Feedback Control Strategies for Diesel Engine Emissions Compliance

by

Mitchell S. Bieniek

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Mechanical Engineering)
in the University of Michigan
2021

Doctoral Committee:

Professor Anna Stefanoupoulou, Chair
Professor Andre Boehman
Associate Research Scientist Emeritus John Hoard
Assistant Research Scientist Robert Middleton
Professor Jeffrey Scruggs
Lecturer Michiel van Nieuwstadt

Mitchell S. Bieniek
bieniekm@umich.edu
ORCID iD: 0000-0001-5869-9190

To my family, for their endless love and support.

ACKNOWLEDGMENTS

I would like to start off by acknowledging my advisors Dr. Anna Stefanopoulou and John Hoard. Dr. Stefanopoulou, I will forever be grateful for providing me the opportunity to pivot my graduate studies towards controls, and helping me develop the skills critical for solving challenging modeling and control problems. John, your mentorship has been instrumental in my growth as a researcher over the last nine years. Thank you for extending the life-changing offer for me to join your lab freshman year.

I would also like to thank the members of my committee, who have all contributed to my growth beyond the actual dissertation process. Dr. Andre Boehman, thank you for all of your support in the medum-duty diesel test cell as well as general engine research techniques. Dr. Jeffrey Scruggs, thank you for being an excellent instructor (especially in Linear Systems Theory, my first graduate controls course!) and providing valuable theoretical insight into tackling this work. Dr. Michiel van Nieuwstadt, thank you for helping me learn the importance of developing theory, while also focusing on real-world applications. Dr. Robert Middleton, thank you for your help with my research, as well as generally helping me navigate the process of finishing a PhD.

I am also thankful for the support of the Ford Motor Comany employees I've worked with over the last nine years. Brien Fulton, thank you for your help setting up the lab and vehicle test setups, teaching me about diesel calibration and controls, and showing me the best lunch spot in Wyandotte. Hoon Cho, thank you for helping me understand the cylinder pressure algorithms, and providing great examples to inform my own work. Mike Hopka, thank you for your insight and help on the connected vehicle optimization work. Dan Styles, Dan Kantrow, and Paul Tennison, thank you for your help getting the engine installed, running, and keeping it running. Thank you to Chris House, Keith Plagens, Chih-kuang Kuan, Sathish Atluri, Chris Wicks, Wolfram Buschhaus, Greg McConville, Tom Brewbaker, Eric Kurtz, Joe Dodyk, Chris Frankenfield, and all of the other awesome Ford employees that taught me so much over the years.

There are a large number of people I must thank from the Auto Lab. Bill Kirkpatrick, Jim Elkins, Chuck Solbrig, Ken Pruss, Melissa McGeorge, Alan Phlipot, William Lim: thank you for providing the mechanical and logistical support needed to keep the Auto Lab running. I have really enjoyed getting to know you all over the years, and I will miss our

random conversations. Big thanks to everyone I've been able to work with and/or alongside over the years: Dr. Di Chen, Dr. Miriam Figueroa, Dr. Chunan Huang, Dr. Peyman Mohtat, Dr. Suhak Lee, Sravan Pannala, Omar Ahmed, Dr. Ting Cai, Eunjeong Hyeon, Dr. Bryan Maldanado, Dr. Shima Nazari, Dr. Rani Kiwan, Dr. Niket Prakash, Dr. Rasoul Salehi, Dr. Tulga Ersal, Dr. Vassilis Triantopoulos, Rinav Pillai, Courtney Videchak, Dr. Jonathan Martin, Dr. Taehoon Han, Dr. Jordan Easter, Dr. Ripudaman Singh, Daniel Knister, and Elizabeth Bernhardt. It's been a pleasure working with and learning from you all.

I would not be here without the support of my close friends and family. Dr. Ashwin Salvi - thank you for your incredible mentorship over the last nine years. Saravanan Duraiarasan and Joe Drallmeier - thank you for being awesome labmates, classmates, and project partners. Ian McEachern, Clay Mikush, Keri Mikush - thank you for helping me vent, relax, and have fun. Sally Dillon - thank you for your endless love, and supporting me through the highs and lows of graduate school. Finally, a huge thanks to my family for helping me chase my dreams; especially my parents Richard and Roberta Bieniek, my sister Veronica, and my brother Max.

Preface

This work was funded and intellectually supported by Ford Motor Company.

Table of Contents

LIST OF FIGURES

LIST OF TABLES

ABSTRACT

Modern diesel engines are equipped with aftertreatment systems which are effective at reducing tailpipe hydrocarbon and oxides of nitrogen (NO_X) emissions when the system's catalysts are lit-off, meaning they are warmed-up to temperatures near 200°C. During engine cold-starts, combustion phasing retard is typically used to provide additional heat to the aftertreatment system to achieve faster light-off. Analysis of emissions cycle data has shown that improved heating during cold-starts could achieve further emission reductions, however combustion phasing retard heating strategies can be limited by combustion variability issues. Aftertreatment temperature issues can also occur after the engine is warmed-up, as real-world driving behaviors like extended idling and low-load operation can result in exhaust temperatures that are insufficient for maintaining catalyst light-off, resulting in emission increases. This thesis presents novel control solutions to achieve emissions reductions during cold-starts and real-world driving.

For cold-start emissions, the concept of closed-loop variance control was analyzed and applied to combustion control, which enables more aggressive combustion phasing retard exhaust heating to achieve faster aftertreatment light-off while avoiding excessive combustion variability issues. Diesel combustion variability was characterized experimentally, and the data was used to identify feedback metrics. Conventional linear controls analysis and statistical theory were used to develop a better understanding of variance feedback control, and the understanding was applied to the engine problem. Closed-loop combustion variability control was performed during both steady-state and transient operation and enabled higher exhaust temperatures while avoiding excessive degradation of engine combustion.

For real-world driving emissions, a model predictive control (MPC) framework was developed that uses long horizon engine speed and load preview along with onboard NO_X measurements to control the engine for good fuel economy subject to emission constraints. To reduce computational complexity, the controller output is a decision variable selecting between two engine calibrations, one with low brake-specific fuel consumption (BSFC) but high brake-specific NO_X ($BSNO_X$), and one with high BSFC, low $BSNO_X$, and increased exhaust heat to aid aftertreatment conversion efficiencies. The

onboard NO_X measurements are used to inform the optimization problem formulations, which include constraining NO_X based on windowed limits. Software-in-the-Loop (SIL) experimental results show that the controller has the ability to track a windowed emissions target, and appropriately responds to noise factors such as aftertreatment temperatures and emission rate errors.

CHAPTER 1

Introduction

1.1 Diesel Engine Emission Challenges

Diesel (compression ignition) engines are widely used in medium/heavy-duty vehicles, as well as off-road construction and farming equipment, due to their ability to produce high torque at lower engine speeds, higher fuel efficiency (lower carbon dioxide emissions), and increased reliability relative to gasoline (spark-ignited) engines. The biggest challenge associated with diesel engines is dealing with the harmful emissions generated during the combustion process. The primary regulated emissions for diesel powertrains include carbon dioxide (CO_2), hydrocarbons (HCs), nitrogen oxides (NO and NO_2, also known as NO_X), and particulate matter (PM, also known as soot). Carbon dioxide is an inevitable consequence from combusting a hydrocarbon fuel and is a known greenhouse gas (GHG) linked to climate change. NO_X contributes to the formation of acid rain, and when combined with HCs and other volatile organic compounds (VOCs) in the atmosphere can lead to the formation of smog. PM emissions can be inhaled and deposited into the lungs leading to long-term respiratory damage and other health consequences.

Due to those concerns, modern diesel powertrain systems are equipped with an aftertreatment system, or a collection of catalysts and other supporting hardware, that either capture harmful exhaust species or convert them to something less harmful. Reductions in engine out emissions are possible, but typically incur power and/or fuel economy (CO_2) penalties that are disadvantageous for consumers – use of a proper aftertreatment system enables improved fuel economy and power while averting excessive tailpipe emissions. A modern medium- or heavy-duty diesel aftertreatment system is typically comprised of three devices: a diesel oxidation catalyst (DOC), used to oxidize hydrocarbons; a diesel particulate filter (DPF), used to capture particulate matter for later oxidation; and a selective catalytic reduction (SCR) system to reduce NO_X to harmless nitrogen and water. The SCR system reduces NO_X using ammonia, often created by injecting diesel exhaust fluid

(DEF, liquid solution containing urea) upstream of the SCR. This combination of devices is extremely effective at reducing tailpipe HC and NO_X emissions upon warming up to temperatures near 200°C, however at lower or higher temperatures they suffer from poor conversion efficiencies, resulting in higher tailpipe emissions [2].

1.1.1 Challenge 1: Cold-Start Diesel Emissions

Analysis of cycle emissions data (example: EPA FTP-75) from current engines show that a majority of the total cycle emissions are made at the beginning of the cycle during the cold-start phase, prior to aftertreatment light-off (aftertreatment temperatures reaching $\approx 200°C$). This data suggests that a reduction in cold-start feedgas emissions and/or earlier operation of the aftertreatment system is critical for continued regulation compliance [3]. Modifying current catalyst structures and operating principles could reduce activation temperatures (and therefore light-off times), but some form of exhaust heating will likely still be necessary. Moving the physical location of the catalyst closer to the engine can reduce the heat losses and the thermal inertia associated with heating the exhaust system, but may be infeasible with vehicle packaging constraints. Again, this would just reduce the amount of heating required, not eliminate it.

1.1.1.1 Exhaust Heating Strategies

There are several opportunities available to improve exhaust heating of aftertreatment systems [3]. Potential exhaust system modifications include adding a burner or electric heater directly upstream of the catalyst for more efficient heating (avoid heating exhaust plumbing between engine and catalysts) [3, 4]. More complex valvetrain systems can enable cylinder deactivation, negative valve overlap, and early exhaust valve opening strategies to increase engine exhaust temperatures [5, 6]. Modern diesel engines typically have an exhaust gas re-circulation (EGR) system, a variable geometry turbocharger (VGT), and fuel injectors capable of multiple injections per cycle, whose usage could be modified during cold-starts to alter combustion behavior to increase engine exhaust temperatures [2, 3, 7, 8]. Leveraging techniques like negative valve overlap, increased exhaust gas re-circulation, and later injection timings achieve higher exhaust temperatures by phasing combustion later into the expansion stroke ("retarding combustion phasing") [7].

Combustion phasing retard is an effective technique for increasing exhaust gas temperatures, but there are practical limits on the amount of phasing retard that can be used, as it can increase combustion variability which can cause noise, vibration, and harshness (NVH) and emission issues. It is typically evaluated using data from an in-cylinder pressure

sensor, which allow for the measurement of cylinder pressure as a function of crankshaft position. When coupled with knowledge of engine geometry, thermodynamics, and heat transfer, the measurements can be used to estimate fuel combustion rates and torque production (characterized by Indicated Mean Effective Pressure, IMEP) by individual cylinders, cycle-to-cycle. Figure 1.1 illustrates the trade-off between combustion phasing retard exhaust heating and combustion variability, using the calculated IMEP values from cylinder pressure feedback. As combustion phasing is retarded via injection timing retard, the variability in IMEP increases in magnitude, which can be observed both from the population statistics of IMEP (standard deviation) and the sequence behavior shown in the bottom subplot. Note the increasing slope of the σ(IMEP) response versus injection timing retard - combustion variability is also sensitive to many parameters (ambient temperatures, fuel quality, engine aging, etc.), and so uncertainties in other variables that impact combustion variability could dramatically destabilize combustion to unacceptable

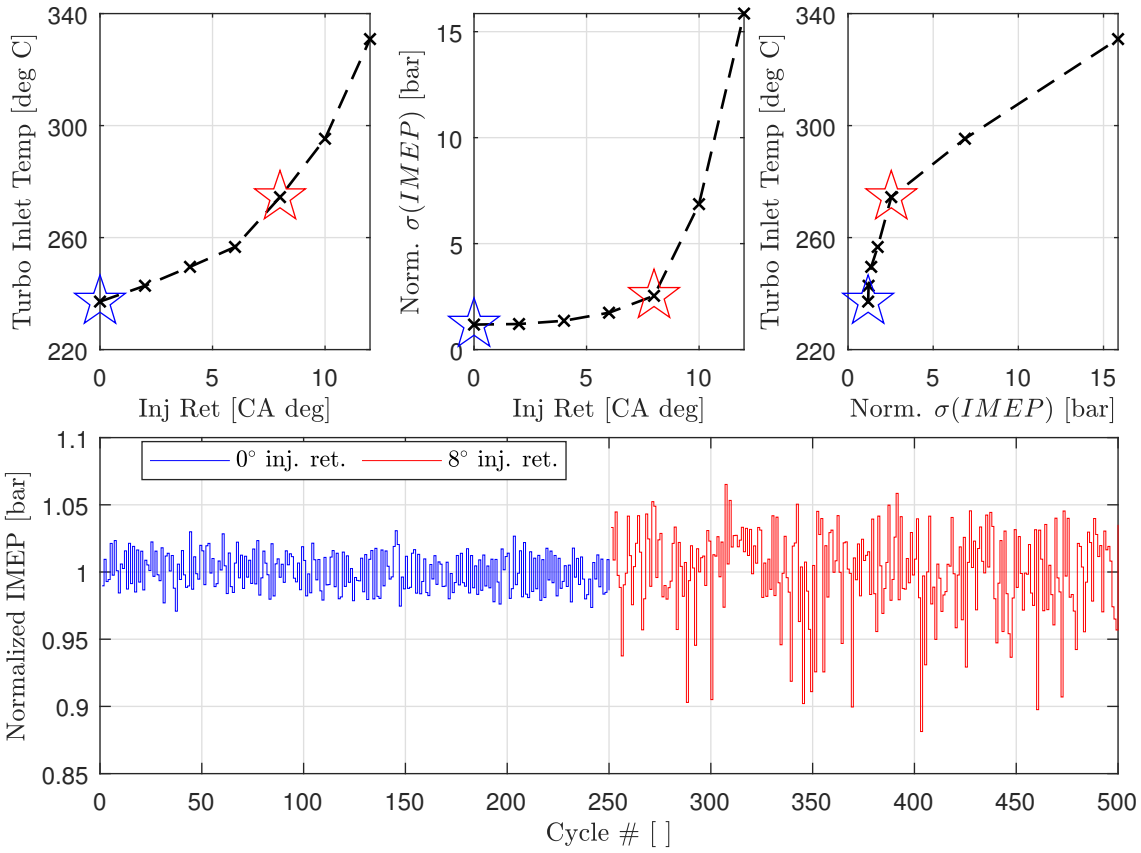


Figure 1.1: Impact of injection timing retard on exhaust temperatures (Turbo Inlet Temp) and combustion torque (IMEP). Note that retarding injection timing can increase exhaust temperatures at the cost of increased combustion variability (Chapter 2) [1].

levels. This suggests that without a technique for monitoring combustion variability, combustion phasing retard exhaust heating strategies must be inherently conservative to avoid unacceptable performance issues due to noise factors [2, 7].

1.1.1.2 Cylinder Pressure Feedback Opportunities

Closed-loop combustion control using feedback from in-cylinder pressure sensors has been discussed and explored on gasoline and diesel engines since the 1950s [9]. Early control architectures mostly focused on achieving max brake torque, including controlling peak combustion pressure position, the apparent combustion phasing (CA50), and control of combustion phasing while avoiding engine knocking issues. Recent work for a diesel application has demonstrated the ability to achieve real-time simultaneous control of IMEP, CA50, and max rate of heat release (ROHR), which allows for reductions in cylinder-to-cylinder variability, as well as reductions in NO_X and particulate matter emissions [10–12]. Although laboratory experiments demonstrated promising results, issues with sensor cost, reliability, and physical placement restricted their typical usage to lab settings.

Increasingly stringent emission requirements, coupled with advances in sensor technology and packaging, have led to OEMs implementing cylinder pressure sensors in consumer engines. One example is Volkswagen's implementation of pressure sensing glow plugs (PSG) on their 1.4, 1.6, and 2.0-liter diesel engines [13]. Mazda's Skyactive-X engine is leveraging cylinder pressure sensors to enable advanced combustion modes (Spark-Assist Compression Ignition, or SACI) to improve fuel economy [14]. Besides offering the opportunity to directly control parameters like cylinder torque and combustion phasing, production cylinder pressure sensors allow for the implementation of more sophisticated models to enhance operation. Real-time monitoring of combustion can be used to estimate fuel quality/properties so that calibration maps can be properly adjusted to ensure drive quality. Cylinder pressure measurements can be used with appropriate models to estimate cyclic NO_X emissions, allowing for better control of the aftertreatment system with respect to a system only using downstream NO_X sensors [15]. Further improvements in sensor cost/reliability, stricter emission regulations, advanced combustion modes, and growing capabilities of cylinder pressure based modeling will likely lead to increased adoption of cylinder pressure sensors for consumer engines.

1.1.2 Challenge 2: Real-world Driving Diesel Emissions

Emission standards are constantly evolving, and some emission regulatory bodies are making a push towards Real-world Driving Emissions (RDE) tests for more vehicle

classifications. A majority of light-duty emission certification tests are conducted in climate-controlled vehicle dynamometer test cells, against known test cycles published by the regulators (ex: FTP-75, LA4, HWFET). This framework led to the infamous Volkswagen "Dieselgate" scandal, where special software was used to detect if the vehicle was on an emission certification test and control the engine and aftertreatment systems accordingly [16]. RDE testing is conducted using Portable Emission Measuring Systems (PEMS) that are able to record tailpipe emission data while driving on public roads. The nature of the tests means that there is more test variability due to ambient conditions, traffic, and other factors that are normally excluded using laboratory test cells. Vehicle manufacturers will be responsible for developing a vehicle that is robust to these uncertainties [17].

In the USA, heavy-duty diesel vehicles are already subject to RDE constraints in the form of Not-To-Exceed (NTE) emission regulations (CFR 40 Part 86 Subpart N) [18]. As of 2021, the heavy-duty NTE constraints are unique in that they only apply when a heavy duty engine has operated for a minimum of 30 seconds in the NTE zone (portion of speed-load range where enforcement is desired) subject to certain exclusions (ex: low ambient temperature, low coolant temperature). When the NTE conditions are met, a time average of the emissions is compared to the standard to check compliance. Rule changes proposed by the US Environmental Protection Agency (USEPA) and California Air Resources Board (CARB) for 2024-2026 include transitioning to a windowed emission compliance standard, similar to the EUROVI standard. These changes would include elimination of the NTE zone residence time, and instead adopt a minimum window power limit, increasing the amount of time vehicles must be compliant during real-world scenarios [19, 20]. One real-world scenario that is particularly challenging for diesel vehicles is extended idle or low-load operation, as the exhaust heat coming from the engine can be insufficient to keep the aftertreatment temperature at a functional temperature, resulting in an increase in tailpipe emissions [21].

1.1.2.1 Vehicle Connectivity Opportunities

One avenue for reducing real-world driving emissions includes leveraging traffic and other forms of preview for powertrain and vehicle control. Researchers have leveraged dynamic programming and model predictive control techniques to achieve better fuel/energy economy while reducing emissions by modifying vehicle velocity profiles, modifying engine actuator setpoints, and in the case of hybrid electric vehicles (HEVs), optimizing the interaction between the internal combustion engine and electric motor [22,23]. Some of the approaches leverage full engine airpath and aftertreatment models with nonlinear dynamics

and seek to use the predicted emissions and fuel consumption to achieve improved compliance. In the case of altering vehicle velocity profiles, these optimization routines can lead to undesirable driving behaviors (from an occupant perspective) unless constraints and cost functions are properly structured. Approaches leveraging full airpath models tend to be computationally expensive, and if leveraging multiple actuators (like injection timing, boost setpoints, exhaust gas recirculation), the optimization problem grows quickly with control steps and preview duration [24–26]. Furthermore, the model predictive control architectures will attempt to optimize fuel and emissions over the preview horizon, but do not include historical emissions feedback data when initializing the model. Ignoring historical emission data in the optimization can yield non-compliant trajectories, as noise-factors impacting the actual emission rates could go unobserved by the controller.

1.1.3 Dissertation Contributions

This thesis presents techniques to address the challenges outlined above by augmenting existing control systems with additional feedback and other information to help improve system performance.

1.1.3.1 Contribution 1: Combustion Variability Feedback Controller

The first challenge, cold-start diesel engine emissions, was approached from the perspective of trying to improve existing engine-based exhaust heating methods. The magnitude at which these methods can be used are restricted by their potential impacts on combustion variability (NVH consequences). The potential for increased cylinder pressure sensor adoption on consumer engines motivated an investigation into techniques to generate and use online combustion variability estimates for feedback control purposes. Specifically, this dissertation explores the premise of using online combustion variability estimates in a feedback control structure to retard combustion phasing until a target combustion variability setpoint is reached. By retarding combustion phasing as late as possible while respecting variability constraints, phasing-based exhaust heating can be maximized while avoiding concerns associated with combustion variability noise factors.

1.1.3.2 Contribution 2: Preview-Based MPC Controller With Windowed Emissions Feedback

The proposed solution for the second challenge, real-world driving emissions, includes combining route preview information from a vehicle's navigation system with feedback from an onboard NO_X sensor to design a novel MPC (Model Predictive Control) control

architecture that optimizes fuel economy subject to emission constraints motivated by the proposed windowed emission regulations (2024-2026 EPA/CARB). This architecture is unique due to its use of historical onboard NO_X measurements to inform the optimization constraints, something not observed in existing literature or patents. The approach uses simplified engine and aftertreatment models to form a simpler computational problem which enables longer horizon preview (>10 minutes), critical for managing aftertreatment thermal dynamics, while leveraging feedback to correct issues posed by model and preview errors. The output of the controller is a decision variable, choosing from one of the following engine calibrations with the following characteristics:

- Low feedgas emissions and higher exhaust heating at the expense of fuel economy - useful when the aftertreatment system is not functional or lower tailpipe emissions are required
- High feedgas emissions and lower exhaust heating at the benefit of improved fuel economy - useful when the aftertreatment is functional, or higher quantities of tailpipe emissions can be tolerated

The controller uses preview of engine speed and load (assumed to be known, but in real life would be synthesized using a vehicle longitudinal model and traffic data) to identify which calibration to use at each control interval to maximize fuel economy while remaining in the aforementioned emission constraints.

1.2 Dissertation Organization

Chapter 2 presents experimental observations regarding diesel engine combustion behavior at a variety of speeds and loads at low coolant temperatures (20-50°C). Observed trends for combustion torque and fast heat release parameters were analyzed from a controls perspective. A potential hydrocarbon emission indicator using cylinder pressure sensor and fuel flow feedback is also shown. Based on the observed behavior, a closed-loop combustion variability controller is proposed, and the experimental data is used to develop a control-oriented combustion variability model based on actuators of interest (injection phasing offset, main fuel quantity offset). The work in this chapter was presented and published in the American Society of Mechanical Engineers Internal Combustion Engine Fall Conference (ASME ICEF) in 2018 [1], and published in the International Journal of Powertrains (IJPT) [27].

The theoretical aspects of variance estimation and control are explored in Chapter 3. Expectation analysis of the windowed variance operator was used to calculate the expected

value and variance of windowed estimator (for zero-mean, varying-mean, and detrended sequences). The analysis was then used to study the windowed variance estimator as a "virtual sensor", looking at distribution and other behaviors to understand the implications on feedback control. The expectation analysis was combined with linear control design principles to design linear time-invariant (LTI) controllers; the ensemble average of the stochastic realizations agreed with LTI predicted behavior when non-linearities were not present. For applications with limited computational resources, a simpler variance estimation strategy using exponential weighted moving averages is presented and compared to the windowed variance estimator. The work in this chapter was published in the International Journal of Powertrains [27], presented and published in the American Automatic Control Council American Control Conference (A2C2 ACC) in 2020 [28], and submitted to Institute of Electrical and Electronics Engineers Transactions on Control Systems Technology (IEEE TCST) [29].

Chapter 4 presents the design and impact of the combustion variability controller proposed in Chapter 2. The content of Chapters 2 and 3 were combined to identify steady-state and transient combustion variability metrics that exhibited desirable responses to the chosen actuator (injection phasing) and a noise factor (coolant temperature). The combustion variability model was then used to design and evaluate a combustion variability controller. Following successful simulation results, the controller was implemented on an engine dynamometer and in multiple test vehicles, and demonstrated that closed-loop combustion variability control is feasible during steady-state and transient operation. The controller achieved the desired exhaust temperature benefit while avoiding excessive degradation of combustion quality. The various tests conducted highlighted the value of closed-loop control in the presence of noise factors (different ambient conditions, different fuels). The work in this chapter was presented and published in the International Federation of Automatic Control Advances in Automotive Control Conference (IFAC AAC) in 2019 [30], and submitted to the Institute of Electrical and Electronics Engineers Transactions on Control Systems Technology [29].

Chapter 5 presents a novel preview-based model predictive engine controller focused on emissions compliance. The optimization objective was to minimize fuel consumption while observing windowed emission constraints. To enable long preview horizons, a simple model and actuator structure was used. The original optimization problem was a computationally expensive nonlinear mixed-integer problem, and so techniques were presented and evaluated to instead solve a continuous problem and discretize the solution using simple conversion policies. Feedback from an onboard NO_X sensor was used to correct for errors due to modeling, noise factors, and preview uncertainty. Simulation-

in-loop testing shows the feedback structure averts issues poised by model uncertainty, actuator discretization scheme, and noise factors. An invention disclosure based on this idea has been submitted to the US Patent and Trademark Office for review [31]. This work has also been submitted for presentation/publication in the American Automatic Control Council Modeling, Estimation and Control Conference in 2021(A2C2 MECC2021).

Conclusions and opportunities for future work are highlighted in Chapter 6. Both control architectures would benefit from further experiments, exploring additional engine actuators, and integrating the control schemes into the rest of the engine control structure for improved synergy.

CHAPTER 2

Characterization of Retarded Combustion Behavior

2.1 Introduction

Modern diesel vehicles are equipped with aftertreatment systems effective at reducing tailpipe hydrocarbon, oxides of nitrogen, and particulate matter emissions upon warming-up to temperatures near 200°C [2]. Due to catalyst ineffectiveness at low temperatures, production engines increase exhaust gas enthalpy during cold-starts, allowing for quicker heating of the aftertreatment system to operational temperatures as early as possible. Upcoming emission regulations will require either a reduction in engine out emissions or better usage of the aftertreatment system to meet targets. Analysis of emissions data from current engines show that a majority of the emissions are released at the beginning of the cycle, prior to aftertreatment light-off, suggesting that earlier operation of the aftertreatment system would allow for regulation compliance [32].

Quicker aftertreatment operation during cold-starts could be achieved by modifying current catalyst structures and operating principles to extend conversion efficiencies at lower temperatures and/or increasing the amount of exhaust heating during the warm-up phase. Note that even with improved catalysts, supplemental heating will likely still be required to maximize the operating time. Therefore, this chapter will focus on techniques available for improving exhaust heating during engine cold-starts to hasten aftertreatment operation.

Existing and novel actuators in literature were studied to understand their inherent advantages and disadvantages. An existing actuator, injection control (timing and quantity) was selected for detailed experimental study for use in more aggressive exhaust heating. Engine experiments revealed predictable shortcomings relating to combustion (torque) variability when using injection control, and so a feedback control architecture using an in-cylinder pressure sensor was proposed to address those shortcomings. Based on the

proposed architecture, engine experiments were performed to generate a control-oriented stochastic model and to understand impacts of the actuator on other engine attributes.

2.2 Background

Background on diesel engine combustion control, indicated combustion analysis, and exhaust heating mechanisms are provided to inform later sections of the Chapter.

2.2.1 Diesel Engine Control and Behavior

Modern diesel engines are controlled by engine control units (ECUs) that are able to leverage sensor feedback and state observers to control actuators in a way that satisfies power, NVH, durability, and emission requirements. Based on a speed/torque (power) request to the engine, values like boost pressure, EGR (Exhaust Gas Recirculation) rate, and injection timing/quantities are controlled using the various actuators available [7].

Unlike spark-ignited (SI) engines (typically gasoline engines), which rely on a plasma ignition source (spark plug) to ignite the air-fuel mixture, diesel engines rely on elevated pressures and temperatures to cause a portion of the injected fuel to autoignite (premixed burn phase) and the rest to burn as it mixes with remaining air in the cylinder (diffusion burn). Note that since modern diesel engines typically use direct-injection, they can inject fuel in the cylinder at any point during the cycle, and often inject at different points in the cycle based on the scenario to achieve different combustion profiles. Combustion profiles have a strong bearing on engine efficiency, emissions, and NVH, so it is critical to identify engine parameter configurations that achieve profiles that meet criteria. When designing and calibrating engines, combustion profiles are often experimentally observed using in-cylinder pressure sensors, described in the following subsection [7].

2.2.2 In-Cylinder Pressure Sensors and Indicated Combustion Analysis

In-cylinder pressure sensors allow for direct measurement of cylinder pressure, p_{cyl}, as a function of crankshaft position, θ_{crank}. In traditional internal combustion engines, the crankshaft and piston are connected using a crank-slider mechanism, and so knowledge of the physical geometry can be used to translate the crankshaft position into the volume of the cylinder, V_{cyl}. Combined with thermodynamic relations, the cylinder pressure and

volume information can be used to calculate and model a variety of parameters useful for characterizing the combustion profile and measuring engine efficiency [7].

Integrating the cylinder pressure and change in volume over the cycle can be used to calculate the Indicated Mean Effective Pressure (IMEP), a measure of the amount of torque a particular combustion event produced normalized by the displacement volume. Shown in (2.1), the two values are integrated as a function of crank angle (θ_{crank}). For a 4-stroke engine (two crankshaft revolutions per power stroke), the integration bounds are from -360°, defined as the start of intake stroke (air entering the cylinder), to 360°, defined as the end of the exhaust stroke (exhaust exiting the cylinder), which encompasses the complete cycle [7].

$$IMEP = \int_{-360^{\circ}}^{360^{\circ}} p_{cyl}\left(\theta_{crank}\right) dV\left(\theta_{crank}\right) \tag{2.1}$$

In multi-cylinder engines, IMEP can be used to identify cylinder-to-cylinder variability issues that may be hard to identify if only using a single crankshaft brake torque measurement, or Brake Mean Effective Pressure (BMEP) [7].

Analyzing the cylinder pressure-volume (P-V) trace at the crank angle level can be used to estimate how the injected fuel combusted in the cylinder. Mass Fraction Burned (MFB) traces vary from 0 to 1, where 0 indicates 0% of the fuel has burned, and 1 indicates 100% of the fuel has burned. A common technique for estimating the MFB trace is the method developed by Rassweiler and Withrow, where the measured pressure in the cylinder is compared to the expected pressure due to compression/expansion, which can be used to estimate the energy released due to combustion as a function of crank angle. The estimation developed by Rassweiler and Withrow is shown in (2.2), where x_b is the mass fraction burned as a function of crank angle θ_{crank}, $p_{cyl,0}$ and V_0 are the cylinder pressure and volume at the start of combustion, $p_{cyl,f}$ and V_f are the cylinder pressure and volume at the end of combustion, p_{cyl} and V are the cylinder pressure and volume at the current crankshaft position, while n is the polytropic coefficient [7].

$$x_b\left(\theta_{crank}\right) = \frac{p_{cyl}\left(\theta_{crank}\right)^{\frac{1}{n}} V\left(\theta_{crank}\right) - p_{cyl,0}^{\frac{1}{n}} V_0}{p_{cyl,f}^{\frac{1}{n}} V_f - p_{cyl,0}^{\frac{1}{n}} V_0} \tag{2.2}$$

Although the presented model is simple, it can be augmented with additional models for heat transfer, vaporization of fuel, and flows during the combustion process to obtain a more accurate result. In diesel combustion, mass fraction burned traces can be used to identify the ignition delay (τ_{ign}, time between start of fuel injection and start of autoignition),

the ratio of autoignition combustion to diffusion combustion, and the overall combustion phasing (when combustion is occuring in the cycle). Combustion phasing is sensitive to many parameters, including parameters that can be directly controlled (boost pressure, EGR rate, injection timing/quantities), environmental parameters (ambient air temperature and relative humidity), engine specific parameters (engine aging, injector drift), and fuel properties (cetane rating, chemistry) [7].

A common metric to characterize combustion phasing is the value CA50, the crank angle at which 50% of the fuel in a cycle has burned ($x_b(\theta_{crank}) = 0.5$). Engine experiments and modeling show that a CA50 value of about 10° after top dead center (aTDC) is desirable from a fuel efficiency perspective. Phasing combusion earlier causes the cylinder pressures during the compression stroke to increase which increases the work required to finish compressing the gas, and also causes an increase in the peak cylinder temperatures leading to increased heat transfer losses, both of which lead to a reduction in net work. Phasing combustion later reduces the cylinder pressure during the expansion stroke, reducing the amount of work extracted which leads to a reduction in net work and higher exhaust gas enthalpies at exhaust valve opening (EVO) [7].

