, and the second is that due to the transformation, the inputs do not represent pump, valve and fan directly. Therefore the manipulated variable rate has no direct physical meaning.
The control horizon $n_c$ is compared between 3, 5, 10, 15, 25 and 50 on random square signal disturbance as shown in Figure 8.12 with the known future inputs. The result shows that larger control horizon has better signal tracing. This is clearly seen when compared to the Root-Mean-Square Error (RMSE) of the actual wall temperature to the reference target (as in Figure 8.13). The RMSE reduces significantly as the $n_c$ is larger, and it begins to settle after 25. This can be explained by the fact that the feedback linearization results in a predominantly integral system (it is no longer self-stable) that needs to be actively controlled, and a longer horizon makes the controller more effective. This means that the control horizon in this integrating system requires a longer control horizon to achieve a good performance.
Figure 8.12: Control outputs and manipulated variables of Feedback Linearization MPC with different control horizons.
Another observation is that for all control horizon settings, the controller does not achieve good wall temperature reference tracking when the target temperature suddenly drops to very low values. The coolant temperature increases more than coolant target temperature, 80°C. This can be seen at 250 seconds to 400 seconds and 1250 seconds to 1350 seconds. This can be explained by the constraint imposed on the pump as in equation (67). The equation shows that the $\dot{Q}_{\text{conv}}$ can be further reduced to increase the wall temperature by higher coolant out temperature. However, the coolant out temperature increase is not fast enough for the wall temperature to be near the target temperature. This is a clear indicator that the constraints imposed create a connection between the pump and the valve to achieve the controller objective.

Overshoot and undershoot tend to happen especially for smaller $n_c$ (except for $n_c = 50$). The overshoots happen before reference temperature step up while undershoots happen before reference temperature stepped down. This can clearly be observed at 1160 seconds to 1200 seconds (for overshoot) and 1220 seconds to 1240 seconds (for undershoot) as shown in Figure 8.14.
The reason is the MPC with known future reference target wall temperature and disturbance generates optimized move that creates the undershoots and overshoots. The optimization move can be seen in Figure 8.15. After the end of the control horizon, the prediction horizon continues with the same manipulated variable, and the nature of integrating system causes the wall temperature to keep decreasing or increasing when there is unbalance to the control input and disturbance (\( \dot{Q}_{\text{conv}} \) and \( \dot{Q}_{\text{comb}} \)). This happens whenever the control horizon is not the equal length to the prediction horizon. This causes the unwanted overshoot to satisfy the MPC optimization problem. The overshoot starts...
decreasing (after 1190 seconds) when the disturbance and the changes of reference temperature enter the control horizon range.

Figure 8.15: MPC prediction trajectory and optimized control movement trajectory at 1180 seconds that create the overshoot in wall temperature ($n_p = 50, n_c = 10$).

This shows that the control horizon is preferable to be equal to the prediction horizon. However, having a long control causes a heavily computational burden. Depending on the nature of the problem, the computation burden increases exponentially in each additional control horizon step.

**Control Input Blocking**

A method that can achieve control performance near the performance of MPC with a long control horizon without this computational complexity is called control input blocking. Control input blocking is a method to “block out” the input moves at selected steps from the calculation by setting them to be equal to the prior value. This creates a reduction in the number of input moves that needs to be computed through the optimization, but it keeps the system under active control until the end of the prediction horizon. The selection of blocking location is critical to achieve the desired behaviour. There is no specific technique in
selecting the blocking location but in general, less blocking is being used for immediate moves and more of the distant.

Five control horizons with blocking are being observed in the same random steps response. The control horizon steps sequences are as shown in Table 8.3 below:

Table 8.3: List of variable control horizon sequences.

<table>
<thead>
<tr>
<th>Control horizon sequence</th>
<th>Number of steps</th>
<th>Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>1, 3, 6, 10, 15, 21, 28, 36, 45</td>
<td>9</td>
<td>Seq 1</td>
</tr>
<tr>
<td>1, 2, 4, 6, 9, 12, 16, 20, 25, 30, 36, 42, 49</td>
<td>13</td>
<td>Seq 2</td>
</tr>
<tr>
<td>1, 2, 3, 5, 7, 9, 12, 15, 18, 22, 26, 30, 35, 40, 50</td>
<td>15</td>
<td>Seq 3</td>
</tr>
<tr>
<td>1, 2, 3, 4, 6, 8, 10, 12, 15, 18, 21, 24, 28, 32, 36, 40, 45, 50</td>
<td>18</td>
<td>Seq 4</td>
</tr>
<tr>
<td>1, 2, 3, 4, 5, 7, 9, 11, 13, 15, 18, 21, 24, 27, 30, 34, 38, 42, 46, 50</td>
<td>20</td>
<td>Seq 5</td>
</tr>
</tbody>
</table>

Overall result in Figure 8.16 shows that all variables for control horizon sequence performance are very good, almost comparable to the MPC with 50 steps control horizons and the difference between the control horizon sequences is also not significant. Nevertheless, the control horizon sequences are a bit unstable especially with lower number of steps as marked in Figure 8.16 at 580 seconds and 1150 seconds. It is obvious for Seq 1 and Seq 2. The reason is the big interval blocking towards the end causes the wall temperature to keep decreasing or increasing when there is an unbalance to the control input and disturbance ($\dot{Q}_{conv}$ and $\dot{Q}_{comb}$) due to integrating system. This is almost the same undershoot and overshoot problem as the short control horizons but it happens at each big blocking intervals with lesser effects. This can be understood as in Figure 8.17. However, the MPC performance with low number of control horizon step is better compared to the same number of control horizon in normal sequence. The undershoot and overshoot are also very minimal. For this purpose, Seq 3 is selected since the performance does not significantly improve compared to Seq 4 and Seq 5.
Figure 8.16: Feedback Linearization MPC with variable control horizon sequences.
Figure 8.17: MPC prediction trajectory and optimized control movement trajectory at 580 seconds that create the interval undershoot (Seq 1).

Prediction Horizon

Prediction horizon lengths $p$ of 10, 30, 50, 70 and 90 are investigated with the control horizon sequence Seq 3. As expected, the result clearly shows that a longer prediction horizon has better control performance. $p = 10$ is clearly too low due to the fluctuation occurs on coolant temperature throughout the test duration. Figure 8.18 below is the RMSE result across prediction horizon. RMSE is reduced as the number of control horizon steps increases (the exception is $p = 10$ which does not achieve acceptable control).
Figure 8.18: Comparing RMSE across number of prediction horizons steps.

The controller performance improves with a higher number of prediction horizon steps which can clearly be seen when the temperature reference drops very low (at 200 seconds to 420 seconds and 1250 seconds to 1350 seconds as shown in Figure 8.19 and Figure 8.20). This is due to coolant temperature that increases quickly for longer prediction horizon. This enables the wall temperature to reduce the temperature drop significantly.
The reason coolant temperature rises quickly can be seen in Figure 8.20. Coolant temperature increases quickly for larger prediction horizon because the valve is able to respond earlier prior to the reference changes. Higher coolant temperature creates hotter wall temperature which helps it to reach the target wall temperature better.
Figure 8.20: Larger prediction horizon trigger quicker coolant temperature increases thus improve wall temperature accuracy.

Even though selecting very large prediction horizon may be desired as it gives better controller performance, it is not feasible from a computational perspective. Larger prediction horizon increases the computation burden. Therefore, $p = 50$ is selected, as it represents a good compromise between performance (which is not much different compared to $p = 90$) and computational cost.
**Coolant Temperature Target**

It is interesting to note that target coolant temperature influences the amount usage of pump and fan. High coolant temperature requires the pump to run faster to remove the same amount of heat produced by the engine to the coolant while the fan can run slower and vice versa. A balanced calibration of the coolant temperature could help to minimize the usage of total power to run the pump and the fan; thus giving a modest reduction in fuel consumption.

The optimization can be done in detail by considering the relative power of pump and fan power, ambient temperature and vehicle velocity. To stay concise, only variations of fan and pump usage is considered in this work. Since the fan generally consumes more power, this work focuses on minimal fan usage as long as the incremental work of the pump is minimal. The calibration is done by observing MPC move throughout different wall temperature targets with a coolant temperature target at 70°C, 80°C, 90°C, 100°C, 110°C and 120°C as shown in Figure 8.21. A coolant temperature of 70°C could not be achieved throughout the simulation due to physical constraints. At 160°C to 170°C wall temperature, the constraint is so significant that the minimum coolant temperature needs to increase up to 110°C. After 197°C, the maximum temperature is limited due to constraints. The fan does not turn on for wall temperatures below 203°C.

To reduce the necessary flow rate and therefore the power of the pump, the coolant temperature should be set as low as possible as long as the fan is not active, and this is shown in Figure 8.21. The calibrated coolant temperature increases slightly from 203°C to 235°C wall temperature to minimize the required fan power. This requires a slightly higher flow rate and therefore more pump power, but the increase remains very small. The calibrated coolant temperature target over wall temperature reduces the usage of pump and fan, which leads to a reduction in fuel consumption.
Figure 8.21: Fan and pump signal at various coolant out target temperatures and the optimum target temperature.

The calibrated coolant temperature target is also observed in the random square signal disturbance. A few weight setups are also implemented to compare the effect of the coolant out temperature target implementation. The weights setting being compared are as follows:

Table 8.4: Weight setting between wall temperature and coolant out temperature with calibrated coolant out temperature target.

<table>
<thead>
<tr>
<th>Configuration</th>
<th>Weight setting</th>
<th>Coolant temperature target</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Wall temperature</td>
<td>Coolant temperature</td>
</tr>
<tr>
<td>NomWgh 1</td>
<td>10</td>
<td>0.1</td>
</tr>
<tr>
<td>Wgh 1</td>
<td>10</td>
<td>0.1</td>
</tr>
<tr>
<td>Wgh 2</td>
<td>10</td>
<td>5</td>
</tr>
<tr>
<td>Wgh 3</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>Wgh 4</td>
<td>5</td>
<td>10</td>
</tr>
</tbody>
</table>
The calibrated coolant out temperature target (Wgh 1) has the same weight setting that shows improvement in controller performance compared to for a fixed coolant out temperature target (NomWgh 1). The performances can be further improved by increasing the weight on the coolant out temperature for the calibrated target temperature as shown in Figure 8.23. Increasing the weight on the coolant improves the coolant out temperature tracking to the calibrate target. This is significant at 250 seconds to 380 seconds and 1250 seconds to 1400 seconds (Figure 8.22). The coolant temperature increases faster and provides better signal tracking for higher coolant out temperature weight setting. Improvement in the coolant out temperature signal tracking also solves the dip problem for the wall temperature. However, increasing the coolant out temperature weight beyond the wall temperature weight tends to make the wall temperature less important in the MPC optimization problem; thus, it reduces the controller performance in wall temperature tracking. The RMSE increases as the weight on the coolant out temperature is further increased. From this observation, the weights shown below demonstrate the best compromise between low RMSE and do not have the wall temperature dip.

Table 8.5 : MPC Coolant out weight setting.

<table>
<thead>
<tr>
<th>Configuration</th>
<th>Weight setting</th>
<th>Coolant temperature target</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Wall temperature</td>
<td>Coolant temperature</td>
</tr>
<tr>
<td>Wgh 2</td>
<td>10</td>
<td>5</td>
</tr>
</tbody>
</table>
Figure 8.22: Calibrated coolant out temperature target under random square signals disturbance.
Figure 8.23: RMSE results comparing different wall temperature and coolant out temperature weight setting.

The pump and fan usage with the calibrated coolant out temperature target is shown in Figure 8.24. It is clear that fan usage reduces significantly compared to the fixed coolant out temperature target, and it reduces further as the weight is being increased. As stated previously, the pump usage may increase slightly as the fan usage is dropped but it is not significant. The highest increase is only 1.12% compared to without calibration. This shows that the MPC with calibrated coolant out temperature target could further reduce fuel consumption by saving power on radiator fan.

Figure 8.24: Changes of pump and fan usage compare to without calibrated coolant out temperature target.
8.5. **MPC Future Prediction**

Since MPC is a receding horizon optimisation, it works best if future disturbances are known. However, in the real world, unpredictable disturbance during driving should be expected, because road conditions and driver behaviour cannot be predicted with certainty. Therefore, the MPC has to be robust enough to handle deviation from the predicted scenarios. Two main approaches can be distinguished:

a) MPC without prediction for future disturbance; and  
b) MPC with incorrect future prediction.

**MPC without Disturbance Predictions**

In most cases, a stable MPC should be able to control a system without knowing future disturbances. Often a status quo prognosis is made, which means that known disturbances are expected to remain constant. This is slightly different from a basic Linear Quadratic Regulator (LQR) controller, which assumes a disturbance of zero, and more like a PI controller, which adjusts to a constant offset.

A comparison is made between MPC with and without known future disturbance knowledge using otherwise identical random square signal disturbance (Figure 8.25). As expected, the MPC with known future disturbances responds faster and performs better than the one without, because it can better manage the slow response of the coolant transient dynamic. This is especially obvious at the intervals from 250 seconds to 380 seconds and from 1250 seconds to 1400 seconds. At these points in time, the coolant temperature is required to be at target coolant temperature in order to achieve the target wall temperature. The MPC with known future disturbances can reach the target coolant temperature earlier, which makes the control overall perform better. Outside of these two intervals, the MPC with known future disturbances still shows a slightly better performance, but the difference is quite small.
Figure 8.25: Comparison between MPC with and without known future inputs.
MPC with Incorrect Future Disturbance Estimates

When predictions are used, fact that they may be incorrect has to be considered. This can happen when estimates of road conditions or driver behaviour models turn out to be wrong, and therefore a demand different from the predicted scenario are seen. The error can be categorised into two severities:

a) False reference 1: a change is not predicted, or a predicted change does not materialise (no correlation); and
b) False reference 2: the prediction is the inverse of the prediction (worst case)
The actual reference in comparison to the incorrect prediction is illustrated as in Figure 8.26 below:

![Figure 8.26: Actual reference compared to incorrect prediction 1 and incorrect prediction 2.](image)

The MPC will still read the actual reference and measure the outputs at current time or in other words the first step of prediction horizon. The rest of prediction horizon (from step 2 till 50) uses the given (incorrect) prediction.