Figure 2.1 shows experimental engine data where the combustion phasing was retarded from the nominal value by retarding the injection timing (by injecting fuel later in the cycle, combustion is delayed later into the cycle) while holding indicated torque (IMEP) constant by injecting more fuel when the phasing was retarded. The Pressure-Crank Angle (P-θ) and Pressure-Volume (P-V) diagrams (subplots a and b) illustrate the combustion phasing effects, where the later combustion phasing causes a reduction in peak cylinder pressures but higher cylinder pressures during the expansion stroke, necessary to extract the same amount of work as the nominal phasing trace. The exhaust gas enthalpy benefit is visible from the higher pressure at the end of the expansion stroke, just prior to exhaust valve opening. The cumulative net heat release (HR) trace (subplot c) illustrates how combustion occurs later in the cycle. Note that this data was estimated without considering heat transfer effects, and so the values are slightly negative prior to start of combustion due to heat transfer losses. The efficiency impact of the retarded combustion phasing is also visible in this plot; more heat must be released (more fuel injected, combusted) to maintain the same indicated torque (IMEP).

In addition to increasing exhaust gas enthalpy, retarded combustion phasing can also increase the probability of partial burns (only a portion of fuel injected combusts) and misfires (fuel does not combust) that can result in large variability of the torque being produced by each cylinder in the engine. Excessive cylinder torque variability can lead to NVH and emission issues that need to be avoided [7]. Referring back to Figure 2.1, subplots

d and e show the increase in the variability of the cylinder torque when combustion phasing is retarded. Combustion variability is often measured using the Coefficient of Variation of IMEP (COV_{IMEP}), the ratio of the standard deviation of IMEP over the mean of IMEP ($COV_{IMEP} = \frac{\sigma}{\mu}$). This ratio is preferable over absolute standard deviation as it provides a measure of relative variability at a given operating condition. It has been shown that vehicle driveability problems appear when cylinder COV_{IMEP} values rise above 2-5 %, depending on the engine operating point [7].

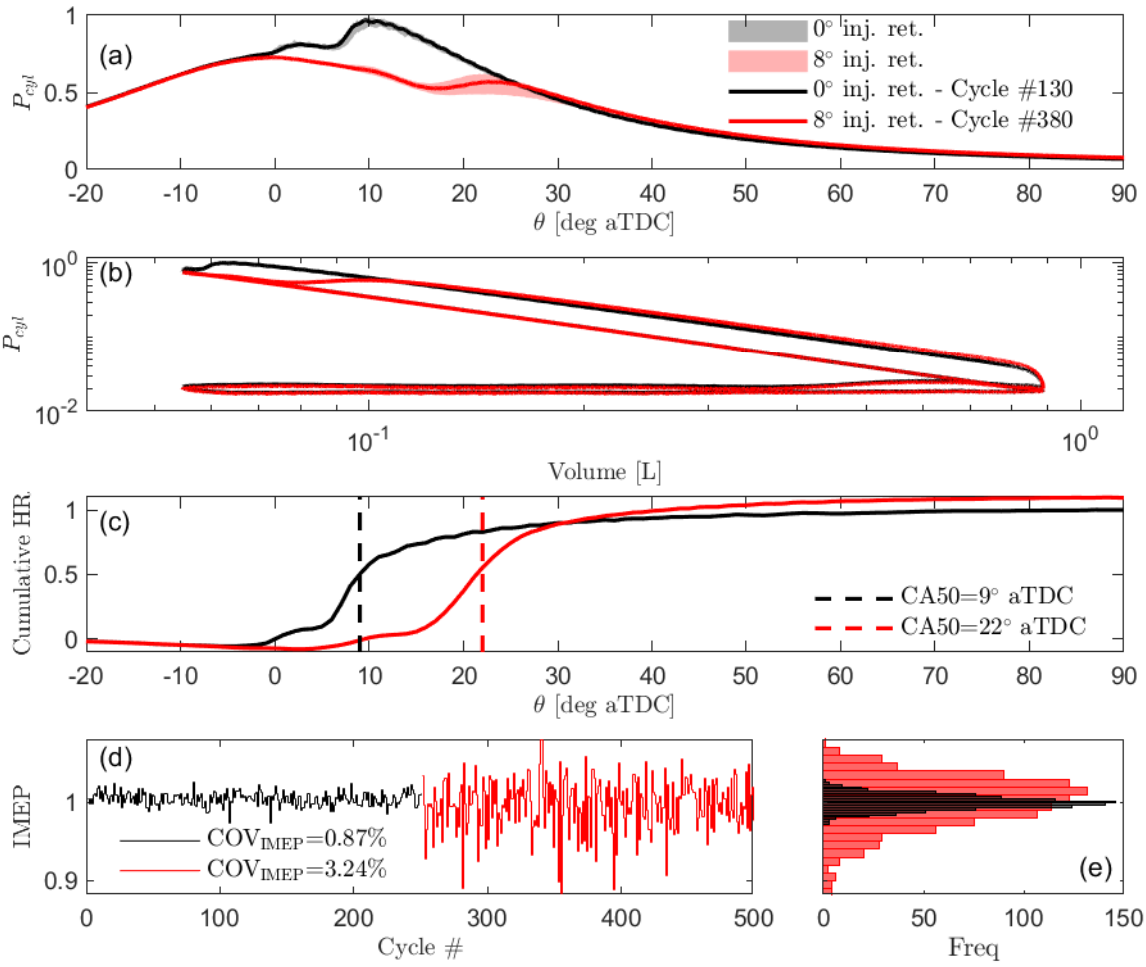


Figure 2.1: Plots of nominal vs. retarded combustion phasing behavior at constant mean IMEP. (a) Normalized P-θ diagram. (b) Normalized P-V diagram. (c) Normalized cumulative heat-release diagram. (d) Normalized IMEP timeseries and (e) histogram plots highlighting cyclic variability. For subplots (a)-(c), the value is normalized by the maximum nominal value (0° inj. ret). For (d) and (e), the values are normalized by the average nominal value.

The combustion features that can be calculated using cylinder pressure sensors can also be used for control. Closed-loop combustion control using in-cylinder pressure

sensor feedback has been explored on gasoline and diesel engine applications since the 1950's [9]. Early research demonstrated the ability to achieve efficiency benefits by controlling cylinder pressure features and combustion phasing to theoretical optimal values [10, 11]. Recent work has demonstrated the ability to perform real-time simultaneous control of cylinder torque (IMEP), combustion phasing (CA50), and max rate of heat release (ROHR, found by differentiating mass fraction burned data) on a multi-cylinder diesel engine, resulting in reductions of cylinder-to-cylinder variability (NVH benefit), NO_x emissions, and particular matter emissions [12]. Cyclic NO_x emission estimation using cylinder pressure and other engine measurements has been demonstrated in a research setting, paving the way for potential improvements in aftertreatment control relative to a system that only uses downstream NO_x sensors [15].

Although research using cylinder pressure sensors for engine control has demonstrated promising results, issues with sensor cost, reliability, and physical placement in the engine has restricted their traditional usage to research and calibration applications. Recent emission standards, advances in sensor technology and packaging, and extended sensor applications (combustion control, emission modeling) have led to limited OEM adoption of in-cylinder pressure sensors in consumer engines [13]. Stricter requirements and improved sensor value (decreased cost, increased reliability, and novel applications) will likely lead to widespread adoption of the sensors.

2.2.3 Exhaust Heating Mechanisms

There are several techniques available for improving exhaust heating of aftertreatment systems. They can be sorted into two categories - techniques which use engine actuators to increase exhaust temperatures, and techniques which directly add heat to the exhaust upstream of the aftertreatment system. The presented techniques are summarized in the pugh chart shown in Figure 2.2.

Engine-based exhaust heating techniques focus on increasing the enthalpy of the exhaust gas exiting the cylinder. Novel engine actuators include variable valve timing (traditionally not found on diesel engines) to increase exhaust gas enthalpy by reducing the amount of expansion work (using early exhaust valve opening, EEVO) and/or retarding combustion phasing (internal EGR via negative valve overlap, NVO) [6, 7]. Although both of these ideas show promise in lab testing, the additional actuators and sensors increase system cost and calibration complexity. Modern diesel engines are typically equipped with piezo-electric injectors capable of multiple injections per cycle, an exhaust gas recirculation (EGR) system, and a variable geometry or wastegated turbocharger (VGT, WGT). As

Technology	Exhaust Exotherm Source	Potential non-CO2 Emission Impacts	Fuel Impacts	Combustion Variability	Required Hardware
Electric Heater	Direct heating Electric	Neutral	Electrical load - alternator eff.	Neutral	Electric heater + wiring
Fuel Burner	Direct heating Fuel	Neutral	Fuel to burner	Neutral	Injector + igniter + wiring
Cylinder De-act. Valvetrain	Inc. cylinder loads Reduce airflow	Increased NOx (Inc. cylinder loads)	Benefit [Gross + Pumping]	Neutral	Cylinder Deactivation System
Cylinder Deact. Fuel Cutoff	Inc. cylinder loads	Increased NOx (Inc. cylinder loads)	Benefit [Gross + Pumping]	Neutral	Existing Actuator
VVT - NVO [iEGR]	Cyl heat to exhaust Phasing retard	Higher HC, CO, & PM (AFR & phasing retard)	Inc. pumping work Lower expansion efficiency	Can increase	Variable valve system
VVT - EEVO [Early EVO]	Cyl heat to exhaust Early EVO	Increased NOx (Inc. cylinder loads)	Lower expansion efficiency	Neutral	Variable valve system
Intake Throttling	Reduce airflow	Higher HC, CO, & PM (AFR)	Inc. pumping work	Neutral	Existing Actuator
EGR	Cyl heat to exhaust Phasing retard	Higher HC, CO, & PM (AFR & phasing retard)	Inc. pumping work Lower expansion efficiency	Can increase	Existing Actuator
Inj. Timing	Cyl heat to exhaust Phasing retard	Higher HC, CO, & PM (AFR & phasing retard)	Lower expansion efficiency	Can increase	Existing Actuator

Figure 2.2: Summary of reviewed exhaust heating methods.

noted in the previous subsection, retarding injection timing can cause combustion to occur later in the cycle, increasing the enthalpy of the exhaust gas. Additional injections placed later in the expansion stroke, commonly referred to as "early late-injections" or "post-injections", can be used to further raise exhaust temperatures in a more efficient manner (avoid generating unwanted torque and high peak temperatures which contribute to heat transfer losses) [8, 33]. Increasing the amount of EGR can also phase combustion later, yielding similar exhaust enthalpy increases. Reducing the amount of work extracted across the turbocharger turbine (by opening wastegate or VGT actuator) can increase the enthalpy of exhaust gas reaching the aftertreatment system [7]. Besides decreasing exhaust gas enthalpy due to work extraction, there is also a reduction of exhaust gas enthalpy across turbochargers due to heat transfer to the turbine housing, which can have a large thermal capacity. As a result, there have also been studies involving turbocharger turbine bypass systems, which redirects some or all of the exhaust flow around the turbine to avoid these losses [3]. Some engine architectures are also equipped with intake throttles for EGR control, which have been investigated to increase exhaust gas enthalpy by reducing total airflow through the engine (due to the excess air acting as a thermal dilluent) [8]. Note that while all of these actuators can be used to increase exhaust gas heating, the heating will come at the expense of fuel consumption and varying impacts on emissions and combustion variability [7,34,35]. Combustion variability is of particular concern due to the previously mentioned vibration and harshness issues that could upset the vehicle occupants. Production control strategies traditionally use special calibrations during cold-start to leverage the mentioned actuators (injection strategy, EGR, VGT/WGT) to increase exhaust

heating while avoiding excessive degradation of other attributes, but the calibration is typically conservative to account for the impacts of the numerous noise factors present in real world driving scenarios.

Direct catalyst heating methods include a fuel burner or electric heater upstream of the catalysts, which allow for direct heating of the catalyst and have the added benefit of avoiding the need to heat the rest of the metal exhaust system (between the engine/turbo and aftertreatment system) [3, 4, 32]. General Motors investigated a combination hardware configuration which used an electrically heated catalyst (EHC) in front of the diesel oxidation catalyst (DOC) with an upstream hydrocarbon (HC) injector and mixer. The EHC enables faster light-off of the DOC, while the upstream HC injector allows for direct HC dosing of the DOC for exotherm generation to heat the components after the DOC (SCR, DPF). Although HC's for DOC exotherm can be generated by the engine (using very late injections), supplying the HC's directly upstream of the catalyst avoids HC condensation issues that can adversely impact engine actuators (EGR valve) [33, 36].

The goal of this work was to develop a more aggressive exhaust heating strategy; to create a containable scope of work, it was decided to focus on a limited number of actuators. Injection timing and quantity modification were an attractive option for further study because they are existing actuators, they have direct impact on combustion phasing (impacts on auto-ignition process), and the engine control software supported changes to the values. Based on the literature review, the two big concerns associated with altering injection parameters include combustion variability and emissions, which were studied experimentally at conditions of interest using the experimental setup presented in the next section.

2.3 Experimental Hadware Configuration

A Ford 6.7L V8 Powerstroke diesel engine was characterized in a dynamometer lab for this study. Information about the engine can be found in Table 2.1. Prototype engine control unit (ECU) software was used which allowed for injection timing and quantity offset modifications (from base values).

Each cylinder in the engine was instrumented with an AVL GH14P direct-mount in-cylinder pressure transducer, capable of measuring cylinder pressures up to 250 bar at average accuracy of ± 0.0649 bar (max individual cylinder error was ± 0.13 bar). A 50 kHz low-pass filter was applied to remove the pressure noise typical of diesel combustion. The transducers were sampled at a crank angle resolution of one-tenth of a degree using an AVL IndiCom system. Air, exhaust, fuel, and coolant temperatures and pressures were

Table 2.1: Diesel engine specifications.

Manufacturer	Ford Motor Company
Production Year	2015
Engine Family	Powerstroke diesel
Cylinders	8
Layout	"Hot" V
Displacement [L]	6.7
Rated power [kW]	328 kW @ 2800 RPM
Rated torque [Nm]	1166 Nm @ 1600 RPM
Bore [mm]	99
Stroke [mm]	108
Connecting rod length [mm]	177
Wrist pin offset [mm]	0.5
Compression ratio	16.2
Aspiration	Turbocharged (Garrett GT37 VGT)
Piston geometry	Bowl-in-piston
Fuel injection system	Bosch high pressure common rail
Injector location	Centrally located
Injector holes	8 nozzles
EGR system	Cooled high-pressure EGR Cooler-bypass Intake throttle valve
Aftertreatment Configuration	DOC Urea SCR DPF

measured throughout the engine. Air flow was measured using a laminar flow element, and coolant and fuel flow were measured using Coriolis mass flow meters. The engine coolant was routed through a coolant conditioning system that allowed for sustained operation at coolant conditions as low as 20°C to allow for characterization of coolant temperature effects. The engine oil is cooled using the engine coolant via an onboard heat exchanger.

Since the focus of this work is exhaust heating via injection retard, only retarded injection timings were considered and because injection retards will only require positive fuel offsets to maintain brake torque, only positive fuel offsets were considered. Unless otherwise noted, one thousand cycles of combustion data were recorded at each characterization point.

2.4 Combustion Characterization

Injection retard characterization experiments were conducted at a variety of speed, load, and coolant temperature conditions to understand impacts on cylinder torque, heat release, and emissions. Observations from the IMEP characterization motivated the development of a closed-loop variability control architecture to maximize heating while avoiding excessive variability. Data from a fixed coolant temperature, speed, and load sweep was used to build a control-oriented model to enable controller design and offline testing for the proposed closed-loop controller.

2.4.1 Challenges of Combustion Variability Characterization

When performing characterization experiments at lower coolant temperatures ($< 60^{\circ}$C) , repeatability/hysteresis issues were observed that did not present during nominal engine operation (coolant temperature near 90°C). In an attempt to understand these issues, a long duration characterization test was conducted at 1200 RPM / 1.5 bar BMEP / 50°C coolant temperature. The test began with $\approx$ 2000 seconds of steady-state data at the nominal injection timing, followed by $\approx$ 6800 seconds of steady-state data at an injection retard of 17.5°. A main injection fuel offset was added to restore the brake torque to 1.5 bar BMEP following the injection retard; this fuel offset was kept constant for the duration of the experiment. Cylinder pressure data was collected every $\approx$ 140 seconds (limitation of data acquisition system). During the experiment, coolant temperature was controlled to $50 \pm 0.25^{\circ}$C and the oil sump temperature was maintained at $62.2 \pm 0.25^{\circ}$C. Brake torque, cylinder torque, and exhaust temperature data are shown in Figure 2.3.

The first plot shows the brake torque, which is very stable at the nominal condition, but noisy with retarded injection timing. The green line shows a moving average of the brake torque, showing the mean value gradually increasing as time passes. The second and third plots show the statistics of IMEP - note the slight increase in the mean IMEP value, μ(IMEP), and reduction in combustion variability, σ(IMEP), that occurs over the 7000 second experiment (2 hours). The gradual decrease in IMEP variability / increase in mean value explains the gradual increase in BMEP as the test progresses. The final plot shows exhaust temperatures during the test - the retarded injection timing is able to achieve a $> 100^{\circ}$C improvement in exhaust temperature, but at the expense of dramatic crankshaft torque variability. The reduction in combustion variability can also be observed by the gradual exhaust temperature increase during the test. Overall, the data presented in Figure 2.3 shows that when characterizing retarded combustion, long time-scale dynamics can be present that can take hours to properly resolve. Note that the coolant, oil, and

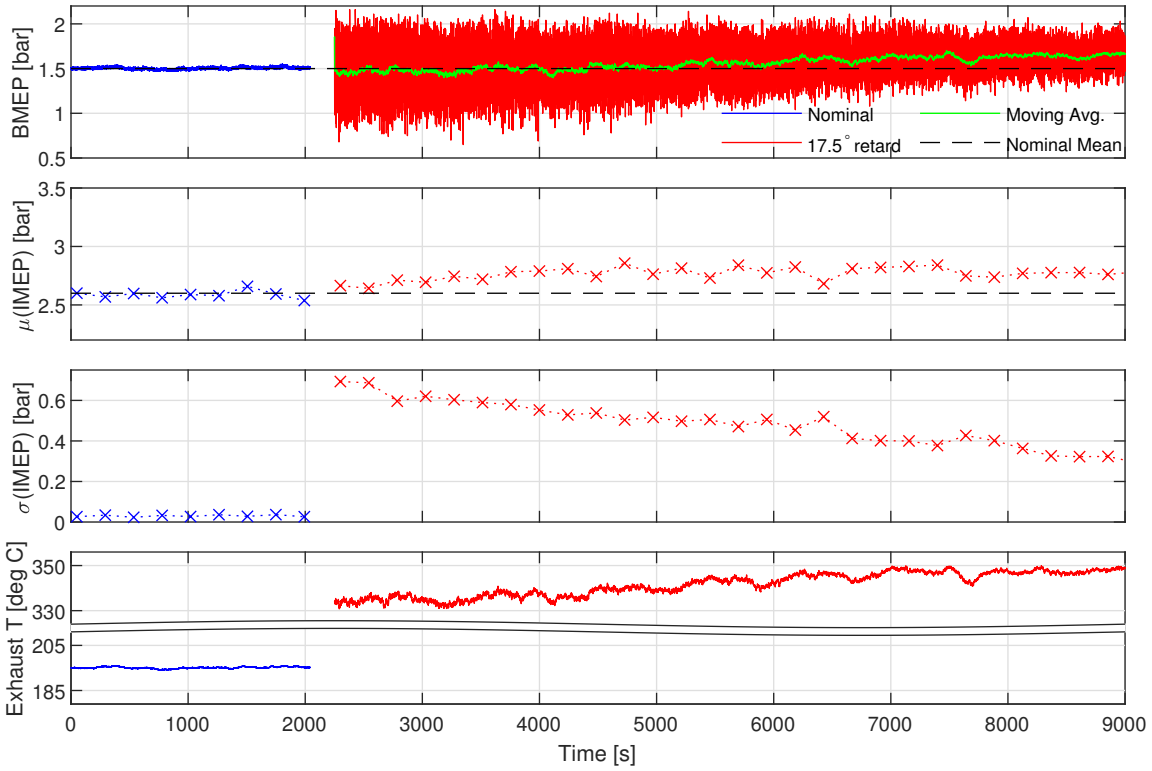


Figure 2.3: Data from long duration characterization test at 1200 RPM / 1.5 bar BMEP / 50°C coolant temperature. First plot shows BMEP; the blue line is at nominal injection timing, red line at retarded injection timing, black line showing the nominal mean value, and the green line is a moving average of the retarded injection timing data. The second and third plots show statistics of IMEP. The fourth plot shows the average exhaust temperature (note Y-axis break).

intake air conditions were closed-loop controlled - based on this, thermal dynamics in the engine block / cylinder bore liner may be responsible for the observed behavior. Regardless, to avoid these issues, experiments were conducted as rapidly as possible (while ensuring proper coolant temperature, speed, load, and injection control) to avoid long-term variability reduction/stabilization behavior, as we are interested in the "worst" variability behavior.

2.4.2 Indicated Mean Effective Pressure

Statistics of IMEP were explored due to the direct impact of cylinder torque on crankshaft torque (BMEP = IMEP − FMEP, where FMEP quantifies friction losses). IMEP is computationally simple to calculate, as it is just the integral of the cylinder pressure over the cycle as shown in (2.1). Aggressive phasing retard strategies can be limited by potential

NVH issues due to unintentional crankshaft torque oscillations; direct monitoring of IMEP using cylinder pressure feedback could allow for detection and mitigation of these issues.

Figure 2.4 shows data from injection retard sweeps at 1200 RPM, 2.5 bar BMEP, and 20°C coolant temperature. The sweep in red was conducted at constant torque, where additional fuel was added to keep BMEP constant ("BMEP*"); the sweep in black was conducted at constant fueling ("Fuel*"). As expected, retarding the injection timing results in combustion occuring later in the cycle, reflected by the increase in CA50 values. The higher CA50 values during the constant BMEP sweep are due to the increased amounts of fuel, as it takes longer for the additional fuel to burn. When additional fuel is added to maintain brake torque, the retarded combustion yields hotter exhaust temperatures, but when fuel is held constant the exhaust temperatures aren't significantly impacted by the phasing retard. The percent increase in fuel required to maintain constant BMEP and the percent decrease in BMEP when fuel is held constant show similar trends. The data also highlights an important aspect of the heating strategy, the fact that larger increases in exhaust temperature require more fuel, and this interaction must be managed to avoid excessive degradation of fuel economy. The $\sigma(IMEP)$ and $\mathrm{CoV_{IMEP}}$ plots show

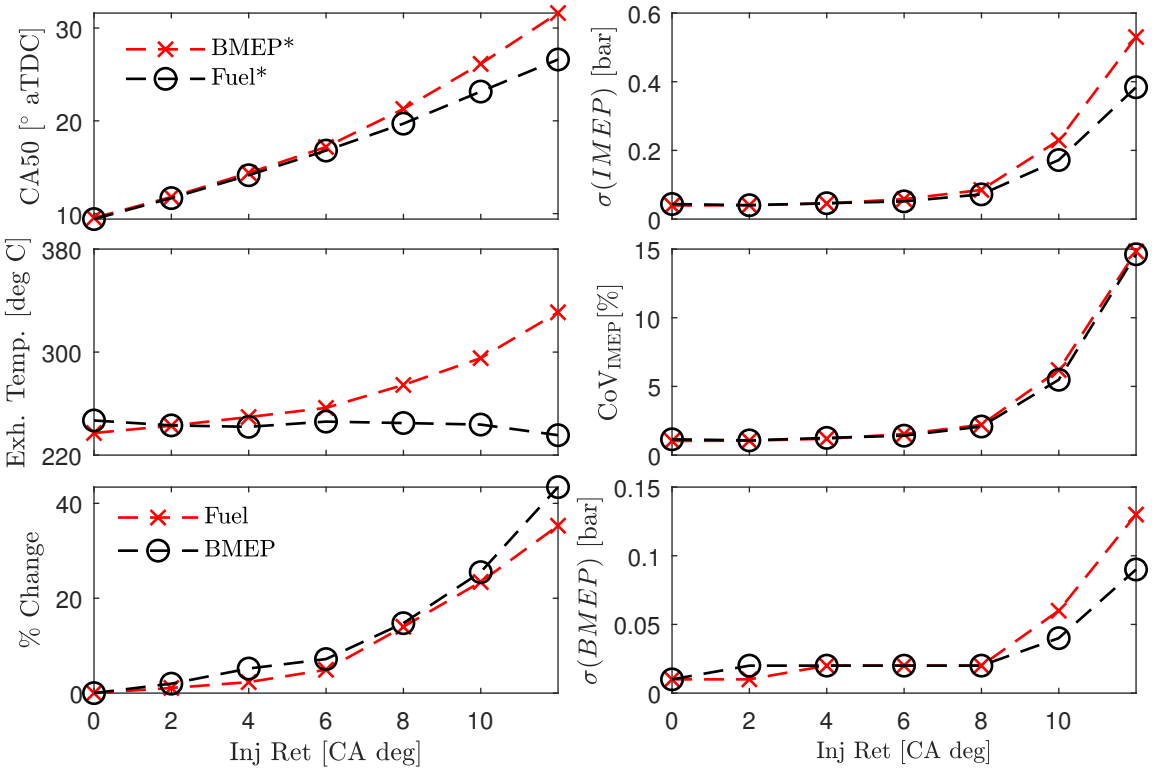


Figure 2.4: Mean value data from injection retard sweep at 1200 RPM, 2.5 bar BMEP, 20°C coolant temperature.

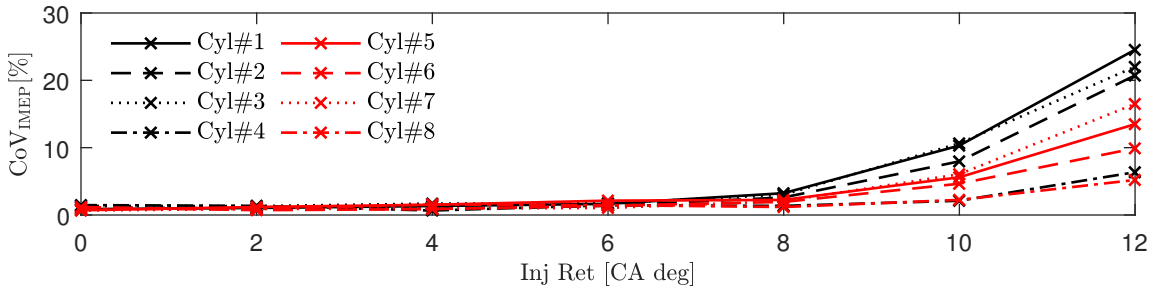


Figure 2.5: Individual cylinder CoV_{IMEP} data from constant BMEP sweep at 1200 RPM, 2.5 bar BMEP, 20°C coolant temperature.

the average cylinder behavior. Note that the variability trends for IMEP and BMEP directionally agree with each other, sensible due to their physical connection.

Also of interest is the exponential-like CoV_{IMEP} response to injection retard near 8 CA degrees, revealing one of the major shortcomings of an open-loop control strategy; it must be inherently conservative and operate away from that regime. Figure 2.5 shows the cylinder-to-cylinder variability present in the same engine at the same operating point. Looking at the data for cylinders #1 and #8, their CoV_{IMEP} values are nearly an order of magnitude apart, likely due to cylinder-to-cylinder air/EGR distribution mechanisms in the engine. This is an example of just one known noise factor; others include fuel properties (cetane number, heating value, viscosity, temperature), ambient/coolant temperature, and engine/injector aging. Even a calibration with plenty of noise factor maps must still leave margin for unknown phenomena.

Figure 2.4 plots $\sigma(\text{IMEP})$ and CoV_{IMEP} $(\frac{\sigma}{\mu})$, where the standard deviation operator best describes normally-distributed data. The IMEP data was further analyzed to determine if IMEP behaves as a Normal distributed random variable. Figure 2.6 plots the statistical behavior of IMEP at three injection timing retard values. The first column of plots shows the timeseries data; the axis are uniform to highlight the increase in variability. The second column of plots show histogram data in grey, with a normal fit in black. Note that at low levels of combustion variability, the experimental data resembles a normal distribution; at higher levels, a heavier left ("bottom") tail develops in addition to the appearance of high IMEP events. The heavier left tail is due to the poor combustion, while the high IMEP events are made possible by additional fueling while trying to compensate for the combustion variability impact on the mean torque value. The third column of plots are return maps, which assist in visualizing the interaction between consecutive combustion cycles to identify potential deterministic behavior, which would have large control implications. The skewed distribution shown in the histogram appears

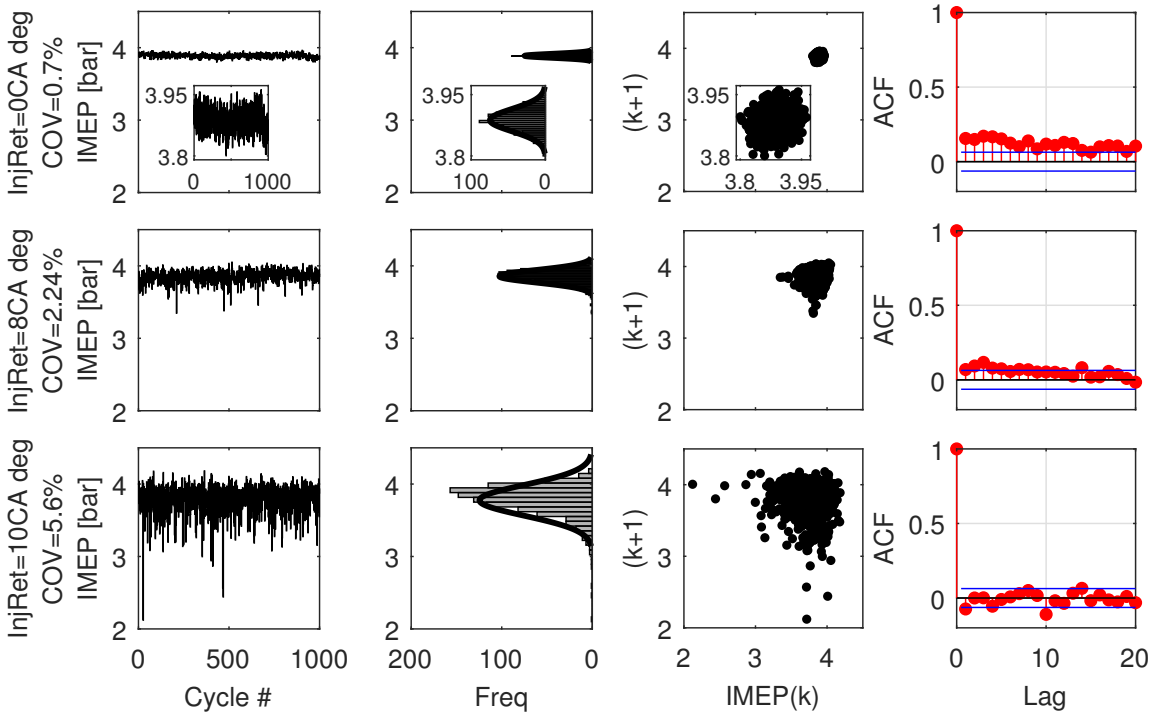


Figure 2.6: Detailed statistical data from constant BMEP sweep at 1200 RPM, 2.5 bar BMEP, 20°C coolant temperature.

as "legs" growing in the return map, but otherwise the plots indicate that there is little-to-no interaction between consecutive cycles. To further ensure deterministic trends are not present, the fourth column of plots show the Sample Autocorrelation Function (ACF). For the nominal and 8° retard points, the ACF values do exceed the 95% confidence bounds; from analyzing the timeseries data, a small change in the mean value over the duration of the test can be observed from start to finish. This was observed over several points, consistent with the observations from 2.4.1.

Overall, the more detailed statistical analysis indicates that in the region of interest ($\text{CoV}_{\text{IMEP}} < 2.5\%$), the IMEP sequence can be described as an independent Gaussian random variable sequence. This understanding is helpful from both a modeling and controls analysis perspective. Although interesting distribution behavior begins to develop at higher levels of variability, the risk of NVH-related consumer complaints from operating in that regime means it should be avoided.

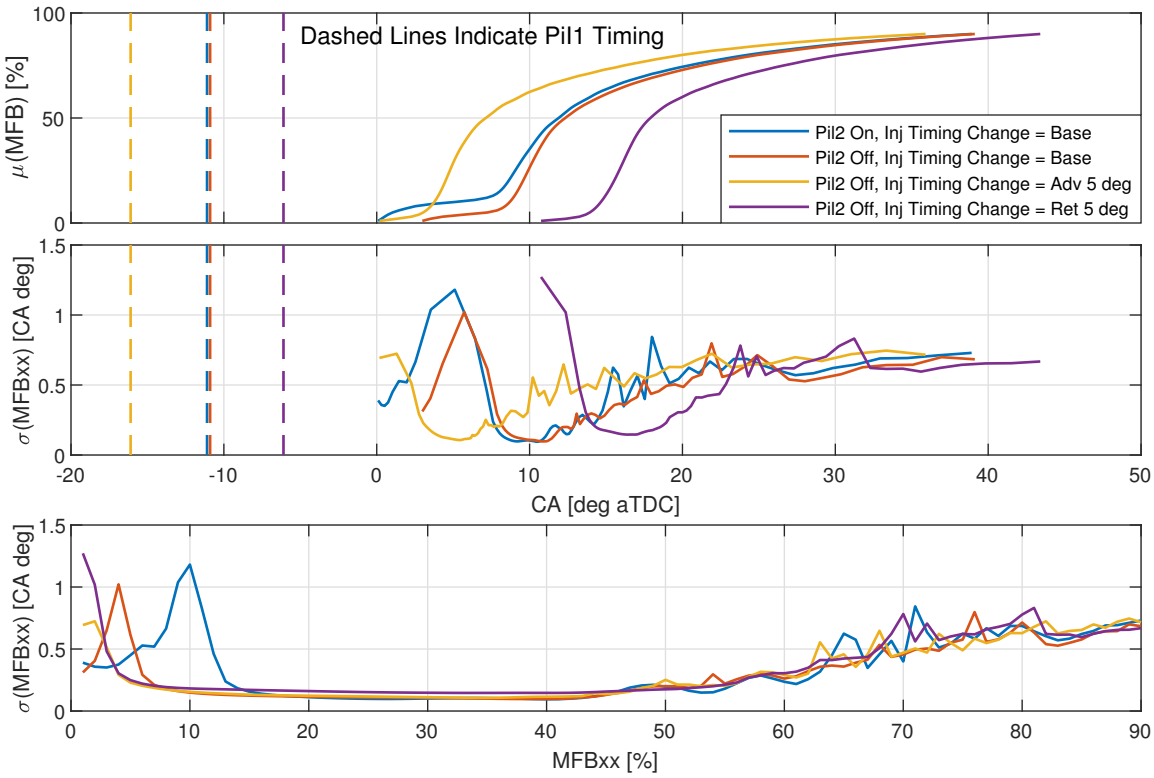


Figure 2.7: Fast heat release behavior at 1200 RPM, 2.5 bar BMEP, 20°C coolant temperature for different injection settings.

2.4.3 Net Heat Release

Fast heat release behavior was also studied, using Equation (2.2). The analysis was limited to the Rassweiler and Withrow method due to its computational simplicity, which is advantageous if trying to implement a real-time cyclic controller. Figure 2.7 shows the statistical response of mass fraction burn (MFB) values due to an additional pilot injection and injection phasing changes at 1200 RPM, 2.5 bar BMEP, 20°C coolant temperature. An interesting result generated by the heat release analysis was the variability in crank angle % burn values (MFBxx, crank angle where xx% of fuel has combusted) due to injection strategy and coolant temperature. The results indicate that since fast heat release does not account for fuel evaporation or heat transfer to the cylinder walls, one should use MFBxx values away from the boundaries (between 20%-80%) as they can be strongly impacted by the presence of pilot injections, changes in injection phasing, and other phenomena.

Designing a feedback controller using MFBxx values outside of this window (for example, MFB10) would require tuning the sensitivity transfer function to reject the noise. Maximizing controller performance would likely involve using techniques like gain-scheduling based on the anticipated impacts of injections on the parameter calculation, but

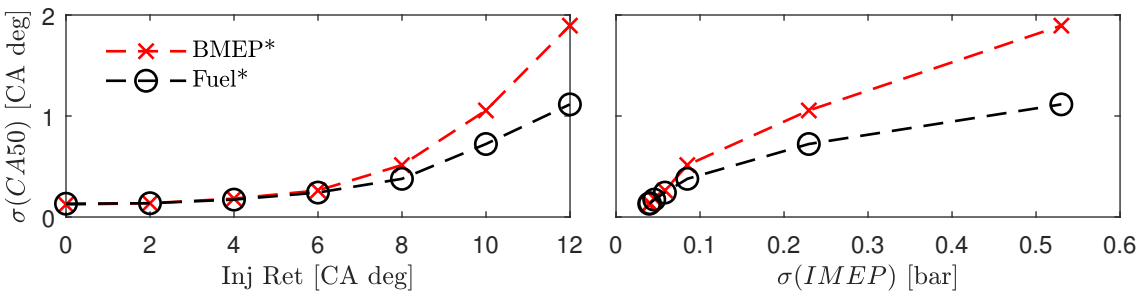


Figure 2.8: CA50 variability response to injection retard, and compared to the variability of IMEP, at 1200 RPM, 2.5 bar BMEP, 20°C coolant temperature.

would also need to account for other noise factors, comparable to the open-loop injection retard problem.

One might consider designing a controller around the variability of MFB50, which exhibit comparable trends to the variability of IMEP, as shown in Figure 2.8. Some data showing the mean-value changes in MFB50 for a variety of speed and load points is shown in Figure 2.9. One issue with such an approach is that retarding injection timing changes the mean value of MFB50, and so if those deterministic changes cannot be predicted, they can excite the variance estimate. The plot suggests that if a combustion metric control strategy were to be implemented, building something around the statistics of IMEP may be preferable compared to MFB50. This is due to the fact that in an ideal scenario, fueling would be compensated so that the engine load tracks the nominal value to avoid drivers observing a loss of engine power as injection timing is retarded. Note that even without fuel compensation, the driver could act as the torque control loop by pressing harder on the accelerator pedal. Another shortcoming stems from the fact that MFB50 is trying to capture the combustion behavior inside the cylinder, whereas IMEP is an actual measure of the torque being produced by the cylinder, potentially more useful when trying to contain torque-related NVH issues.

2.4.4 Hydrocarbon Emission

As noted in Figure 2.2, retarded injection timing can cause an increase in hydrocarbon (total HC, THC), carbon monoxide, and particulate matter (smoke) emissions. As of 2021 in the United States, carbon monoxide is not heavily regulated, and the widespread industry of adoption of diesel particulate filters (DPF) mean feedgas particulate matter emissions are less of a concern. The primary concern is hydrocarbon emissions as these will be created during the cold-start phase when the diesel oxidation catalyst is not functional. A reduction in catalyst light-off times at the expense of dramatic HC emission increases could

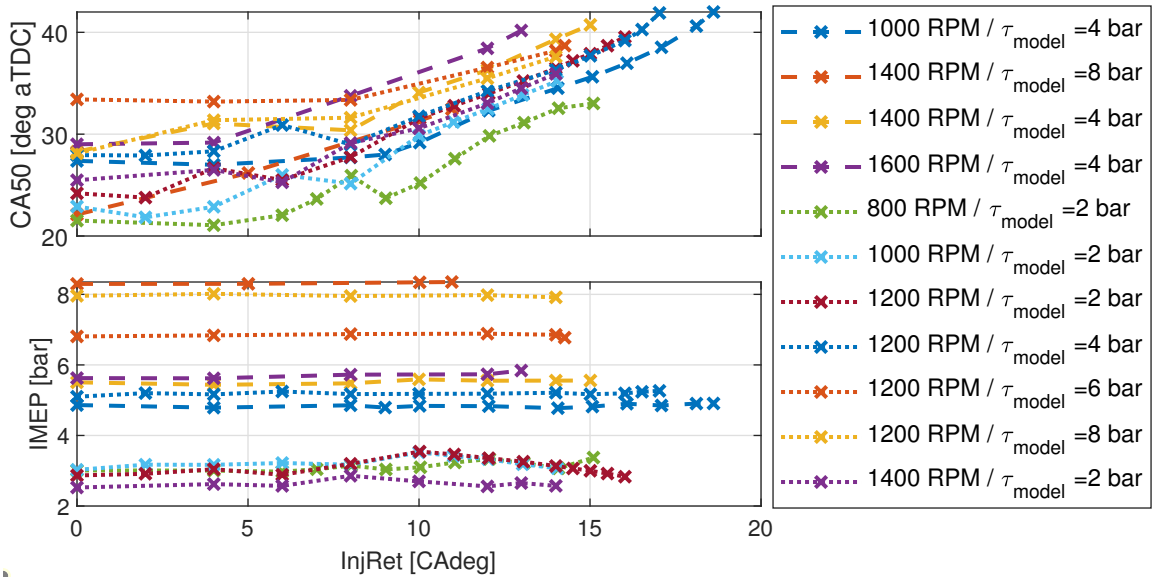


Figure 2.9: Injection retard sweeps conducted at multiple engine speed and loads at 30°C coolant temperature. Note that CA50 responds to speed, load, and injection retard, and these behaviors would need to be understood for feedback controller development.

potentially hurt regulatory compliance.

During the combustion characterization experiments, it was observed that each cylinder can have unique response behaviors (as seen in Figure 2.5), and so it would stand to reason that they would also have different levels of hydrocarbon emissions. To better understand HC emission behavior, mapping experiments were conducted using a Cambustion HFR500 (Fast FID, 1000 Hz response) and AVL SESAM FID (5 Hz response) for fast / slow HC emission measurements. The HFR500 was installed in the exhaust manifold downpipe shown in Figure 2.10, upstream of the turbocharger. Windowing of the fast HC data based on timing of exhaust valve events and exhaust transport delays through the exhaust manifold allow for isolating cylinder-to-cylinder HC emissions. Figure 2.11 shows crank angle resolved cylinder pressure and hydrocarbon emission data at 1000RPM, 4.6 bar BMEP, 5.7 bar IMEP, 30°C coolant temperature. The Air-Fuel Ratio (AFR) during the sweep remained above 20:1, above the 14.5:1 stoichiometric diesel ratio. For the plot on the left, the engine is operating at nominal injection timing, and the engine is emitting minimal HC. In the right plot, cylinder #6's injection timing is shifted 21 crank angle degrees later, causing extremely late phased combustion. While the hydrocarbons emitted by cylinder #6 reach levels exceeding 800 ppm, the CoV_{IMEP} of cylinder #6 was about 0.91%, suggesting that combustion torque variability can not be used as a hydrocarbon emission indicator.

An injection timing sweep on cylinder #6 was conducted at the 1000RPM/4.6 bar

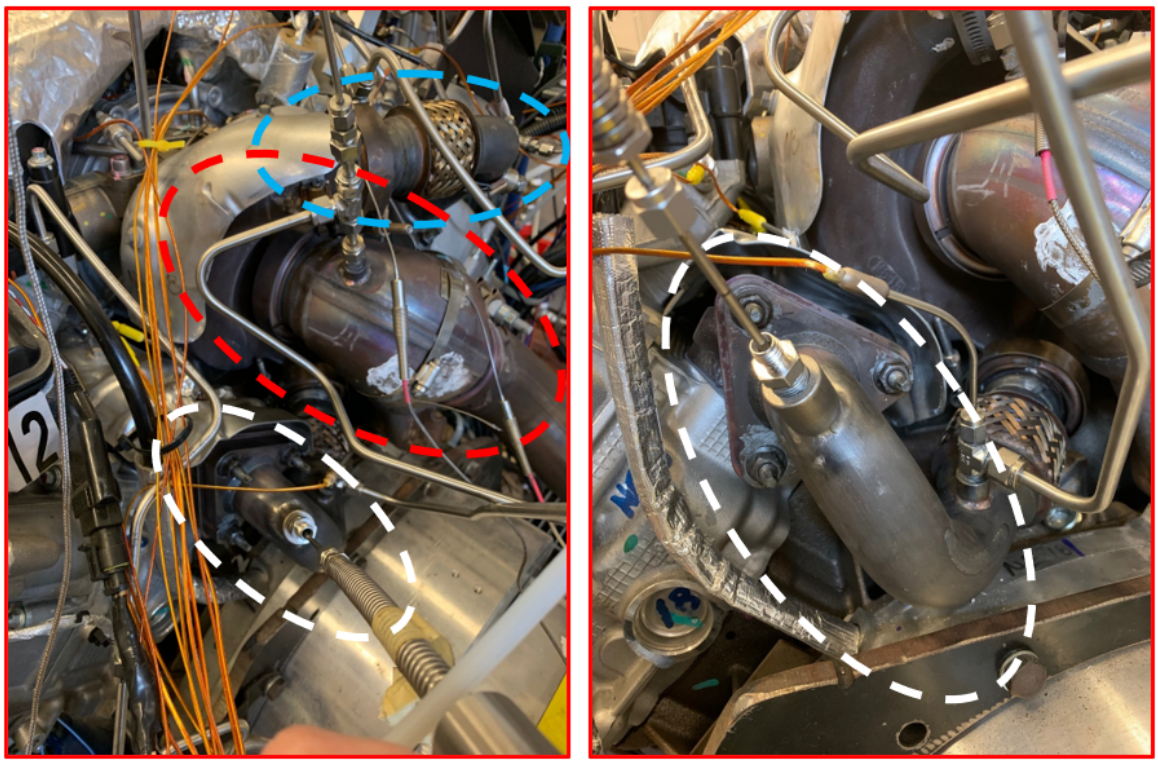


Figure 2.10: HFR500 probe install location. White-dashed line shows the turbo inlet pipe from right bank (cyl 5,6,7,8), blue-dashed line shows turbo inlet pipe from left bank (cyl 1,2,3,4), and red-dashed line shows the turbo downpipe. The probe tip was centered in the exhaust passage.

BMEP point while recording test cell fueling. Analysis of the heat release data revealed that increases in HC emissions sensibly correlate with decreases in apparent combustion efficiency (no heat transfer / fuel evaporation corrections), defined as:

$$\eta_{comb,apparent} = \frac{Q_{tot,apparent}}{V_{fuel}\rho_{fuel}Q_{LHV}} \tag{2.3}$$

where $Q_{tot,apparent}$ is the total apparent heat released, V_{fuel} is the volume of fuel injected for a particular cycle, ρ_{fuel} is the density of the fuel, and Q_{LHV} is the lower heating value of the fuel. Cycle-specific hydrocarbon emissions are plotted against the cycles apparent combustion efficiency in Figure 2.12.

This preliminary result suggests that online estimates of combustion efficiency could possibly be used to estimate HC emission formation to constrain an injection retard strategy, rather than an empirically identified MFB50 / injection offset limit map. To properly use this strategy, the fuel quantity estimation must be accurate – during dyno experiments, the ECU fuel flowrate estimation was found to be inaccurate at lower fueling

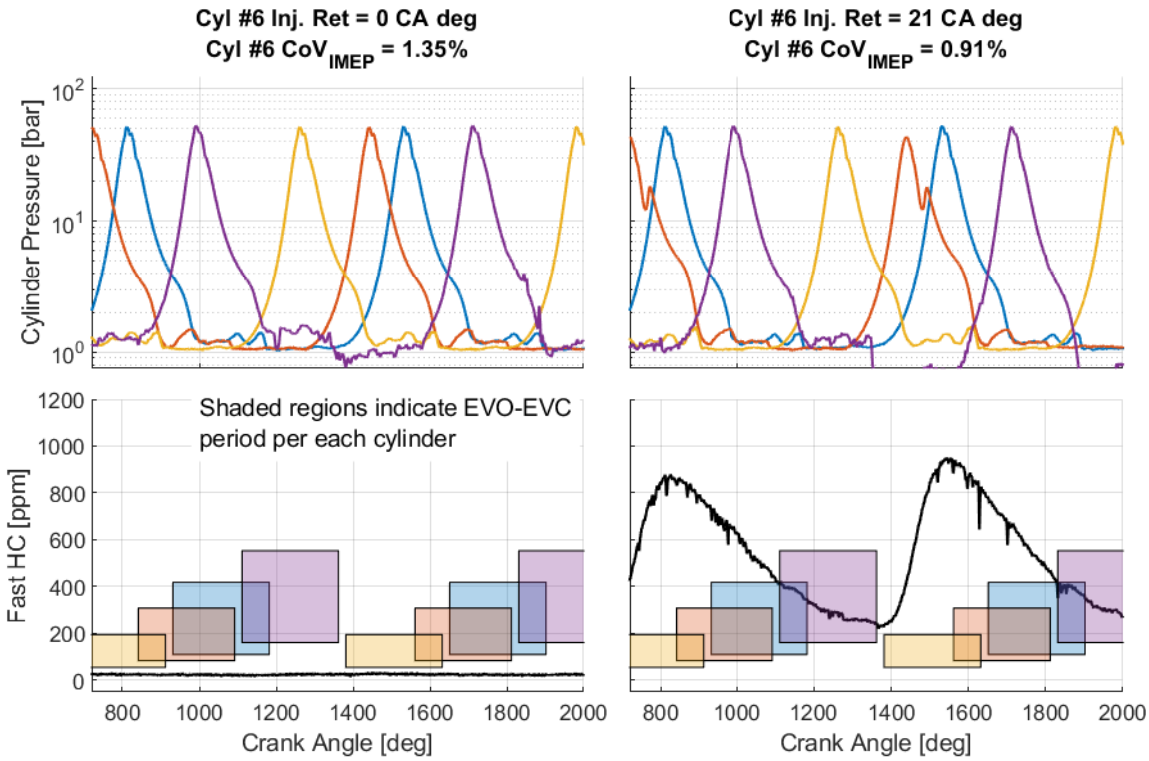


Figure 2.11: Crank-angle resolved hydrocarbon emission data at 1000RPM, 4.6 bar BMEP, 5.7 bar IMEP, 30°C coolant temperature at two injection phasing conditions for cylinder #6. In both cases, no combustion variability issues were present.

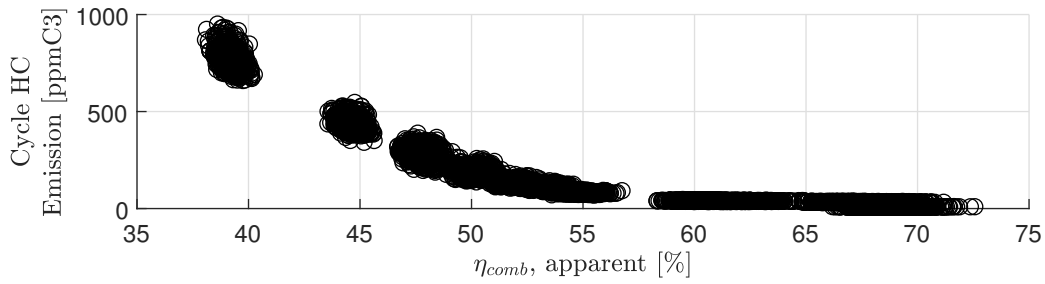


Figure 2.12: Cyclic hydrocarbon emissions plotted against cyclic apparent combustion efficiency. Note that the increase in cyclic hydrocarbons appears correlated with the decrease in apparent combustion efficiency.

rates. It would also be sensitive to variation in the fuel's lower heating value, which could be addressed by online estimation techniques. Another potential issue may be cylinders having unique HC vs. combustion efficiency responses, as well as sensitivity to other noise factors. This may be addressable by trying to empirically identify a cylinder which consistently emits the most HC and model constraints off that cylinder. Further work on

this topic is required to generate an implementable solution.

Engine average hydrocarbon emission data was collected at a variety of speed/load points at 30° C coolant temperature using the previously mentioned AVL SESAM bench. In these experiments, injection timings were offset for all cylinders. Figure 2.13 shows a variety of trends from the experiments. Excessive increases in hydrocarbon emissions were observed when injection timing was retarded more than 10 degrees. For lower loads (when torque model is less than 2 bar), increases in hydrocarbons correlate with increases in CoV_{IMEP}, desirable if one is only able to observe IMEP statistics. But for higher loads, as previously shown, large increases in THC emissions can occur without any significant increase in CoV_{IMEP}. Plotting the turbocharger inlet temperatures against the hydrocarbon emissions reveals that the best opportunities for increasing exhaust temperatures while avoiding hydrocarbons occur at higher engine loads, where the turbocharger inlet temperature versus THC slope is nearly vertical. Looking at the hydrocarbon emission trends versus combustion phasing, at higher loads the large increases in hydrocarbons tend to occur when CA50 is $> 37.5°$ aTDC. The last plot shows additional hydrocarbon emission data versus apparent combustion efficiency calculated using coriolis fuel flow measurements. The trends look encouraging, but more work is required to generate a functional solution.

Based on the empirical observations shown above, for the given engine architecture and calibration, it makes sense to constrain an injection retard strategy to avoid values exceeding 12 degrees, and CA50 values exceeding $\approx 37.5°$ aTDC. These conclusions are only applicable to this engine; changes in combustion chamber geometry (piston bowl), rail pressures, and other parameters could yield different combustion and emission trends, and so the development of a control strategy for a different engine would require similar characterization experiments.

2.5 Proposed Closed-Loop IMEP Statistical Controller

Based on the behaviors observed in Figures 2.3-2.13, pushing the limits of injection phasing retard exhaust heating would require an intense calibration effort, and would still need to be inherently conservative due to unknown phenomena as shown in Figure 2.14. It is instead proposed that an in-cylinder pressure sensor be used to provide IMEP variability feedback, and a controller is used to maintain the variability at the acceptable limit to allow for aggressive exhaust heating while preventing consequences of excessive variability, as shown in Figure 2.15. In addition, a fuel control loop would be implemented to reject the decreases in net indicated work due to the combustion phasing retard and ensure the desired

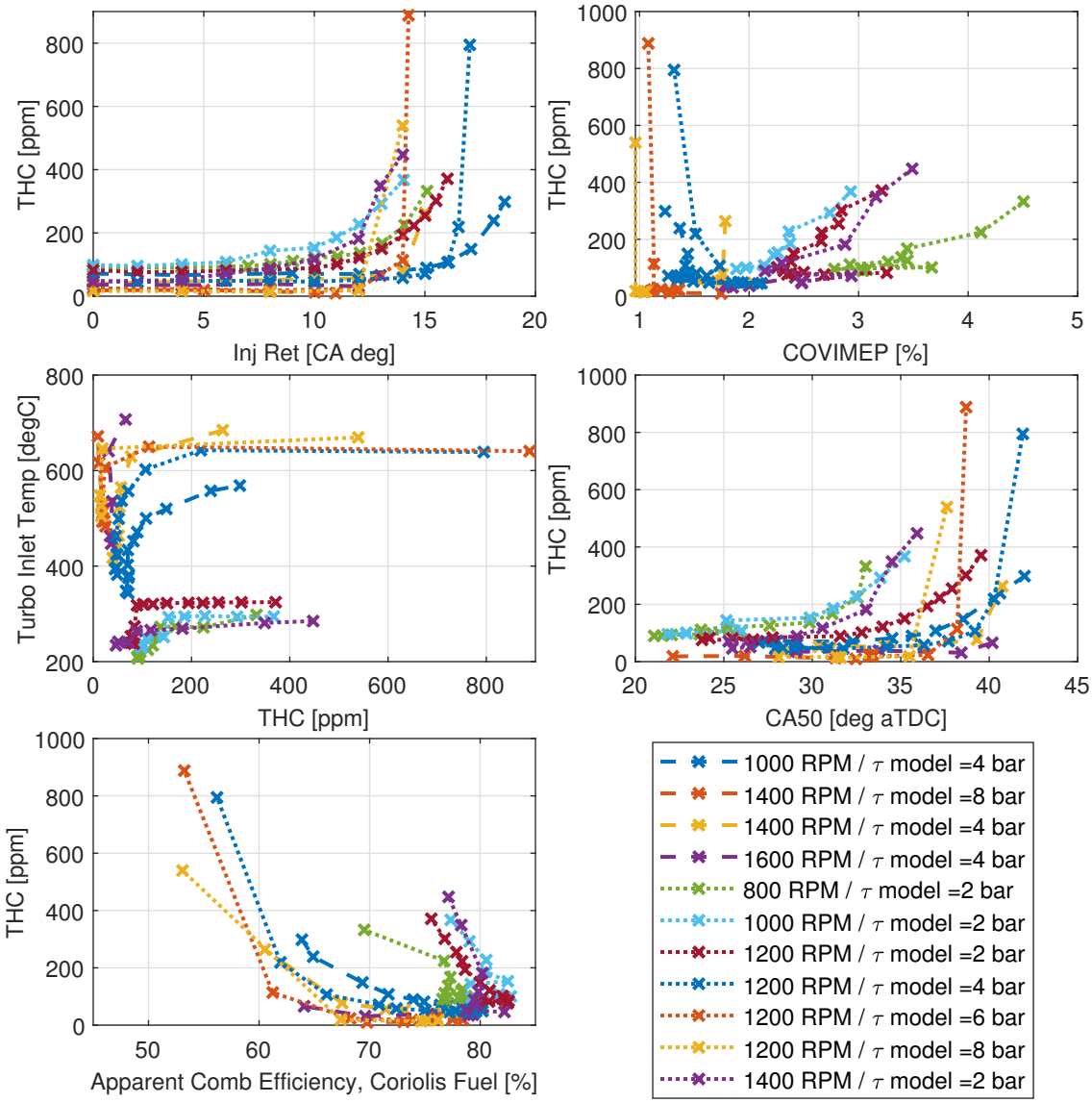


Figure 2.13: Hydrocarbon emission behaviors plotted against various experimental values.

exhaust temperature increase. Limits on MFB50 and/or maximum injection retard would limit excessive HC emissions. The large number of factors and individual cylinder behavior would be rejected by the feedback, significantly decreasing the calibration complexity, and unique cylinder behavior could be managed by controlling each cylinder independently.

Although cylinder pressure transducer prices have reduced to a level where production implementation is feasible, instrumenting each cylinder in a V8 engine is not attractive from a financial perspective. In Chapter 4, techniques for reducing the number of indicated cylinders are presented, with the evaluated vehicle implementation only using one cylinder pressure sensor.

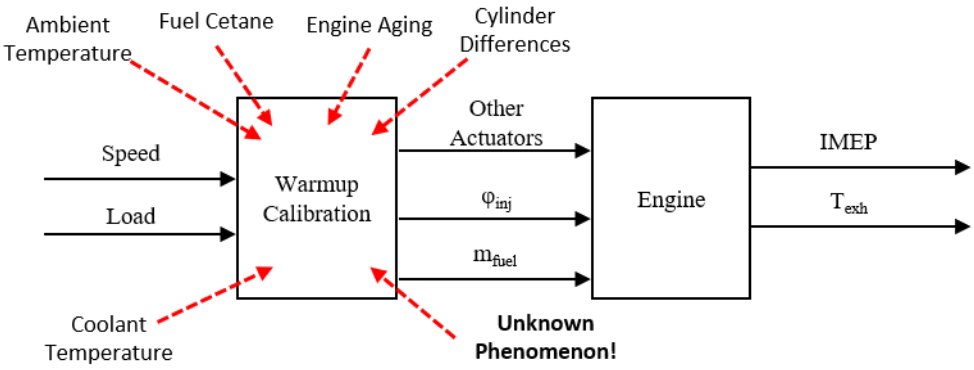


Figure 2.14: Diagram highlighting the challenges of an aggressive open-loop control injection phasing retard strategy.

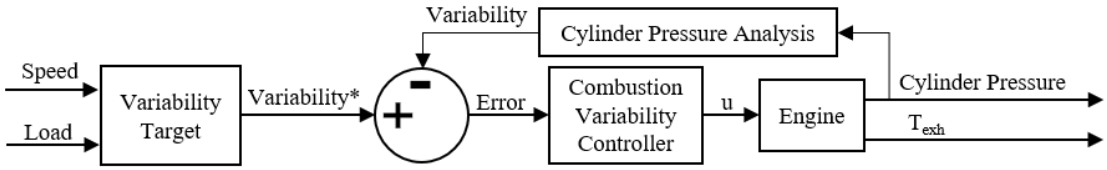


Figure 2.15: Proposed closed-loop structure.

2.6 Control-Oriented Combustion Variability Model

In order to develop a feedback control strategy, a control-oriented stochastic combustion model was created that captures the impacts of injection timing (φ) and fuel quantity (m^{fuel}) on combustion variability. The proposed model focuses on perturbations from an engine's nominal operating condition at a single speed and load point of interest (1200 RPM/2.5 bar BMEP/20°C coolant temperature) for an initial exploration of the feasibility of model-based variability controller design. Empirical engine data from injection timing and fuel quantity offset experiments were used to populate maps for the mean and standard deviation values of IMEP as a function of injection timing retard ($\Delta\varphi$) and fuel quantity offset (Δm^{fuel}), shown in equation form as (2.4)-(2.6).

$$\mu_k^{\text{IMEP}} = f_\mu(\Delta\varphi_k, \Delta m_k^{\text{fuel}}) \tag{2.4}$$

$$\sigma_k^{\text{IMEP}} = f_\sigma(\Delta\varphi_k, \Delta m_k^{\text{fuel}}) \tag{2.5}$$

$$\text{IMEP}_k = \mu_k^{\text{IMEP}} + \sigma_k^{\text{IMEP}} w_k, \qquad w_k \overset{iid}{\sim} \mathcal{N}(0,1) \tag{2.6}$$

Table 2.2: Normalized coefficients for empirical f_μ and f_σ models shown in Equations (2.7) and (2.8).

c_0	c_1	d_0	d_1	d_2
-1.08×10^{-2}	4.51×10^{-2}	9.09×10^{-4}	8.14×10^{-3}	8.63×10^{-1}
d_3	d_4	d_5	d_6	
8.06×10^{-3}	5.40×10^{-5}	9.21×10^{-1}	6.52×10^{-2}	

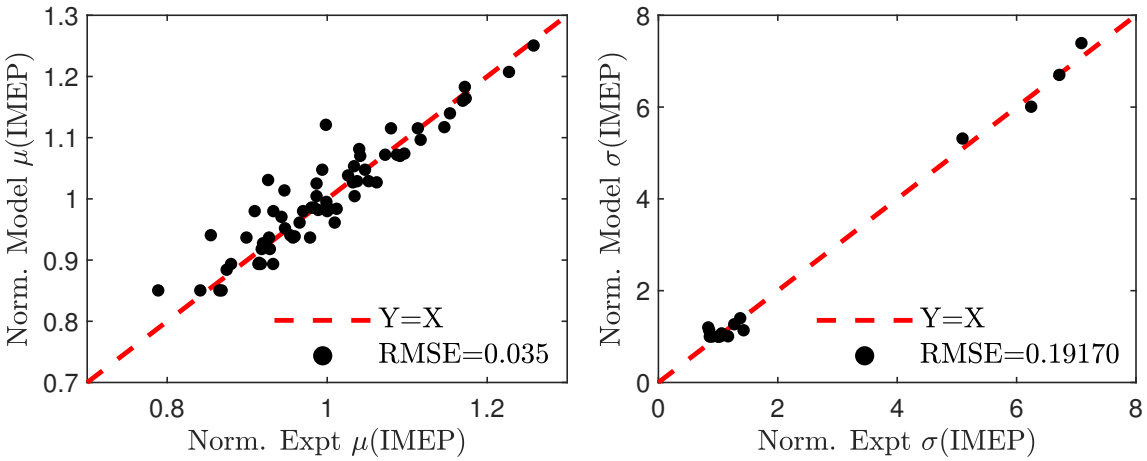


Figure 2.16: IMEP statistical model correlation to experimental data.

Note that w_k represents a sequence of independent identically distributed (*iid*) Gaussian normal random variables; this assumption is supported by the statistical behavior seen in Figure 2.6. A normalized version of the f_μ, f_σ relations is shown in Equations (2.7) and (2.8) below, with the normalized coefficients listed in Table 2.2. Note that $f_\sigma(\cdot)$ in (2.8) was fitted by a piecewise function due to the relatively linear response of σ_k^{IMEP} at small injection timings and its nonlinear response at more retarded conditions.

$$\mu_k^{\text{IMEP}} = 1 + c_0\Delta\varphi_k + c_1\Delta m_k^{\text{fuel}} \tag{2.7}$$

$$\sigma_k^{\text{IMEP}} = \begin{cases} (1+d_0\Delta\varphi_k)\left(1+\Delta m_k^{\text{fuel}}\Delta\varphi_k d_1\right) \text{ if } 0 \le \Delta\varphi < 8 \\ \left(d_2\exp(d_3\Delta\varphi_k) + d_4\exp(d_5\Delta\varphi_k)\right)\left(1+d_6\Delta m_k^{\text{fuel}}\right) \text{ if } \Delta\varphi \ge 8 \end{cases} \tag{2.8}$$

Figure 2.16 compares the experimental data to the model output, indicating the proposed equation structure does an adequate job at capturing the statistical trends. The regressions and accompanying model structure were implemented in MATLAB/Simulink for controls simulations, shown in Figure 2.17. "Random Noise" blocks in Simulink were used to

generate the w_k sequence, which can be parameterized with a random number generation seed value for consistent stochastic sequences. The "FIR/IIR Filtering Strategy" block refers to techniques for generating online variability estimates that are discussed in detail in Chapter 3, where FIR stands for Finite Impulse Response (windowed estimates) and IIR stands for Infinite Impulse Response (recursive estimate). The estimates are generated using values from the IMEP_k sequence.