MPC with false prediction will result in a significant regulation error when the actual reference is different than the predicted reference (Figure 8.27). The False Prediction 1 shows a significant drop at 150 seconds, lasting until 300 seconds. This is because the expected heat generated $\dot{Q}_{\text{comb}}$ to the wall is higher and reference coolant temperature is far lower than the actual wall reference. False Prediction 2 performs slightly better in the same time frame, despite the expected heat generated $\dot{Q}_{\text{comb}}$ being higher than in False Prediction 1. This can be explained by the coolant temperature reference for False Prediction 2 being higher and causing the coolant temperature to increase more. False Prediction 2 shows a higher wall temperature error at 300 seconds to 450 seconds. This is due to the expected $\dot{Q}_{\text{comb}}$ and coolant temperature reference which is far higher compared to the actual reference. The high reference creates high coolant temperature which
increases the wall temperature even when the pump is already at maximum speed. Despite the error, the wall temperature for MPC with False Prediction 2 still remains below the wall temperature limit of 250°C, which means that the control limits are applied correctly.

Figure 8.27: Comparing MPC with false reference prediction.

Both results (Figure 8.26 and Figure 8.27) show that the MPC with an incorrect prediction is worse than an MPC without knowledge of future disturbances. This highlights how important the accuracy of the prediction is for the MPC performance in this system. If the prediction is highly uncertain, it may be advisable to use a controller that ignores it completely.
Controller Comparison

The results have shown that the Feedback Linearization MPC shows the best performance in terms of wall target temperature tracking and stability compared to previous MPC configurations (Figure 8.28). The feedback linearization configuration is the only controller that does achieve a low steady state error. This demonstrates that the engine thermal management nonlinearity problem can be solved by using the feedback linearization approach.

Figure 8.28: Comparing controller performance between MPC from feedback linearization, system identification, MMPC 8 sub regions and MMPC 27 sub regions.
The feedback linearization also shows positive results in terms of computational cost. The computational cost is as low as the MPC with system identification linearization (Figure 8.29), even with additional complex from input and output signals transformation. The average simulation time for the feedback linearization configuration is just under 7% compared to the actual time. This indicates that the feedback linearization approach gives good controller performance without adding significant computational cost.

![Figure 8.29: Computational cost between MPC with feedback linearization, system identification, MMPC 8 and MMPC 27.](image)

However, the implementation is not as easy as the linear MPC based on a system identification approach. The steady state model responses such as the heat transfer coefficient response, coolant mass flow rate response and air flow rate response are required as important elements in the input and output signals transformation. A number of steady state measurements in various conditions are required to create these steady state models. The implementation of the actuator limits in terms of the transformed inputs turned out to be particularly complex. The constraints require comprehensive steady state data and engineering judgement to transform the complex nonlinear surface constraints into linear constraints.
8.7. **Conclusion**

A new approach for dealing with the non-linearity of the cooling system using feedback linearization has been demonstrated. The linear constraints sufficiently allow the MPC to replicate the pump and fan constraints based on the test done. The integral nature of the linearized system requires the MPC to have a longer control horizon, and this is achieved using blocking to prevent excessive computational complexity. The controller performance is further improved by having calibrated coolant out temperature target. The variable coolant out temperature target reduces the total usage of pump and fan; thus, it expected further reduction in fuel consumption, without creating any additional computational burden. Future prediction of inputs changes also plays a major role to the MPC controller performance. Wrong prediction would cause big temperature offset compared to the reference wall temperature.
This chapter explains the experimental validation of the new MPC controller. The experiment is performed on a small-scale test rig, which provides a highly controlled environment for the key components of the cooling system. The test rig represents the actual engine cooling system which consists of a radiator, an electric pump, a servo valve, an aluminium water block and a heater surface.
9.1. **Experiment Setup**

**Introduction**

![Image of test rig](image)

Figure 9.1: The test rig representing engine cooling system.

The test rig (as shown in Figure 9.1) has been designed to experimentally demonstrate the Feedback Linearization MPC control strategy. It consists of the same type of components as an actual engine cooling system, used in functionally identical ways. The MPC control strategy is tested on this rig instead of an actual engine for the following reasons:

- A typical engine ECU is responds closely to changes in coolant temperature, and it will change engine spark timing and air fuel ratio (spark timing and air-fuel ratio correction) when the coolant out temperature is not at a normal working temperature (usually set at 90°C). The changes would distort the results of the experiment.
- The equipment and operating costs of an engine and an engine test cell are much higher, due to the need for a dynamometer, usage of fuel, management of heat and emission, and dedicated control equipment.
- It would be risky to run an engine near to the temperature limits and during control development that would inevitably occur. Every engine has its own design of the water jacket to remove excessive heat from the gas combustion, and without a detailed understanding of the heat flow local heating could lead to engine failure.
- The test rig enables more efficient testing, as it does not need any specific pre-test procedure such as engine warm-up.
- A test rig provides a much more controlled test environment which reduces unwanted influence such as variations in combustion, oil temperature, air temperature and unmeasurable heat loss.

**Experiment setup and components**

The test rig consists of an electric water pump, a motorized 3-way valve, a radiator, a radiator fan and an aluminium water block with a heater. Figure 9.2 shows the schematic diagram of the test rig.

![Figure 9.2: The test rig cooling system schematic diagram.](image)
**Aluminium Water Block**

The aluminium water block is fabricated to represent the engine cylinder wall and water jacket. It has groves on one side of the block to function as engine water jacket and a surface heater placed on the opposite side. The water jacket groove is designed (as shown in Figure 9.3) based on considerations and comparison with a few other designs made in Autodesk® CFD 2015. The design creates 5.3 m/s of maximum coolant velocity at maximum water pump speed. The local maximum velocity is close to the actual local coolant velocity of an engine [104–106]. The wall thickness of 13mm is comparable to an actual engine cylinder wall thickness at the cylinder head [107], leading to very similar thermal dynamics.

Two thermocouples are installed at 2mm and 8mm from the surface heater to measure the water block temperature and calculate the generated heat transfer across the water block. The surface heater and water block are wrapped with insulator to minimize heat loss to atmosphere.

![Aluminium Water Block](image)

Figure 9.3: Water jacket inside the aluminium block design in 3D.

**Surface Heater**

A 1kW surface heater is placed on the opposite side of the water block. The surface heater is an aluminium nitride (AlN) type of ceramic heater from Watlow® ULTRAMIC® (CER-1-01-00002) that can closely replicate the fast-changing heat transfer in combustion engine. The heater can operate up to 400°C with an ultra-fast ramp rate of up to 150°C per second, which again aligns with typical cylinder wall surface temperatures. This good thermal
conductivity of the heater provides rapid heat dissipation, enables the heater to be constructed with this high-power density.

Figure 9.4: Aluminium nitride ceramic heater from Watlow® ULTRAMIC®

The surface heater is coupled with a Watlow® DIN-A-MITE® A power controller (DA10-02F0-0000). It uses a silicon control rectifier with duty cycle based control to regulate the surface heater output power. The variable duty cycle control changes the ratio of alternate current cycles that are active and cycles that are inactive, because the silicon control rectifier can only switch during the zeros of the sign wave signal. The ratio of alternate current turned on to alternate current turned off increases as demand increases. This method of deterministic switching can control the amount of average power produced by the surface heater. Due to the thermal mass of the system, the switching is well filtered, and not visible in the temperature measurement.

However, the switching does not provide infinitely variable control of the demands. It can only implement finite increments of 10%, but no intermediate values. The maximum power is limited to 50% since the surface temperature approaches 200°C at 60% power. Higher temperatures would degrade the thermal compound bonding between the surface heater and the aluminium water block (Figure 9.5). At higher heater power, even maximum pump speed and a fully open mixer valve cannot maintain a safe operating temperature. At the 50% power setting, the surface heater can produce 0.656MW/m² heat flux which is close to the actual heat flux seen in an engine combustion chamber [108].
Figure 9.5: Surface heater generated heat and temperature for every 10% of power demand.

*Water Pump*

A variable speed electric water pump (MCP655-PWM) is used to achieve the convection of the coolant, and the pump power is controlled via a digital Pulse Width Modulators (PWM) input. Two identical pumps are used in series to reach the desired flow rate at the high back pressure of the coolant circuit. Depending on the input, the pumps produce a flow rate between 0.026 kg/s and 0.150 kg/s. The pumps have an operating temperature limit of 60°C, and therefore the coolant temperature is limited to 55°C during the experiment. This temperature is significantly lower than the coolant temperature on a pressurised circuit, which can reach up to 120°C at high engine power. The dynamic behaviour is largely unaffected by this change, and the temperatures are scaled linearly to correspond to the engine situation. In this experiment, the temperature range of 30°C to 55°C is used to represent an engine coolant temperature of 60°C to 120°C.

*Servo Valve*

The servo valve consists of a mechanical 3-way valve and an electric servo actuator. The 3-way valve is the VRG130 valve from ESBE, which is combined with the recommended valve actuator ARA639 (also from ESBE). The actuator consists of an electric motor, gearing, a position sensor, and position control servo electronics, which maintains the actual valve position close to the requested reference position. The valve is symmetric, so at the half
way position, equal flow should go through both circuits. Different flow restrictions can distort this balance, and a flow restriction is added to the hot (bypass) circuit which provides the same flow restriction as the radiator; therefore, it helps to equalise the coolant circuit. This equalisation significantly helps the modelling and control of the system, especially at very uneven flow distributions.

**Data Acquisition**

A CompcatRIO™ cRIO-9074 from National Instrument™ (NI) is used as data acquisition for the test rig. In addition to an embedded CPU, the cRIO-9074 has a reconfigurable field-programmable gate array (FPGA) circuitry in the chassis, which has control of the I/O modules. FPGA is a reprogrammable silicon chip that can implement specified functions directly rather than running in a software application, and it enables the implementation of custom logic at hardware speeds. All high-speed time critical input output tasks are performed in FPGA (including the PWM generation), while the control loop runs as a real-time program on the CPU with a much lower cycle time of several milliseconds.

The FPGA is inherently suitable for parallel processing or multi-tasking since each task has a dedicated piece of hardware; they all run rather in series as in a conventional time share multi-tasking system. The FPGA also runs continuously without potential disturbances such as interrupts [109]. This is an important feature in this experiment, because it means that data can be processed in a deterministic manner. For example, the digital control signal for the pump has to be encoded as PWM with a much higher temporal resolution and determinism than the control algorithm, and the FPGA provides a way of achieving this with little effort.

The cRIO-9074 FPGA is programmed in the NI LabVIEW FPGA software, which provides a graphical programming interface based on the LabVIEW environment [110]. This graphical approach is much easier to learn than the conventional way of programming an FPGA in logic languages such as VHDL of Verilog, also referred to as hardware description languages (HDLs). The block diagram below shows the main LabVIEW FPGA functions that are running in parallel:

- Thermocouples readings
- Main PWM: flow rate meter.
- Secondary PWM: water pumps speed and radiator fans speed.
- Other controller: Valve position, pump controller, valve controller, Heater controller, radiator fan controller.

![LabVIEW FPGA block diagram](image)

Figure 9.6: LabVIEW FPGA block diagram for the test rig consists of thermocouple readings (bottom right), main PWM (top right), secondary PWM (bottom left) and other controller (top left).

**Controller**

The MPC controller is implemented in MathWork™ Simulink®. It runs on the PC, and is coupled to the LabVIEW Host VI using the Transmission Control Protocol (TCP). This arrangement does not require a compilation of the control code each time a change is made in Simulink®, which allows significantly faster development. The Simulink® time based is aligned with real time using “Soft Real Time” block available in File Exchange MATLAB® Central [111]. Given the slow thermal time constant of the rig, the PC is more than fast enough to provide an adequate sample rate. The experiment overview with connections from the user interface computer down to the test rig can be seen in Figure 9.7.
9.2. **Controller implementation**

The controller implementation follows the approach explained in CHAPTER 8 (Figure 9.8). It starts with a model linearization, which uses the same thermodynamics structure as in the model in CHAPTER 6. This leads to the same number of variables ($\dot{Q}_{com}$, $\dot{Q}_{conv}$ and $\dot{Q}_{cool}$), and further details of the implementation can be found in CHAPTER 8.

**Feedback Linearization**

**System Characterisation**

As mentioned previously, the model is being divided into two separate sub models: the wall temperature model and the coolant out temperature model. Feedback linearization
requires knowledge of the value of convection heat transfer coefficient \( Ah \), which is not entirely constant over the operating range. The heat transfer coefficient \( Ah \) can be calculated by convection heat transfer \( \dot{Q}_{\text{conv}} \) equation (70) below. At steady state conditions, both the convection rate \( \dot{Q}_{\text{conv}} \) and the disturbance rate \( \dot{Q}_{\text{comb}} \) are equal because heating losses in the block are minimal and can be neglected.

\[
\dot{Q}_{\text{comb}} = \dot{Q}_{\text{conv}} = Ah(T_{w1} - T_{\text{out}})
\]

\[
Ah = f(T_{\text{out}}, \dot{m}_c)
\]

\( \dot{Q}_{\text{comb}} \) = Surface heater heat transfer rate to wall [W]
\( \dot{Q}_{\text{conv}} \) = Convection heat transfer rate from wall to coolant [W]
\( A \) = Heat transfer area [m\(^2\)]
\( h \) = Convection heat transfer coefficient [W/m\(^2\)K]
\( T_{w1} \) = Wall temperature measured at 2mm distance [K]
\( T_{\text{out}} \) = Coolant out temperature [K]
\( \dot{m}_c \) = Coolant mass flow rate through the block [kg/s]

The heat input or disturbance \( \dot{Q}_{\text{comb}} \) can be calculated by the equation (71) below. The \( T_{w1} \) and \( T_{w2} \) are the temperatures measured by thermocouple at 2mm and 8mm from the surface heater. The distance of the first thermocouple is 2mm from the surface heater since the temperature swing in an actual engine effectively disappeared from 2mm to 3mm distance of the combustion surface [112].

\[
\dot{Q}_{\text{comb}} = \frac{Ak}{d} (T_{w1} - T_{w2})
\]

\( \dot{Q}_{\text{comb}} \) = Surface heater heat transfer rate to wall [W]
\( A \) = Heat transfer area [m\(^2\)]
\( k \) = Wall conductivity [W/mK]
\( d \) = Distance between two measured temperature [m]
\( T_{w1} \) = 2mm thermocouple temperature reading [K]
\( T_{w2} \) = 8mm thermocouple temperature reading [K]

The convection heat transfer coefficient \( Ah \) is mapped throughout the coolant mass flow rate and coolant out temperature by using the MathWork™ Model-Based Calibration Toolbox™. The result is as shown in Figure 9.9 below. The model fitting result shows that the heat transfer coefficient is not constant, and it varies mostly with changes in the coolant flow, while coolant temperature has only a minor effect.
Figure 9.9: Experimental result of heat transfer coefficient throughout the coolant flow rate and temperature.