An example of such an estimator would be the windowed CoV_{IMEP} estimate, $\hat{\text{CoV}}_{\text{IMEP}}$, which can be represented as (2.9)-(2.11), where N is the window size (number of IMEP values used to calculate statistics).

$$\hat{\mu}_k^{\text{IMEP}} = \frac{1}{N}\sum_{i=0}^{N-1} \text{IMEP}_{k-i} \tag{2.9}$$

$$\hat{\sigma}_k^{\text{IMEP}} = \sqrt{\frac{1}{N}\sum_{i=0}^{N-1} \left(\text{IMEP}_{k-i} - \hat{\mu}_k^{\text{IMEP}}\right)^2} \tag{2.10}$$

$$\hat{\text{CoV}}_{\text{IMEP}} = \frac{\hat{\sigma}_k^{\text{IMEP}}}{\hat{\mu}_k^{\text{IMEP}}} \cdot 100 \tag{2.11}$$

Figure 2.18 compares the model performance with respect to comparable engine experiments. Two conditions have been chosen for comparison: 1) stock injection strategy (red color) and 2) 8° injection retard from stock calibration (blue color). The top-left plots show the probability density function for the modeled and experimental IMEP. Note that the assumption of introducing Gaussian-distributed perturbation matches well with the experimental observations. The bottom-left plots show the return maps of IMEP for both injection timing. The "shotgun" pattern observed suggests a lack of cycle-to-cycle correlation in the IMEP time series. This aligns well with the assumption that the perturbation sequence $\{w_k\}_{k\in\mathbb{Z}}$ in the model is *iid*. The top-right plot shows the comparison of the IMEP time series and the bottom-right plot shows the comparison between the

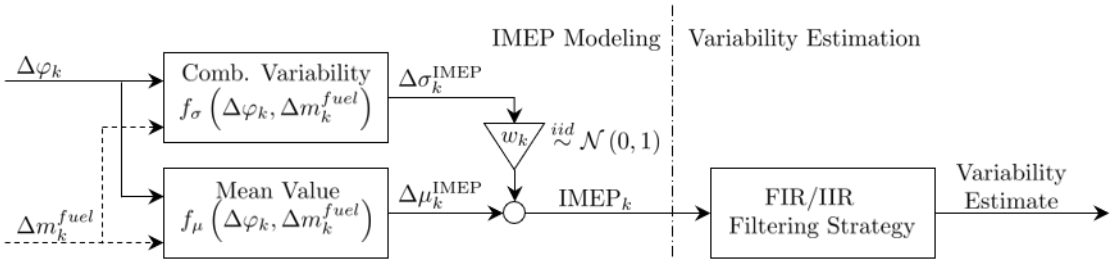


Figure 2.17: Proposed combustion variability model architecture for controller design and simulation.

estimated $\mathrm{CoV}_{\mathrm{IMEP}}$, $\hat{\mathrm{CoV}}_{\mathrm{IMEP}}$, when $N = 10$. Note the variability in $\hat{\mathrm{CoV}}_{\mathrm{IMEP}}$ - this is due to the fact that $\hat{\mathrm{CoV}}_{\mathrm{IMEP}}$ is a function of random variables, and so it is also a random variable, which will be further explored in Chapter 3. Overall, Figure 2.18 demonstrates the close agreement between experiments and simulations for the IMEP statistics.

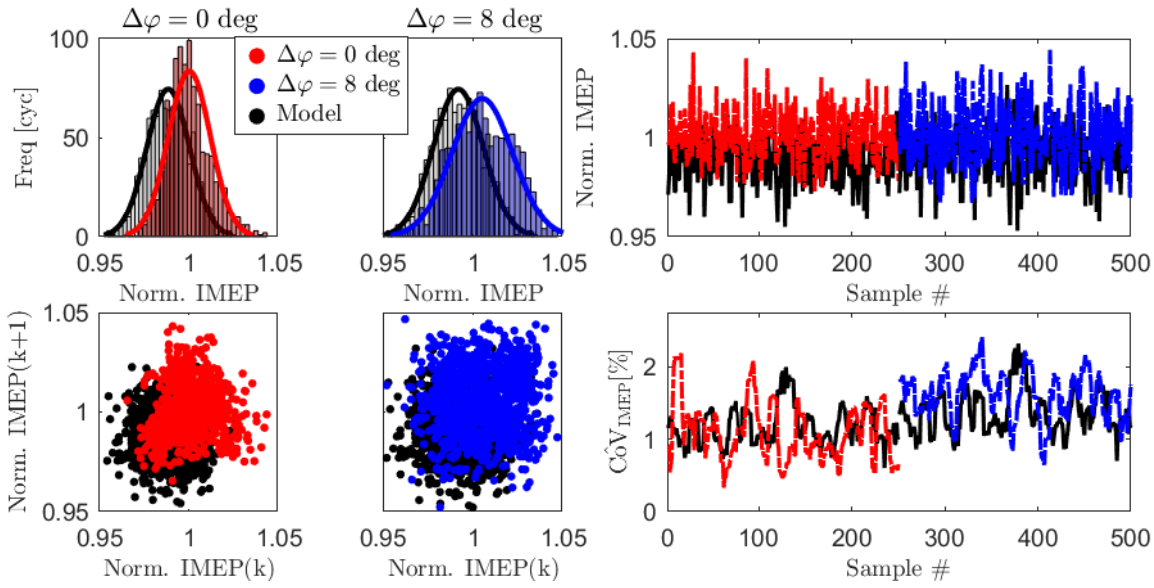


Figure 2.18: Staistical behavior of model versus experiments for two injection timing retard conditions. $\hat{\mathrm{CoV}}_{\mathrm{IMEP}}$ estimated using 10 point window ($N = 10$).

2.7 Summary and Conclusions

Analysis of upcoming emission regulations indicates the necessity for faster aftertreatment operation and/or a reduction feedgas engine emissions - achieving faster aftertreatment light-off allows for a faster transition to more efficient engine operation (lower fuel consumption at cost of increased $\mathrm{NO_X}$ emissions). Background on diesel engine controls, indicated combustion analysis, and exhaust heating mechanisms were presented. Combustion phasing retard via injection timing retard was evaluated using engine dynamometer experiments to identify how it could be used to improve exhaust heating. Impacts on IMEP, heat release, and hydrocarbon emissions were studied at a variety of speed and load points. The experiments showed that more aggressive combustion phasing retard could achieve better aftertreatment heating at the cost of combustion variability, which could cause poor drivability.

It was proposed that combustion variability be controlled to an acceptable level to achieve high exhaust heating while avoiding issues associated with excessive variability.

Experiments and literature suggested that open-loop combustion variability control was calibration prohibitive, and so a novel controller using combustion variability feedback from an in-cylinder pressure sensor was proposed. To aid in controller design, a stochastic combustion variability model was designed, populated with engine experiments, and validated against experimental data. The model is combined with the content presented in Chapter 4 to design and simulate controllers presented in Chapter 4.

CHAPTER 3

Variance Estimation and Control

3.1 Introduction

This Chapter uses statistical and control theory to explore online variance estimation and control using windowed variance estimators to form controller principles for the structure presented in Section 2.5. The stochastic nature of the windowed variance feedback means that root locus and other classical control design techniques cannot be immediately used. Expected value analysis is used to generate transfer functions for the average response of such systems under various assumptions, including when deterministic signal mean changes are occurring in the process to be estimated. This is critical when trying to estimate engine combustion variance, where deterministic changes in the torque occur as the engine torque demand changes. The results are used to design controllers, and stochastic simulations are presented to verify the analytical results and make other observations concerning closed-loop variance control. The windowed variance estimator requires maintaining a data buffer for calculating statistics - an exponentially-weighted moving average (EWMA) variance estimation technique that only requires two states is also studied to understand the advantages and disadvantages of such a technique.

This chapter is organized as follows: Section 3.2 provides background on variability-motivated controller design techniques. Section 3.3 analyzes the windowed variance operator using expectation analysis, and analyzes the open-loop behavior of windowed variance with consideration to usage for closed-loop control. Section 3.4 uses classical LTI design principles to design a feedback controller that achieves a target variance. Section 3.5 presents simulation results of the closed-loop system and provides insight on special considerations that must be made to achieve effective closed-loop variance control. Section 3.6 presents analysis of an EWMA estimation strategy.

3.2 Background

Typical control design methods are focused on minimizing the impact of process and measurement noise on the state covariance of the closed-loop system. For stochastic linear systems, there are a wide variety of techniques based on the properties of the system. Bode loop shaping techniques are helpful when trying to reject sensor or disturbance noise at frequencies of interest [37]. If the process and measurement state covariance is well-understood, linear quadratic Gaussian (LQG) controller design techniques can be used which also have advantages when dealing with Multiple-Input Multiple-Output (MIMO) systems [37]. When controller performance objectives are stated in terms of output state covariance, design techniques presented by Skelton and Hotz can be used to design appropriate controllers by using the target state covariance matrix in the controller design process [37, 38]. When concerned with the expected value and variance of a quadratic cost function, minimal cost variance control design techniques presented by Sain can be used. Unlike LQG, which has a sole focus on minimizing the quadratic cost, this approach allows for designing controllers that limit the observed cost function variability based on the stochastic properties of the target implementation [39, 40].

Although these techniques can constrain the impact of stochastic noise in control systems, they lack a method for direct control of the variance to a target setpoint. If a system's statistical response to the desired inputs is well understood, open-loop control could be used to achieve the desired behavior without exceeding covariance limits. However, in situations where the statistical behavior is influenced by many factors, or worse, unknown phenomena, open-loop control performance could be jeopardized. For such scenarios it would be ideal to be able to use feedback control to track the variance target while being able to reject noise factors that impact the statistical behavior.

3.3 Analysis of Windowed Variance

In this section, statistical analysis is used to identify the analytical behavior of the windowed variance operator for three scenarios. The first scenario involves a process where the mean value of the random variable is constant (constant mean sequence). The second scenario involves a process where the mean value of the random variable is allowed to change with the index variable (varying mean sequence). The third scenario presented involves using predictions of the expected mean sequence behavior to detrend the output and reduce the impact of mean sequence changes on the variance estimate. For all three cases, $\{x_k\}_{k\in\mathbb{Z}}$ is a sequence of independent, normally distributed random variables with

mean and variance sequences that are allowed to vary with index k, μ_k and σ_k^2. In other words, $x_k \sim \mathcal{N}(\mu_k, \sigma_k^2)\ \forall k \in \mathbb{Z}$. In this work, online estimates of the output variance, $\widehat{\sigma_k^2}$, are generated at each sequence index k using a windowed variance function with window size N [38]:

$$\text{Var}\{X\} = \varepsilon\{X^2\} - \varepsilon\{X\}^2 \tag{3.1}$$

$$\widehat{\sigma_k^2} = \frac{1}{N}\sum_{i=0}^{N-1} x_{k-i}^2 - \left(\frac{1}{N}\sum_{i=0}^{N-1} x_{k-i}\right)^2. \tag{3.2}$$

3.3.1 Derivations for Constant Mean Sequence

In the case where the mean sequence is constant ($\mu_k = \mu$), the mean value can be subtracted from the original sequence to generate a zero-mean sequence, z_k ($z_k = x_k - \mu$). With a zero-mean sequence, the windowed variance operator can be simplified:

$$\widehat{\sigma_k^2} = \frac{1}{N}\sum_{i=0}^{N-1} z_{k-i}^2. \tag{3.3}$$

Equation (3.3) shows that $\widehat{\sigma_k^2}$ is a function of the primitive random variables $\{z_k, z_{k-1}, \ldots, z_{N-1}\}$ and therefore itself is a random variable. By leveraging the linearity of the expectation operator and the fact that the $\{z_k, z_j\}$ are independent when $k \neq j$, the expected value of $\widehat{\sigma_k^2}$ can be calculated as follows [41]:

$$\varepsilon\left\{\widehat{\sigma_k^2}\right\} = \frac{1}{N}\sum_{i=0}^{N-1} \sigma_{k-i}^2 \tag{3.4}$$

which shows that the expected value of the windowed variance operator is a low-pass filtered version of the variance sequence σ_k^2, where increasing the window size N increases the amount of filtering (derivation shown in Lemma 1 in Section 3.8). Further usage of expectation analysis can be used to calculate the central second moment of the variance estimator, i.e. $\text{Var}\{\widehat{\sigma_k^2}\}$, shown in (3.5):

$$\text{Var}\left\{\widehat{\sigma_k^2}\right\} = \frac{2}{N^2}\sum_{i=0}^{N-1} \sigma_{k-i}^4. \tag{3.5}$$

Equation (3.5) shows that $\text{Var}\{\widehat{\sigma_k^2}\}$ increases with the square of the variance sequence values, σ_k^2, in the variance estimator buffer (derivation shown in Lemma 2 in Section 3.8).

This result suggests that if attempting to estimate/control a state's variance with a large magnitude, the window size must be increased to avoid excessive feedback noise. The result also allows us to calculate a Signal-to-Noise Ratio (SNR) for the windowed variance estimator, which can be calculated as (3.6):

$$\text{SNR} = \frac{\varepsilon\{(\widehat{\sigma_k^2})^2\}}{\text{Var}\{\widehat{\sigma_k^2}\}}. \tag{3.6}$$

Figure 3.1 provides a visual representation of the variance of the windowed variance estimator based on a constant variance sequence ($\sigma_k^2 = \sigma^2 \, \forall k$). When the input variance sequence is constant, the variance of the variance estimate increases by the square of the variance sequence multiplied by a constant that is a function of the window size, $\alpha(N) = \frac{2}{N}$. Increasing the window size decreases the variance of the variance estimate and therefore increases the SNR, which grows increasingly important as the variance sequence increases in magnitude, as shown in Figure 3.1.

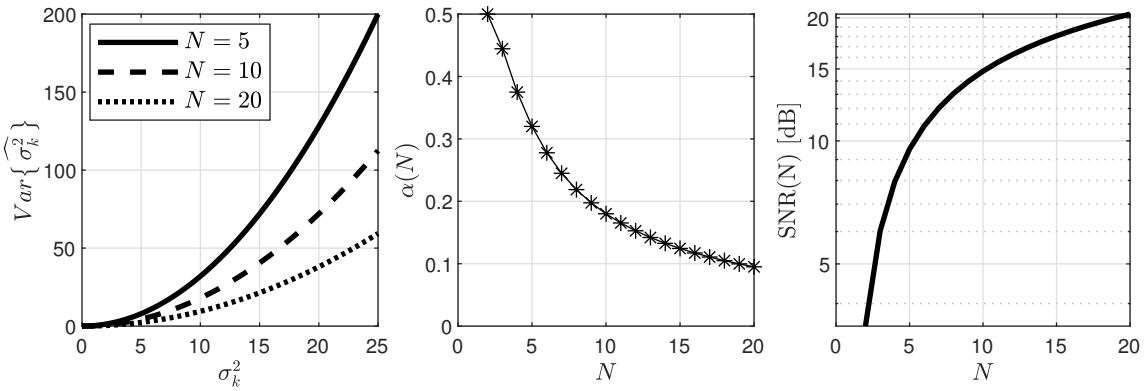


Figure 3.1: Plot illustrating the behavior of the variance of the windowed variance operator for different window sizes, N, and constant variance sequence values, σ_k^2.

To verify the analytical results, random number simulations were conducted to compare the presented equations for the statistical moments to ensemble statistics of the simulation trials. In doing so, convergence of the ensemble statistics to the analytical results presented suggest a proper derivation of the moments. A trial consisted of generating a normally distributed zero-mean number sequence, scaling the number sequence by an input variance sequence, and then applying the windowed variance operator to the number sequence. One thousand simulations were performed to ensure an adequate number of samples were available for calculating ensemble statistics, with different random number seeds used to generate unique trials. Figure 3.2 illustrates the random nature of the variance estimates for each realization (window size $N = 10$) and how the ensemble statistics converge to

the theoretical expected values shown in (3.4)-(3.5). Note that as the variance sequence increases in magnitude, the estimate can at times exceed five times the true value, which can be avoided by using a larger sampling window.

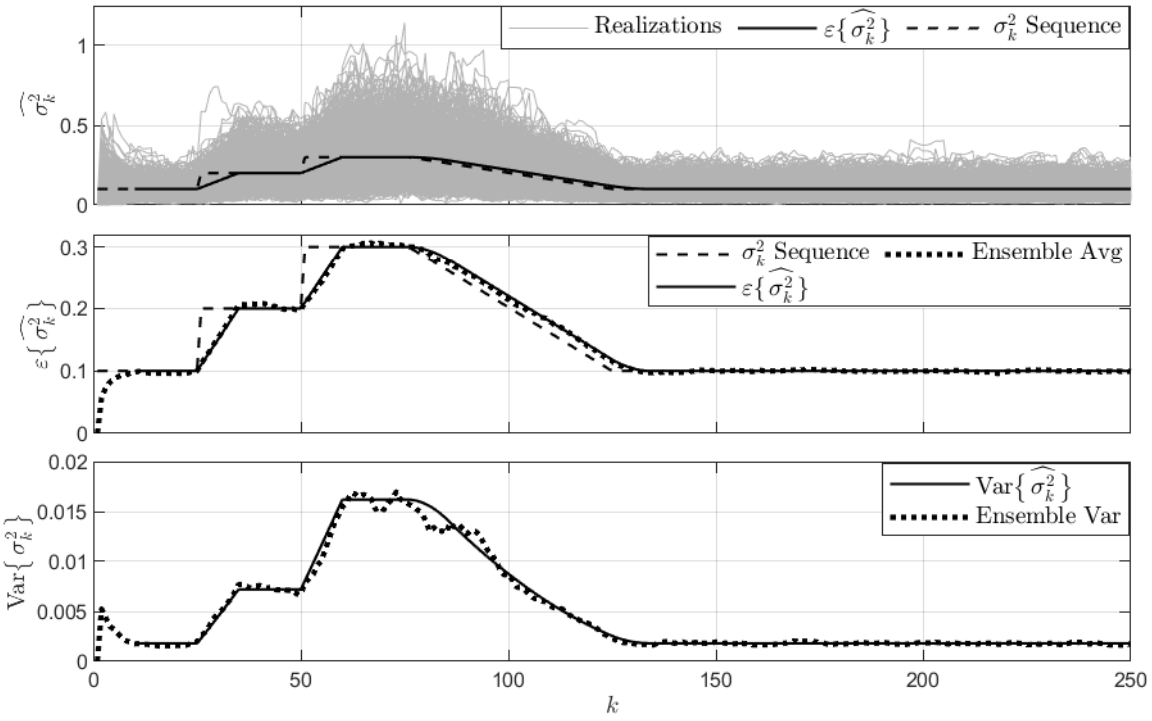


Figure 3.2: Ensemble statistical values of the numerical experiments agree with the derived theoretical values (3.4)-(3.5). The low-pass behavior of the windowed variance operator when trying to estimate an index-varying variance sequence is also visible.

The presented expectation analysis of windowed variance as a random variable was based solely on the first and second moment of the estimation and not the underlying distribution. Since the windowed variance is the weighted sum of squared normal random variables, it is distributed as a chi-squared (χ^2) random variable with N degrees of freedom. If the variance sequence is a constant value, the probability density function for the windowed variance has been shown to be a Pearson type III distribution parameterized by the window size and variance magnitude [42]. Figure 3.3 plots the windowed variance distribution when $\sigma^2 = 1$ alongside a Gaussian distribution with the first moment equal to σ^2 and the second moment equal to $\mathrm{Var}[\widehat{\sigma^2}]$ for varying window sizes N. As N approaches ∞, the windowed variance distribution approaches a normal distribution with the aforementioned first and second moments. Note that at smaller window sizes, the distribution has a long right tail, explaining the behavior observed in Figure 3.2 where the variance estimate occasionally reached five times the expected value.

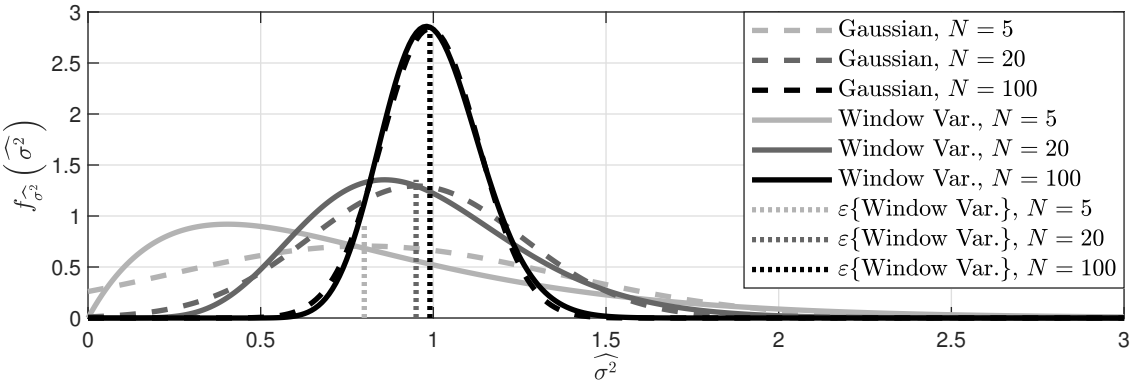


Figure 3.3: Windowed variance distributions plotted alongside Gaussian distributions with equivalent first and second moments. As the window size is increased, the χ^2 distribution approaches an equivalent normal distribution. Note the large right tail at smaller window sizes.

3.3.2 Derivations for Varying Mean Sequence

In the case where the mean sequence varies with index k, changes in the mean value of the sequence excites the windowed variance estimator, increasing its magnitude relative to the true value of the variance sequence. The expected value of the windowed variance for an index-varying mean is shown in (3.7):

$$\varepsilon\left\{\widehat{\sigma^2_{x,k}}\right\} = \frac{N-1}{N^2}\sum_{i=0}^{N-1}\left[\sigma^2_{k-i}+\mu^2_{k-i}\right] - \frac{2}{N^2}\sum_{i=0}^{N-1}\sum_{j>i}^{N-1}\left[\mu_{k-i}\cdot\mu_{k-j}\right]. \tag{3.7}$$

Note that the estimator is biased by a factor of $\frac{N-1}{N}$ even when the mean sequence is constant - this is due to the fact that the definition of windowed variance shown in (3.2) must be used due to the changing mean sequence.

3.3.3 Derivations for Detrended Sequence

As shown in 3.3.2, an index-varying mean sequence excites the variance estimator as the estimator cannot distinguish changes in the random variable caused by stochastic noise from deterministic changes caused by the varying mean sequence. However, in processes where the expected system behavior can be modeled or the output value is tracking a reference, detrending can be used to remove the deterministic trends from the signal, leaving behind the stochastic noise. Let $\widehat{\mu_k}$ be the predicted mean sequence (expected

behavior) - define the detrended signal v_k as the difference between the nominal sequence and predicted mean,

$$v_k = x_k - \widehat{\mu}_k \tag{3.8}$$

and define the mean error sequence e_k as the difference between the true mean sequence and predicted mean sequence,

$$e_k = \mu_k - \widehat{\mu}_k. \tag{3.9}$$

Expectation analysis can be used to find the expected value of the windowed variance of the detrended signal v_k, shown in (3.10):

$$\begin{aligned} \varepsilon\left\{\widehat{\sigma^2_{v,k}}\right\} = \frac{N-1}{N^2}\sum_{i=0}^{N-1}\left[\sigma^2_{k-i} + e^2_{k-i}\right] - \\ \frac{2}{N^2}\sum_{i=0}^{N-1}\sum_{j>i}^{N-1}\left[e_{k-i}\cdot e_{k-j}\right]. \end{aligned} \tag{3.10}$$

Note that (3.10) has a similar form as (3.7) due to the similarity of the problems - (3.7) can be recovered by setting $\mu_k = 0$ in (3.10), in which case $\varepsilon\left\{\widehat{\sigma^2_{x,k}}\right\} = \varepsilon\left\{\widehat{\sigma^2_{v,k}}\right\}$. However, (3.10) can be used to determine how accurate the deterministic model $\widehat{\mu}_k$ must be to achieve variance estimation that is adequate for closed-loop control.

3.4 Controller Design for Windowed Variance

Using the first moment of the windowed variance for a constant mean sequence derived in Section 3.3.1, a discrete transfer function relating the expected value of the windowed variance estimator to the input variance sequence can be developed, shown in (3.11).

$$\frac{\varepsilon\left\{\widehat{\sigma^2}\right\}}{\sigma^2} = \frac{1}{N}\sum_{i=0}^{N-1} z^{-i} \tag{3.11}$$

Leveraging the result for feedback controller design requires knowledge of the system's variance response to the modeling parameters of interest, presented for the diesel combustion variability example in Section 2.6. An additional modeling example can found in [43]. In this chapter, simple mathematical variance models will be leveraged to observe how the closed-loop response behavior is influenced by the open-loop variance response and

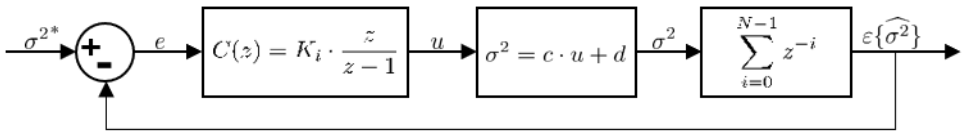


Figure 3.4: Block diagram of the proposed closed-loop variance control architecture.

controller tuning. For initial observations, consider the linear variance response presented in (3.12).

$$\sigma^2 = c \cdot u + d \tag{3.12}$$

Note that in (3.12), the resulting σ^2 must be positive, therefore when implementing the model for simulation the input u must be bounded (based on the model parameters c, d) or the model must have saturation to avoid negative σ^2 values. Referring back to Figures 3.1 and 3.3, the windowed variance estimate contains a lot of noise at lower window sizes. Therefore, a simple discrete integral controller is used to avoid propagating the feedback noise and enable set-point tracking. Figure 3.4 illustrates the proposed controller architecture. For initial controller analysis a window size of $N = 50$ was used for feedback generation, as it provides an excellent signal-to-noise ratio and feedback distribution that resembles a Normal random variable. Equation (3.13) shows the linearized open-loop transfer function for the system that was used to design the controller using root locus techniques, with an example root locus plot shown in Figure 3.5 when $N = 20$ [37].

$$L(z) = K_i \cdot \frac{z}{z-1} \cdot c \cdot \frac{1}{N} \sum_{i=0}^{N-1} z^{-i} \tag{3.13}$$

3.5 Closed-Loop Windowed Variance Control Observations

The controllers designed using root locus technique were tested in numerical simulations using MATLAB. Figure 3.6 shows how the proposed control structure was adapted for numerical simulation. Note that the input u (and integrator state) was restricted to be > 0 (input saturation). Various phenomena were observed and studied to understand the principal causes, explained in the following subsections.

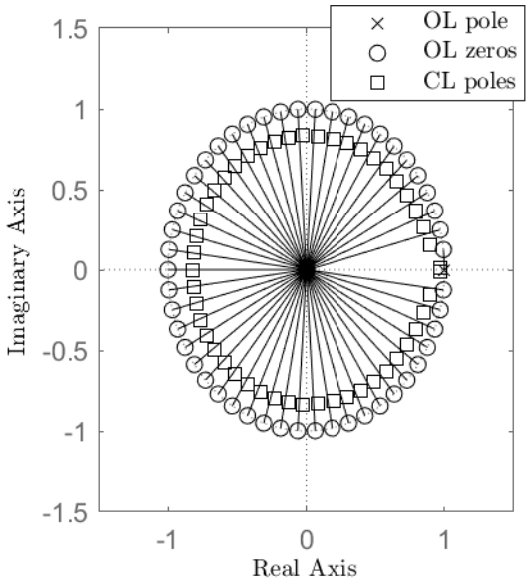


Figure 3.5: Example root locus plot when $N = 50$; the resulting closed-loop system had an effective damping ratio of $\zeta = 0.95$.

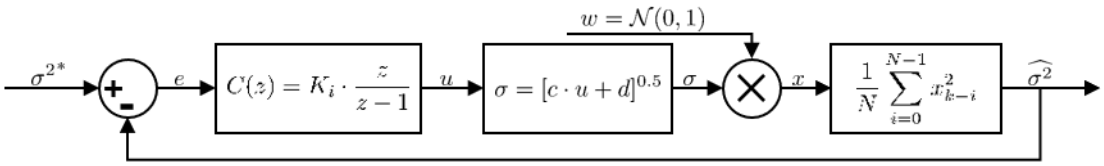


Figure 3.6: Diagram illustrating how the closed-loop variance control architecture was implemented in MATLAB to conduct numerical experiments.

3.5.1 Ensemble Analysis and Observed Linear Behavior

Recall that the transfer function presented in (3.13) is based on the expected behavior of the windowed variance estimate, therefore any comparisons to a modeled linear system response should be made against the ensemble average behavior of the system, which requires running many simulations to ensure the ensemble statistical values are representative of the system behavior. Unless noted otherwise, one thousand stochastic simulations were used to analyze the ensemble time response and output distribution behavior.

The first test case presented is based on a variance system with a slope $c = \frac{1}{4}$ and offset $d = 0.01$, with a $N = 50$ point estimator window and a controller tuned to be slightly underdamped ($\zeta = 0.95$). The variance setpoint σ^{2*} was set to 1.5 to avoid actuator non-linear behavior caused by actuator saturation. Figure 3.7 illustrates that not only is closed-loop variance control feasible, but that in fact the ensemble behavior of

the closed-loop numerical experiments agrees with the step response generated using the closed-loop transfer function of the expected behavior. The figure also highlights why one thousand simulations were used to generate ensemble statistics, as the stochastic nature of the feedback variable results in plenty of variability in the closed-loop trials - simulating less test cases may hide potential undesirable actuator saturation behavior, presented in a later figure.

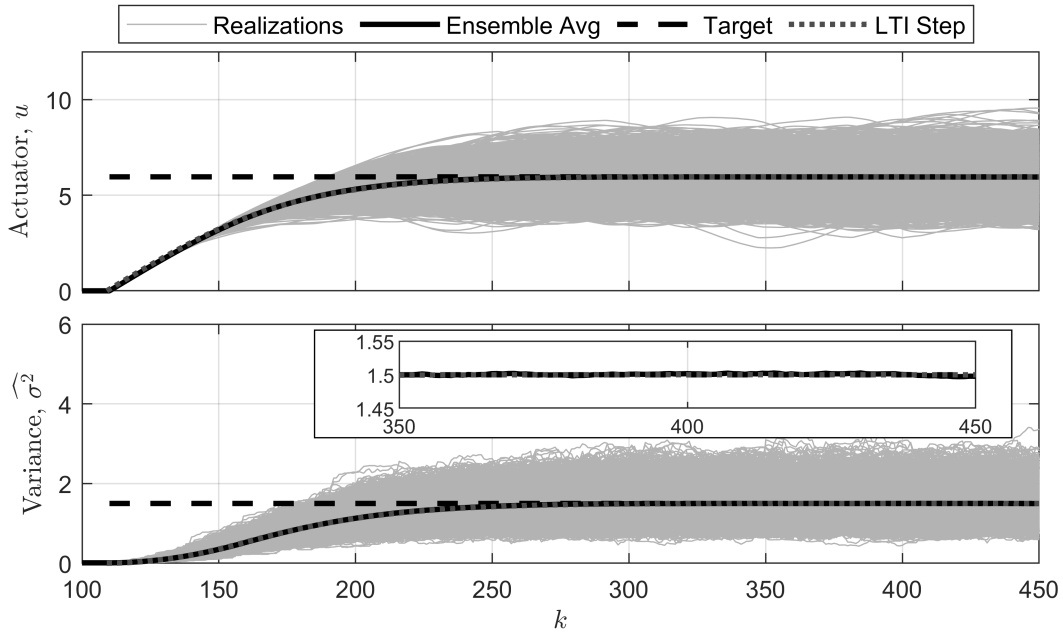


Figure 3.7: Controller simulations indicate that closed-loop variance control is feasible, as the stochastic realizations (grey lines) and ensemble average of windowed variance (solid black line) is driven to the target variance (dashed black line). Note that the LTI analysis (dotted black line) agrees with the ensemble behavior.

In classical linear systems, increasing the gain of the controller decreases the damping ratio of the closed-loop system, which results in overshoot and oscillations. This behavior is also observable with the windowed variance control, shown in Figure 3.8 where the model parameters were kept the same, but the controller was tuned to reduce damping ($\zeta = 0.55$) and excite oscillatory behavior. Not only is it possible to produce overshoot and oscillatory behavior, but the step response of the expected closed-loop variance continues to agree with the ensemble average of the closed-loop realizations. A few of the realizations had control signals at the saturation limit ($u \not\leq 0$), but not enough to perturb the ensemble statistics.

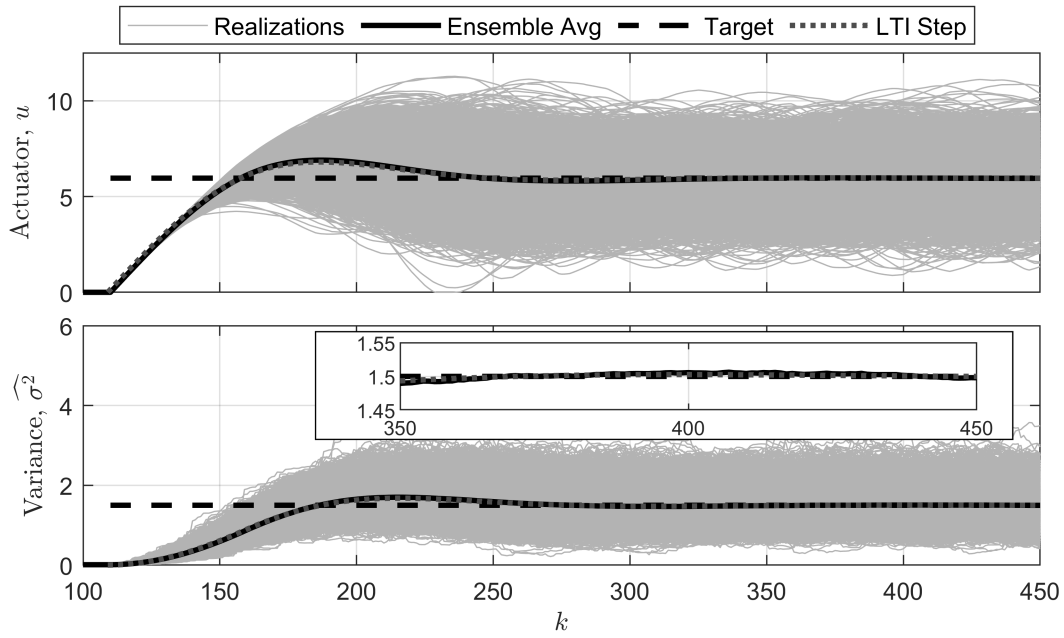


Figure 3.8: Tuning the controller to further decrease system damping causes oscillatory realization and ensemble behavior that agrees with LTI analysis.

3.5.2 Nonlinear Impacts of Actuator Saturation

In traditional control systems, physical limits on actuators can limit the performance and operating range of the system. When the desired actuator state deviates outside of these limits, the feedback loop is broken and nonlinearities are introduced into the system. In Figures 3.7 and 3.8, the plant, estimator, and controller were designed to ensure that the actuator operated away from the low actuator saturation bound and ensure linear behavior.

To explore the impact of actuator saturation on closed-loop variance control, the plant was updated such that $d = 1.4$ while using the slightly underdamped controller tuning ($\zeta = 0.95$). Increasing the nominal variance, d, resulted in the controller requiring less actuator effort to drive the system to the desired variability ($\sigma^{2*} = 1.50$), as seen in Figure 3.9. To better illustrate the actuator saturation behavior, an additional thin black line has been added to show a single realization - note that noise in the estimator causes the actuator to saturate at zero, breaking the feedback loop and introducing a nonlinearity. The consequences of the nonlinearity can be seen in the ensemble actuator plot, where the nominal actuator usage is higher than the theoretical amount required, and the ensemble variance plot, where now the ensemble variance is higher than the desired value.

From an implementation perspective, this behavior must be avoided to ensure that the system's output variance is not overexcited, as it may cause system failure depending on

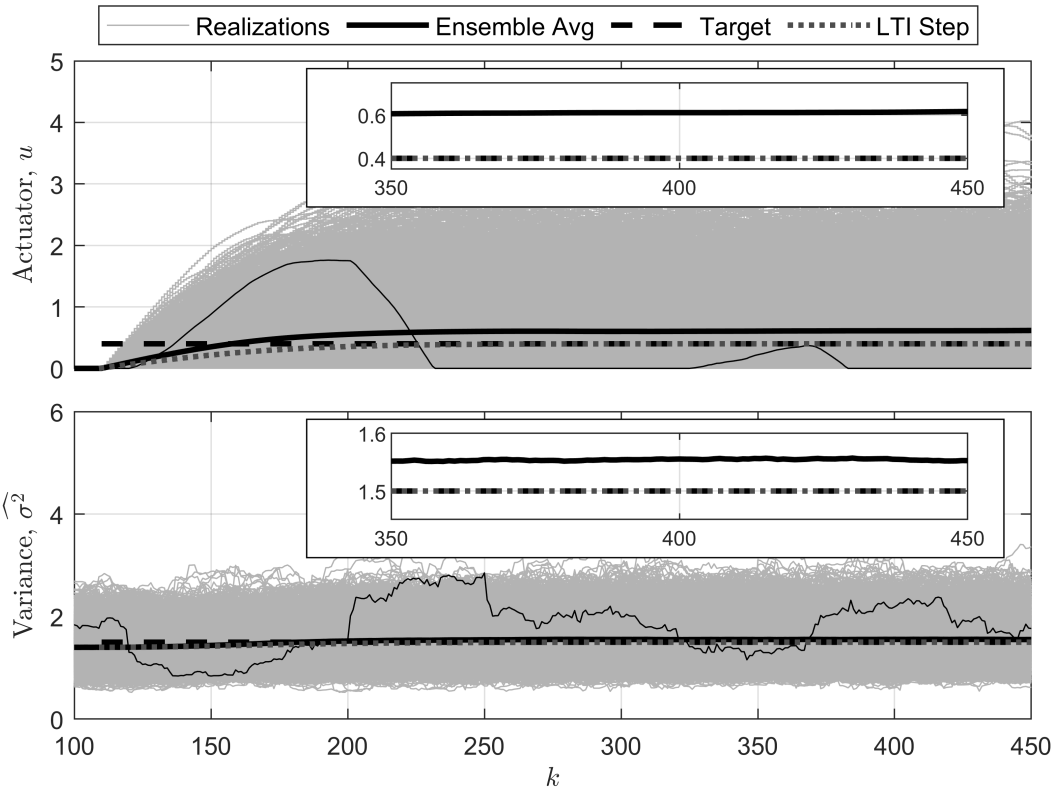


Figure 3.9: Increasing the nominal variance causes actuator saturation at the lower bound ($u \nleq 0$). A single realization is plotted for clarity as a think black line. When actuator saturation is present, the ensemble behavior no longer agrees with the predicted LTI behavior, and the ensemble variance exceeds the desired target value.

the functional requirements. Note that since the linearized analysis does not include the nominal variance, d, or the setpoint σ^{2*}, and that the saturation issue itself is due to the variance estimator noise in the system, this behavior should be observed/accounted for via stochastic simulations. Ideally, one should run many simulations to ensure that the controller is properly designed for the infinite number of realizations possible.

3.5.3 Impact of Control on Output Distributions

The previous sections have highlighted that closed-loop variance control is possible, agrees with linear system analysis when actuator saturation is not present, and that actuator saturation can cause over-excitation of the system's variance (Figure 3.9). To further understand the impacts of closed-loop variance control, the first four statistical moments (mean μ, variance σ^2, skewness γ_1, and kurtosis κ) were analyzed with respect to the

integral gain used in the windowed variance controller. To clarify, one thousand simulations were run using a particular integral gain, and the system was allowed to reach stationarity in each simulation.

Statistics of the closed-loop stationary x_k sequence were calculated, and then plotted against the integral gain K_i shown in Figure 3.10. The mean value ($\mu(x)$) of the x_k sequence remains 0 and the distribution of the realization means is consistent across controller gains, which is a sensible outcome considering the mean of the random sequence was set to zero for all simulations. The mean variance of the x_k realizations remained at the 1.25 target until a gain of about $K_i = 0.3$, which also correlates with when actuator saturation is found to be present in the simulations, shown in the $\%u$ saturation plot. However, increasing the gain is shown to reduce the variance of the realization population variance distribution (shown using the error bars), indicating that higher control gains can be beneficial for tighter variance control until actuator saturation is present. The skewness ($\gamma_1(x)$) remains near zero, indicating the closed-loop distribution remains near-symmetric, however the spread of the realization skewness values grows larger as the control gain increases. Finally, the kurtosis ($\kappa(x)$) grows in magnitude with increasing control gain, indicating that the

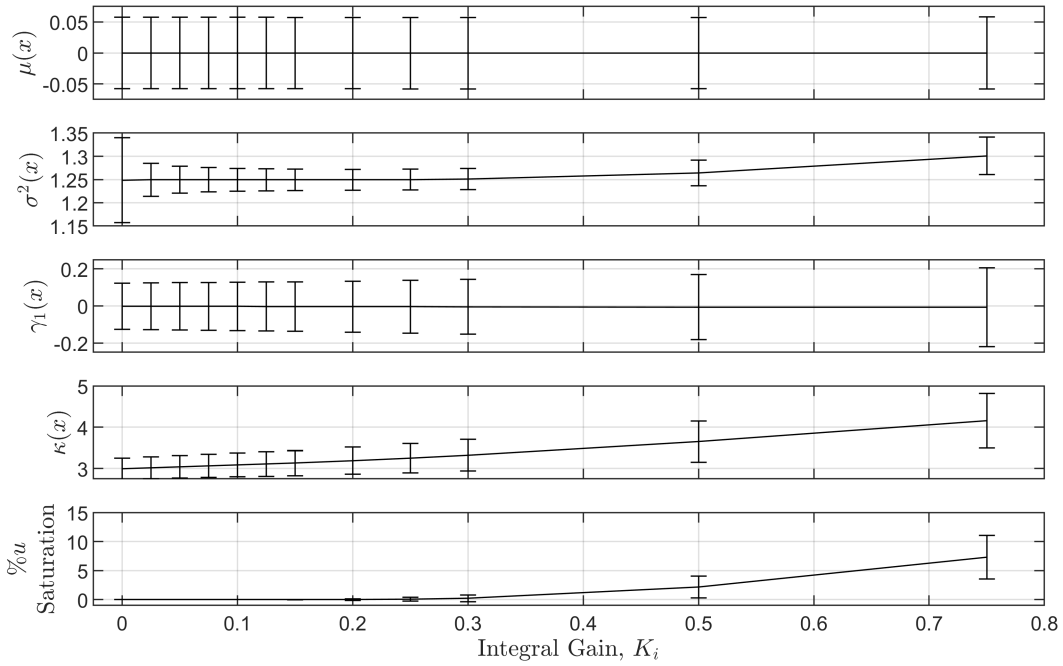


Figure 3.10: One thousand simulations were run at each gain, and the first through fourth moments of the closed-loop stationary x_k sequence are plotted against the controller integral gain K_i. The plot lines indicate the mean of the observed moment values per gain K_i, where the error bars indicate $\pm 2\sigma$ of the observed moment values.

stationary distribution is more concentrated at the mean value, which may be a desirable property depending on the system being controlled.

3.6 IIR/EWMA-based Variance Estimation

A disadvantage of the finite impulse response (FIR) windowed variance estimator shown in (3.2) is that storage and computational requirements of the algorithm scale linearly with N. Although available processing power and memory in embedded devices continues to improve, limited operating time of the control architecture using the estimator may make it a target for complexity reduction. An alternative to the FIR based estimator is an an infinite impulse response (IIR) based variance estimation strategy using exponentially weighed moving averages.

Let $\{x_k\}_{k\in\mathbb{Z}}$ be a sequence of independent, identically distributed random normal variables mean μ and variance σ^2. Consider the exponentially weighted moving average (EWMA):

$$\widehat{\mu}_{\mathrm{IIR},k} = \alpha\widehat{\mu}_{\mathrm{IIR},k-1} + (1-\alpha)x_k, \quad 0<\alpha<1, \tag{3.14}$$

which can also be written as an infinite series:

$$\widehat{\mu}_{\mathrm{IIR},k} = (1-\alpha)\sum_{i=0}^{\infty}\alpha^i x_{k-i}, \quad 0<\alpha<1. \tag{3.15}$$

Note that increasing values of α implies more smoothing done by the IIR filter. Using properties of expectation and variance we deduce the statistical properties:

$$(1-\alpha)\sum_{i=0}^{\infty}\alpha^i = (1-\alpha)\frac{1}{1-\alpha} = 1, \quad 0<\alpha<1 \tag{3.16}$$

$$\mathrm{E}[\widehat{\mu}_{\mathrm{IIR},k}] = (1-\alpha)\sum_{i=0}^{\infty}\alpha^i\mathrm{E}[x_{k-i}] = \mu \tag{3.17}$$

$$\mathrm{Var}[\widehat{\mu}_{\mathrm{IIR},k}] = (1-\alpha)^2\sum_{i=0}^{\infty}(\alpha^2)^i\mathrm{Var}[x_{k-i}] = \frac{1-\alpha}{1+\alpha}\sigma^2. \tag{3.18}$$

Moreover, note that $\widehat{\mu}_{\mathrm{IIR},k}$ is a linear combination of Gaussian random variables. Hence,

$$\widehat{\mu}_{\mathrm{IIR},k} \sim \mathcal{N}\left(\mu, \frac{1-\alpha}{1+\alpha}\sigma^2\right). \tag{3.19}$$

Define the random variable $z_k = x_k - \widehat{\mu}_{\mathrm{IIR},k}$. The sequence $\{z_k\}_{k\in\mathbb{Z}}$ corresponds to a zero-

mean Gaussian random process with the following variance:

$$\begin{aligned}\mathrm{Var}[z_k] &= \mathrm{Var}[x_k] + \mathrm{Var}[\widehat{\mu}_{\mathrm{IIR},k}] - 2\mathrm{Covar}[x_k, \widehat{\mu}_{\mathrm{IIR},k}] \\ &= \left(1 + \frac{1-\alpha}{1+\alpha} - 2(1-\alpha)\right)\sigma^2 = \frac{2\alpha^2}{1+\alpha}\sigma^2. \end{aligned} \tag{3.20}$$

Second, define the variable $w_k = \left(\sqrt{(1+\alpha)/2\alpha^2}\right) z_k$ which scales z_k to achieve the same variance as x_k. It immediately follows that $w_k \sim \mathcal{N}(0, \sigma^2)$. Hence, $\mathrm{E}[w_k^2] = \sigma^2$ and $\mathrm{Var}[w_k^2] = 2\sigma^4$. At this point, the variable w_k^2 suffices for the estimation of σ^2. This estimator follows a Chi-squared distribution with one degree of freedom: $w_k^2 \sim \sigma^2 \chi_1^2$. However, unlike the windowed variance estimate which is a distributed as a χ_N^2, the mode is located at 0. As a consequence, although on average w_k^2 stays at σ^2, most of the time the realization of w_k^2 will produce near-zero values. To avoid this and locate the mode closer to the mean, a second filtering stage can be used. Consider a second EWMA filter for variance estimation:

$$\widehat{\sigma^2}_{\mathrm{IIR},k} = \beta \widehat{\sigma^2}_{\mathrm{IIR},k-1} + (1-\beta) w_k^2 \quad 0 < \beta < 1. \tag{3.21}$$

Similar to α, increasing β increases the amount of smoothing performed. Under stationarity conditions $\mathrm{E}[\widehat{\sigma^2}_{\mathrm{IIR},k}] = \sigma^2$, showing that the variance estimator is unbiased for any feasible pair (α, β).

If the IIR estimation strategy is to be used to replace an existing FIR strategy, one method technique for tuning the (α, β) is to select values which cause the IIR filters to have the same variance as the FIR filters. The parameter α should be set to match the variance of the FIR and IIR mean estimates.

$$\mathrm{Var}[\widehat{\mu}_k] = \mathrm{Var}[\widehat{\sigma^2}_{\mathrm{IIR},k}] \tag{3.22}$$

$$\frac{\sigma^2}{N} = \frac{1-\alpha}{1+\alpha}\sigma^2 \tag{3.23}$$

$$\alpha = \frac{N-1}{N+1} \tag{3.24}$$

The parameter β should be set to match the variance of the FIR and IIR variance estimates. The variance of the FIR variance estimate, after correcting for the $\frac{N-1}{N}$ estimation bias, can be computed as:

$$\mathrm{Var}[\widehat{\sigma^2}_{\mathrm{FIR},k}] = \frac{1}{N-1} 2\sigma^4, \tag{3.25}$$

and the variance of the IIR variance estimate (after the second EWMA filter) can be computed as:

$$\mathrm{Var}[\widehat{\sigma^2}_{\mathrm{IIR},k}] = \frac{1-\beta}{1+\beta}\left(1+\frac{\beta(1-\alpha)^2}{2(1-\alpha^2\beta)}\right)2\sigma^4. \tag{3.26}$$

Tedious algebraic manipulations can be used to show that the $\beta \in (0,1)$ value satisfies the following equation:

$$a_N\beta^2 + b_N\beta + c_N = 0 \tag{3.27}$$

$$\begin{aligned} & & a_N &= (N+1)(N-1)(N-2) \\ \text{where,} & & b_N &= -N\left[(N+1)^2+(N-1)(N-3)\right] \\ & & c_N &= (N+1)^2(N-2) \end{aligned}$$

As $N \to \infty$, the variance matching (α,β) approaches 1. This is a sensible conclusion, as the variability in the mean and variance estimates decreases with window size, and so a high level of filtering through the EWMAs is required to achieve the same variance.

Table 3.1 compares the computational requirements of the FIR and IIR based estimation strategies. Note that for the FIR strategy, storage and computation requirements scale linearly with N. The IIR strategy only requires storage of two states (for each EWMA) and 9 calculations to complete the algorithm.

Table 3.1: Calculation requirements of FIR vs IIR windowed variance estimator.

	Add/Subtract	Multiply/Divide	Total
Calculate Second Moment	N-1	N+1	2N
Calculate First Moment Squared	N-1	2	N+1
Second Moment - First Moment Squared	1	0	1
Estimate Variance - FIR	2N-1	N+3	3N+2
Estimate Mean (EWMA1) - $\widehat{\mu}_k$	1	2	3
Generate Zero-Mean Sequence - z_k	1	0	1
Rescale z_k - w_k	0	1	1
Estimate Variance (EWMA2) - $\widehat{\sigma^2}_{\mathrm{IIR},k}$	1	3	4
Estimate Variance - IIR	3	6	9

3.7 Summary and Conclusions

To understand and enable closed-loop variance controller design, the windowed variance operator was presented as a potential source of online process variance estimation, and a combination of statistical and classical controls analysis was leveraged to design controllers. The designed controllers were tested using numerical experiments, and the closed-loop behavior was examined from an ensemble and distribution perspective. The results showed that closed-loop covariance control was feasible, that the ensemble behavior agreed with the predicted linear response results (when nonlinearities such as actuator saturation were not present), and that care must be taken to properly tune the controller based on window size, the system response, and the controller setpoint.

3.8 Derivations

Lemma 1 *Let $\{z_k\}_{k\in\mathbb{Z}}$ be a sequence of zero-mean, independent, random variables with a variance sequence that is allowed to change with index k, σ_k^2. Define the windowed variance operator as in Eqn.* (3.3)*, then:*

$$\varepsilon\left\{\widehat{\sigma_k^2}\right\} = \frac{1}{N}\sum_{i=0}^{N-1}\sigma_{k-i}^2$$

Proof: By leveraging properties of expectation based on stated assumptions:

$$\varepsilon\left\{\widehat{\sigma_k^2}\right\} = \varepsilon\left\{\frac{1}{N}\sum_{i=0}^{N-1}z_{k-i}^2\right\}$$

$$\varepsilon\left\{\widehat{\sigma_k^2}\right\} = \frac{1}{N}\sum_{i=0}^{N-1}\varepsilon\left\{z_{k-i}^2\right\}$$

$$\varepsilon\left\{\widehat{\sigma_k^2}\right\} = \frac{1}{N}\sum_{i=0}^{N-1}\sigma_{k-i}^2$$

□

Lemma 2 *Let $\{z_k\}_{k\in\mathbb{Z}}$ be a sequence of zero-mean, independent, random variables with a variance sequence that is allowed to change with index k, σ_k^2. Define the windowed*

variance operator as in Eqn. (3.3)*, then:*

$$Var\left\{\widehat{\sigma_k^2}\right\} = \frac{2}{N^2}\sum_{i=0}^{N-1}\sigma_{k-i}^2$$

Proof: By leveraging properties of expectation, independence of z_k, z_j when $k \neq j$, and $\mu_k = 0$:

$$\begin{aligned}
\mathrm{Var}\left\{\widehat{\sigma_k^2}\right\} &= \mathrm{Var}\left[\frac{1}{N}\sum_{i=0}^{N-1} z_{k-i}^2\right] \\
\mathrm{Var}\left\{\widehat{\sigma_k^2}\right\} &= \frac{1}{N^2}\mathrm{Var}\left[\sum_{i=0}^{N-1} z_{k-i}^2\right] \\
\mathrm{Var}\left\{\widehat{\sigma_k^2}\right\} &= \frac{1}{N^2}\left(\sum_{i=0}^{N-1}\mathrm{Var}\left\{z_{k-i}^2\right\} + \sum_{i\neq j}^{N-1}\mathrm{Cov}\left\{z_{k-i}^2, z_{k-j}^2\right\}\right) \\
\mathrm{Var}\left\{\widehat{\sigma_k^2}\right\} &= \frac{1}{N^2}\sum_{i=0}^{N-1}\mathrm{Var}\left\{z_{k-i}^2\right\} \\
\mathrm{Var}\left\{\widehat{\sigma_k^2}\right\} &= \frac{1}{N^2}\sum_{i=0}^{N-1}\left[\varepsilon\left\{z_{k-i}^4\right\} - \varepsilon\left\{z_{k-i}^2\right\}^2\right] \\
\mathrm{Var}\left\{\widehat{\sigma_k^2}\right\} &= \frac{1}{N^2}\sum_{i=0}^{N-1}\left[3\sigma_{k-i}^4 - \sigma_{k-i}^4\right] \\
\mathrm{Var}\left\{\widehat{\sigma_k^2}\right\} &= \frac{2}{N^2}\sum_{i=0}^{N-1}\sigma_{k-i}^4
\end{aligned}$$

□

CHAPTER 4

Control of Engine Combustion Variance

4.1 Introduction

The experimental combustion variability observations and model from Chapter 2 were combined with the theoretical windowed variance control theory from Chapter 3 to design combustion variance controllers. Transient combustion data from FTP-75 traces were analyzed to generate a transient-capable variability metric. Initial controller design and testing was focused on steady-state operation of the dyno engine. After experimental validation of the concept at steady-state, the transient metric was implemented in the controller, and was tested on dyno against the cold-start phase of the FTP-75, generating the desirable exhaust temperature increase without excessive degradation of combustion.

The indicated analysis equipment (cylinder pressure sensors and data acquisition hardware) used during the dyno testing is sophisticated and expensive, intended for engine development in lab settings. To ensure the proposed technique would function in the real-world, the controller was adapted to a production vehicle using production-intent pressure-sensing glow plugs. Implementing the controller in the vehicle also permitted studying the controller impacts on occupants during real-world driving scenarios. After addressing vehicle-specific engine behaviors not present during dyno testing, the controller was repeatedly tested against a synthesized drive cycle consisting of low-load driving. The impact of different fuel types were investigated during this testing to observe the noise factor response. Following validation of the controller on the synthesized drive-cycle, the controller was evaluated against the FTP-75 emission certification cycle in an emissions lab to observe impacts on aftertreatment heating, fuel economy, and emissions.

4.2 Application of Windowed Variance Control Theory

To design combustion variance controllers, the theory developed in 3.4 is combined with the model and experimental observations from Chapter 2. Challenges and solutions associated with instrumentation strategies (every cylinder vs. limited number of cylinders) are also presented.

4.2.1 Model-Based Controller Design

Linear control design techniques were used to design the combustion variability controller framework. The stochastic combustion model presented in (2.4)-(2.6) was linearized about an actuator pairing that resulted in a CoV_{IMEP} of 2% at the nominal IMEP value prior to retarding combustion; the actuator values at this point were $\Delta\varphi_0 = 9.5$ CA$^\circ$ and $\Delta m_0^{\text{fuel}} = 2.5$ mg/str. The 2% CoV_{IMEP} value was selected as an acceptable level of combustion variability based on literature and vehicle testing observations [33]. The linearized form of the model is shown in (4.1), where $\Delta^*\varphi$ and $\Delta^* m^{\text{fuel}}$ represent injection parameter deviations from the linearization point.

$$\begin{bmatrix} \Delta\sigma_k^{\text{IMEP}} \\ \Delta\mu_k^{\text{IMEP}} \end{bmatrix} = \begin{bmatrix} \frac{\partial\sigma^{\text{IMEP}}}{\Delta\varphi}\big|_{\Delta\varphi_0,\Delta m_0^{\text{fuel}}} & \frac{\partial\sigma^{\text{IMEP}}}{\Delta m^{\text{fuel}}}\big|_{\Delta\varphi_0,\Delta m_0^{\text{fuel}}} \\ \frac{\partial\mu^{\text{IMEP}}}{\Delta\varphi}\big|_{\Delta\varphi_0,\Delta m_0^{\text{fuel}}} & \frac{\partial\mu^{\text{IMEP}}}{\Delta m^{\text{fuel}}}\big|_{\Delta\varphi_0,\Delta m_0^{\text{fuel}}} \end{bmatrix} \begin{bmatrix} \Delta^*\varphi_k \\ \Delta^* m_k^{\text{fuel}} \end{bmatrix} \tag{4.1}$$

$$\begin{bmatrix} \Delta\sigma_k^{\text{IMEP}} \\ \Delta\mu_k^{\text{IMEP}} \end{bmatrix} = \begin{bmatrix} 0.019 & 0.004 \\ -0.040 & 0.168 \end{bmatrix} \begin{bmatrix} \Delta^*\varphi_k \\ \Delta^* m_k^{\text{fuel}} \end{bmatrix} \tag{4.2}$$

The DC gain matrix values in (4.2) show that perturbations in injection timing and quantity impact both the mean value and variability of IMEP, resulting in a coupled two-input two-output system. Note that retarding injection timing to control combustion variability will result in a torque decrease - with a driver in the loop they could simply increase the accelerator pedal position to compensate for the torque reduction, but the driver may perceive this behavior as the engine losing power. To avoid this issue, it is instead proposed to also control the mean value of IMEP. Although the inputs and outputs of the system are coupled, which suggest usage of a MIMO control structure, observing the relative magnitudes of the actuator impacts on each value suggest that a decoupled two-input two-output controller should be suitable for control. This observation is reaffirmed by relative gain array (RGA) analysis, where the diagonal entries of the RGA matrix were approximately 1 ($\lambda = 0.999$), indicating that injection timing offset $\Delta\varphi$ should be used to control $\Delta\sigma^{\text{IMEP}}$ and fuel quantity offset Δm^{fuel} should be used to control $\Delta\mu^{\text{IMEP}}$ [44]. Use of a simpler decoupled control structure yields both analysis and

experimental implementation benefits.

Based on experimental observations of IMEP and windowed variance feedback, two pure integral controllers were used to control the statistics of IMEP to avoid propagating feedback noise and to enable set-point tracking. Note that in (4.1), the output is $\Delta\sigma_k^{\text{IMEP}}$, the linearized version of the population standard deviation sequence. There are two adapations that must be made. First, the linear controller design analysis presented in Section 3.4 is based on windowed variance feedback, and so the standard deviation equation needs to be converted to variance by squaring the standard deviation sequence terms. Second, feedback control will instead be performed using the windowed variance estimate; for the linearized system presented, the expected value of the estimator is shown in (4.3), where N is the estimation window size and σ_0 is the nominal combustion variability at the linearized point.

$$\varepsilon\left\{\widehat{\sigma_k^2}\right\} = \frac{1}{N}\sum_{i=0}^{N-1}\left(0.0191\cdot\Delta^*\varphi_{k-i} + \Delta^* m_{k-i}^{\text{fuel}}\cdot 0.004 + \sigma_0\right)^2 \tag{4.3}$$

Sequential loop closing techniques were used to design the two Single-Input Single-Output (SISO) integral controllers [45]. A block diagram of the open-loop system can be seen in Figure 4.1. The $\Delta\mu^{\text{IMEP}}$ loop was closed first due to its faster response time, versus the slower response time of the variance estimation due to the sliding window. The simple nature of the IMEP loop, a pure gain augmented with an integrator, results in the max stable gain being a function of the sensitivity of IMEP to a fuel quantity offset. Equation (4.3) was used to generate the transfer functions g_{21} and g_{22}, which represent the variance estimators sensitivity to the deviation in injection parameters and contain the window estimation dynamics. The transfer function from $\Delta\varphi_k$ to $\Delta\varepsilon\left\{\widehat{\sigma_k^2}\right\}$ with the fuel controller (c_1) in place was calculated to consider the interactions between the two loops when designing the variance controller.

$$\frac{\Delta\varepsilon\left\{\widehat{\sigma_k^2}\right\}}{\Delta^*\varphi} = g_{21}(z) - \frac{g_{22}(z)\cdot c_1(z)\cdot g_{11}(z)}{1+g_{12}(z)\cdot c_1(z)} \tag{4.4}$$

This analysis shows that the variance controller stability, from a theoretical perspective, is only impacted by window size, the derivatives of the statistical responses at the variance setpoint, the fuel controller gain, and the variance controller gain. Note that the data presented in Chapter 2 showed that statistical response behavior can vary as a function of speed and load. Therefore, one approach for controller tuning could include extensive

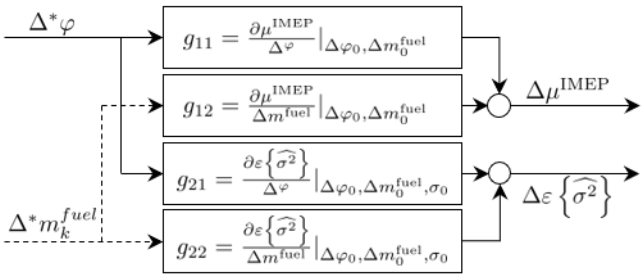


Figure 4.1: Open-loop linearized block diagram used for controller design.

mapping of the statistical response behavior across speeds, loads, and various noise factors, and using that data to implement gain-scheduling to maximize controller performance at the various operating conditions. A short-coming of this approach would include the large amount of experimental testing required to populate the map, including noise factor testing critical for ensuring proper operation in consumer environments. Another approach would be to map the statistical behavior of the engine where combustion stability is known to be problematic (typically low speeds, loads), and tune the controller based on the least stable point, including adequate gain margin to ensure robustness against noise factors. Although there may be some loss of performance, it reduces the amount of experiments needed for controller design and reduces the controller complexity, which can be critical when dealing with resource (computational, storage space) shortages in controls hardware.

4.2.2 Fully-Instrumented Control Approach

A fully-instrumented control approach refers to using feedback from in-cylinder pressure sensors in every cylinder in the engine (”fully-instrumented”). The primary advantages of such an approach include superior noise-factor rejection, as cylinder-to-cylinder variability differences as seen in Figure 2.5 can be directly monitored and controlled. Although variability controllers could be wrapped around each individual cylinder, the increased frequency of unique combustion data (for the V8, 8 times faster vs. a single cylinder) suggests instead building feedback windows that leverage data from all cylinders, so that the variability feedback updates 8 times faster. When using a window size of 8 (one event from each cylinder), this will be referred to as ”segment statistics”, where ”segment” is used to refer to the last 720 CA° of combustion events (for the 8-cylinder engine used in this work, 8 unique cylinder combustion events). This does not yield any controller gain tuning benefits, where stable gains will still be limited by the number of points in this statistical window, but it does increase the sampling time of the discrete system, yielding

temporal performance benefits. A potential shortcoming of such an approach is that based on the variance estimation theory, controlling the segment variance to a target value means the average variance across the cylinders should reach the target, meaning some cylinders' variability could exceed the setpoint, potentially in a debilitating way. This could be addressed by applying a safety factor to the setpoint or performing control intervention based on monitored statistics from individual cylinders.

Another issue with using a statistical window with feedback from multiple cylinders is the impact of mean cylinder torque biases on the calculated statistics. The impact can be quantified by using the result shown in (3.10), where the error sequence e_k is instead each cylinder's mean torque bias from the overall segment average. Ideally these imbalances would be corrected via a cylinder load balancing strategy, but the experimental configuration lacked this feature. The effect of the biases on statistics can be addressed at steady-state conditions by averaging the combustion events for each cylinder and comparing them to the overall average of the eight cylinders. When attempting closed-loop control, the injection retard and fuel quantity offset can induce time-varying mean behavior and the cylinders can have unique reactions to the offsets, causing inflation of the statistics not reflective of actual combustion variability issues.