**Pump Signal**

The pump signal determines the pump speed and therefore the coolant flow. In the feedback linearization control scheme, the coolant flow is used to achieve the required heat flow, which is the manipulated variable $\dot{Q}_{\text{conv}}$ in the MPC controller as in equation (70). In the previous simulation model, the coolant flow rate is considered to be proportional to the pump signal and to the valve opening. However, in the experiment, the valve opening influences the overall system pressure head; it is high whenever the valve is at fully close and fully open conditions, which reduces the resulting coolant flow rate. To compensate this characteristic, the pump signal is mapped throughout the required coolant flow rate and current valve position as in Figure 9.10.
The required coolant flow rate is chosen using equation (70) according to the desired heat flow. Therefore, the required coolant flow rate is also mapped throughout varying heat transfer convection and coolant out temperature as shown in Figure 9.11 below.
**Valve Signal**

The valve signal is used by the feedback linearization equation (62) to achieve the desired coolant heat flow $\dot{Q}_{cool}$, which is a manipulated variable of the MPC controller after feedback linearization. The flow rate across the valve depends on the pressure difference between the inlet and the outlet of the valve. In the previous model, the flow characteristic was assumed to be a linear split, but under real condition this is only an approximation. The valve signal depends not just with the desired radiator flow rate, but also on the pump flow rate. The valve signal is mapped throughout the required radiator flow rate and the pump flow rate as shown in Figure 9.12.

![Figure 9.12: Valve signal throughout required radiator flow rate and current pump flow rate.](image)

**Model Fitting**

Wall and coolant temperature are the key variables of the transient behaviour in this test rig. They are critical for the identification process, which is more complicated than a simple first order model that, would provide a poor fit. The wall temperature rise time from a heat input step response (51.03 seconds) is significantly lower than coolant temperature rise
time for the same input step response (325.97 seconds) (Figure 9.13). The coolant temperature rise time is so long, because the rig is completely uncontrolled, whereas a typical cooling system always has a passive thermostat that will maintain the temperature within a narrow band. Another contributing factor is the relative large amount of coolant in this rig compared to the rated power; the maximum heat transfer to coolant is 0.5kW with 750ml of coolant in the test rig. As such, the ratio for a typical engine cooling systems is 0.073L/kW while the test rig ratio is 0.325L/kW, leading to longer time constant. The coolant temperature requires a large and long input step change to capture a step response sufficient for system identification. So the goal of the identification is to separate out the dynamics of the coolant temperature from the wall temperature, and to achieve a more accurate model especially for the wall.

![Graph showing rise time for wall and coolant temperature](image.png)

Figure 9.13: Rise time for the wall and coolant temperature.

The wall temperature model achieves a fit of 84.6%. The result is lower than in the simulation, but this is to be expected due to the presence of noise and heat loses into the environment. However, the dynamics are captured well, and the overall accuracy is considered reasonable.
The coolant temperature model accuracy is very low (at 21.4%), but it is still considered reasonable for the MPC controller, because the differences are mainly in the stationary characteristics, the dynamics are captured reasonably well, and the deviations are reasonably small on an absolute scale with no more than ±3°C (Figure 9.15). The coolant temperature can be measured, and the controller can accommodate any steady differences to the model.

The lower fit of coolant out temperature model can be explained by the existence of nonlinearity in the system, and by the fact that every single system component has an influence on the coolant temperature through heat loss into the environment or heat production. Unmodelled heat loss happens in the components and in the long and thin pipe. The coolant also experiences heat gain from the running pump in the system. Changes of room temperature can also influence the results. The valve position control has limited accuracy: it is subject to a quantisation of 0.01, and the positioning error is of a similar size. This corresponds to only 1% of the total valve total range, which may seem small, but it can have a large impact when the radiator flow rate is very low: a 1% valve position change can increase the radiator flow by 7% of its maximum value (Figure 9.16).
MPC Constraints

One of the key advantages of MPC is that it can deal with system constraints gracefully. In order for this to work, the constraints need to be identified accurately. Two key difficulties are that the constraints need to be linear; they need to be expressed in terms of the linearized inputs and the manipulated variables, not the physical system inputs. This section explains how the constraints of the experiment are found and formalised.
**Water pump constraint**

The pump constraint is identified based on the calculated convection heat transfer using equation (72) for the maximum constraint and equation (73) for the minimum constraint. The maximum heat transfer coefficient \( A_{h_{w(max)}} \) is measured at maximum coolant flow rate and current coolant out temperature \( T_{out} \) while the minimum coefficient \( A_{h_{w(min)}} \) is determined at minimal values. The coolant flow rate is also dependent on the valve position (Figure 9.17), which results in slight variation of the maximum and minimum flow limit. Unfortunately, the valve position is not one of the MPC model variables, and therefore it is not possible to take it into account when determining the current pump constraints. The key equations are:

\[
\dot{Q}_{conv}^{(max)} = A_{h_{w(max)}} \cdot (T_w - T_{out})
\]

And,

\[
\dot{Q}_{conv}^{(min)} = A_{h_{w(min)}} \cdot (T_w - T_{out})
\]
The depended coolant flow rate problem is solved by assuming a constant valve position which is used to calculate the limit. A few considerations need to be taken in determining this position, using an understanding of the function of the system. Generally, the maximum coolant flow rate coupled with low coolant temperature is required at high $\dot{Q}_{comb}$ and vice versa (shown as red area in Figure 9.18). This can be implemented by having the valve at fully open position when low coolant temperature is required to further reduce the wall temperature and vice versa. The flow rate when the valve is fully opened is 0.136 kg/s, slightly below the maximum 0.148 kg/s achieved at medium valve position (Figure 9.18). This difference is 8.11% of the maximum flow rate. The minimum pump signal flow rate at fully closed position is 0.029 kg/s as compared to 0.036 kg/s at the medium position (Figure 9.18). This may seem like a more significant difference, but in absolute terms it is only 0.007 kg/s, which is 4.73% of the maximum flow rate. Therefore, this difference has only a minor impact on the wall temperature, and can be neglected.
Another consideration is to look at the worst-case error between the actual and the modelled limit. These occur for the minimum flow at the maximum pump signal and the maximum flow at the minimum pump signal. The minimum flow with the maximum pump signal happens when the valve is at the minimum position. The flow rate is 0.115 kg/s which are 22.30% reductions from maximum flow rate. The variation is too big to assume a constant limit, because it would make the pump flow rate band quite narrow. It is a rare situation for the valve position and pump signal to end up in that area (shown as red area in Figure 9.18). The minimum pump signal gives a much less variable flow rate compared to the maximum pump signal. The minimum pump signal flow rate difference between at fully closed position and maximum flow rate is not significant i.e. only 0.007 kg/s. From these two considerations, the valve position value for maximum pump signal is at fully opened position while minimum pump signal is at its maximum flow rate.

The calculated upper and lower constraints of $\dot{Q}_{conv}$ are in equations (74) and (75). The constraints are two linear surfaces which are dependent both on the wall temperature and the coolant temperature, but the latter is within a much narrower range, so it is less significant (Figure 9.19).

Upper constraint:

$$\dot{Q}_{conv} \leq -6.07 \cdot T_{out} + 6.125 \cdot T_w - 2.204$$  \hspace{1cm} (74)

$$\dot{Q}_{conv} = \text{Convection heat transfer rate [W]}$$
\[ T_{out} = \text{Coolant engine out temperature [K]} \]
\[ T_{wall} = \text{Cylinder wall temperature [K]} \]

Lower constraint:

\[
\dot{Q}_{conv} \geq 4.828 \cdot T_{out} + 4.965 \cdot T_{w} - 5.499
\]  \hfill (75)

\[
\dot{Q}_{conv} = \text{Convection heat transfer rate [W]} \\
T_{out} = \text{Coolant engine out temperature [K]} \\
T_{wall} = \text{Cylinder wall temperature [K]} \\
\]

Figure 9.19: Upper constraint and lower constraint for \( \dot{Q}_{conv} \).

**Valve Constraint**

The limited range of putting a constraint on the coolant heat transfer \( \dot{Q}_{cool} \) is according to equation (76) below.
\[
\dot{Q}_{\text{cool}} = \dot{m}_{\text{rad}} \cdot c_p \cdot (T_{\text{out}} - T_{\text{rad}})
\]

\[\dot{m}_{\text{rad}} = \text{Coolant mass flow rate through radiator [kg/s]}\]
\[c_p = \text{Coolant specific heat capacity [J/kgK]}\]
\[T_{\text{out}} = \text{Coolant engine out temperature [K]}\]
\[T_{\text{rad}} = \text{Coolant radiator out temperature [K]}\]

The maximum and minimum constraints are determined by calculating \(\dot{Q}_{\text{cool}}\) throughout water pump speed at maximum and minimum valve positions. The radiator flow rate \(\dot{m}_r\) also depends on the valve position and current pump flow rate (Figure 9.20). The radiator coolant out temperature \(T_{\text{rad}}\) is essential for calculating the heat transfer, and it is modelling in dependence of the radiator flow rate \(\dot{m}_r\) and the coolant out temperature \(T_{\text{out}}\) (Figure 9.21). Some of these conditions had to be extrapolated as they could not be achieved in stationary close loop cooling system, especially at the top of the range of \(T_{\text{out}}\) and \(\dot{m}_r\). These conditions can only be achieved at high heater setting which causes the surface heater to exceed the maximum temperature limit.

Figure 9.20: Radiator flow rate throughout valve position and pump coolant flow rate.
As explained in CHAPTER 8, the radiator flow rate $\dot{m}_r$ depends on both the valve position and the pump flow rate $\dot{m}_c$, and therefore it depends indirectly on the $\dot{Q}_{conv}$ value. The $\dot{Q}_{cool}$ constraint can be plotted over $\dot{Q}_{conv}$ and $T_{out}$ (Figure 8.9). The $\dot{Q}_{conv}$ value is determined by the target wall temperature $T'_w$ as equation (77) below:

$$\dot{Q}_{conv} = Ah_w \cdot (T'_w - T_{out})$$  \hspace{1cm} (77)

- $\dot{Q}_{conv}$ = Maximum convection heat transfer rate [W]
- $A$ = Heat transfer rate area [m$^2$]
- $h_w$ = Heat transfer coefficient at maximum flow [W/m$^2$K]
- $T'_w$ = Target wall temperature [K]
- $T_{out}$ = Coolant out temperature [K]

This constraint cannot be handled in the same way as before by using only the target $T'_w$, because $\dot{Q}_{conv}$ covers a rather narrow range on the test rig, and therefore the $\dot{Q}_{cool}$ constraints would become unreasonably tight (Figure 9.22). The narrow range is caused by a narrow pump operating range compared to the simulation and the use of a constant fan speed. Knowing that the transient response of wall temperature $T_w$ is much faster than
coolant temperature $T_{out}$, this may cause the MPC controller to operate outside of the constraint during transients. Operating points are outside the constraint band with a wall temperature deviation as low as 10°C (Figure 9.22).

Figure 9.22: $\dot{Q}_{cool}$ constraints with throughout wall target temperature $T_{wall}$ (solid surface) and $T_{wall} - 10^\circ C$ (faded surface)

The wall temperature is studied between 70°C and 110°C to address the narrow band throughout different wall temperatures (Figure 9.23). The result in Figure 9.23 shows that moving constraints along the wall temperature solves the narrow range constraints. The working condition outside the constraints area should not occur since the constraints are based on current wall and coolant temperature. This clearly proves that the wall temperature is also a factor for consideration for constraints.
The upper $\dot{Q}_{\text{cool}}$ constraint can be represented by a single big constraint without depending on wall temperature $T_w$ (Figure 9.24). The upper constraint does not depend on $T_w$ as it has no impact to the MPC controller. The new imposed constraint is still far from the maximum operating point. The constraint result is as in equation (78).

$$
\dot{Q}_{\text{cool}} \leq -0.009805 \cdot \dot{Q}_{\text{conv}} + 11.09 \cdot T_{\text{out}} - 97.94
$$

\begin{align*}
\dot{Q}_{\text{cool}} &= \text{Removed heat transfer rate at radiator [W]} \\
\dot{Q}_{\text{conv}} &= \text{Convection heat transfer rate [W]} \\
T_{\text{out}} &= \text{Coolant engine out temperature [K]} 
\end{align*}
The lower constraint differs from the upper constraints. The operating point at 20% $\dot{Q}_{comb}$ is very close to the lower constraint (unlike the simulation due to difference in the fan behaviour). It is critical to establish the constraint as precisely as possible, because dynamics of the system are slow at these points due to low flow rates. Using the full operating range helps to improve the warm-up time of the coolant. The lower constraint forms a more complex surface compare to the simulation. It requires two linear constraints to replicate the “L” shape of the actual constraint surface (Figure 9.25). Equation (79) below is the linear equation of the imposed lower constraints:

$$
\dot{Q}_{cool} \geq 53.91 \cdot \dot{Q}_{conv} - 39.31 \cdot T_{wall} + 38.62 \cdot T_{out} + 128.63
$$

$$
\dot{Q}_{cool} \geq 848.54 \cdot \dot{Q}_{conv} - 676.55 \cdot T_{wall} + 204.01 \cdot T_{out} + 175.00
$$

\begin{align*}
\dot{Q}_{cool} &= \text{Removed heat transfer rate at radiator [W]} \\
\dot{Q}_{conv} &= \text{Convection heat transfer rate [W]} \\
T_{out} &= \text{Coolant out temperature [K]} \\
T_{wall} &= \text{Wall temperature [K]}
\end{align*}
MPC Setup

The MPC setup is similar to the MathWork™ Simulink® simulation as shown below:

<table>
<thead>
<tr>
<th>Variable</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sample time, $t_s$</td>
<td>0.5 seconds</td>
</tr>
<tr>
<td>Prediction horizon, $n_p$</td>
<td>50</td>
</tr>
<tr>
<td>Control horizon, $n_c$</td>
<td>Seq. 3 (as in Table 8.2)</td>
</tr>
<tr>
<td>Output weight – $T_w$</td>
<td>10</td>
</tr>
<tr>
<td>Output weight – $T_c$</td>
<td>5 (Calibrated)</td>
</tr>
<tr>
<td>Input rate weight – $\dot{Q}_{conv}$</td>
<td>0.001</td>
</tr>
<tr>
<td>Input rate weight – $\dot{Q}_{cool}$</td>
<td>0.001</td>
</tr>
</tbody>
</table>

A blocked control horizon is used (as long as the prediction horizon) following the same structure as in the Simulink® simulation. The weight is putting the main emphasis on the controller objective to achieve the best performance with variable coolant target.
temperature (Figure 9.26). The prediction horizon is the same length as in the Simulink® simulation, since the transient response of the wall temperature is considered similar. Figure 9.27 below is the comparison of both transient responses from both experiment and simulation in a step response.

![Coolant temperature target throughout wall temperature](image1)

**Figure 9.26:** Coolant temperature target throughout wall temperature.

![Rise time comparison between experiment and simulation](image2)

**Figure 9.27:** Rise time comparison between experiment and simulation at max flow rate, half valve position and step heat input from 20% to 80%.
9.3. **MPC Performance Analysis**

The MPC controller is tested using a random square signal of $\dot{Q}_{\text{comb}}$ disturbance as shown in Figure 9.28. The wall and coolant temperature references are set based on the correlation of $\dot{Q}_{\text{comb}}$ disturbance changes as shown in Figure 9.26.

![Graphs showing random $\dot{Q}_{\text{comb}}$ disturbance and temperature tracking performance](image)

**Figure 9.28:** Random $\dot{Q}_{\text{comb}}$ disturbance and for MPC performance analysis.

**Initial Result**

The MPC controller shows good tracking performance as shown in Figure 9.29. The wall temperature tracking performance is reasonable good and stable, with deviations small and diminishing quickly. The MCP controller also shows that it is able to utilise the known future disturbance and to act accordingly. The coolant temperature tracking is not quite as good, with some serious deviations that diminish slowly. There are two reasons for this:

- coolant temperature has slower transient dynamic characteristic; and
- model fitting for the MPC controller is low (only 21.4%).
It is also noticeable that the fluctuation of coolant out temperature occurs between 180 seconds to 350 seconds, which indicates that the controller is not entirely stable.

![Graphs showing temperature and signal fluctuations](image)

**Figure 9.29:** MPC controller performance tracking target temperature in experiment.

To check the accuracy of the constraints and the impact they have on the control scheme, the limits for the manipulated variable ($\dot{Q}_{\text{conv}}$ and $\dot{Q}_{\text{cool}}$) are added to the plot (Figure 9.30). The $\dot{Q}_{\text{conv}}$ upper and lower constraints represent the water pump actuator limit very well. The water pump signal does not show any sign of being constrained except when it hits the actual pump physical limit. On the other hand, as expected the $\dot{Q}_{\text{cool}}$ constraint does not perform as well, because it uses a linear surface constraint to represent a much more
complex shape. This creates limitation due to the imposed $\dot{Q}_{cool}$ constraints that sometimes fall short of the actual valve actuator limits. Still, the performance of coolant out temperature control is reasonable, the deviation is maintained within ±3°C, and it shows no negative effect on the wall temperature tracking performance. Further improvement to stability can be made by tuning the MPC parameters such as input rate weight and output weight.

![Figure 9.30: MPC controller constraints for manipulated variables.](image_url)
Input Rate Weight

Standard practise to reduce the fluctuation is to make the controller less aggressive by increasing the $\dot{Q}_{\text{cool}}$ input rate weight. Increasing values of 0.001, 0.01, 0.1 and 1 are assessed, while the $\dot{Q}_{\text{conv}}$ input rate weight is kept constant at 0.001 (Figure 9.31). The $\dot{Q}_{\text{cool}}$ input rate weight of 0.1 begins to show a stabilising effect on the coolant out temperature, but the control performance deteriorates at the value of 1. The input rate weight of 0.1 is therefore considered the best trade-off between reduced fluctuations and still a reasonable control performance.

![Graph showing temperature fluctuations with different input rate weights]

Figure 9.31: Comparison of MPC controller input rate weight of 0.001, 0.01, 0.1 and 1.
Coolant Out Temperature Weight Rate

The coolant temperature tracking is not the main concern of the controller, because it is a secondary or intermediate goal. Still, the tracking can be improved by raising the respective output weight (Figure 9.32). As it can be seen, this benefits the overall performance of the MPC controller. The weight is higher than in the Simulink® simulation, which can be explained by the more constrained range here (from 44°C to 47°C) compared to the simulation (80°C to 120°C), which is 13 times bigger. This means a slight coolant temperature error for example of 1°C is not significant in the simulation, but it is in the experiment. Increasing the weight in the experiment should improve MPC controller tracking response despite the more limited range. However, fluctuations of the coolant temperature appear for weights above 10, which can clearly be seen at 600 seconds to 750 seconds. A coolant weight of 10 results in a stable control with slight improvements in coolant target temperature tracking compared to a weight of 5. The changes made to the coolant temperature output weight have no significant effect on the wall target temperature tracking performance.
Future Disturbance Knowledge

The MPC controller can benefit from predictions of future disturbances, and the improvements compared to an assumed constant disturbance can be seen in the experimental results (Figure 9.33). As expected, the wall temperature reacts much faster if the MPC controller has knowledge of future disturbances (Figure 9.34). The error is reduced because control actions can be taken before the disturbance changes, and this reduces the peak deviation, for example from 18°C to 13°C at 775 seconds. However, it is noticeable that the controller without the disturbance prediction shows a more reproducible and
relaxed response; this could be because the controller with prediction is trying too hard to achieve a response that resembles the step change in reference value, relying too much on an approximate model of the system. The performance of coolant temperature tracking is also better without disturbance prediction. This could be explained by the fact that the controller without disturbance prediction relies on feedback rather than feedforward.

Figure 9.33: Comparing MPC controller performance between with and without known future disturbance.
Figure 9.34: Comparing MPC controller performance between with and without known future disturbance during wall temperature warm-up.

9.4. Conclusion

This chapter demonstrates the implementation and performance of new MPC controller strategy using developed test rig. The test rig was designed to be similar to a conventional engine cooling system: removing heat generated in the engine block via the coolant and the radiator. The setup is considered comparable with the actual engine cooling system setup.
The new controller is implemented as mentioned in CHAPTER 8. The main differences are the identified characteristics of the valve and the pump, which are not linear as in simulation. Data of the mass flow rate for both actuators are mapped to ensure the model of the flow rates is accurate. The wall and coolant temperature model for the MPC controller are fitted at 84.6% and 21.4%. The linear surface constraints are also constructed to trace the $\dot{Q}_{\text{conv}}$ and $\dot{Q}_{\text{conv}}$ bound for the actuators constraints.

The MPC controller performance to track the wall temperature target is considered good. The tracking error is less than ±2°C and very stable. Deviations for the coolant temperature (a secondary goal) are slightly larger at ±3°C. A key reason is that the coolant temperature model is difficult to fit, and the dynamic behaviour is slower and more complex than the wall temperature behaviour. Imposing the linear constraint surfaces required by the MPC leads to an incomplete approximation of the actual complex constraint surface of $\dot{Q}_{\text{cool}}$. Tuning of the MPC controller weights improves the tracking performance significantly.
CHAPTER 10  Benefits over a Drive Cycle

The new MPC controller equipped with feedback Linearization offers good wall temperature tracking performance. The robustness has been demonstrated and established in CHAPTER 8, and further experimental validation is shown in CHAPTER 9. This chapter presents the final part of the MPC controller analysis: it demonstrates the application to on-the-road driving conditions, where the engine runs over a wide operating range subject to quick changes in torque. The implementation differences are discussed relative to the previous experiments. The simulation study is based on a GT-SUITE engine, cooling system and vehicle model from CHAPTER 3 that are adopted for this purpose.

10.1. Implementations

The Feedback Linearization MPC controller in MathWork™ Simulink® is connected with the GT-SUITE engine model (Figure 10.1) to demonstrate the controller performance in real-world engine and driving conditions. The GT-SUITE model is the model constructed in CHAPTER 3; consisting of four cylinders with its own wall structure replicating real four cylinders engine wall structure and water jacket.
The implementation of the MPC controller is based on the previous simulation and experiment work, with the following differences, that account for the slightly different setup and the lack of un-modelled disturbances.

Figure 10.1: Engine and cooling system model build in GT-SUITE.

**Additional Implementation Steps**

The main flow of the MPC implementation is the same as that in CHAPTER 9 with two extensions in detail of the model:

- Combustion heat transfer to wall
- Linear mathematical model

*Combustion Heat Transfer to Wall*

Previously, the disturbance of the combustion heat transfer rate $\dot{Q}_{comb}$ to wall is considered to be coming from a direct heat source for both the simulation and experiment, which is proportional to the heater signal controller demand. The $\dot{Q}_{comb}$ in an engine depends on the engine load and speed, and therefore on the driving conditions. Further correction may also need to be applied, since S. Salbrechter et al. (2014) stated that the correction of the heat transfer from the gas combustion is dependent on the coolant temperature [113].
The model in this work uses the mean heat transfer coefficient and gas temperature data from the previous engine model in CHAPTER 3. These in-cylinder gas boundary conditions generate heat and transfer it to the cooling system model via the cylinder structure. The local heat at the cylinder head is measured throughout the engine speed and load as the measured disturbance $\dot{Q}_{comb}$ variable in the MPC controller (Figure 10.2). The coolant temperature during $\dot{Q}_{comb}$ measurement is chosen at the optimum wall target temperatures as shown in Figure 3.13.

![Graph showing local combustion heat transfer rate $Q_{comb}$ at optimized coolant out temperature.](image)

**Figure 10.2:** Local combustion heat transfer rate $\dot{Q}_{comb}$ at optimized coolant out temperature.

**Linear Mathematical Model**

The linear mathematical model for coolant out temperature $T_{out}$ is different compared to both the previous simulation and experiment. The term of convection heat transfer rate $\dot{Q}_{conv}$ and heat transfer rate to the environment $\dot{Q}_{cool}$ are the main terms of the thermal model, and were previously considered sufficient to simulate the $T_{out}$ transient behaviour as in equation (80) below (neglecting other heat losses).

$$C_c \frac{dT_{out}}{dt} = \dot{Q}_{conv} - \dot{Q}_{cool}$$

$$\dot{Q}_{conv} = \text{Convection heat transfer rate [W]}$$
\[ \dot{Q}_{cool} = \text{Radiator heat transfer rate [W]} \]
\[ T_{out} = \text{Coolant out temperature [K]} \]
\[ C_c = \text{Coolant heat capacity [J/K]} \]

The engine model used here is a bit more sophisticated. The local convection heat transfer at the cylinder head \( \dot{Q}_{\text{conv}} \) is not linearly proportionate to the total convection heat from the cylinder wall structure to the engine cooling system \( \dot{Q}_{cool} \). The complex shape of the combustion wall structure creates three dimensions heat flux direction \cite{114}. The oil temperature and friction also contribute to the heat flow complexity. This creates the local convection heat transfer rate \( \dot{Q}_{\text{conv}} \) that is not linearly proportionate to the total removed heat at radiator \( \dot{Q}_{cool} \) during the steady state measurement (Figure 10.3).

![Figure 10.3: Nonlinearly proportional relation between \( \dot{Q}_{\text{conv}} \) and \( \dot{Q}_{cool} \).](image)

The nonlinearly relations between the \( \dot{Q}_{\text{conv}} \) and the \( \dot{Q}_{cool} \) can cause a performance degradation of the MPC controller due to a poor model fit. The engine friction heat, heat flow from the oil and other convection heat transfer could be included in the model to avoid this. However, it will not be effective to define the heat transfer only as a disturbance variable in the MPC controller. The local heat transfer in the cylinder head \( \dot{Q}_{\text{conv}} \) as compared to the total heat transfer is only about 1%. This amount is too small for the MPC controller to use the \( \dot{Q}_{\text{conv}} \) to control the coolant out temperature \( T_{out} \). The assumption of the remaining 99% to be an only disturbance for \( T_{out} \) will not be ideal since the coolant mass flow rate \( \dot{m}_c \) or which indirectly represents \( \dot{Q}_{\text{conv}} \) has a big influence to the coolant out temperature \( T_{out} \).
The approach used to solve this problem start by using a fitting line or regression line of the \( \dot{Q}_{\text{conv}} \) and \( \dot{Q}_{\text{cool}} \) relationship. The regression line generates a linear relationship between the variables (Figure 10.3). The remaining deviation are modelled using a third heat transfer rate \( \dot{Q}_{\text{other}} \) (combining friction heat transfer, heat transfer to oil and others) introduced to balance the equation. The equation shows that the total of \( \dot{Q}_{\text{conv}} \) and \( \dot{Q}_{\text{other}} \) is equal to \( \dot{Q}_{\text{cool}} \) in steady state. Therefore, the coolant temperature \( T_{\text{out}} \) transient behaviour is represented by equation (81) below and \( \dot{Q}_{\text{other}} \) is being mapped throughout both the engine speed and load (Figure 10.4).

\[
C_c \frac{dT_{\text{out}}}{dt} = \dot{Q}_{\text{conv}} - \dot{Q}_{\text{cool}} + \dot{Q}_{\text{other}} 
\]

\( \dot{Q}_{\text{conv}} \) = Convection heat transfer rate [W]
\( \dot{Q}_{\text{cool}} \) = Radiator heat transfer rate [W]
\( \dot{Q}_{\text{other}} \) = Other heat transfer [W]
\( T_{\text{out}} \) = Coolant out temperature [K]
\( C_c \) = Coolant heat capacity [J/K]

Figure 10.4: \( \dot{Q}_{\text{other}} \) throughout the engine speed and load.
Model Fitting

The model fitting for the wall and coolant temperature model is performed separately using the MathWork™ System Identification toolbox™. The wall temperature model is fitted with an accuracy of 80.9% (Figure 10.5). The wall temperature model fit is less accurate compared to the previous chapters, because the unmodelled heat flow are more significant here. The wall structure in the cylinder head causes 2 dimensional heat flux, and it also shows nonlinear behaviour, both reducing the model accuracy during model fitting [107,108].

![Wall temperature model fitting for GT-SUITE engine thermal management.](image)

Figure 10.5: Wall temperature model fitting for GT-SUITE engine thermal management.