To address this issue, a real-time bias estimation algorithm was constructed using multiple moving average filters to unbias the statistics without masking the underlying combustion variability. To estimate the cylinder bias, an IMEP buffer is maintained for each cylinder. A visual representation of the buffer is shown in Figure 4.2. Maintaining a buffer for each cylinder allows for the calculation of running averages of each cylinder torque, as well as the running average of the segment. Comparing the running average of the cylinder torque to the running average of the segment allows for estimation of the

		Segment #				Cylinder Windowed Statistics		
		k	k-1	…	k-(N-1)	Std	Mean	COV
Cylinder #	1	2.8	2.7	…	2.6	0.08	2.70	3.02
	3	3	2.9	…	2.85	0.06	2.92	2.14
	7	2.7	2.6	…	2.5	0.08	2.60	3.14
	2	2.8	2.75	…	2.6	0.08	2.72	3.13
	6	2.85	2.75	…	2.65	0.08	2.75	2.97
	5	2.9	2.8	…	2.8	0.05	2.83	1.66
	4	2.85	2.75	…	2.65	0.08	2.75	2.97
	8	2.8	2.7	…	2.6	0.08	2.70	3.02
Segment Statistics	Std	0.08	0.08	…	0.11			
	Mean	2.84	2.74	…	2.66			
	COV	2.89	2.94	…	4.04			

Figure 4.2: Visual representation of segment vs. cylinder statistics and bias estimation buffer.

cylinder bias from the mean. The use of moving averages helps to smooth the stochastic engine behavior and prevent the bias estimator from hiding exceptionally bad combustion events by learning them out as a bias. The algorithm is expressed mathematically in (4.5)-(4.10), where IMEP$[i,k]$ is used to denote the k-th IMEP event from cylinder i, RA stands for "running average", $b[i,k]$ is used to denote the k-th bias estimate for cylinder i, k_b is the estimator gain, and N describes the length of the buffer.

$$\mu_{seg}[k] = \frac{1}{8}\sum_{i=0}^{7} \text{IMEP}[i,k] \tag{4.5}$$

$$\mu_{RA-seg}[k] = \frac{1}{N}\sum_{j=k-n+1}^{k} \mu_{seg}[j] \tag{4.6}$$

$$\mu_{RA-cyl}[i,k] = \frac{1}{N}\sum_{j=k-n+1}^{k} IMEP[i,j] \tag{4.7}$$

$$e_b = \big(\big(\mu_{RA-cyl}[i,k] - b[i,k]\big) - \mu_{RA-seg}[k]\big) \tag{4.8}$$

$$b[i,k+1] = b[i,k] + k_b e_b \tag{4.9}$$

$$\text{IMEP}_{UB}[i,k] = \text{IMEP}[i,k] - b[i,k] \tag{4.10}$$

In Figure 4.3, the statistical consequences of the cylinder mean biases can be seen in the left set of plots, and the advantages of removing them can be seen in the right set of plots. The dashed lines indicate the population statistics for three different cylinders, whereas the black line indicates the segment statistics. Looking at the left set of plots, note that although the cylinder population CoV_{IMEP} values vary from 0.75-1.5%, the segment CoV_{IMEP} value hovers about 2.5%, which is slightly higher than the target value chosen in the previous section and could be perceived as excessive variability. The right set of plots show that the online estimation algorithm hovers about the population-based bias values, and applying the bias corrections collapses the segment statistics to values that hover about the original population values of the individual cylinders in agreement with theoretical results.

The algorithm was also applied to transient combustion data collected from FTP-75 cycles to verify functionality at two initial coolant temperatures and three injection phasing conditions to understand impacts on the online variability estimate behaviors. A snippet of the test data is plotted in Figure 4.4, where each line labeled $\tau_{cyl,X}$ represents combustion data from a trial with a constant X° timing offset applied during the entire test. The ECU torque model, τ_{model}, an estimate of the torque being generated by the engine, is also shown to help identify commanded changes in torque versus combustion variability. Control limitations prevented compensating for the torque reduction as the timing offset

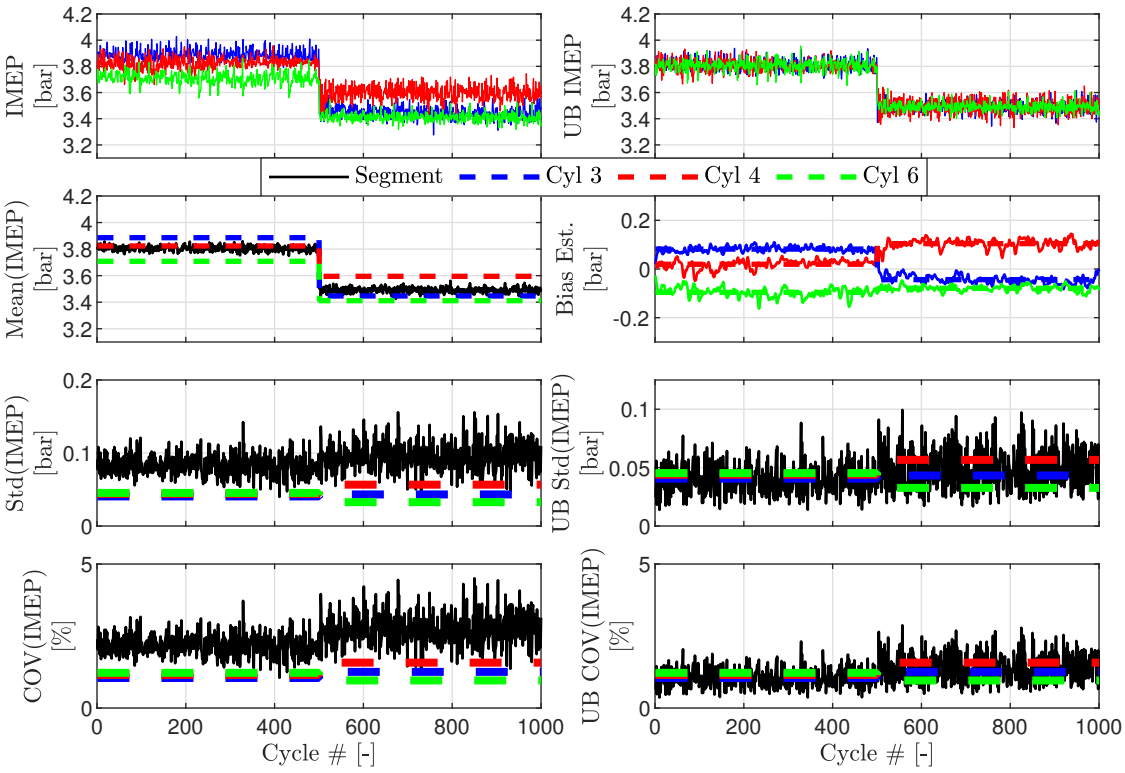


Figure 4.3: Plots illustrating impact of mean biases on segment statistics, and behavior and statistical impacts of online bias estimation.

was increased. The left set of plots show data from a "Cold Start", where the engine coolant was at ambient temperature at the start of the test, and the right set of plots show data from a "Warm Start", where the engine coolant temperature was at 90°C at the start of the test. The top row of plots show the uncorrected torque data (τ) from all cylinders at the various conditions. Note that at nominal phasing ($\tau_{cyl,0}$), there is a visible spread of cylinder torque traces, but they are otherwise stable and lack cycle-to-cycle variability. However, looking at the second row of plots plotting the segment standard deviation, even at the steady-state operation occuring between 190-200 seconds, the magnitude of the standard deviation estimate and mean torque corresponds to a CoV_{IMEP} of 2.9%, indicative of bad combustion that clearly isn't present at nominal timing. By applying bias estimation, the individual cylinder offset behaviors can be removed, as seen in the third row of plots (UB τ) where the traces have collapsed to the overall mean behavior. Performing statistics on these unbiased traces yields the bottom plots, where the magnitude of the standard deviation estimate corresponds to a CoV_{IMEP} of 0.9%, more in line with what's considered acceptable and also what is being observed in the cyclic torque behavior at nominal timing. Looking at the offset injection timing cases, it is easy to see through

both the biased and unbiased torque values that combustion variability was increased by the injection timing retard for the cold start due to the increased width of lines; and by looking at the unbiased segment statistics, there is a clear trend that as the injection offset was increased, the observed segment standard deviation exhibited a consistent increase. Running the same tests when the engine is warmed-up ("Warm Start") generates cylinder torque traces that lack the variability seen during the "Cold-Start", which is reflected in both segment standard deviation plots. These experimental results are positive from a controls perspective, because they show that proper processing of combustion feedback can yield online estimation results that agree with existing variability guidelines ($< 2\%\text{CoV}_{\text{IMEP}}$), respond to available actuators like injection timing, and also to noise factors present during

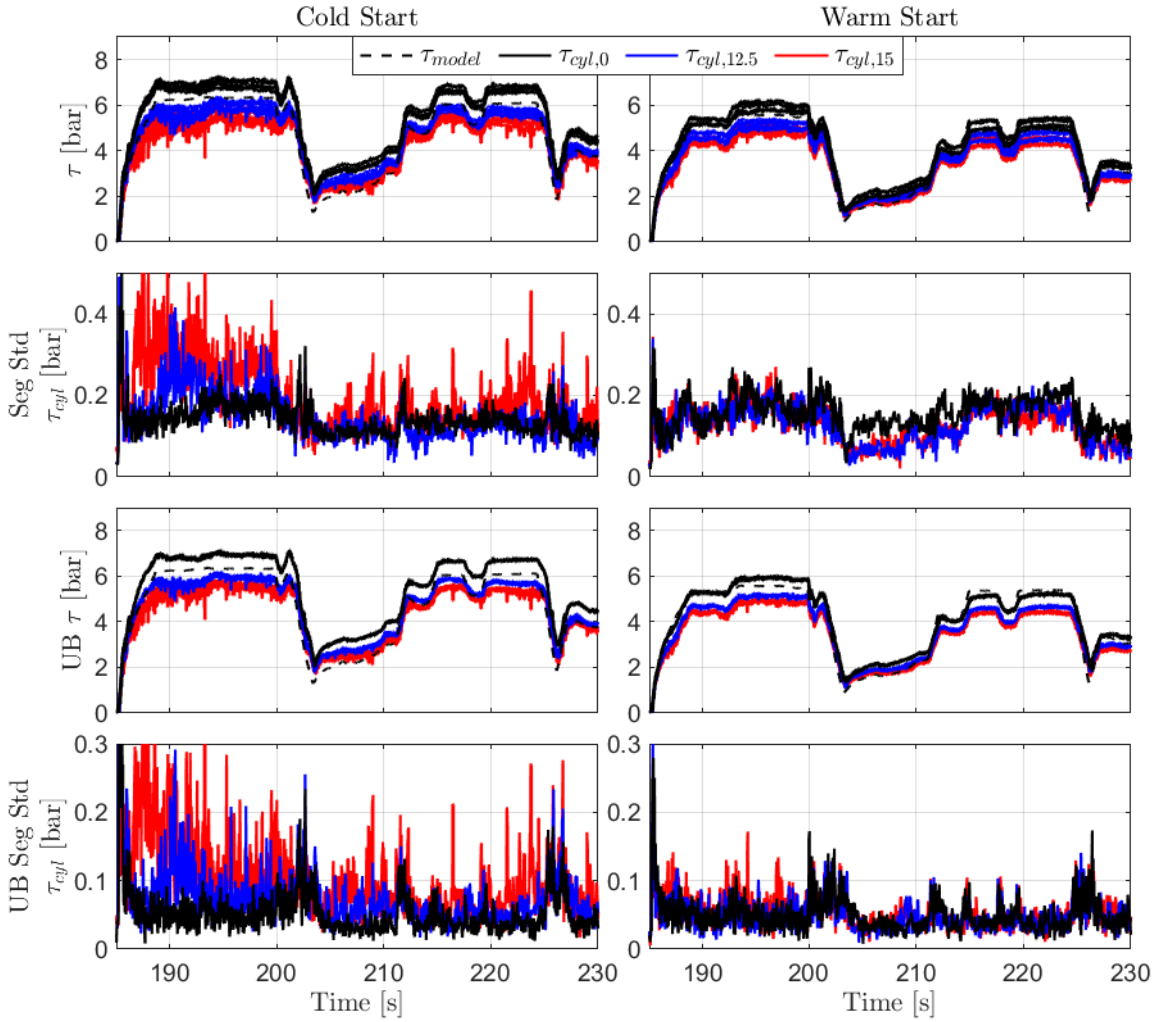


Figure 4.4: Torque and segment statistic data from FTP-75 data collected at two initial coolant temperature conditions ("Cold Start","Warm Start") and three injection offset phasings.

cold-starts like coolant temperature.

4.2.3 Single-Cylinder Approach

Although production-intent cylinder pressure sensors have reduced in price enough to be viable for production implementation (Volkswagen TDI's, Mazda Skyactive-X), they are still quite expensive, and so it may be economically infeasible to instrument every cylinder. Referring back to Figure 2.5, each cylinder in an engine may have a unique variability response to an injection phasing retard, and so it is critical to design the feedback structure in such a way that avoids causing excessive non-observable variability. These unique variability responses can be due to physical attributes of the engine design (block temperature distribution, individual cylinder breathing characteristics, EGR distribution), and therefore extensive testing on multiple engines (of the same architecture) could be performed to see if certain cylinders tend to be more variable than others to select which cylinder is used for feedback control, and/or to investigate if cylinder variability patterns exist that could be used to model/bound the other cylinders' variability based on the

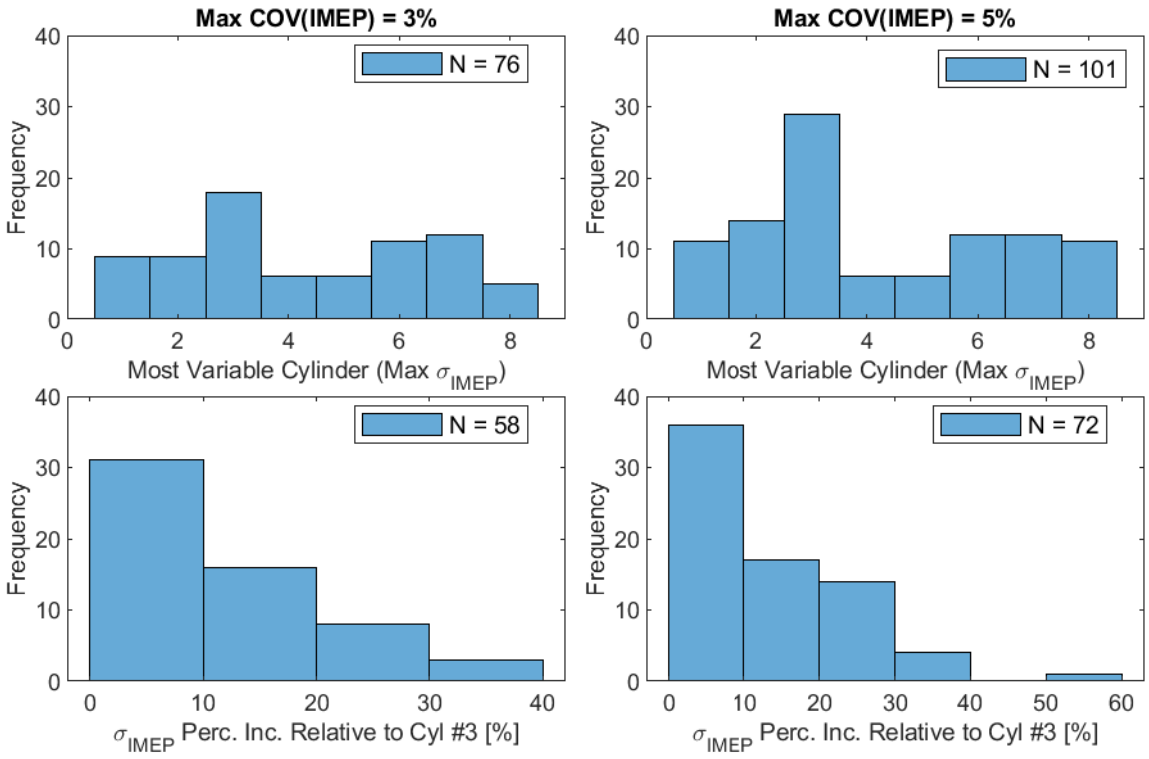


Figure 4.5: Histogram plots highlighting cylinder variability trends at a variety of engine speed and load points.

instrumented cylinder.

Experimental data presented in Chapter 2 was further analyzed based on this premise, with the results shown in Figure 4.5. The top plots show the frequency at which a particular cylinder had the highest σ_{IMEP}, with the left column considering data points where the maximum CoV_{IMEP} was 3% and the right column considering data points where the maximum CoV_{IMEP} was 5%. This down-sampling was performed to see the impact of increasing variability on the distribution. Note that if we assumed each cylinder had an equal probability of being the most variable cylinder, the expected frequency for each cylinder would be 9.5 when CoV_{IMEP} is bounded by 3% and 12.6 when CoV_{IMEP} is bounded by 5%. Instead, cylinder #3 was found to be the most variable cylinder 18 times when CoV_{IMEP} was bounded by 3% and 29 times when CoV_{IMEP} was bounded by 5%. At the points where cylinder #3 was not the most variable, it's standard deviation was still within 60% of the maximum cylinder's standard deviation, but more often within $0-10\%$ of the maximum cylinder's standard deviation for both CoV_{IMEP} bounds (N=31, N=36). Based on these observations, cylinder #3 was chosen as the sole cylinder for variance feedback control when using a single-cylinder architecture.

4.2.4 Transient Metric Validation

When using a single cylinder for variability feedback, the deterministic (requested) changes in torque can have a huge impact on windowed statistics as cylinder torque feedback is collected 8 times slower versus the fully-indicated approach. For example, to maintain a window size of $N = 8$, the combustion data in the buffer is now collected over 16 engine revolutions, which at 1000 rpm corresponds to just under a second temporally. To address the issues posed by deterministic torque changes, the detrended variance estimation analysis presented in 3.3.2 was applied to enable transient variance estimation using feedback from a single cylinder. Note that the analysis requires a predicted mean sequence, $\hat{\mu}_k$; for this work a pre-existing engine torque model value, τ_{model}, in the ECU was used as the predicted mean sequence for generating the detrended torque sequence, $\tau_{detrended}$. It's underlying structure is unavailable (proprietary), but is used for a variety of control purposes in the ECU. If such a model was unavailable, there are many examples of non-proprietary engine models capable of predicting torque, typically leveraging a combination of physics-based dynamical models and maps that are tuned/fitted with experimental data [46, 47]. An example of a engine model developed for a real-time embedded implementation can be found in [26].

The FTP-75 data described in 4.2.2 and shown in Figure 4.4 was further analyzed

with respect to the proposed single-cylinder detrended variability metric, with a snippet of the test data plotted in Figure 4.6. Similar to the previous plot, each line labeled $\tau_{cyl,X}$ represents combustion data from a trial with a constant X° timing offset applied during the entire test, and the columns represent data from different initial coolant temperatures. The manner in which injection timings were altered in the ECU software meant that the ECU torque model, τ_{model}, was unaffected by the change, and so it is the same for all three timing offsets. The top row of plots highlight how the torque model does a good job at predicting deterministic changes in engine torque - although the model doesn't perfectly capture the measured torque, it does a sufficiently good job capturing the deterministic trends, and note that as long as the model error is a constant offset it will not impact the variability estimate.

The second row of plots show the windowed standard deviation of τ_{cyl} for a window size $N = 8$ - note the sharp peaks in the standard deviation estimate when aggressive transients occur, mathematically explained by the mean sequence term μ_k in (3.7). The bottom row of subplots highlight the impact of performing statistics on the detrended signal, $\tau_{detrended}$

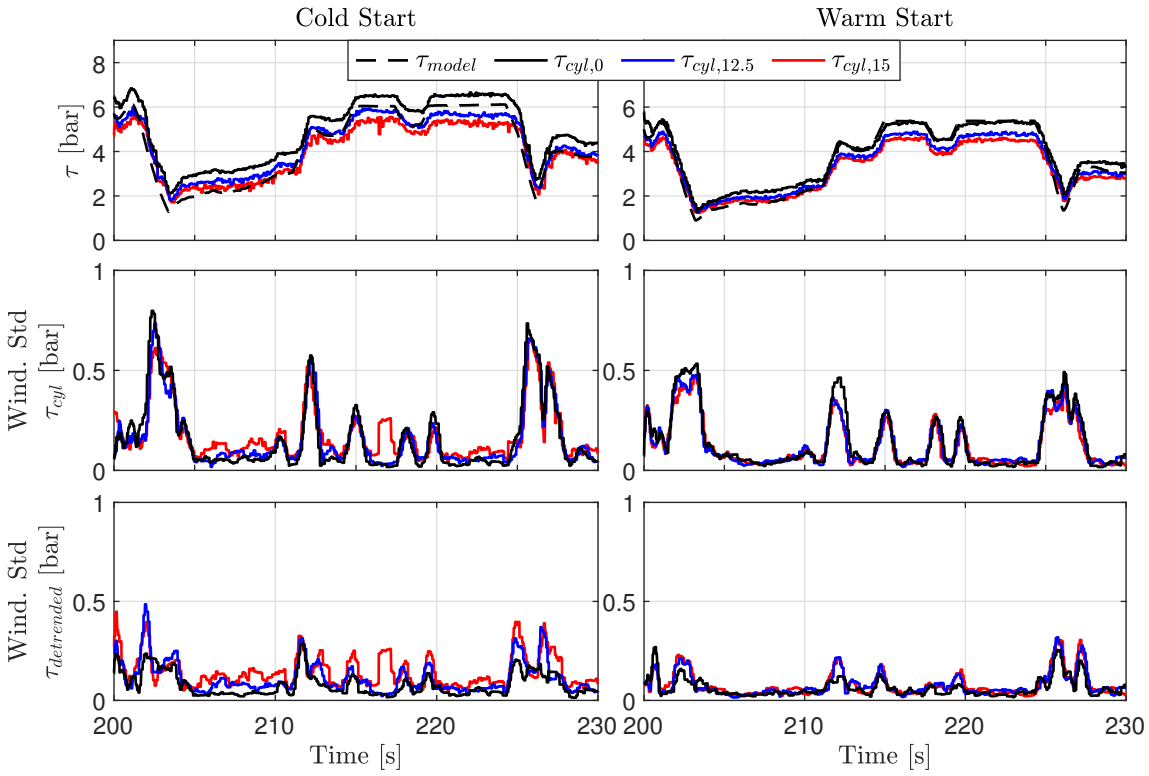


Figure 4.6: Cylinder torque, windowed standard deviation of cylinder torque, and windowed standard deviation of detrended cylinder torque data from FTP-75's collected at two initial coolant temperature conditions ("Cold Start","Warm Start") and three injection offset phasings.

- the magnitude of the estimator peaks are dramatically decreased during transients. By reducing the estimator sensitivity to transients, it is easier to observe actual combustion variability in the engine torque signal τ_{cyl} when the engine is cold and injection timing is highly retarted ("Cold Start", $\Delta\varphi = 15$, red line). Note the visible increase in the magnitude of the detrended variability estimate relative to the other two trials.

4.2.5 Proposed Control Architecture

The model-based control design work was combined with the transient metric validation to form the control architecture shown in Figure 4.7. Starting from the right, cylinder pressure data is analyzed to calculate the IMEP, τ_{cyl}, a measure of the torque an individual cylinder is producing. This torque is detrended by subtracting off the predicted/modeled value from the ECU, τ_{model}, to generate the detrended torque value, $\tau_{detrended}$. The detrended torque is used to generate an online estimate of the combustion variance, $\widehat{\sigma^2}$, which is subtracted from a target/reference variance, σ^{2*}, and fed into the variability controller. To retard combustion phasing for higher exhaust gas enthalpy and combustion variability, the injection timing of the engine is offset by the controller output $\Delta\varphi$. Retarded combustion phasing reduces the efficiency of the combustion cycle, as energy that would normally be converted to shaft work instead raises the enthalpy of the exhaust gas leaving the cylinder. To compensate for the torque loss that will occur from the less efficient combustion cycle, additional fuel is added via two mechanisms. The first is a feed-forward fuel offset, $\Delta m_{fuel,FF}$, that is added to maintain torque based on the injection timing offset applied. The command is a function of engine speed, load, and injection timing offset, and was calibrated using experimental engine data. The second fuel command, $\Delta m_{fuel,cont}$, is the output of a pure integral controller that is trying to minimize the difference between τ_{cyl}

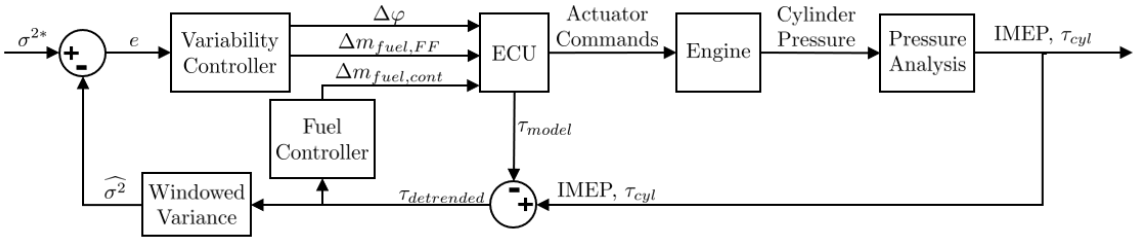


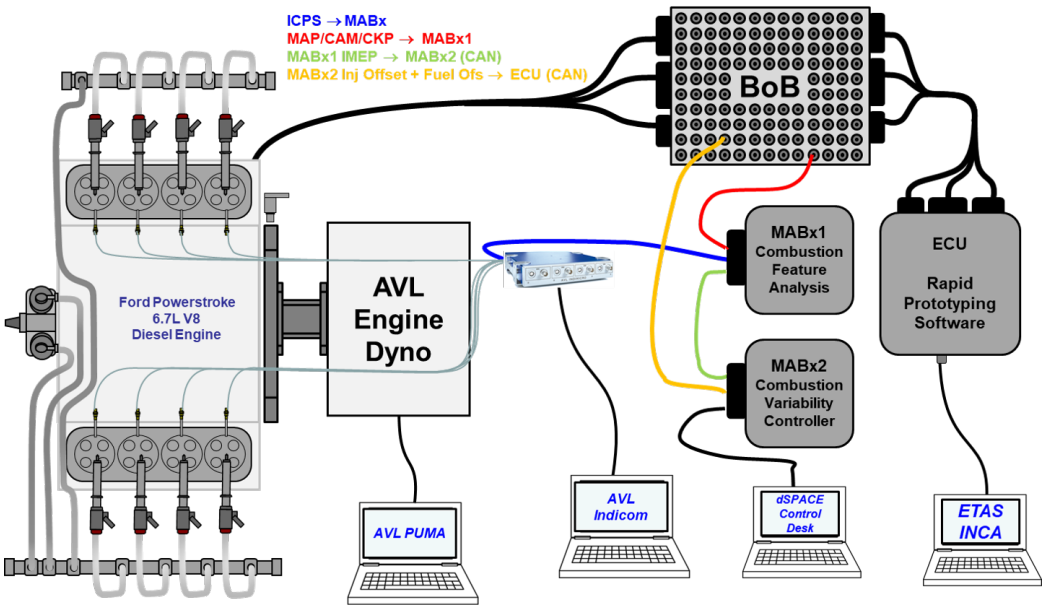
Figure 4.7: Diagram showing the cylinder torque variance controller implemented on test engine setup. The Variability and Fuel controller are simple integral controllers to avoid excessive actuator variability due to random cyclic combustion behavior. The actuator commands are applied uniformly to all cylinders in the engine, while only one cylinder is used for variability feedback.

and τ_{model}. The presence of the feed-forward command avoids torque dips due to injection retard that would need to be observed by the fuel controller before corrective action could be applied.

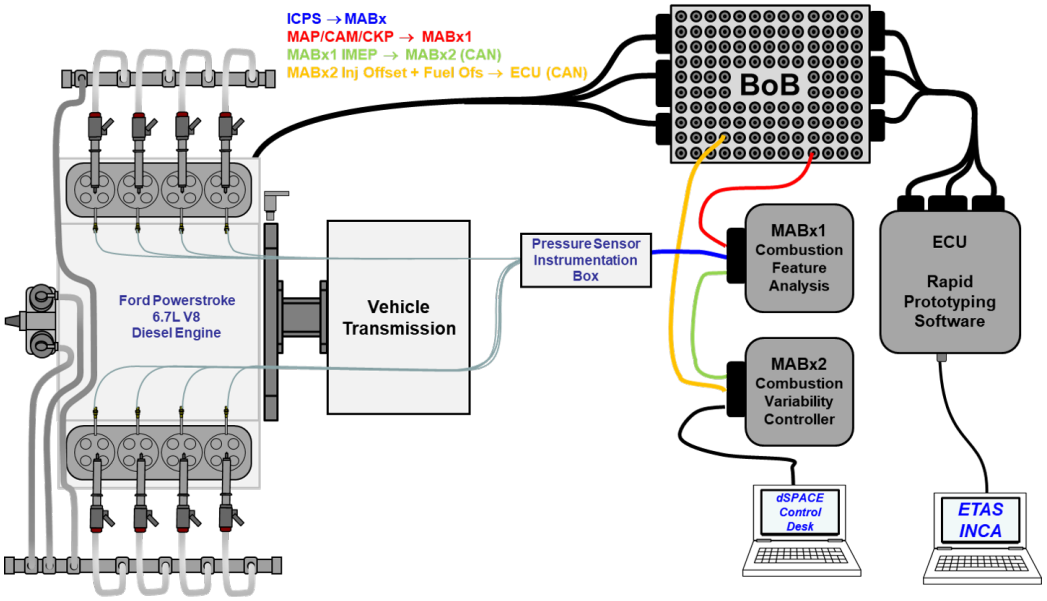
4.3 Experimental Setups

The engine dyno setup used for the characterization experiments presented in Chapter 2 was upgraded to support cyclic closed-loop combustion control, which can be seen in the test setup diagram presented in Figure 4.8(a). The AVL IndiCom system associated with the dyno cell lacked real-time communication support, which motivated the implementation of "MABx1, Combustion Feature Analysis", a dSPACE MicroAutoBox system responsible for combustion feature analysis that could transmit calculated features via CAN in real-time. The system sampled manifold pressure, crank position, cam position, and cylinder pressure sensor output to apply the analysis described in Section 2.2.2. The output from "MABx1" was validated against output from the AVL IndiCom system during the characterization experiments to ensure proper functionality. Combustion features of interest were sent over CAN to "MABx2, Combustion Variability Controller", which ran the combustion variability controller and sent injection timing and quantity offsets to the ECU via CAN. A computer running dSPACE ControlDesk was used to control and tune MABx2. The ECU used for experiments ran rapid prototyping software that applied the offsets received via CAN to the injection strategy and was controlled using ETAS INCA software.

Experimental testing was also conducted using two production medium-duty vehicles equipped with the same Ford Powerstroke diesel engine tested in the dyno lab. A diagram of the vehicle test configuration can be seen in Figure 4.8(b). The primary difference versus the dyno test setup is the absence of development-grade indicating analysis equipment (AVL IndiCom, piezoelectric presusre sensors) that aren't well-suited for real-world testing environments. The piezoelectric pressure sensors were replaced with production-intent ratiometric pressure-sensing glow plugs. The sensors were connected using an instrumentation box which provided a voltage reference and ground, and routed the sensors signals back to MABx1 for combustion analysis.



(a) Dyno Configuration



(b) Vehicle Configuration

Figure 4.8: Experimental test configurations. In-cylinder pressure sensor (ICPS), manifold absolute pressure (MAP), camshaft position (CAM), and crankshaft position (CKP) sensors were connected to MicroAutoBox 1 (MABx1) using a breakout box (BoB) to enable cylinder pressure analysis. The ICPS signals were also read in by an AVL IndiCom system. CAN was used to interface the MABx's and ECU to transmit data and controller actuator values.

4.4 Steady-State Control

Steady-state control experiments were conducted at the modeled speed/load point (1200 rpm/2.5 bar BMEP) to verify concept feasibility, enable comparison to the stochastic engine model, and observe controller impacts on system statistics. These experiments were conducted using the dyno test configuration shown in Figure 4.8(a).