The coolant temperature model achieves a model fit of 77.1% (Figure 10.6). Interestingly, the coolant temperature model is better than in the test rig despite the complex cooling circuit system used here in the GT-SUITE model. The separated head and block water jacket do increase the complexity, but they do not cause any significant nonlinearities of the coolant temperature transient behaviour. This shows that the valve discrete movement behaviour in the test rig produces high nonlinearity to the coolant out temperature transient behaviour.
Constraints

The $\dot{Q}_{\text{conv}}$ Constraint

The constraints for both $\dot{Q}_{\text{conv}}$ and $\dot{Q}_{\text{cool}}$ are imposed in the same way as in the previous chapters for the simulation and the experiment. The $\dot{Q}_{\text{conv}}$ limits are determined by equation (57) and (58). It is based on the coolant out temperature $T_{\text{out}}$ and the wall temperature $T_w$ with $A h_{\text{min}}$ for the lower limit and $A h_{\text{max}}$ for the upper limit. The imposed linear lower constraint is similar to the actual lower limit calculated from the equations (the lower surface in Figure 10.7). However, the calculated upper limit creates a notable curved surface (the upper surface in Figure 10.7) which cannot be properly approximated using only one linear surface as an upper linear constraint. Instead, two linear constraint surfaces are created that make for a much better fit for the curved surface.

The upper limit in the GT-SUITE engine model shows a much stronger curvature compared to the previous simulation and experiment. This indicates that the value of $A h_{\text{max}}$ in the GT-SUITE engine model reduces significantly as the coolant temperature reduces, which indicates that the nonlinearity in the GT-SUITE engine model for the wall temperature model is significant. The resulting upper and lower linear surface constraints are as follows:

Upper constraint:
\[ \dot{Q}_{\text{conv}} \leq -503.9 \cdot T_{\text{out}} + 727.5 \cdot T_{\text{wall}} - 2.683 \times 10^4 \]  

\[ \dot{Q}_{\text{conv}} \leq 32.59 \cdot T_{\text{out}} + 548.6 \cdot T_{\text{wall}} - 4.65 \times 10^4 \]

\[ \dot{Q}_{\text{conv}} = \text{Convection heat transfer rate [W]} \]

\[ T_{\text{out}} = \text{Coolant engine out temperature [K]} \]

\[ T_{\text{wall}} = \text{Cylinder wall temperature [K]} \]

Lower constraint:

\[ \dot{Q}_{\text{conv}} \geq -18.95 \cdot T_{\text{out}} + 65.22 \cdot T_{\text{wall}} - 4112 \]

\[ \dot{Q}_{\text{conv}} = \text{Convection heat transfer rate [W]} \]

\[ T_{\text{out}} = \text{Coolant engine out temperature [K]} \]

\[ T_{\text{wall}} = \text{Cylinder wall temperature [K]} \]

Figure 10.7: \( \dot{Q}_{\text{conv}} \) upper and lower linear constraints in the GT-SUITE model simulation.
The $\dot{Q}_{\text{cool}}$ constraint

As noticed before, the $\dot{Q}_{\text{cool}}$ limits in the GT-SUITE engine model create a more complex surface (Figure 10.8). The upper limit is a convex curve that is heavily twisted. It shows very much same characteristics as the upper limit in the previous simulation work (CHAPTER 8). The concave curvature makes it impossible to fit linear constraints with perfect accuracy. Instead, the placement of approximating linear constraints focuses on the critical areas for control as explained in CHAPTER 8.

Figure 10.8: The $\dot{Q}_{\text{cool}}$ upper and lower limits in GT-SUITE simulation.

This leads to three separate linear upper constraints as shown in Figure 10.9. The upper constraint 1 focuses on the low coolant temperature $T_{\text{out}}$ and high convection heat transfer $Q_{\text{conv}}$ while upper constraint 2 focuses on the lower part of $Q_{\text{conv}}$ at the same $T_{\text{out}}$. The upper constraint 3 covers the most of high $T_{\text{out}}$. The largest error of this approximation is in the area of high value of $T_{\text{out}}$ and high $\dot{Q}_{\text{conv}}$ which will limit the maximum cooling the control thinks it can achieve, but this error is unavoidable. The upper $\dot{Q}_{\text{cool}}$ linear constraint equations are:
\[ \dot{Q}_{\text{cool}} \leq 3.215 \cdot \dot{Q}_{\text{conv}} + 1466 \cdot T_{\text{out}} - 1.57e + 05 \]
\[ \dot{Q}_{\text{cool}} \leq 0.7794 \cdot \dot{Q}_{\text{conv}} + 2847 \cdot T_{\text{out}} - 2.247e + 05 \]
\[ \dot{Q}_{\text{cool}} \leq 1.829 \cdot \dot{Q}_{\text{conv}} + 1782 \cdot T_{\text{out}} - 1.659e + 05 \]  

(84)

\[ \dot{Q}_{\text{cool}} = \text{Removed heat transfer rate at radiator [W]} \]
\[ \dot{Q}_{\text{conv}} = \text{Convection heat transfer rate [W]} \]
\[ T_{\text{out}} = \text{Coolant engine out temperature [K]} \]

Figure 10.9: Three linear upper constraints representing the complex surface of the \( \dot{Q}_{\text{cool}} \) upper limits.

The lower \( \dot{Q}_{\text{cool}} \) limit is also considered a complex surface due to the fact that the twisting characteristic is visible especially at high coolant temperature \( T_{\text{out}} \) (Figure 10.10), but for control purposes it is much less sensitive. The lower constraint 2 represents the surface at 120°C \( T_{\text{out}} \) while lower constraint 1 covers the remaining surface. The lower \( \dot{Q}_{\text{cool}} \) constraint equations are:

\[ \dot{Q}_{\text{cool}} \geq 0.1195 \cdot \dot{Q}_{\text{conv}} + 82.52 \cdot T_{\text{out}} - 7433 \]  

(85)
\[ \dot{Q}_{\text{cool}} \geq 1.353 \cdot \dot{Q}_{\text{conv}} + 2766 \cdot T_{\text{out}} - 3.325e + 05 \]
\begin{align*}
\dot{Q}_{\text{cool}} &= \text{Removed heat transfer rate at radiator [W]} \\
\dot{Q}_{\text{conv}} &= \text{Convection heat transfer rate [W]} \\
T_{\text{out}} &= \text{Coolant engine out temperature [K]}
\end{align*}

Figure 10.10: Two linear lower constraints representing the complex surface of the \( \dot{Q}_{\text{cool}} \) lower limits.

\subsection*{10.2. MPC Setup}

\textbf{Initial Setup}

The MPC controller used successfully in the previous Simulink® simulation is being replicated for the GT-SUITE model (Table 10.1).

\begin{table}[h]
\centering
\begin{tabular}{|l|l|}
\hline
\textbf{Variable} & \textbf{Value} \\
\hline
Sample time, \( t_s \) & 0.5 seconds \\
\hline
\end{tabular}
\caption{Feedback Linearization MPC setup for GT-SUITE simulation.}
\end{table}
The coolant reference temperature throughout the engine speed and load in Figure 10.11 below is used. The reference temperature is determined by comparing the MPC controller results at steady state conditions in a few coolant reference temperatures.

The initial setup result shows significant oscillations with both the overshoot and undershoot for both wall and coolant temperature (Figure 10.12). This could potentially create unwanted thermal stress to engine components, and it is caused by the nonlinearity of the model, specifically the variable transport delay of the system caused by changing coolant flow rates. The low coolant flow rate also caused low heat rejection in the radiator, as seen in Figure 10.13 at 195 seconds where the valve opens early as it tries to reduce the
increasing coolant temperature. The opening valve causes the wall and coolant temperature undershoot when the coolant flow rate suddenly increases.

Figure 10.12: MPC result with the initial setup in GT-SUITE simulation.
This is clear that the temperature undershoots can be avoided less aggressive control tuning, which reduces the sudden flow rate changes. This can be done by increasing the $\dot{Q}_{\text{conv}}$ input rate weight: a higher input rate weight will make $\dot{Q}_{\text{conv}}$ less volatile, leading to slower flow rate changes.

As expected, the pump signal is less aggressive with higher $\dot{Q}_{\text{conv}}$ input rate weight. The pump signal tends to increase earlier and be more relaxed as the input rate weight

Figure 10.13: Illustration of MPC inputs and outputs before the temperature undershoot.

$\dot{Q}_{\text{conv}}$ Input Rate Weight
increases. This reduces the sudden flow rate changes. The coolant temperature undershoot is obviously reduced as the input rate weight increases, but the consequence is that the wall temperature rise time also reduces slightly. The input rate weight of 0.06 gives the best trade-off between the coolant temperature undershoot and wall temperature rise time.

Figure 10.14: Comparison of the $\dot{Q}_{\text{conv}}$ input rate weight to the pump signal response.
10.3. Drive Cycle Performance

The new MPC controller with feedback Linearization has been shown to deliver good wall temperature tracking performance and stability in the previous simulation (CHAPTER 8) and the experimental work (CHAPTER 9). However, the results were based on artificial random square signal under very controlled circumstances, and looking at the control performance only. An actual engine runs at a wide operating range and is subject to very quick and strong changes based on real driving conditions. Eight drive cycles are being used to analyse the MPC controller with the feedback linearization performance under realistic conditions, and to assess the gains in terms of fuel economy, the original goal of thermal management. The cycles used here are:

- New European Drive Cycle (NEDC),
- Worldwide harmonized Light vehicles Test Procedures (WLTC),
- US Federal Test Procedure 75kph (FTP75kph),
- US Highway Fuel Economy Driving test Schedule (HWY),
- US Supplemental Federal Test Procedure (US06),
- Artemis Urban Cycle,
- Artemis Rural Road Cycle, and
- Artemis Motorway Cycle.

Comparison to the Conventional Cooling System

New European Drive Cycle

Figure 10.15 shows the comparison between the conventional cooling system and the new MPC performance in the NEDC. As expected, the MPC controller gives higher wall temperature compared to the conventional cooling system as it tracks the wall temperature reference. This is due to the state of the conventional cooling system design (CHAPTER 2). Another expected result is that the average pump speed is slightly lower (13.5%), and the fan speed is much lower (74.0%) with the MPC controller compared to the conventional cooling system. The fuel consumption reduce 1.90% compared to the conventional cooling system (Figure 10.16).
Figure 10.15: MPC controller performance compared to the conventional cooling system in the New European Drive Cycle.
Figure 10.16: Fuel consumed throughout the NEDC by the conventional cooling system compared to MPC and a perfect controller of ideal head temperature.

*US Supplemental Federal Test Procedure*

The results trend also shows the same trend even in the higher volatility of temperature reference as in US06 drive cycle (Figure 10.17). The average temperature is higher and closer to the temperature reference than the conventional cooling system. This creates fuel consumption reduction of 1.49% compared to the conventional cooling system (Figure 10.18). The water pump average speed also shows reduction up to 13.0% and fan up to 64.9%. So in addition to the more efficient combustion and reduced friction, this indicates that the MPC controller can also save energy by reducing the pump and fan power.
Figure 10.17: MPC controller performance compared to the conventional cooling system in the US Supplemental Federal Test Procedure.
Figure 10.18: Fuel consumed throughout the US06 by the conventional cooling system compared to MPC and a perfect controller of ideal head temperature.

**MPC with and without Future Prediction**

*New European Drive Cycle*

The MPC controller performance is compared with and without the future knowledge of disturbance in the NEDC (Figure 10.19). The MPC controller without future knowledge of disturbance shows slightly higher deviations in the wall temperature tracking test. It has higher RMSE result of 11.4°C compared to with the future knowledge of disturbance 9.9°C; this is a 15.0% increase. The average pump speed and valve movement are also higher, indication slightly more energy is being used; it is 9.45% increase in the average pump speed and 84.1% increase in the average valve movement. Interestingly, the fan average speed is reduced up to 26.3% in the MPC controller without the future prediction. This can be explained by the fact that cooling down the coolant is one of the key advantages of the MPC controller with the future prediction, and this requires higher fan speed to improve the control authority of $\dot{Q}_{\text{cool}}$. The fuel consumption still shows sign of reduction (0.08%) but it is not that significant (Figure 10.20).
Figure 10.19: MPC controller performance with, and without the future knowledge of disturbance in NEDC.
Figure 10.20: Fuel consumed throughout the NEDC by the MPC without the known future disturbance compared to MPC with the known future disturbance and a perfect controller of ideal head temperature.

*US Supplemental Federal Test Procedure*

The results trend also shows the same trend even in the higher temperature reference volatility as in US06 drive cycle (Figure 10.21). The temperature tracking performance shows higher RMSE of 23.7°C compared to 19.9°C in the MPC controller with the future prediction; this is 19% increase. This also creates slight fuel efficiency improvement of 0.02% (Figure 10.22). The pump average speed and valve movement average increase up to 8.2% and 143.2% but with 3.3% of fan average speed reduction. This indicates that the prediction of future demand unlocks significant performance improvement in the MPC scheme (even with slight increase in the radiator fan average speed), and therefore a reasonably accurate prediction of the future demand is an important aspect of an optimal cooling system.
Figure 10.21: MPC controller performance with, and without the future knowledge of disturbance in US Supplemental Federal Test Procedure.
Figure 10.22: Fuel consumed throughout the US06 by the MPC without the known future disturbance compared to MPC with the known future disturbance and a perfect controller of ideal head temperature.

**Overall Results**

Figure 10.23, Figure 10.25, Figure 10.26 and Figure 10.27 are the wall temperature RMSE, average pump speed signal, average valve movement and average fan speed signal results that have been compared with the conventional, MPC with and without the future knowledge of disturbance in all 8 drive cycles. The drive cycles are sorted from high to low wall target temperature autocorrelation based on the result established in CHAPTER 4.

**Wall Temperature Tracking Performance**

The temperature error significantly reduces with the MPC controller in all drive cycles (average of 44.2% drop). The error slightly increases with the MPC controller without future disturbance knowledge (average of 20.5% increase), but it is still better than the conventional controller. As expected, the NEDC shows the highest improvement of all since it has the most constant temperature reference in a long time. The drive cycles sorted by the CHAPTER 4 autocorrelation result do not show any connection to the MPC controller performance to track the reference temperature. This is probably due to the fact that the statistical information could not really give an accurate interpretation of the temperature reference and disturbance behaviour in a drive cycle.
Fuel Consumption

The fuel consumption predicted in CHAPTER 3 is based on a static model with no delay whatsoever, and obviously in a dynamic system it is not possible to achieve the same improvements. But the simulations here show that the MPC controller manages to achieve an average of 1.55% fuel economy improvement over the tested cycles, assuming knowledge of future disturbance. The MPC controller without the future knowledge disturbance is achieves nearly the same result, with an improvement of 1.51%. The US Highway Fuel Economy Driving test Schedule and Artemis Motorway Cycle is two the closest fuel consumption reduction compared to the predicted result in CHAPTER 3 (14.2% and 18.7% lower than the predicted result). The Artemis Rural Road Cycle and the Artemis Urban Cycle is the two farthest fuel consumption reduction compared to the predicted result (32.9% and 31.5% lower than the predicted result).
Figure 10.24: Fuel consumption reduction full potential reduction, MPC with and without the future knowledge of disturbance of compared to conventional cooling system.