4.4.1 Control Simulations and Experiments

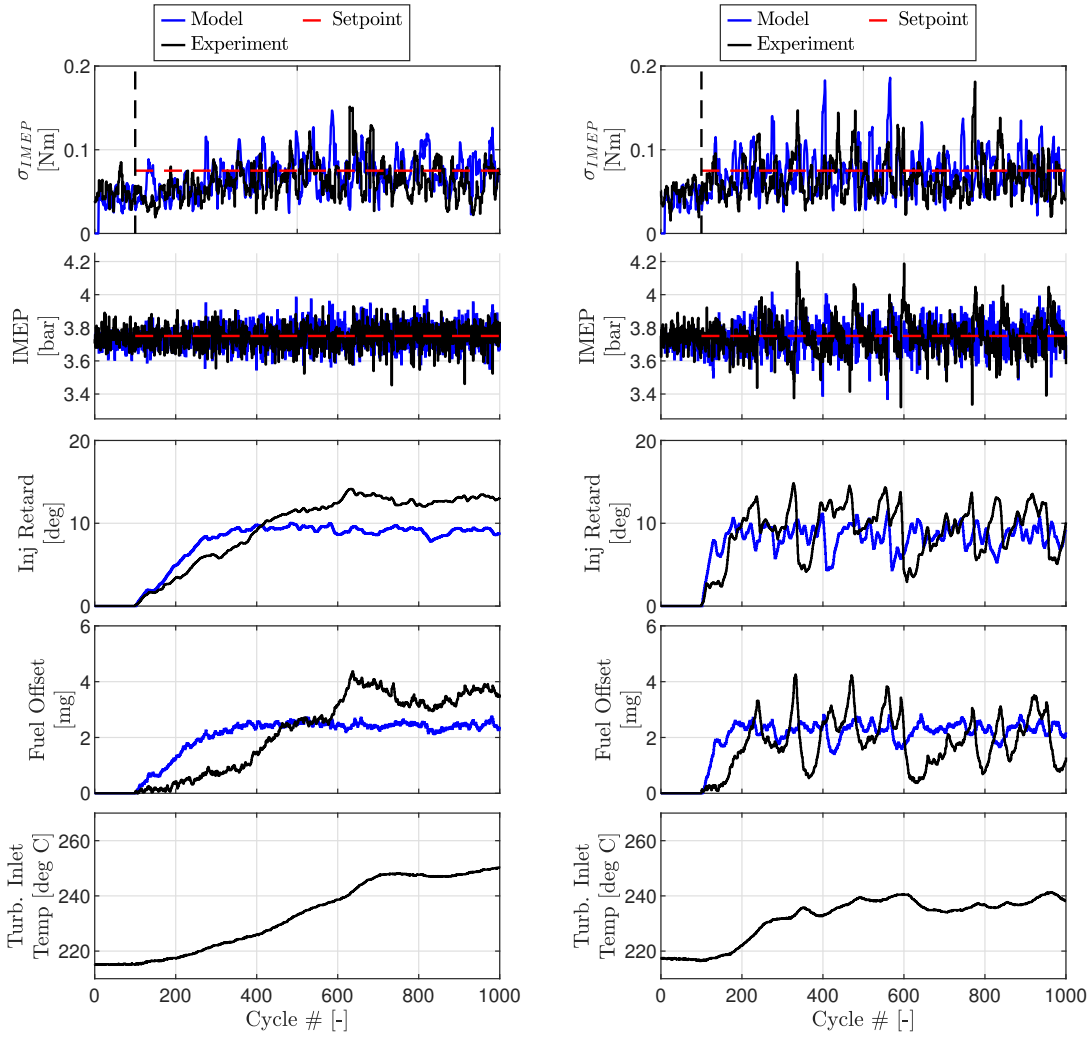
A variety of controllers were tuned with the engine model using window sizes that varied from $N = 4$ to $N = 24$, with intended damping ratios (ζ) based on a pure N point feedback delay due to the windowed statistics. The theoretical variance estimation and controls analysis revealed this to be a conservative estimate, as it shows the windowed statistics instead behave like a moving average, and so it is instead noted in quotation marks ("ζ"). The variance target σ^{2*} for all trials was $(0.075)^2$ bar, which would achieve a 2% $\mathrm{CoV}_{\mathrm{IMEP}}$.

Experimental data for the window and tuning configurations can be found in Table 4.1. Four experimental trials were run for each window/gain configuration, and one thousand cycles of closed-loop steady-state combustion/actuator data were used for calculating statistics. The experimental data confirms that simultaneous control of both the feedback variance and mean value of IMEP for a single cylinder is feasible using the proposed feedback structure. Initial review of the closed-loop statistics showed that the fuel controller did an excellent job regulating μ_{IMEP} to the nominal torque of 3.75 bar, but the

Table 4.1: Steady-state closed-loop control statistics for various window and gain configurations. One thousand cycles of data were used to calculate statistics.

N	"ζ"	$\frac{K_i}{K_{i,0}}$	μ_{IMEP}	σ_{IMEP}	σ^*_{corr}	$\mu(\Delta\varphi)$	$\sigma(\Delta\varphi)$
[-]	[-]	[-]	[bar]	[bar]	[bar]	[CA deg]	[CA deg]
4	1	11.76	3.75 ± 0.001	0.117 ± 0.004	0.100	11.7 ± 0.3	1.7 ± 0.2
4	0	23.42	3.75 ± 0.001	0.151 ± 0.009	0.100	8.9 ± 0.2	3.6 ± 0.3
8	1	2.65	3.75 ± 0.000	0.082 ± 0.002	0.086	12.7 ± 0.3	0.7 ± 0.1
8	0.707	4.01	3.75 ± 0.001	0.084 ± 0.002	0.086	12.5 ± 0.3	0.8 ± 0.2
8	0	12.41	3.75 ± 0.001	0.112 ± 0.004	0.086	9.9 ± 0.4	3.0 ± 0.2
16	1	1.47	3.75 ± 0.000	0.075 ± 0.000	0.080	12.5 ± 0.4	0.5 ± 0.1
16	0.707	2.05	3.75 ± 0.001	0.079 ± 0.001	0.080	13.0 ± 0.2	0.6 ± 0.1
16	0	6.37	3.75 ± 0.002	0.095 ± 0.005	0.080	12.1 ± 0.2	1.8 ± 0.2
24	1	$K_{i,0}$	3.75 ± 0.001	0.077 ± 0.001	0.078	13.0 ± 0.3	0.5 ± 0.1
24	0.707	1.38	3.75 ± 0.001	0.077 ± 0.001	0.078	13.6 ± 0.2	0.5 ± 0.0
24	0	4.29	3.75 ± 0.001	0.085 ± 0.004	0.078	13.0 ± 0.3	1.6 ± 0.9

variance control resulted in population σ_{IMEP} values exceeding the 0.075 bar target, even when the average variance estimate $\widehat{\sigma^2}$ was equal to the setpoint. The discrepancy can be explained by the varying-mean sequence derivations in 3.3.2, where equation (3.7) shows that the expected value of the variance estimator is biased by a factor of $\frac{N-1}{N}$. The controller was amended to avoid this issue in subsequent testing, but for this data set the bias factor was used to calculate σ^*_{corr}, the population standard deviation target after accounting for this estimation bias. With this correction in place, the closed-loop statistics when "$\zeta = 1$" show good agreement with the setpoints. Increasing the gain in nearly all cases causes the average injection retard to decrease, the variability in injection retard to increase, and the closed-loop σ_{IMEP} to increase as well - this behavior can be seen temporally in Figure 4.9.



(a) $N = 8$, "ζ"$= 1$ (b) $N = 8$, "ζ"$= 0$

Figure 4.9: Steady-state experimental data

Figure 4.9 shows experimental and model data for feedback controllers designed using an 8-point statistical window for "ζ" values of 1 and 0. Comparing the experimental and model results demonstrates the value of using feedback - if an open-loop injection retard strategy was designed using the model, it would only retard injection timing by about 10 CA°. By using a feedback controller that is directly controlling variability, Figure 4.9(a) shows it is possible to get another ≈ 2.5 CA° of injection retard while avoiding excessive degradation of combustion. Figure 4.9(b) shows the consequences of aggressive controller tuning, which can result in hunting behavior - not only does it cause large variability in IMEP, but it does so while generating a lower exhaust temperature increase with respect to the "$\zeta = 1$" case. This highlights a fundamental tuning trade-off between response-time and steady-state performance when using a fixed controller gain; however, gain scheduling and/or a controller deadband could be used to simultaneously improve both aspects of the closed-loop behavior.

One thousand points of the steady-state closed-loop IMEP and windowed variance estimate $\widehat{\sigma^2}$ data at the various gains for $N = 8$ were used to generate the histograms along with the sample autocorrelation function (ACF) of IMEP plots shown in 4.10. The red lines in the IMEP histograms show the normal distributions using the sample mean and standard

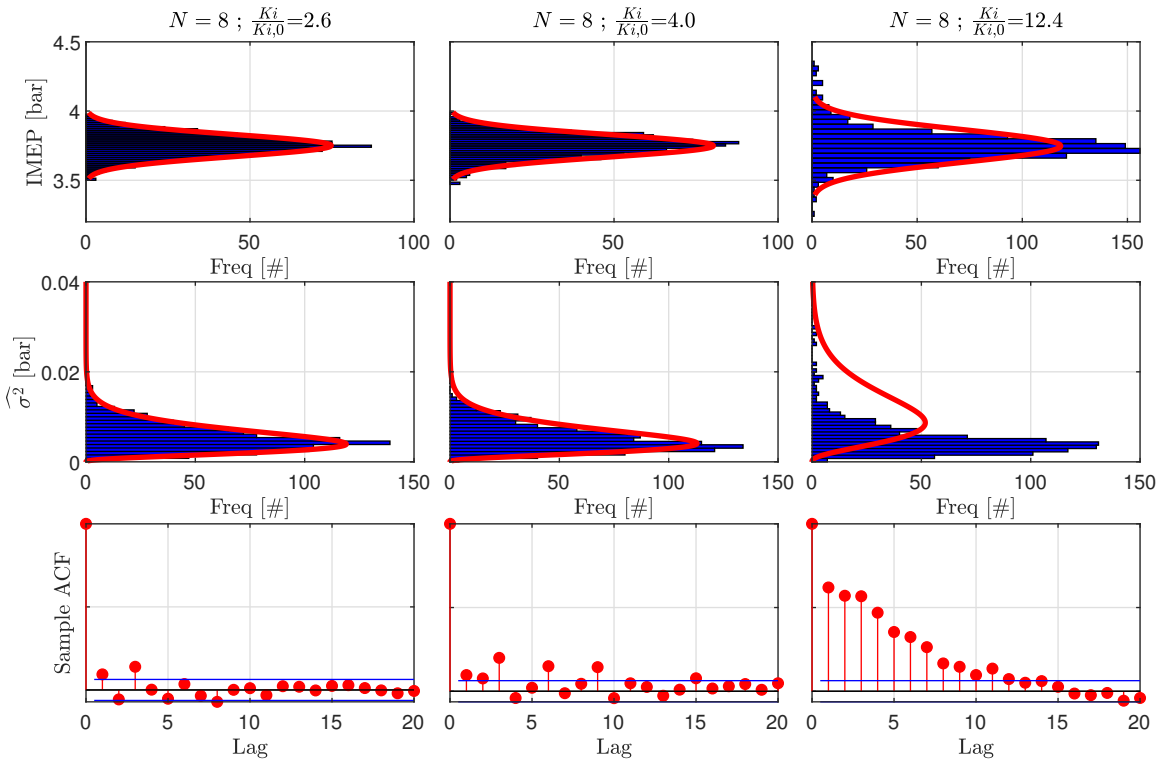


Figure 4.10: Closed-loop IMEP distribution behavior when $N = 8$.

deviation, and the red lines in the $\widehat{\sigma^2}$ histograms show the theoretical windowed variance estimate distribution (χ^2) assuming independent, identically-distributed events. At lower gains, the closed-loop statistical distributions show rough agreement with theoretical distributions. However as the gain is increased, IMEP events tend to be more concentrated around the mean, indicating an increase in the kurtosis of the distribution, which was also observed in the numerical experiments presented in Figure 3.10. For the largest gain test, the increase in the standard deviation of IMEP causes the theoretical $\widehat{\sigma^2}$ to grow wider, whereas the observed closed-loop distribution behavior still resembles what was seen at lower gains. This disagreement is sensible, as the theoretical distribution is based on a iid normally distributed sampling pool; the IMEP outliers at the higher control gain result in an inflated σ_{IMEP} that causes a wider theoretical $\widehat{\sigma^2}$ distribution, and the sample ACF function shows that consecutive events are now highly correlated.

4.4.2 Long Duration Steady-State Setpoint Sweep

Based on the long-term thermal dynamics observed in 2.4.1, a long-duration setpoint sweep experiment was conducted at the modeling point of 1200 rpm / 2.5 bar BMEP. A fully-indicated control approach using segment statistics was used for the experiment. The gain from the $N = 8$, "$\zeta = 0.707$" test case in Table 4.1 was used in the controller. NO_X and total hydrocarbon (THC) emissions were sampled using the AVL SESAM i60 emissions bench. Figure 4.11 shows the results of the sweep, where the black lines show experimental data, the red-dashed lines indicate the segment IMEP and segment CoV_{IMEP} setpoints (controller variance setpoint calculated using CoV_{IMEP} target), and the vertical blue-dashed lines separate the different setpoint regions. The CoV_{IMEP} target of 1% yields little controller activity, however adjusting the setpoint to 1.25% causes about a 10 CA° injection retard. The initial behavior at the onset of the step at about 1250 seconds is particularly interesting due to the rapid retard, advance, followed by retard of injection timing. The fact that the controller returns and remains stable at the injection retard value reached prior to the event suggests that the initial retard maneuver occurs faster then the previously hypothesized variability stabilization dynamics could occur. The controller observes the excessive variability, advances injection timing, but then re-approaches the previously commanded value, where now the dynamics have resolved and the timing can be tolerated at the CoV_{IMEP} of interest.

As the CoV_{IMEP} target was further increased, injection timing, exhaust temperatures, and THC emissions continued to increase, and NO_X continued to decrease, as expected. Note that the increases in THC outpaced reductions in NO_X, unfavorable when being

evaluated against combined THC+NO_X standards. There was also an observable increase in the variability of those signals, with the exhaust temperature and THC emissions variability likely being due to some combination of the combustion variability, as well as the injection retard actuator variability Note the long timescale changes in injection retard, turbocharger inlet temperature, and emissions as the residence time at CoV_{IMEP} setpoints $> 1.5\%$ increases, again indicating the presence of long timescale combustion variability dynamics originally discussed in 2.4.1.

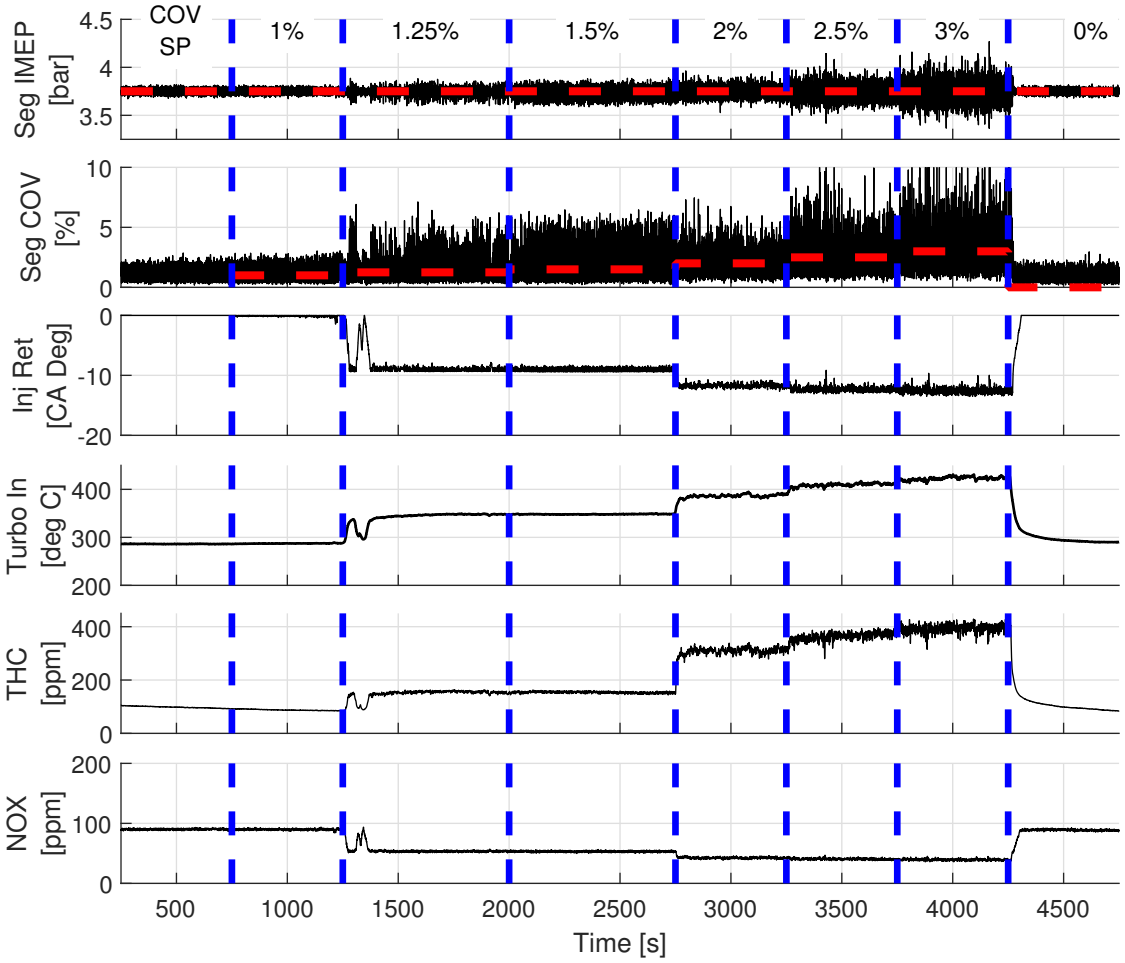


Figure 4.11: Experimental data from long-duration fully-indicated segment variability setpoint sweep.

4.5 Transient Control

After validating both the variance control theory and experimental dyno test set-up at steady-state operation, the controller was enabled during transient FTP-75 cycles on dyno.

Positive transient control results in the engine dyno cell led to controller implementation on the first test vehicle, which was evaluated against low-load real-world driving behaviors. The controller was then adapted to another test vehicle, which was tested in an emissions-equipped chassis dyno cell against the FTP-75 emissions certification cycle to understand contoroller impacts on fuel economy and emissions.

4.5.1 Transient Control, Engine Dyno

Transient control experiments on dyno were conducted over the first 505 seconds of the FTP-75 cycle. For this testing, the fuel controller was disabled and only a feed-forward fuel command was used due to the accuracy of the feed-forward fuel command. The variance target, σ^{2*}, was calculated based on a CoV_{IMEP} target of 2% using the current average torque. Experimental data showing the impacts of the controller can be seen in Figure 4.12. "Nominal" refers to a test without any external intervention (nominal cold-start exhaust temperature behavior) while "CVC" is a plot with the Combustion Variability Controller active. Note how the controller is able to retard injection timing (third plot) to achieve hotter exhaust temperatures (fourth plot) and maintain a similar cylinder torque trajectory (first plot) to the nominal test case. Similar to the other figures showing online variance estimation feedback, the variability feedback hovers about the setpoint, due to the noise present in the variance estimate. Analyzing the variability data over the cycle shows that the average windowed standard deviation of the detrended torque was increased 0.015 bar (0.79 Nm) by the controller, relative to the average commanded increase of 0.017 bar (0.91 Nm). The DOC light-off time (outlet temperature $> 200^{\circ}$C) was reduced by 14 seconds, while SCR light-off time was reduced by 74 seconds. The dramatic SCR light-off time reduction was made feasible by the aggressive injection retard during the second hill of the FTP test (150-300s). Emissions data was unavailable for this testing, but based on observations from the characterization experiments, the controller activity during the second hill of the FTP ($\approx$160-300 seconds) likely generated large amounts of hydrocarbons due to the large injection retard ($> 15\text{CA}^{\circ}$) at higher loads.

4.5.2 Transient Control, Vehicle Low Load Driving, Fuels

The first batch of vehicle test data was focused on evaluating controller performance against low-load real-world driving behaviors, fuel controller configurations, and sensitivity to different fuels. The test vehicle was a medium-duty pickup truck with a Ford 6.7L Powerstroke engine with the experimental configuration shown in Figure 4.8(b). Limitations of the engine control software at the time of testing restricted engine operation

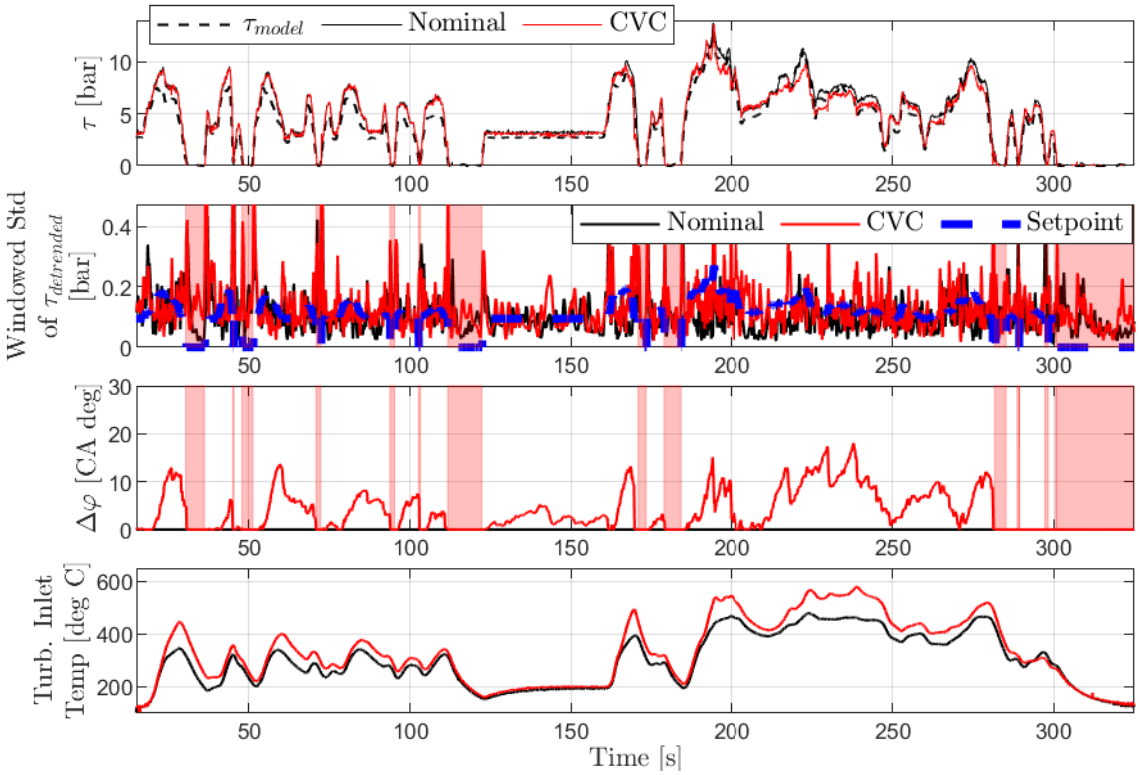


Figure 4.12: Plots highlighting the functionality of the Combustion Variance Controller (CVC) during engine transients. The highlighted portions indicate sections of the transient test where the engine was not injecting fuel, so the controller was reset to zero.

to the nominal calibration, where fuel economy is a large priority, versus aftertreatment heating in the cold-start warm-up calibration. Data collection was performed while the vehicle was warmed-up, necessary to generate a sufficiently large sample pool with the limited amount of test time available. Although the data is not directly relevant to the cold-start issue trying to be addressed, it provides valuable insight into the controllers response to noise factors, as well as the impact of the feed-forward fuel command. Statistics were calculated for values of interest (injection retard, turbo inlet temperatures, combustion phasing) to observe the impact of the different parameters.

Three controller configurations were evaluated during this testing; “No Control”, where no feedback control was used during the tests; “Var”, where only the variance controller was used during the tests; and “Var & FF-Fuel”, where the variance controller and a feed-forward fuel controller were used during the tests. Three fuels were also tested; “ULSD” (Ultra-Low Sulfur Diesel), conforming to the ASTM D975 D2 specification (thought to be around 42 cetane); “B15”, which was ASTM D975 mixed with 15% biodiesel meeting the ASTM specification; and “EURO”, which was European diesel fuel conforming to

the EN590 specification (thought to be around 52 cetane). Twenty trials of each fuel and controller configuration were collected, for a total of 180 samples. Low load drive cycle timeseries data while testing with ULSD can be seen in Figure 4.13. The vehicle speed profile consists of an acceleration to 20 MPH, followed by a short cruise, followed by an acceleration to 35 MPH, followed by a long cruise; note the data was collected on a real road (not lab environment), and so trials were not perfectly repeatable. Note that when the feed-forward fuel command was disabled (blue line), the pedal position, which correlates to an engine torque request, required to maintain the vehicle trajectory increased, showing the impact of injection timing retard on torque without any open-loop or closed-loop fuel control. It also demonstrates how a human driver-in-the-loop can act as a controller to maintain the desired torque, although the driver will likely be able to perceive the controller's impact on the vehicle's driving characteristics due to the increased pedal required. The normalized torque traces, τ_{cyl}, highlight the low load nature of the cycle (using less than 1/3 of the total torque of the engine), and how the torque traces were consistent between the trials, sensible due to the goal of maintaining similar velocity profiles for each trial. Injection offsets for both control cases trended towards 20 CA°, more than observed in steady-state dyno experiments. This can be attributed to the warmed-up coolant and also operating in the nominal calibration, where the nominal combustion phasing is earlier in the cycle and so more combustion phasing retard can be tolerated prior to variability issues. The exhaust temperature profiles demonstrate that large increases in exhaust temperatures can be achieved even during low-load driving, with 100°C temperature benefits visible throughout the cycle.

Statistics of values of interest which included injection retard, turbocharger inlet temperatures, combustion phasing, and combustion variability were calculated and plotted using histograms to study the impacts of different fuels and controller configurations and are shown in Figure 4.14. Over the cycle, the controller was able to retard the nominal injection timing about 15-20 CA° past its nominal value. Note that the controller was able to retard injection timing $\approx$ 2.5 CA° more when the engine was running on the European diesel, while maintaining a similar if not lower level of combustion variability - this is hypothesized to be due to the increased cetane rating of the fuel, as observed in literature [2]. The injection retard drove average MFB50 values past 30 deg aTDC without misfires or other negative vehicle behaviors - the higher average injection retard values for the European diesel are reflected in the later MFB50 values as well. Average turbocharger inlet temperature benefits approaching 100°C were observed for the ULSD and European diesel tests. The B15 fuel tests show the lowest net temperature increase, but with similar injection retard and MFB50 values as ULSD - this may be due to lower fuel energy content

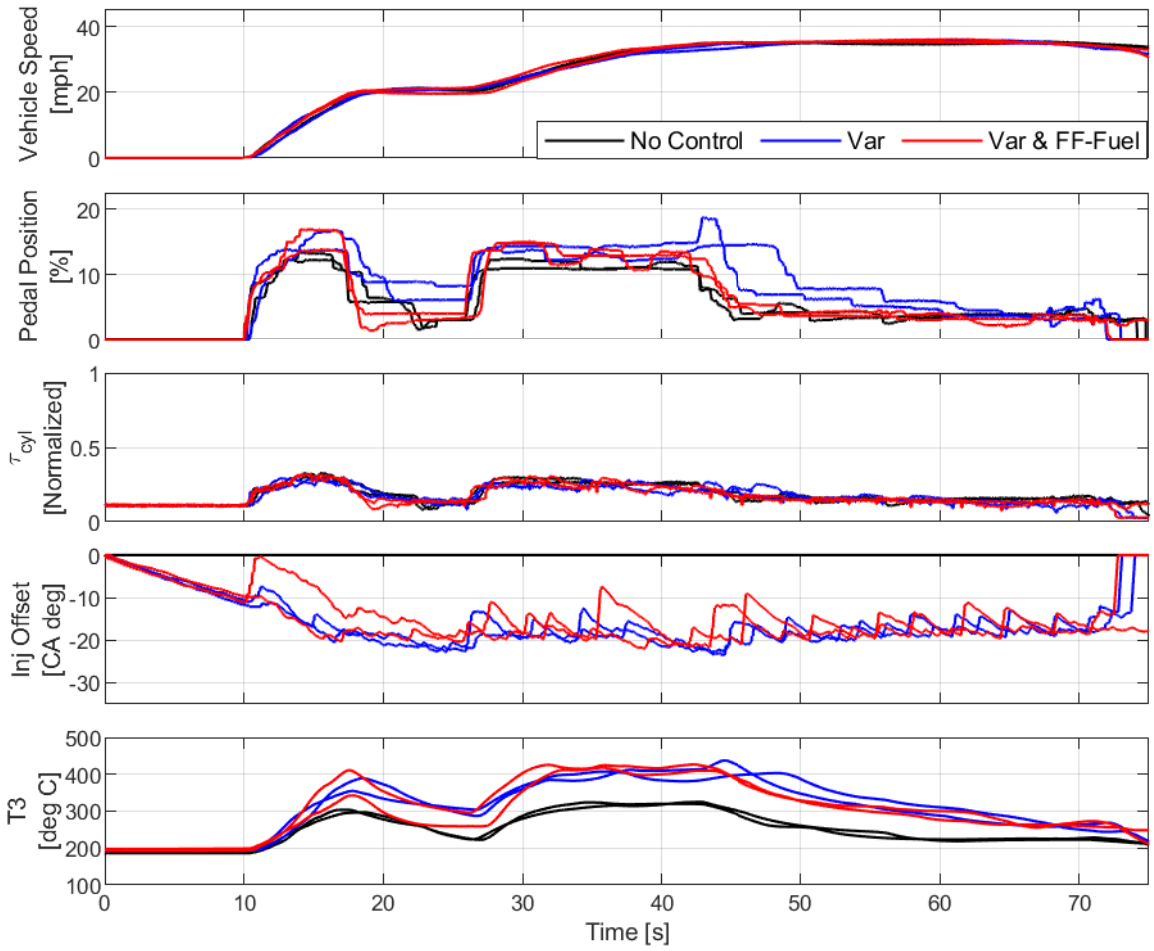


Figure 4.13: Selection of low-load drive cycle traces when using ULSD fuel.

of B15 fuel.

4.5.3 Transient Control, Vehicle Low Load Driving, Injection Strategy

Additional controller testing opportunities using the same vehicle and hardware as discussed in 4.5.2 became available the following year. The engine control software issues were addressed, permitting operation in the aftertreatment heating ("AT Heating") mode that the vehicle operates in during cold-starts, rather than the nominal ("Nominal") mode that the vehicle operates in once warmed-up. To enable comparisons to the previous data, the same low-load drive cycle was used and was tested using four configurations; "ULSD/Nominal/No Control", using ULSD fuel in the nominal calibration without variance control; "ULSD/AT Heating/No Control", using ULSD fuel in the aftertreatment

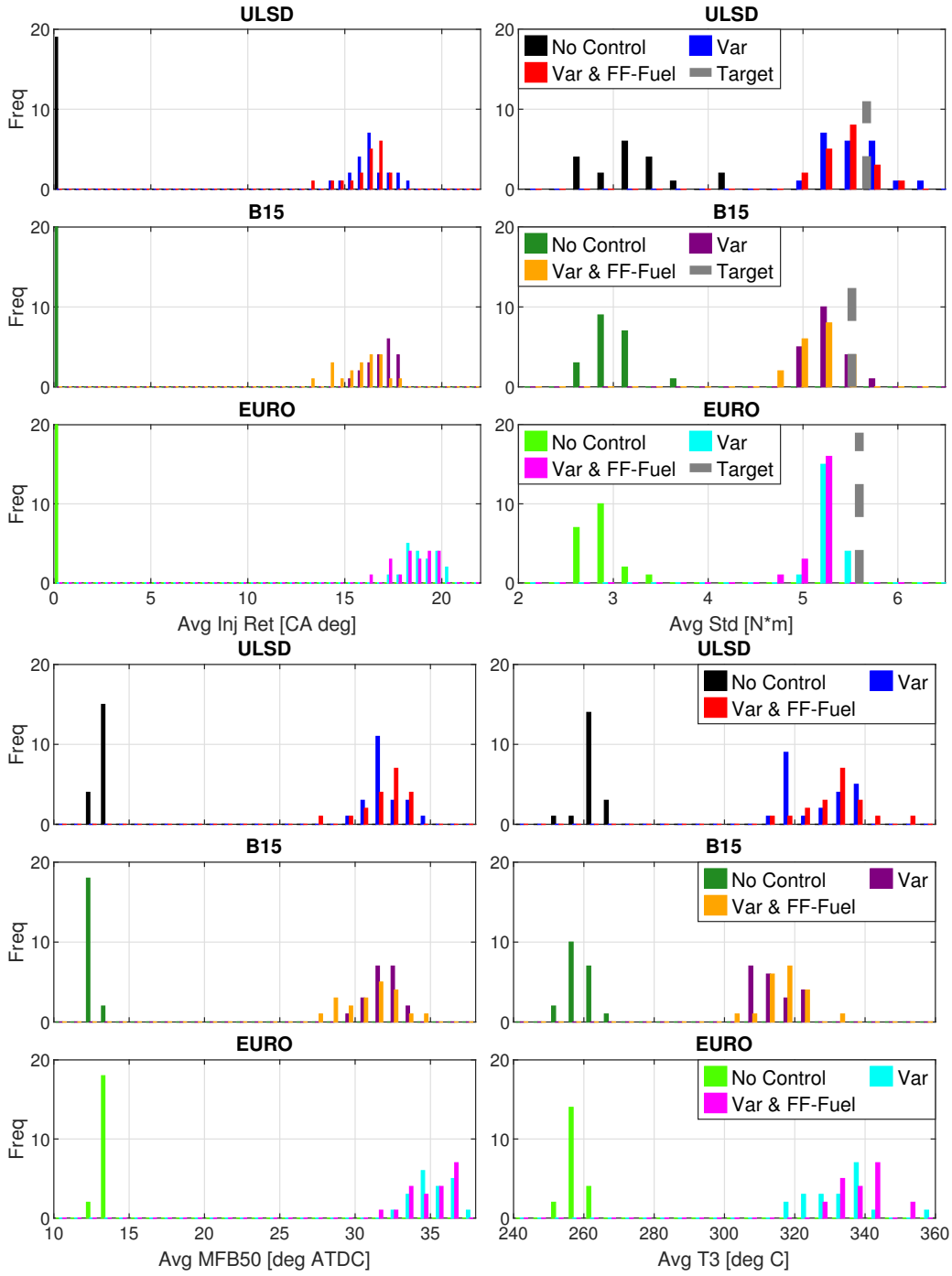


Figure 4.14: Low-load drive cycle statistics, including impacts of different fuels and fuel control strategies.