**Water Pump Average Speed**

The water pump average speed, valve average movement and fan average speed throughout a drive cycle can give an indication of the power required to run the actuators, and this is not included in the results above, because the influence depends heavily on the details, such as the auxiliary loads and the alternator efficiency. Therefore, the electricity consumption is analysed separately: a lower reading is an indication that less power is required, which suggest reduce auxiliary losses and therefore a potentially more efficient engine.

It is clear that the MPC controller with future disturbance knowledge runs the lowest water pump average speed. This is a 16.1% overall reduction compared to the conventional cooling system. The MPC controller without anticipation also runs at a lower speed with a 4.2% overall reduction compared to the conventional cooling system. However, in the WLTC, Artemis Rural Road Cycle and Artemis Urban Cycle, it can cause higher water pump average speed if the MPC controller runs without future disturbance knowledge. These means the water pump in the MPC controller without anticipation features can be more aggressive to control the fast changing wall temperature reference in certain driving conditions. The MPC weight tuning might reduce this behaviour.
Valve Average Movement

The valve average movement reduces by almost half (45.6%) when the MPC controller runs with, as compared to without the future disturbance knowledge. This indicates that the MPC controller with the anticipation feature can reduce the power in the valve actuator movement. The average valve movement is only compared between MPC with, and without the future disturbance knowledge, since the wax thermostat in the conventional cooling system does not require any additional power to operate it. For future research, it would be interesting to analyse the effect of more realistic prediction that is often correct, but not always, because a miss-prediction could cause additionally actuator movements.
Figure 10.26: Average valve movement comparison between the conventional cooling system, MPC with and without the future knowledge of disturbance.

**Fan Average Speed**

Again, the average fan speed is also lower with the MPC controller (60.9% reduction) compared to conventional cooling system with high speed driving style (Artemis Motorway and US06 Drive Cycle) give the highest reduction in the fan average speed. This can simply be because the radiator works more efficient in the MPC controller by running higher average coolant temperature than conventional cooling system when feasible. There is very little influence of the future prediction on the average fan speed.

Figure 10.27: Fan average speed comparison between the conventional cooling system, MPC with and without the future knowledge of disturbance.
10.4. Conclusion

This chapter demonstrates the MPC controller implementation in the GT-SUITE cooling system model and its performance under actual driving conditions. The implementation in the GT-SUITE is closely representative of an actual engine cooling system, and it contains a number of relevant effects. The implantation is more complex than previous models in that the heat combustion $Q_{comb}$ is dependent on both the engine load and speed. A heat correction term depending on engine speed and load is required for the coolant out temperature model to include the heat generated from the frictions, the head from the engine oil and imbalanced heat transfer due to the complex engine wall structure.

The performance of the MPC controller is not quite as good as in the previous simulation and experiment, but they are more realistic and representative of a real engine. Reasons for the deterioration in performance include the thermal inertia of the wall structure, and the complex non-linear dynamics of the GT-SUITE model. Deviation between the simulation and control model mean that the wall temperature weight output needs to be increased significantly, which makes the controller less aggressive. This improves stability and reduces oscillations, but it also costs some performance.

Overall, this chapter has demonstrated that the MPC controller achieves good wall temperature tracking performance in actual driving conditions, which leads to significant fuel economy saving. The auxiliary losses are also reduced, because average pump and fan speeds are significantly lower with MPC controller. To achieve the best results, the MPC controller requires advanced knowledge of future disturbances, which can be an area of future research. But even without this prediction, the improvements of thermal management are significant and cost effective.
CHAPTER 11  Conclusion and Future Work

11.1. Summary and Conclusion

This thesis presents a thermal management strategy for internal combustion engines using two inputs (coolant flow and temperature) MPC controller with a novel feedback linearization approach to solve the nonlinearity issue. The nonlinear behaviours are from the variable heat transfer coefficient between the combustion wall and the coolant; variable heat transfer rate in the radiator and the variable transport delays. The usage of both the coolant flow and temperature as its control inputs in the same time offers better response and wider control range, but it also amplifies the nonlinearity.

The beginning of the thesis considered how thermal management can be used to improve engine thermal efficiency using a steady state approach. The engine thermal management is optimized by calibrating the cylinder wall temperature throughout the engine speed and load to give the best engine output trade-off.

In addition, eight legislative and academic drive cycles were selected to evaluate this wall temperature set points volatility in actual driving conditions. The results indicate that the wall temperature in urban style driving is the most volatile; it required a responsive and anticipative thermal management controller as it tended to be random after 5.4 seconds as in the autocorrelation result and estimated 0.09Hz of the corner frequency in the PSD result. This is a clear indication of the requirement to use both the coolant flow and temperature as it gives more range as well as better response. The MPC controller should be the control strategy as it is known to be good at handling the MIMO system which can also anticipate and exploit future disturbances.

The feedback linearization approach for the MPC was being used to solve the nonlinearity issue in the system. The feedback linearization method goes back to the first principle of
thermodynamics, which is a linear equation. The system is divided into two parts; the wall temperature model and coolant temperature model. The wall temperature model dynamic behaviour is based on the combustion heat $Q_{\text{comb}}$ and the wall to coolant convection heat transfer rate $Q_{\text{conv}}$ interaction, while the coolant temperature model is based on the $Q_{\text{conv}}$ and the removed heat transfer rate in radiator $Q_{\text{cool}}$. This creates an integral system behaviour which will cause the temperatures to keep decreasing or increasing when there is an imbalance to the control input and disturbance, so it is not inherently stable. An MPC controller with a control blocking sequence is used to find a viable and stable solution without excessive computational complexity.

The feedback linearization approach can linearize the engine thermal management system with good model fitting results; 97.4% for the wall temperature model and 81.5% for the coolant temperature model in the MathWork™ Simulink® simulation. The transformation creates nonlinear constraints, and this applies especially to the $Q_{\text{cool}}$ upper limit, which has a twisting surface characteristic that is difficult to approximate using linear constraints without sacrificing some parts of the operating range. Multiple linear constraints are being used produce an appropriate approximation.

The coolant temperature set points also plays a significant role in affecting both the water pump and fan speed. A high coolant temperature can reduce the fan speed but will increase the water pump speed slightly, and vice versa. Optimizing the coolant temperature set points will result in a lower overall water pump and fan speed, and in return more efficient engine output from the lower water pump and fan power consumption. It is shown in simulation that at some points the fan speed can be reduced up to 59.7% with only 0.66% increase of the water pump speed. This figure predicts significant auxiliary power saving, especially since the fan has a higher power rating than the water pump.

Applying the Feedback Linearization MPC controller to a scaled test rig confirms the same characteristic already seen in simulation. The wall temperature tracking performance is good and stable but with just slight fluctuations in the coolant temperature since the coolant temperature model only fits at 21.4%. There are a number of reasons for this, but the effect on the control performance is reasonably limited.
Finally, the feedback linearization approach is applied to eight legislative and academic drive cycles, using a full engine model in GT-SUITE. The results show that the wall temperature set point volatility is very high for the MPC controller, but the tracking performance is still acceptable and an average of 1.55% fuel consumption reduction. In addition, further reductions in the average water pump and fan speed were recorded in all drive cycles compared to the conventional cooling system. This indicates that the MPC controller can reduce the engine fuel consumption not just by controlling the cylinder wall temperature, but it can also reduce the auxiliary power consumption of the cooling system.

The MPC controller wall temperature tracking performance deteriorates when the MPC controller could not anticipate the future disturbance. The wall temperature tracking performance drops by 18.5% compared to the MPC controller with the future disturbance knowledge. The pump average speed and valve movement average increased by 14.2% and 45.6% on average. However, the average fan speed does not give a consistent trend in which some of the drive cycles have fans with lower speed.

11.2. Future Work

The main interesting extension of this work should be to develop this new engine thermal management as a commercial solution for road vehicles. At the same time, there are also further academic questions related to thermal management that can be investigated, and some of these may be useful for the commercialisation:

- The optimized coolant temperature was done in the GT-SUITE engine model with natural aspirated 2.0L base engine. Optimization temperature studies on real engine with different configurations such as turbocharged engine, diesel engine or hybrid engine should give information on the challenges to the proposed engine thermal management. The information should also benefit by means of optimising the water pump and radiator fan speed.
- The autonomous vehicle is a good platform for this engine thermal management to have a good future disturbance prediction. The autonomous vehicle uses the route
pre-set by the passenger and absorbs all information of the vehicle environment before the vehicle reacts to the environment. The vehicle is also equipped with the required sensors and high capability computer to implement the new controller. Any related future works relevant to this will provide good demonstration and input for the study of the engine thermal management, since they will use the processed information prior to the action as future disturbance prediction from real vehicles.

- Another factor that should be considered in future studies is the degradation of the engine component as the vehicle reaches high mileage and age to the controller’s performance. The MPC controller performance relies on the accuracy of the model. For an example; the rust build-up inside the water jacket will alter the heat transfer between the water jacket and engine wall. The carbon build up inside the combustion chamber will also affect the heat transfer from the combustion gas to the engine wall.

- Engine components durability after the implementation of the controller should be one of the aspects of the future studies before the implementation on-the-road vehicle. The new controller runs with variable coolant temperature and higher average wall temperature as compared to the conventional cooling system.
Reference


78. Su, H., “A Road Load Data Processing Technique for Durability Optimization of


100. Bello, O., Hamam, Y., and Djouani, K., “Multiple Model Predictive Control Based on Fuzzy Switching Scheme of a Coagulation Chemical Dosing Unit for Water Treatment Plants**This work was supported by Tshwane University of Technology, Pretoria, South Africa.,” IFAC-PapersOnLine 48(11):180–185, 2015, doi:10.1016/j.ifacol.2015.09.180.


A.1. **Engine Gasoline Natural Aspirated 2.0L engine list**

A list of 2.0L natural aspirated gasoline engines is being made as an indicator to determine GT-SUITE engine model’s specification. The engines detail specifications are being referred in various websites such as automaker official websites, Wikipedia.org and Carfolio.com. The specific criteria are taken into account are the engine should be:

- natural aspirated,
- using port fuel injector, and
- displacement from 1950cc to 2040cc.

The engine list and its detail are as in table below:

<table>
<thead>
<tr>
<th>Maker</th>
<th>Year</th>
<th>Engine Name</th>
<th>Displacement (cc)</th>
<th>bore</th>
<th>stroke</th>
<th>Compression ratio</th>
<th>No. of valve</th>
<th>Power (kW) @ RPM</th>
<th>torque (Nm) @ RPM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Audi &amp; VW</td>
<td>2.0 R4</td>
<td>1984</td>
<td>82.5</td>
<td>92.8</td>
<td>8.9</td>
<td>4</td>
<td>110</td>
<td>6000</td>
<td>180</td>
</tr>
<tr>
<td>Ford</td>
<td>1992-2004</td>
<td>Zetec-E</td>
<td>84.8</td>
<td>88</td>
<td>10</td>
<td>4</td>
<td>97</td>
<td>5750</td>
<td>174</td>
</tr>
<tr>
<td>Manufacturer</td>
<td>Models</td>
<td>Years</td>
<td>VIN Code</td>
<td>Brake Horsepower</td>
<td>Torque</td>
<td>Year</td>
<td>RPM</td>
<td>Horsepower</td>
<td>Torque</td>
</tr>
<tr>
<td>-----------------</td>
<td>-------------------------</td>
<td>---------------</td>
<td>----------</td>
<td>------------------</td>
<td>--------</td>
<td>-------</td>
<td>-----</td>
<td>-------------</td>
<td>--------</td>
</tr>
<tr>
<td>General Motors</td>
<td>20XE</td>
<td>1998</td>
<td>86</td>
<td>86</td>
<td>10.8</td>
<td>4</td>
<td>117</td>
<td>198</td>
<td></td>
</tr>
<tr>
<td>General Motors</td>
<td>C20XE</td>
<td>1998</td>
<td>86</td>
<td>86</td>
<td>10.8</td>
<td>4</td>
<td>112</td>
<td>196</td>
<td></td>
</tr>
<tr>
<td>General Motors</td>
<td>B204i</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>10.1</td>
<td>97</td>
<td>5500</td>
</tr>
<tr>
<td>General Motors</td>
<td>B026</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>99</td>
<td></td>
</tr>
<tr>
<td>Honda</td>
<td>A20A4</td>
<td>1955</td>
<td>82.7</td>
<td>91</td>
<td>3</td>
<td>89</td>
<td>5500</td>
<td>165</td>
<td>4000</td>
</tr>
<tr>
<td>Honda</td>
<td>B20B</td>
<td>1973</td>
<td>84</td>
<td>89</td>
<td></td>
<td>94</td>
<td>5400</td>
<td>180</td>
<td>4800</td>
</tr>
<tr>
<td>Honda</td>
<td>B20Z</td>
<td>1973</td>
<td>84</td>
<td>89</td>
<td></td>
<td>107</td>
<td>6200</td>
<td>180</td>
<td>5600</td>
</tr>
<tr>
<td>Honda</td>
<td>B20A1</td>
<td>1958</td>
<td>81</td>
<td>95</td>
<td>9.4</td>
<td>4</td>
<td>119</td>
<td>6300</td>
<td>186</td>
</tr>
<tr>
<td>Honda</td>
<td>B20A5</td>
<td>1958</td>
<td>81</td>
<td>85</td>
<td>9</td>
<td>4</td>
<td>101</td>
<td>6200</td>
<td>172</td>
</tr>
<tr>
<td>Honda</td>
<td>B20A9</td>
<td>1958</td>
<td>81</td>
<td>85</td>
<td>9.3</td>
<td>4</td>
<td>104</td>
<td>6000</td>
<td>175</td>
</tr>
<tr>
<td>Honda</td>
<td>K20Z2</td>
<td>1998</td>
<td>86</td>
<td>86</td>
<td>9.8</td>
<td>4</td>
<td>114</td>
<td>6000</td>
<td>188</td>
</tr>
<tr>
<td>Honda</td>
<td>K20A2</td>
<td>1998</td>
<td>86</td>
<td>86</td>
<td>11</td>
<td>4</td>
<td>147</td>
<td>7400</td>
<td>193</td>
</tr>
<tr>
<td>Honda</td>
<td>F20C</td>
<td>1997</td>
<td>87</td>
<td>84</td>
<td>11.7</td>
<td>4</td>
<td>179</td>
<td>8300</td>
<td>207</td>
</tr>
<tr>
<td>Honda</td>
<td>F20B5</td>
<td>1997</td>
<td>85</td>
<td>88</td>
<td>11.1</td>
<td>4</td>
<td>118</td>
<td>7000</td>
<td>137</td>
</tr>
<tr>
<td>Hyundai</td>
<td>Beta G4GF (2.0D)</td>
<td>1975</td>
<td>82</td>
<td>93.5</td>
<td></td>
<td>4</td>
<td>102</td>
<td>6000</td>
<td>180</td>
</tr>
<tr>
<td>Hyundai</td>
<td>Theta 2.0L</td>
<td>1997</td>
<td>86</td>
<td></td>
<td>10.5</td>
<td>4</td>
<td>107</td>
<td>6000</td>
<td>190</td>
</tr>
<tr>
<td>Mazda</td>
<td>FS-DE</td>
<td>1991</td>
<td>83</td>
<td>92</td>
<td>9.1</td>
<td>4</td>
<td>97</td>
<td>6000</td>
<td>183</td>
</tr>
<tr>
<td>Mazda</td>
<td>MZR</td>
<td>1999</td>
<td>87.5</td>
<td>83.1</td>
<td>10</td>
<td>4</td>
<td>110</td>
<td>6500</td>
<td>187</td>
</tr>
<tr>
<td>Mercedes-Benz</td>
<td>M102.961 W201 190E</td>
<td>1997</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>90</td>
<td></td>
</tr>
<tr>
<td>Mercedes-Benz</td>
<td>M111.940</td>
<td>1998</td>
<td>89.9</td>
<td>78.7</td>
<td>9.6</td>
<td>4</td>
<td>100</td>
<td>190</td>
<td></td>
</tr>
<tr>
<td>Mitsubishi</td>
<td>4G63</td>
<td>1997</td>
<td>85</td>
<td>88</td>
<td></td>
<td></td>
<td></td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>Nissan</td>
<td>FJ20E</td>
<td>1990</td>
<td>89</td>
<td>80</td>
<td>9.1</td>
<td>4</td>
<td>110</td>
<td>6000</td>
<td>181</td>
</tr>
<tr>
<td>Nissan</td>
<td>SR20DE</td>
<td>1998</td>
<td>86</td>
<td>86</td>
<td>9.5</td>
<td>4</td>
<td>104</td>
<td>6400</td>
<td>179</td>
</tr>
<tr>
<td>Nissan</td>
<td>QR20DE</td>
<td>1997</td>
<td>89</td>
<td>80.3</td>
<td>10</td>
<td>4</td>
<td>110</td>
<td>6400</td>
<td>200</td>
</tr>
<tr>
<td>Nissan</td>
<td>MR20DE</td>
<td>1997</td>
<td>84</td>
<td>90.1</td>
<td>10</td>
<td>4</td>
<td>104</td>
<td>5100</td>
<td>193</td>
</tr>
<tr>
<td>Peugeot (PSA)</td>
<td>EW10 J4(RFR)</td>
<td>1997</td>
<td>85</td>
<td>88</td>
<td></td>
<td>4</td>
<td>105</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Peugeot (PSA)</td>
<td>EW10 J4(RFN)</td>
<td>1997</td>
<td>85</td>
<td>88</td>
<td></td>
<td>4</td>
<td>100</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Renault</td>
<td>F4R 2.0L</td>
<td>1998</td>
<td>82.7</td>
<td>93</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Renault</td>
<td>F7x</td>
<td>1998</td>
<td>82.7</td>
<td>93</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Subaru</td>
<td>EJ20D</td>
<td>1994</td>
<td>92</td>
<td>75</td>
<td></td>
<td>4</td>
<td>110</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Toyota</td>
<td>1RZ</td>
<td>1998</td>
<td>86</td>
<td>86</td>
<td>9</td>
<td>79</td>
<td>5400</td>
<td>167</td>
<td>2800</td>
</tr>
</tbody>
</table>
The GT-SUITE engine model’s bore and stroke specification is selected based on the highest frequency engine bore and stroke from table above as it gives more possible data reference for other specification.

Figure A.1: Engine bore and stroke frequency based on Table A.1.
A.2. **Conventional Cooling system**

The fan strategy in conventional cooling system starts to run at 90.5°C coolant temperature and the fan runs at maximum when the coolant temperature is above 93.5°C. The fan controller strategy has a 2°C deadband and a first order transfer function to make a smooth transition. The strategy in MathWork™ Simulink® is as below:

![Fan strategy diagram](image)

Figure A.2: Fan strategy in the GT-SUITE cooling system model.

A.3. **Cylinder Wall Temperature Throughout Drive Cycle**

Figure A.3 to Figure A.10 are calibration and actual head temperature throughout the drive cycles.
Figure A.3: Calibrated head temperature (red), actual head temperature (blue) and vehicle speed (green) throughout the NEDC.

Figure A.4: Calibrated head temperature (red), actual head temperature (blue) and vehicle speed (green) throughout the WLTC.
Figure A.5: Calibrated head temperature (red), actual head temperature (blue) and vehicle speed (green) throughout the Artemis Urban Cycle.

Figure A.6: Calibrated head temperature (red), actual head temperature (blue) and vehicle speed (green) throughout the Artemis Rural Road Cycle.
Figure A.7: Calibrated head temperature (red), actual head temperature (blue) and vehicle speed (green) throughout the Artemis Motorway Cycle.

Figure A.8: Calibrated head temperature (red), actual head temperature (blue) and vehicle speed (green) throughout the FTP-75kph.
Figure A.9: Calibrated head temperature (red), actual head temperature (blue) and vehicle speed (green) throughout the US06.

Figure A.10: Calibrated head temperature (red), actual head temperature (blue) and vehicle speed (green) throughout the HWY.
A.4. **Fuel Consumption Comparison Calculation in Test Cycle**

The fuel consumption during test cycles is calculated as in Figure A.11 below. It requires the current engine speed, load and cylinder wall temperature to generate the current BSFC output. The fuel flow rate can be calculated from the BSFC output.

The BSFC is modelled using MathWork™ Model-Based Calibration (MBC) Toolbox™ with the three inputs (Head temperature, engine speed and engine load). The model is a Gaussian Process Model with 0.607 PRESS RMSE and 0.458 RMSE. The model throughout engine speed and load at 170°C, 200°C and 230°C are as in Figure A.12, Figure A.13 and Figure A.14 below.

![Diagram](image)

Figure A.11: Total fuel consumed throughout the test cycle calculation.

\[
\text{Fuel flow } [g/s] = \frac{\text{BSFC} \times \text{Engine Power} [kW]}{60 \times 60}
\]
Figure A.12: BSFC throughout engine speed and load at 170°C.

Figure A.13: BSFC throughout engine speed and load at 200°C.
The cylinder wall temperature for conventional cooling system is simulated in GT-SUITE with full cooling system layout, while the optimized cylinder wall temperature is based on steady state temperature data based on current engine speed and load.

### A.5. Engine Conditions Time Spend

Figure A.15 to Figure A.22 below are the results of engine condition time spent in each cycle for the use of simulating current optimized cylinder wall temperature:
Figure A.15: Time spend of engine speed and load in Federal Test Procedure 75kph Test Cycle.

Figure A.16: Time spend of engine speed and load in US06 Supplemental Federal Test Procedure.
Figure A.17: Time spend of engine speed and load in The Highway Fuel Economy Test Cycle.

Figure A.18: Time spend of engine speed and load in New European Driving Cycle.
Figure A.19: Time spend of engine speed and load in Worldwide harmonized Light vehicles Test Cycle.

Figure A.20: Time spend of engine speed and load in Artemis Urban Test Cycle.
Figure A.21: Time spend of engine speed and load in Artemis Rural Road Test Cycle.

Figure A.22: Time spend of engine speed and load in Artemis Motorway Test Cycle.
B.1. **Engine Cooling System Model in MathWork™ Simulink®**

The engine cooling system model for MPC development is as Figure B.1 below. It consists of wall, water jacket, coolant mixture in thermostat and radiator heat transfer models. The specification is as in Table B. below.

![Engine Cooling System Model](image)

*Figure B.1: Engine cooling system model in the MathWork™ Simulink®.*
Table B.1: Specification of Simulink® engine cooling system model parameter.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wall dimension</td>
<td></td>
</tr>
<tr>
<td>Wall heat specific</td>
<td>949 J/kgK</td>
</tr>
<tr>
<td>Wall density</td>
<td>2702 kg/m³</td>
</tr>
<tr>
<td>Area</td>
<td>0.0025 m²</td>
</tr>
<tr>
<td>Thickness</td>
<td>0.01 m</td>
</tr>
<tr>
<td>Water Jacket</td>
<td></td>
</tr>
<tr>
<td>Coolant heat specific</td>
<td>3548 J/kgK</td>
</tr>
<tr>
<td>Coolant density</td>
<td>1113.2 kg/m³</td>
</tr>
<tr>
<td>Area</td>
<td>0.0025 m²</td>
</tr>
<tr>
<td>Thickness</td>
<td>0.015 m</td>
</tr>
<tr>
<td>Heat ratio to wall</td>
<td>20</td>
</tr>
<tr>
<td>Bypass pipe</td>
<td></td>
</tr>
<tr>
<td>Pipe radius</td>
<td>0.01 m</td>
</tr>
<tr>
<td>Coolant density</td>
<td>1113.2 kg/m³</td>
</tr>
<tr>
<td>Pipe length</td>
<td>1 m</td>
</tr>
<tr>
<td>Pipe to radiator</td>
<td></td>
</tr>
<tr>
<td>Pipe radius</td>
<td>0.015 m</td>
</tr>
<tr>
<td>Coolant density</td>
<td>1113.2 kg/m³</td>
</tr>
<tr>
<td>Pipe length</td>
<td>1 m</td>
</tr>
<tr>
<td>Pipe from radiator</td>
<td></td>
</tr>
<tr>
<td>Pipe radius</td>
<td>0.015 m</td>
</tr>
<tr>
<td>Coolant density</td>
<td>1113.2 kg/m³</td>
</tr>
<tr>
<td>Pipe length</td>
<td>1 m</td>
</tr>
<tr>
<td>Radiator</td>
<td></td>
</tr>
<tr>
<td>Coolant heat specific</td>
<td>3548 J/kgK</td>
</tr>
<tr>
<td>Coolant density</td>
<td>1113.2 kg/m³</td>
</tr>
<tr>
<td>Coolant volume</td>
<td>0.056 m³</td>
</tr>
</tbody>
</table>

B.2. Linear Model Estimation Results

Four steady state models and five transfer function models fitting results are being compared as in Figure B.2 and Figure B.3. This fitting is based on inputs and outputs data in Figure 7.2 as a linear model of the cooling system model in Appendix B.1. The State Space model “SS 4” is the best fit compared to others.
Figure B.2: State Space models fitting result.

Figure B.3: Transfer function models fitting result.
B.3. **The System Identification Model Behaviour**

Figure B.4 to Figure B.6 are the step response and Bode plot of the “SS 4” state space model.

![Step response and Bode plot](image)

*Figure B.4: Step response result from Figure B.5: Bode Plot result of wall temperature.*
B.4. Jacobian Linearization

The linearizations are being made at predetermined equilibrium points for the Jacobian Linearization. The list of equilibrium points are as in Table B.2 and Table B.3. The linearizations and the MPC setup are made as Matlab® coding below:

```matlab
%% Initial Setting

% Opening Simulink
open_system('ModelSYSID.slx')
model = 'ModelSYSID';
% Initial setup linearization
io(1)=linio('ModelSYSID/Heat',1);
io(2)=linio('ModelSYSID/Pump',1);
io(3)=linio('ModelSYSID/Valve',1);
io(4)=linio('ModelSYSID/Mux1',1,'openoutput');
op = findop(model,5000);
% Setup MPC
Ts=0.5;
p=50;
m=10;
Weights=struct('ManipulatedVariables',[0 0],...
'ManipulatedVariablesRate',[0.00001 0.00001],...
'OutputVariables',[10 0.1]);
```

Figure B.6: Bode Plot result of coolant out temperature.
MV(1)=struct('Min',0,'Max',1,'RateMin',-Inf,'RateMax',Inf,'Target',0,'MinECR',0,'MaxECR',0);
MV(2)=struct('Min',0,'Max',1,'RateMin',-Inf,'RateMax',Inf,'Target',0,'MinECR',0,'MaxECR',0);
OV(1)=struct('Min',150,'Max',235,'MinECR',0.001,'MaxECR',0.001);
OV(2)=struct('Min',50,'Max',120,'MinECR',0.001,'MaxECR',0.001);

%% Linearization and MPC Builder
DIVISION=3;%% Division selection (2 or 3)
sigALL=linspace(0,1,DIVISION*2+1);
Cool_targetlin=linspace(70,120,DIVISION*2+1);
XXX=zeros(8,3);
i=0;
% Generate Linearization and MPC
for Heatsig=1:DIVISION;
    for Pumpsig=1:DIVISION;
        for Valsig=1:DIVISION;
            Heat=sigALL(2*Heatsig);
Pump=sigALL(2*Pumpsig);
            Cool_target=Cool_targetlin(2*Valsig);
            options = optimoptions('fmincon','FunctionTolerance',0.01,...
                'Display','off','StepTolerance',0.0025);
            [x,fval]=fmincon(@(Valve)Coolout_target(Valve,Heat,Pump,...
                Cool_target),0.25,[1;-1],[1;0.001],[],[],[],[],[],options);
i=i+1
            set_param('ModelSYSID/Heat','Value',num2str(sigALL(2*Heatsig)));
            set_param('ModelSYSID/Pump','Value',num2str(sigALL(2*Pumpsig)));
            set_param('ModelSYSID/Valve','Value',num2str(x));
            cp = findop(model,5000);
            modelnew= mylinearize(io,cp);
            XXX(i,1:3)=[sigALL(2*Heatsig),sigALL(2*Pumpsig),x];
            set(modelnew,'InputGroup', struct('MV',2:3,'MD',1))
            eval(['MPCobj' num2str(i) ...
                '=mpc(modelnew,Ts,p,m,Weights,MV,OV);']);
        end
    end
end

Table B.2: Equilibrium points list on 8 sub-regions configuration.

<table>
<thead>
<tr>
<th>No.</th>
<th>Heat signal</th>
<th>Pump signal</th>
<th>Valve signal</th>
<th>Fan</th>
<th>Coolant temp</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.25</td>
<td>0.25</td>
<td>0.500</td>
<td>0.1</td>
<td>80</td>
</tr>
<tr>
<td>2</td>
<td>0.25</td>
<td>0.25</td>
<td>0.250</td>
<td>0.1</td>
<td>110</td>
</tr>
<tr>
<td>3</td>
<td>0.25</td>
<td>0.75</td>
<td>0.165</td>
<td>0.1</td>
<td>80</td>
</tr>
<tr>
<td>4</td>
<td>0.25</td>
<td>0.75</td>
<td>0.084</td>
<td>0.1</td>
<td>110</td>
</tr>
<tr>
<td>5</td>
<td>0.75</td>
<td>0.25</td>
<td>1.000</td>
<td>1.0</td>
<td>97</td>
</tr>
<tr>
<td>6</td>
<td>0.75</td>
<td>0.25</td>
<td>0.950</td>
<td>0.1</td>
<td>110</td>
</tr>
<tr>
<td>No.</td>
<td>Heat signal</td>
<td>Pump signal</td>
<td>Valve signal</td>
<td>Fan</td>
<td>Coolant temp</td>
</tr>
<tr>
<td>-----</td>
<td>-------------</td>
<td>-------------</td>
<td>--------------</td>
<td>-----</td>
<td>--------------</td>
</tr>
<tr>
<td>7</td>
<td>0.75</td>
<td>0.75</td>
<td>0.900</td>
<td>0.1</td>
<td>80</td>
</tr>
<tr>
<td>8</td>
<td>0.75</td>
<td>0.75</td>
<td>0.315</td>
<td>0.1</td>
<td>110</td>
</tr>
</tbody>
</table>

Table B.3: Equilibrium points list on 27 sub-regions configuration.
<p>| | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>23</td>
<td>0.8333</td>
<td>0.5000</td>
<td>0.9200</td>
<td>0.1</td>
<td>95.000</td>
</tr>
<tr>
<td>24</td>
<td>0.8333</td>
<td>0.5000</td>
<td>0.5500</td>
<td>0.1</td>
<td>111.667</td>
</tr>
<tr>
<td>25</td>
<td>0.8333</td>
<td>0.8333</td>
<td>1.0000</td>
<td>0.2</td>
<td>78.333</td>
</tr>
<tr>
<td>26</td>
<td>0.8333</td>
<td>0.8333</td>
<td>0.5500</td>
<td>0.1</td>
<td>95.000</td>
</tr>
<tr>
<td>27</td>
<td>0.8333</td>
<td>0.8333</td>
<td>0.3300</td>
<td>0.1</td>
<td>111.667</td>
</tr>
</tbody>
</table>

**Figure B.7**: Wall temperature Bode plot for Config B (MMPC27).

**Figure B.8**: Coolant temperature Bode plot for Config B (MMPC27).
Figure B.9: MMPC27 with input rate weight at 0.00001, 0.1, 1 and 5.
C.1. **Water Block Groove Design**

10 types of water block groove design are being compared for the experiment (Figure C.1). The design selection is based on the simulation in Autodesk® CFD 2015 software results. Two man criteria are being observed:

- Maximum wall temperature should not exceed 200°C. This is limited by the thermal compound working temperature (200°C).
- The Reynold Number of coolant flow should be as close as possible to the actual Reynold Number in an actual engine near the exhaust valve bridge.

![Figure C.1: 10 types of water block groove design for the experiment.](image)

The best groove design according to the stated criteria is the “3 line cut square”. The design has maximum temperature just less than 200°C and the cold side is not more than 45°C (Figure C.2). The coolant flow rate creates 5240 Reynold Number with maximum velocity of 5.34m/s.
Figure C.2: The water block temperature result from Autodesk® CFD during maximum heat and coolant flow rate being applied.

Figure C.3: The coolant flow result from Autodesk® CFD during maximum heat and coolant flow rate being applied.

Table C.1 below is the result of all the groove design comparing highest temperature recorded, wall dimension and highest velocity recorded.

Table C.1: The groove designs result from Autodesk CFD simulation.

<table>
<thead>
<tr>
<th>Water block</th>
<th>Highest temperature [°C]</th>
<th>Highest velocity [m/s]</th>
<th>Hydraulic diameter [mm]</th>
<th>Re Number</th>
<th>Wall thickness [mm]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pinned</td>
<td>142.2</td>
<td>5.88</td>
<td>1.043</td>
<td>2143</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---------</td>
<td>-------</td>
<td>-------</td>
<td>-------</td>
<td>-------</td>
<td>-------</td>
</tr>
<tr>
<td>9 block</td>
<td>152.3</td>
<td>3.80</td>
<td>3.429</td>
<td>4551</td>
<td>10</td>
</tr>
<tr>
<td>2 mm</td>
<td>295.5</td>
<td>4.33</td>
<td>3.625</td>
<td>5484</td>
<td>10</td>
</tr>
<tr>
<td>4 mm</td>
<td>343.9</td>
<td>2.36</td>
<td>6.630</td>
<td>5466</td>
<td>10</td>
</tr>
<tr>
<td>6 mm</td>
<td>335.0</td>
<td>1.99</td>
<td>9.160</td>
<td>6368</td>
<td>10</td>
</tr>
<tr>
<td>9 circle</td>
<td>393.4</td>
<td>2.36</td>
<td>3.000</td>
<td>2473</td>
<td>10</td>
</tr>
<tr>
<td>3 line</td>
<td>174.4</td>
<td>6.23</td>
<td>3.000</td>
<td>6529</td>
<td>11</td>
</tr>
<tr>
<td>3 line big</td>
<td>1803</td>
<td>5.85</td>
<td>3.000</td>
<td>6131</td>
<td>12</td>
</tr>
<tr>
<td>3 line cut</td>
<td>207.7</td>
<td>6.19</td>
<td>2.442</td>
<td>5280</td>
<td>15</td>
</tr>
<tr>
<td>3 line square</td>
<td>199.6</td>
<td>5.34</td>
<td>3.000</td>
<td>5596</td>
<td>13</td>
</tr>
</tbody>
</table>

The Reynold Number is calculated using equation (86) below with 25% coolant mixture viscosity. The hydraulic diameter is calculated at the highest coolant velocity location.

\[
Re = \frac{\rho_{cool}v_{cool}L}{\mu_{cool}}
\]

- \(Re\) = Reynold Number
- \(v_{cool}\) = Coolant velocity [m/s]
- \(L\) = Hydraulic diameter [m]
- \(\mu_{cool}\) = Coolant (25% mixture) dynamic viscosity [m²/s]
- \(\rho_{cool}\) = Coolant (25% mixture) density [kg/m³]

### C.2. Steady State Model for Experiment

These are the steady state model response details being used during the experiment:
<table>
<thead>
<tr>
<th>Model</th>
<th>PRESS RMSE</th>
<th>RMSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>RBF- recmultiquadric-41</td>
<td>87.451</td>
<td>84.786</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Model</th>
<th>PRESS RMSE</th>
<th>RMSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Poly-6</td>
<td>7.357e-4</td>
<td>6.384e-4</td>
</tr>
<tr>
<td>Model</td>
<td>PRESS RMSE</td>
<td>RMSE</td>
</tr>
<tr>
<td>----------</td>
<td>------------</td>
<td>---------</td>
</tr>
<tr>
<td>Poly-6</td>
<td>9.285e-4</td>
<td>8.605e-4</td>
</tr>
<tr>
<td>Mean-RBF</td>
<td>8.636e-5</td>
<td>8.164e-5</td>
</tr>
<tr>
<td>Model</td>
<td>PRESS RMSE</td>
<td>RMSE</td>
</tr>
<tr>
<td>-------------</td>
<td>------------</td>
<td>--------</td>
</tr>
<tr>
<td>Liner-RBF</td>
<td>0.027</td>
<td>0.01</td>
</tr>
<tr>
<td>Mean-RBF</td>
<td>0.124</td>
<td>4.634e-3</td>
</tr>
</tbody>
</table>
C.3. **LabVIEW Software Block Diagram**

Figure C.4 and Figure C.5 are the LabVIEW block diagram made in the LabVIEW Software and LabVIEW FPGA for the experiment.

![LabVIEW Software Block Diagram](image-url)

**Figure C.4: LabVIEW Software block diagram**
Figure C.5: LabVIEW FPGA block diagram.
C.4. **Data for Model Fitting**

This is the estimation data and validation data for experiment model fitting (Figure C.6 and Figure C.7).

![Figure C.6: Estimation and validation data for experiment wall model fitting.](image-url)
System identification result for the experiment is a state space model as follows:

\[
\begin{bmatrix}
\dot{x}_1 \\
\dot{x}_2 \\
\dot{x}_3 \\
\dot{x}_4 \\
\end{bmatrix} =
\begin{bmatrix}
-0.0057 & 0.0038 & 0 & 0 \\
0.0231 & -0.0189 & 0 & 0 \\
0 & 0 & -6.1281e-04 & 3.2985e-05 \\
0 & 0 & 8.2990e-05 & -0.0016 \\
\end{bmatrix}
\begin{bmatrix}
x_1 \\
x_2 \\
x_3 \\
x_4 \\
\end{bmatrix} +
\begin{bmatrix}
7.9342e-05 & -6.9746e-05 & 0 \\
-2.8807e-04 & 2.4140e-04 & 0 \\
0 & 1.0318e-05 & -8.7589e-06 \\
0 & -4.4974e-04 & 3.6887e-04 \\
\end{bmatrix}
\begin{bmatrix}
\dot{Q}_{comb} \\
\dot{Q}_{conv} \\
\end{bmatrix}
\]

\[
\begin{bmatrix}
T_w \\
T_{out} \\
\end{bmatrix} =
\begin{bmatrix}
347.6599 & -0.1331 & 0 & 0 \\
0 & 0 & -118.2046 & -3.5290 \\
\end{bmatrix}
\begin{bmatrix}
x_1 \\
x_2 \\
x_3 \\
x_4 \\
\end{bmatrix}
\]

C.5. Coolant Capacity Ratio

The average actual engine energy heat loss to coolant and coolant capacity ratio is based on Table C.2 below. The energy heat loss to coolant is assumed equal to the engine maximum
power output. The engine power rating and coolant capacity data is taken from “http://cararac.com/coolant/”.

Table C.2: List of engine power rating and coolant capacity.

<table>
<thead>
<tr>
<th>Vehicle</th>
<th>Engine category</th>
<th>Power rating</th>
<th>Coolant capacity</th>
<th>Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Peugeot 206</td>
<td>2.0L</td>
<td>99kW</td>
<td>7.8L</td>
<td>0.0788</td>
</tr>
<tr>
<td>Honda CR-V</td>
<td>2.0L</td>
<td>110kW</td>
<td>6.2L</td>
<td>0.0563</td>
</tr>
<tr>
<td>Mitsubishi Lancer</td>
<td>2.0L</td>
<td>101kW</td>
<td>7.0L</td>
<td>0.0693</td>
</tr>
<tr>
<td>Toyota Camry</td>
<td>2.0L</td>
<td>96kW</td>
<td>8.5L</td>
<td>0.0885</td>
</tr>
<tr>
<td>Nissan X-Trail</td>
<td>2.0L</td>
<td>104kW</td>
<td>7.4L</td>
<td>0.0712</td>
</tr>
<tr>
<td>BMW 3-series</td>
<td>2.0L</td>
<td>105kW</td>
<td>7.5L</td>
<td>0.0714</td>
</tr>
<tr>
<td><strong>Average</strong></td>
<td></td>
<td></td>
<td></td>
<td>0.0726</td>
</tr>
</tbody>
</table>
Appendix D  

MPC Performance in Drive Cycles

D.1. Results

The MPC performance with and without future knowledge of disturbance compared to conventional cooling system throughout the drive cycles are shown in Figure D.1 until Figure D.8. The results include wall temperature, coolant temperature, water pump signal, valve signal and radiator fan signal.
Figure D.1: MPC controller performance with the future knowledge of disturbance, without the future knowledge of disturbance and conventional cooling system in Artemis Urban Cycle.
Figure D.2: MPC controller performance with the future knowledge of disturbance, without the future knowledge of disturbance and conventional cooling system in Artemis Rural Road Cycle.
Figure D.3: MPC controller performance with the future knowledge of disturbance, without the future knowledge of disturbance and conventional cooling system in Artemis Motorway Cycle.
Figure D.4: MPC controller performance with the future knowledge of disturbance, without the future knowledge of disturbance and conventional cooling system in NEDC.
Figure D.5: MPC controller performance with the future knowledge of disturbance, without the future knowledge of disturbance and conventional cooling system in WLTC.
Figure D.6: MPC controller performance with the future knowledge of disturbance, without the future knowledge of disturbance and conventional cooling system in FTP-75kph.
Figure D.7: MPC controller performance with the future knowledge of disturbance, without the future knowledge of disturbance and conventional cooling system in HWY.
Figure D.8: MPC controller performance with the future knowledge of disturbance, without the future knowledge of disturbance and conventional cooling system in US06.

D.2. Fuel Consumption throughout Drive Cycle

Figure D.9 to Figure D.16 are the results of fuel consumed throughout the drive cycles from a perfect wall temperature controller potential compared to conventional cooling system, MPC with and without the future knowledge of the disturbances.
Figure D.9: Fuel consumed throughout the NEDC by a perfect controller of ideal head temperature compared to the conventional cooling system, MPC with and without the known future disturbance.

Figure D.10: Fuel consumed throughout the WLTC by a perfect controller of ideal head temperature compared to the conventional cooling system, MPC with and without the known future disturbance.
Figure D.11: Fuel consumed throughout the Artemis Urban Cycle by a perfect controller of ideal head temperature compared to the conventional cooling system, MPC with and without the known future disturbance.

Figure D.12: Fuel consumed throughout the Artemis Rural Road Cycle by a perfect controller of ideal head temperature compared to the conventional cooling system, MPC with and without the known future disturbance.
Figure D.13: Fuel consumed throughout the Artemis Motorway Cycle by a perfect controller of ideal head temperature compared to the conventional cooling system, MPC with and without the known future disturbance.

Figure D.14: Fuel consumed throughout the FTP 75kph by a perfect controller of ideal head temperature compared to the conventional cooling system, MPC with and without the known future disturbance.
Figure D.15: Fuel consumed throughout the US06 by a perfect controller of ideal head temperature compared to the conventional cooling system, MPC with and without the known future disturbance.

Figure D.16: Fuel consumed throughout the HWY by a perfect controller of ideal head temperature compared to the conventional cooling system, MPC with and without the known future disturbance.