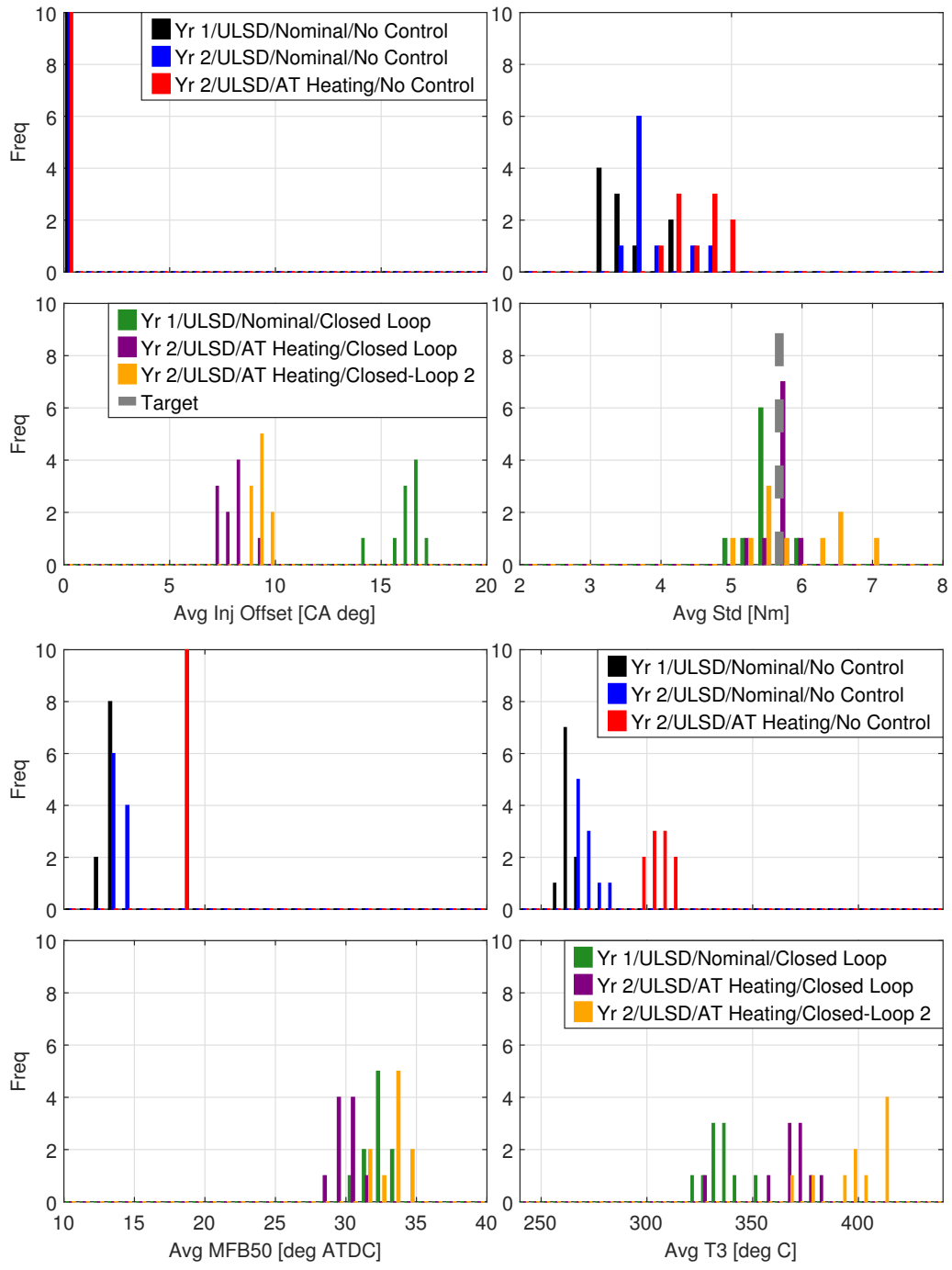


Figure 4.15: Additional low-load drive cycle statistics, comparing impacts of active engine calibration and injection offset scheme.

heating calibration without variance control; "ULSD/AT Heating/Closed-Loop", using ULSD fuel in the aftertreatment heating calibration with variance control retarding all injections (pilots, main, post); and "ULSD/AT Heating/Closed-Loop 2", using ULSD fuel in the aftertreatment heating calibration with variance control retarding only the pilot and main injections (post-injection timings left undisturbed). Ten trials of each configuration were tested.

Statistics for the trials can be found in Figure 4.15. Ten datapoints from the previous batch of testing ("Yr 1/ULSD/Nominal/No Control" and "Yr 1/ULSD/Nominal/Closed-Loop") were included for comparison. The "Nominal/No Control" trials for both years show good agreement, indicating that there were no dramatic changes in the vehicle or engine during the pause in testing. Enabling the aftertreatment heating calibration yields a $\approx$1 Nm increase in average variability, an average MFB50 retard of $\approx 5^{\circ}$, and an average exhaust temperature increase of $\approx 30^{\circ}$C. Enabling the controller while in the aftertreatment heating mode results in a lower average injection offset, but comparable average MFB50s compared to the nominal operating mode. Exhaust temperature benefits up to 100°C are still observable (with respect to the nominal values in each calibration) Note that decoupling the post-injections from the controller ("ULSD/AT Heating/Closed-Loop 2") allows the controller to retard a few additional degrees, reflected in both the average MFB50 and average exhaust temperature statistics. Achieving similar to hotter exhaust temperatures while not disturbing the post-injection timings likely resulted in lower feedgas hydrocarbons (on-road data, so THC emissions unavailable), which would be critical during cold-starts when the diesel oxidation catalyst is not converting hydrocarbons.

4.5.4 Transient Control, Vehicle Dyno FTP

Observations from both the characterization and closed-loop control experiments were used to further refine the single-cylinder combustion variability controller. The changes included limitations on the maximum injection retard value to avoid excessive hydrocarbon emissions, unbiasing the variability setpoint to achieve the correct population statistics, disabling the controller at a target aftertreatment temperature to avoid overheating, and decoupling the post injections from the controller to avoid retarding already late injections in the cycle.

The refined controller was implemented on another medium-duty pickup truck with a Ford 6.7L Powerstroke engine, and was evaluated against the FTP-75 cycle in an chassis dynamometer cell equipped with a constant volume sampling (CVS) emissions system for measurement of true bag emissions. The controller was configured to turn

off when the diesel oxidation catalyst (DOC) outlet temperature reached 200°C, when it is approximately lit-off. The transition from the aftertreatment heating calibration to the nominal operating calibration was manually controlled based on a threshold DOC outlet temperature of either 250 or 275°C - manual control was necessary due to software constraints, with the temperature being changed to observe the impacts on cycle behavior. The emissions and temperature data was post-processed to evaluate the change in parameters of interest. Unfortunately, initial analysis of the results revealed the SCR catalyst was no longer functional (severely degraded), and so the measured NO_X emissions were effectively engine-out. To project the impacts of the controller in an application with a functional SCR system, a simple temperature-based SCR conversion efficiency model was combined with the SCR catalyst temperature data to estimate the average SCR conversion efficiency from the start of the cycle to SCR light-off (at ≈ 200°C) [48].

Table 4.2 summarizes the impact of the controller on the Bag 1 (cold-start phase) FTP-75 data. The unexpected result during the testing was the decrease in hydrocarbon (HC) emissions coupled with the increase in feedgas NO_X - as retarding combustion phasing typically induces the opposite behavior, a reduction in NO_X but increase in HC. The reduction in HC emissions is likely due to the improved heating of the DOC and the restricted authority of the controller - note that DOC conversion efficiency functions typically resemble sigmoid functions with the transition occuring between 100-200°C, and so any improvement in temperature can yield significant increases in conversion efficiency, and controller operation was restricted to avoid the excessive HC increases observed during the characterization experiments [49].

The large NOx increase for the 250°C calibration switch test can be explained by the switch point - the improved heating during the closed-loop trial caused the calibration transition to occur during the idle portion between hills 1 and 2, whereas during the baseline trial the transition occurred ≈ 30 seconds into hill 2. The beginning of the second hill includes a large load transient, which yields a large amount of NO_X when operating in the nominal operating mode. Normally this NO_X would be converted by the SCR, but as mentioned, the SCR was non-functional during testing. Changing the transition point to 275°C delayed the calibration switch into a more mild portion of the second hill, causing a reduction in the NO_X disparity.

In both cases, the controller contributed to a reduction in fuel economy, expected due to the efficiency impacts of retarding combustion phasing. It was able to improve catalyst light-off times (defined as time to 200°C) by 12-14 seconds for the DOC, and by 1-5 seconds for the SCR. Although SCR light-off time improvement was minimal, note that the estimated average SCR conversion efficiency (η_{SCR}) improvement ranged from 3-5%,

critical as NO_X standards continue to get stricter.

Table 4.2: Bag 1 FTP-75 controller impacts. Positive numbers indicate an improvement due to the controller.

Calibration Transition Temperature (DOC Out) [°C]	250	275
% Reduction HC [%]	12.3	10.3
% Reduction NO_X [Feedgas] [%]	-35.2	-22.2
% Reduction Fuel Economy [%]	-20.3	-18.4
Δ DOC light-off [s]	14	12
Δ SCR light-off [s]	5	1
$\Delta\mu(\eta_{SCR})$ (Estimated) [%]	5.50	3.09

4.6 Summary and Conclusions

Control theory, statistical theory, and engine experimental observations were combined to design combustion variance controllers, and illustrate feedback generation techniques using one or multiple cylinder pressure transducers. Initial steady-state dyno experiments verified controller functionality and investigated the impact of window size, controller gains, and variance setpoints. Transient dyno experiments were then conducted over the FTP-75 cycle and demonstrated a DOC light-off time (outlet temperature $> 200°C$) reduction of 14 seconds and a SCR light-off time reduction of 74 seconds when not considering emission impacts.

The controller was adapted from the ideal conditions of the laboratory environment and implemented in a test vehicle using production-intent sensors. The test vehicle was used to evaluate the controller against real-world driving conditions and noise factors including fuel properties, where it continued to function properly and showed the ability to adapt to different fuel properties. The controller was implemented in another test vehicle, where it was evaluated against the FTP-75 emissions cycle. The desirable aftertreatment heating benefit was observed, although the fuel economy, HC, and NO_X trade-off data was unavailable due to testing issues.

CHAPTER 5

Preview-Based Model Predictive Engine Control for Emissions Compliance

5.1 Introduction

Modern diesel engines are equipped with aftertreatment systems, a sequence of catalysts and supporting hardware, which are effective at reducing tailpipe hydrocarbon and oxides of nitrogen (NO_X) emissions when the catalysts are warmed-up to temperatures near 200° Celsius [2, 3]. Aftertreatment systems have been critical for enabling efficient diesel engine combustion while averting the harmful pollution impacts. This is because diesel combustion profiles that achieve high fuel efficiency, or low Brake Specific Fuel Consumption (BSFC), do so at the expense of increased feedgas, or pre-aftertreatment system, NO_X emissions due to the in-cylinder dynamics that govern NO_X formation [7]. When the NO_X catalysts are up to temperature, conversion efficiencies exceeding 90% can be achieved, resulting in both low CO_2 and NO_X emissions at the tailpipe. While oxidation catalysts for oxidizing hydrocarbons tend to have monotonic conversion efficiency curves with respect to temperature, some NO_X catalysts (ex: Cu/Zeolite) suffer from reduced conversion efficiency at both lower and higher temperatures, indicating that proper thermal conditioning of the catalyst is critical for minimizing CO_2 and NO_X tailpipe emissions [48, 50].

5.1.1 Emission Regulations

In the US, medium- and heavy-duty diesel engines are typically certified on engine dynamometers against the HDFTP (heavy duty federal test procedure) cycle, supplemental emission test (SET, steady-state 13-mode test), and In-Unit Testing (IUT), also known as Not-to-Exceed (NTE) testing [18]. NTE testing requirements were added to identify defeat-devices after the 1998 United States Environmental Protection Agency (USEPA) and diesel

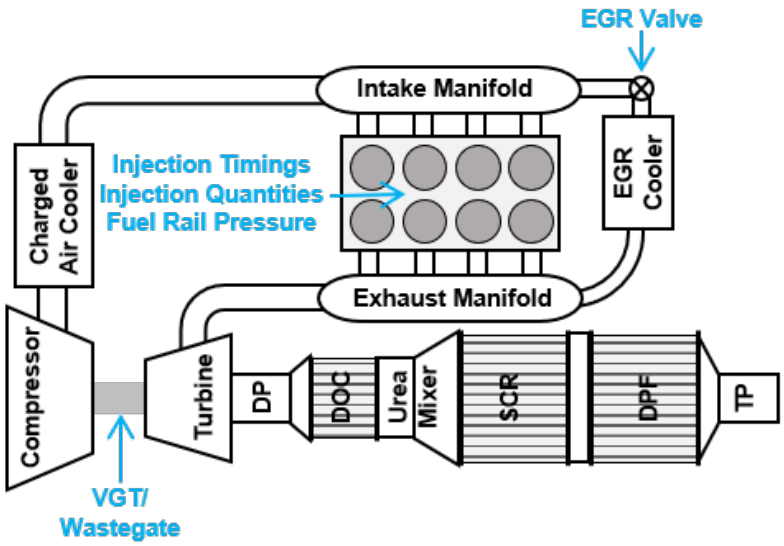


Figure 5.1: Diagram of medium-duty diesel engine with exhaust aftertreatment system. Blue text highlights potential actuators that would be affected by a calibration change.

manufacturers consent decree [51]. The testing involves using portable emission analyzer equipment to measure emissions during real-world driving scenarios. Emission data is only considered when the truck/engine operates in the "NTE Zone" (portion of the speed/load map used for normal vehicle operation) for a minimum of 30 seconds, and the valid data is compared to the emission standard adjusted by a conformity factor. The 2021 NTE zone (as of 2021) excludes the portion of the speed-load map where the engine power is less than 30% of the rated power of the engine. Diesel engines are intended to be run lean; at idle and low-load (power) the abundance of excess air can cause exhaust temperatures to dip below 200° Celsius, meaning that without external intervention the aftertreatment system can be cooled below its ideal operating temperature, resulting in poor conversion efficiencies and therefore higher observed emissions [21]. Due to the 30% minimum power exclusion for the current NTE zone, insufficient exhaust temperature for aftertreatment thermal conditioning is less of a concern.

As environmental concerns continue to drive more stringent emission standards (CO_2, HC, NO_X), there is greater emphasis on emissions generated by real-world driving behaviors [17]. In Europe, the EUROVI In-Service Conformity (ISC) testing (Europe "equivalent" of USEPA NTE) leverages a Moving Average Window (MAW) technique [20]. The length of the moving window is a function of either the reference work or CO_2 from a transient certification cycle (WHTC). For EUROVI steps A-C, windows with an average power less than 20% of the rated power or coolant temperature less than

70° Celsius were excluded. The current EUROVI-D standard reduces the minimum power to 10%, and the minimum coolant temperature to 30° Celsius, dramatically expanding the operating range at which vehicles need to be in-service compliant. When comparing in-use emission data from Europe and the US, it was found that US heavy-duty engines emit 3 times more NO_X than their European counterparts at speeds less than 50 mph - likely due to the fact that with the 30% minimum power exclusion and the minimum 30 second zone residence requirement, US heavy-duty engines undergoing NTE testing have only been subject to the regulation 17% of the time [19, 52].

Increased pressure to reduce real-world emissions have led the USEPA and California Air Resources Board (CARB) to develop a new IUT method based on the EUROVI MAW technique. The proposed regulations for MY 2024-2026 include a window size based on the work or CO_2 equivalent on a FTP cycle and a window minimum power threshold of 10%. Unlike the EUROVI-D requirements, there will still be a cold-start engine coolant temperature exclusion [19]. The increased operating range (both temporally and in the engine map) at which engines must comply with regulations will require improved management of both the engine and the aftertreatment system. Based on historical regulation trends, minimum power and coolant temperature exclusions could disappear entirely in the next decade.

5.1.2 Advanced Control Opportunities

Improvements in computational power, theoretical controls understanding, and vehicle connectivity has created opportunities for many advanced engine and vehicle control opportunities. The concept of using traffic and other forms of preview for powertrain and vehicle control has been approached in a variety of ways. With respect to emissions, researchers have investigated using dynamic programming and model predictive control techniques to achieve better fuel/energy economy while reducing emissions by modifying vehicle velocity profiles, altering engine actuator or electric motor setpoints, and in the case of hybrid electric vehicles (HEVs), optimizing the interaction between the internal combustion engine and electric motor / battery system [22, 23]. Some of the approaches leverage full engine airpath and aftertreatment models with nonlinear dynamics and seek to use the predicted emissions and fuel consumption to achieve improved compliance [24–26]. In the case of modifying vehicle velocity profiles, these optimization routines can lead to undesirable driving behaviors (from an occupant perspective) unless constraints and cost functions are properly structured. Approaches leveraging full airpath models tend to be computationally expensive to execute, and if trying to leverage multiple actuators (like

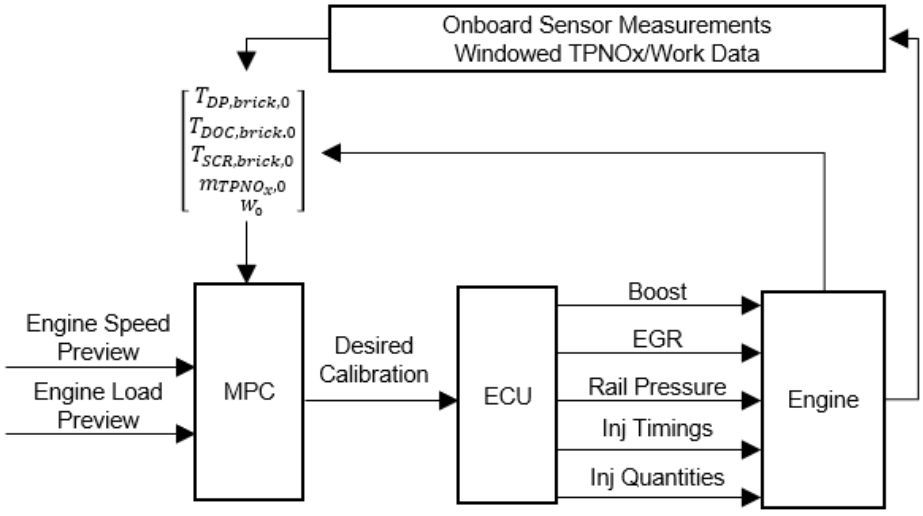


Figure 5.2: Proposed feedback-based model predictive control architecture. The goal of using the historical NO_X and work data in the controller is to avert errors caused by model and preview uncertainty.

injection timing, boost setpoints, exhaust gas recirculation) the optimization problem grows quickly with control steps and preview duration. Relying on modeled values for constraints requires formulations need to be inherently conservative to avoid noise factor issues.

To address the challenges posed by stricter transient testing, this work presents a novel MPC framework that uses engine speed and load preview along with onboard NO_X measurements to control the engine for best fuel economy subject to emission constraints. To reduce simulation computational complexity, the controller output is a decision variable selecting between two engine calibrations (which set injection, EGR, boost, etc. setpoints), one which prioritizes reduced fuel consumption at the expense of increased NO_X emissions, and one which prioritizes reduced NO_X emissions and increased exhaust heating at the expense of increased fuel consumption. Steady-state mapping data from each calibration is used to build empirical models to simulate the aftertreatment system to estimate NO_X and fuel consumption. The reduced model complexity enables long preview horizons, which is helpful when trying to optimize aftertreatment thermal dynamics that have long time constants. The emission constraints are enforced using a time-based moving window, and the use of historical NO_X sensor feedback and engine power values to set optimization problem states averts issues caused by modeling error and unmodeled phenomena. The proposed structure is shown in Figure 5.2.

5.2 Model Structure

A model for a MY2017 Ford F-250 Super Duty truck with a 6.7L V8 PowerStroke Diesel engine was developed that uses the engine speed and load to estimate fuel consumption and tailpipe NO_X emissions over driving maneuvers of interest. The model consists of steady-state engine data maps for the two calibrations, an SCR NO_X conversion efficiency model, and an aftertreatment thermal model. In this work, it is assumed that GPS route preview data (from an onboard GPS unit, or fleet logistics software) for a vehicle is used to estimate a future speed and road grade trajectory. That trajectory would be passed to a vehicle longitudinal model to estimate the required engine speed and load to follow the estimated trajectory. Examples of vehicle longitudinal modeling can be found in [53, 54]. For this work, it is assumed that the speed and load preview trajectory is directly available.

5.2.1 Regression-based Models

Fuel mass flowrate, $\dot{m}_{fuel}$, exhaust mass flowrate, $\dot{m}_{exh}$, turbocharger outlet temperature, T_4, and feedgas NO_X mass flowrate, $\dot{m}_{NO_x}$ were calculated using look-up tables as a function of engine speed, N, engine load, τ, and the desired calibration, u. The look-up tables were generated using steady-state engine mapping data. Let $u = 0$ denote when the engine is running the low BSFC (fuel efficient) but high $BSNO_X$ calibration, and $u = 1$ denote when the engine is running the high BSFC, low $BSNO_X$, increased exhaust heating calibration.

$$\left[\dot{m}_{fuel}, \dot{m}_{exh}, T_4, \dot{m}_{NO_x}\right] = f(N, \tau, u) \tag{5.1}$$

Figure 5.3 shows the steady-state benefits of running in the low $BSNO_X$ calibration for exhaust temperature, fuel consumption, and NO_X emission rate as a function of normalized engine speed and torque. Note the 20% increase in turbocharger outlet temperature and 100% reduction in NO_X emissions in the middle of the speed-load map. These calibration maps were based on existing controls strategies, and so they do not perfectly adhere to the ideal calibration difference guidelines. A production implementation would likely include investing more time designing calibrations around the proposed optimization structure.

The SCR NO_X conversion efficiency, η_{SCR}, is calculated using the SCR brick temperature, $T_{brick,SCR}$ [48]. The function is plotted as a function of normalized temperature in Figure 5.4. T_0 corresponds to the temperature at which $\eta_{T_{SCR}}$ is maximized, and is used for normalizing the other temperature traces in this paper. The conversion efficiency is used to

calculate the mass flowrate of NO_X exiting the tailpipe, $\dot{m}_{NO_x,TP}$.

$$\dot{m}_{NO_x,TP} = \left(1 - \eta_{SCR}\left(T_{brick,SCR}\right)\right) \cdot \dot{m}_{NO_x} \tag{5.2}$$

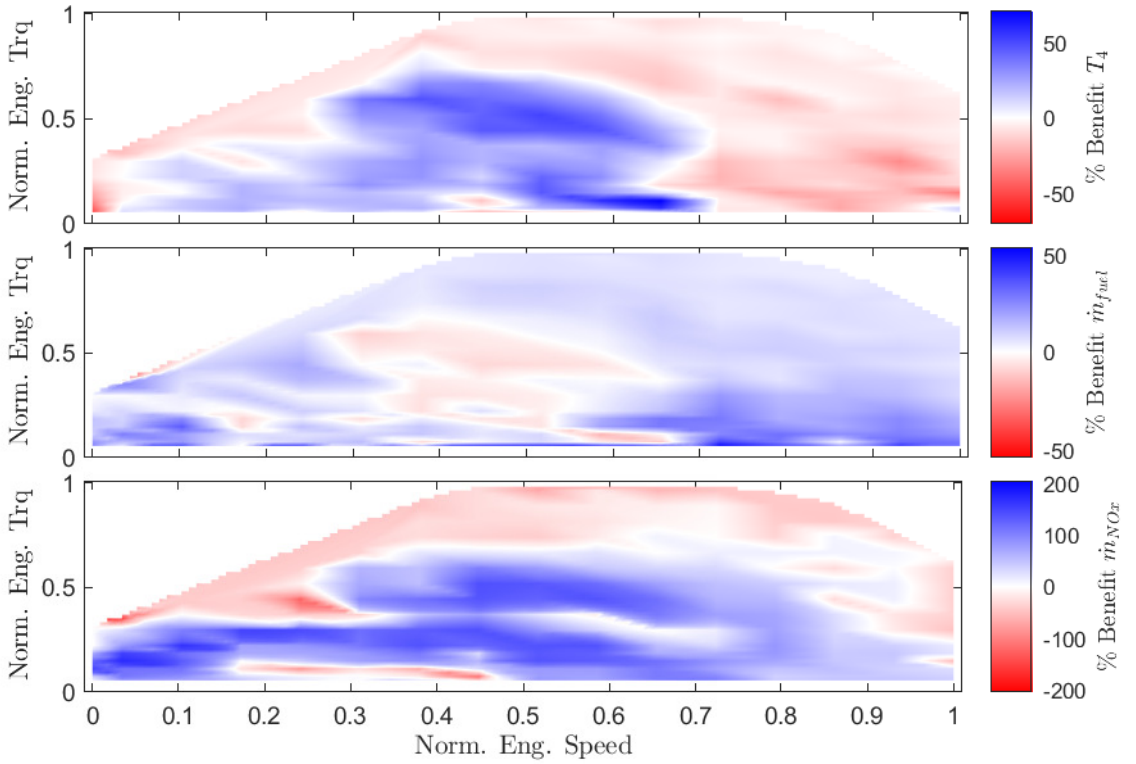


Figure 5.3: Surface plot highlighting exhaust temperature, fuel economy, and NO_X differences between calibrations (normalized, $\frac{x_{u=1}-x_{u=0}}{x_{u=0}}$). Blue indicates improvement.

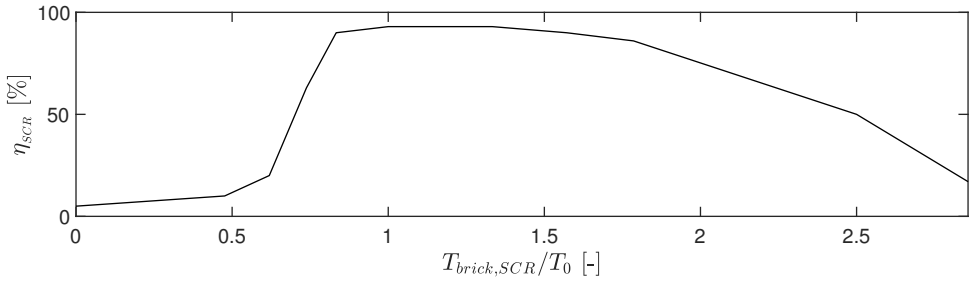


Figure 5.4: SCR NO_X conversion efficiency as a function of temperature. At sufficiently high temperatures efficiency degrades, forming an optimal operating range.

5.2.2 Aftertreatment thermal dynamics

The SCR conversion efficiency is estimated using a thermal model of the aftertreatment system and the relationship shown in Figure 5.4. The model consists of transport delay and lumped thermal mass sub-models for the exhaust downpipe (pipe from turbocharger to DOC), DOC catalyst, and SCR catalyst. The model structure assumes the aftertreatment configuration presented in Figure 5.1, commonly used on medium-duty diesel vehicles - DOC, followed by SCR, followed by the DPF. The thermal dynamics of the DPF are not modeled as they do not impact SCR NO_X conversion.

Based on literature and experimental data, temperature transport delay models were used to account for the time it can take for exhaust gas and heat to flow through the aftertreatment system components, shown in Equations (5.3) and (5.4) [55]. The subscript i represents the component of interest (downpipe, DOC, SCR), $T_{\text{gas,delayed-in},i}$ is a time-delayed version of the original temperature trace, $T_{gas,in,i}$ is the unshifted temperature trace, $\Delta\tau_{delay,i}$ is the delay time, $\dot{m}_{exh}$ is the mass flowrate of exhaust, and L_i is the component delay constant that captures the component length/size impact on thermal delays.

$$T_{\text{gas,delayed-in},i}(t) = T_{gas,in,i}\left(t - \Delta\tau_{delay,i}\right) \tag{5.3}$$

$$\int_{t-\Delta\tau_{delay,i}}^{t} \dot{m}_{exh}\, ds = L_i \tag{5.4}$$

The equations for the generic lumped thermal mass models used are shown in Equations (5.5)-(5.6). Where $T_{gas,out,i}$ is the gas outlet temperature of the component, $C_{p,g}$ is the specific heat of the exhaust gas, $(A\cdot\Delta x)_i$ represents the volume of the component (area times length), $(h_1a_1)_i$ is a lumped term consisting of the convective heat transfer coefficient and surface area for the interior of the component (heat transfer between brick and exhaust gas), $(h_2a_2)_i$ is a lumped term consisting of the convective heat transfer coefficient and component surface area for the exterior of the component (heat transfer between brick and ambient air), $T_{brick,i}$ is the temperature of the brick, T_{amb} is the ambient temperature, and $\left(m_{brick}\cdot c_{p,brick}\right)_i$ is the thermal capacity of the brick [26, 55, 56].

$$T_{gas,out,i} = \frac{\frac{\dot{m}_{exh}\cdot C_{p,g}}{(A\cdot\Delta x)_i}\cdot T_{in,delayed-in,i} + (h_1a_1)_i \cdot T_{brick,i}}{(h_1a_1)_i + \frac{\dot{m}_{exh}\cdot C_{p,g}}{(A\cdot\Delta x)_i}} \tag{5.5}$$

$$\frac{\mathrm{d}T_{brick,i}}{\mathrm{d}t} = \frac{(h_1a_1)_i\left(T_{gas,out,i} - T_{brick,i}\right) - (h_2a_2)_i\left(T_{brick,i} - T_{amb}\right)}{\left(m_{brick}\cdot c_{p,brick}\right)_i} \tag{5.6}$$

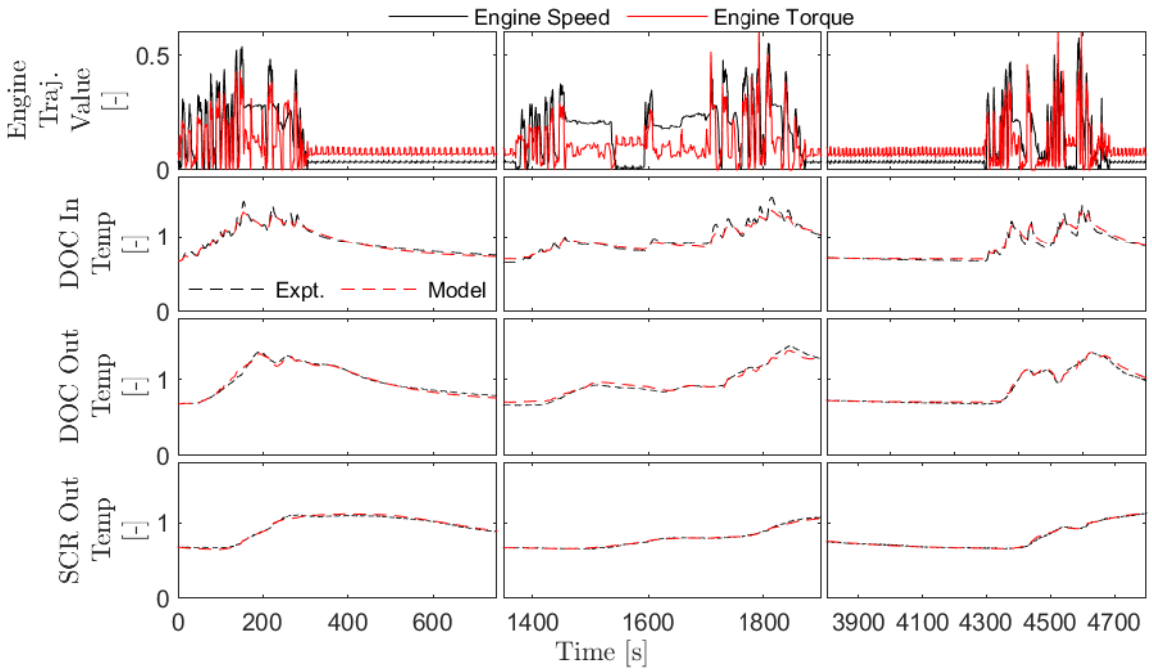


Figure 5.5: Thermal model validation using experimental low-speed, low-load vehicle data. The thermal models were tuned using this experimental data shown in black. Plotted values are normalized (speed / load by max values; temperatures by T_0).

The model parameters were identified using a combination of component measurements, literature values, and least-squares tuning using experimental data. The experimental tuning data was collected using a warmed-up test vehicle conducting low-load driving maneuvers with long idles that are interesting from a thermal management and emissions standpoint. Figure 5.5 shows the thermal model for the low load maneuvers, showing excellent agreement between the experimental results and model predictions.

To ensure model overfitting did not occur, it was also checked against experimental FTP-72 cycle data where the engine and aftertreatment system are initially at ambient temperature, shown in Figure 5.6. The model mismatch during the first 400 seconds of the test is to be expected, as there are cold-start dynamics present that are not properly addressed in the chosen modeling scheme. However, after about 400 seconds, the model agrees well for all three aftertreatment components without any adjustment of the parameters identified using the low-load vehicle data.

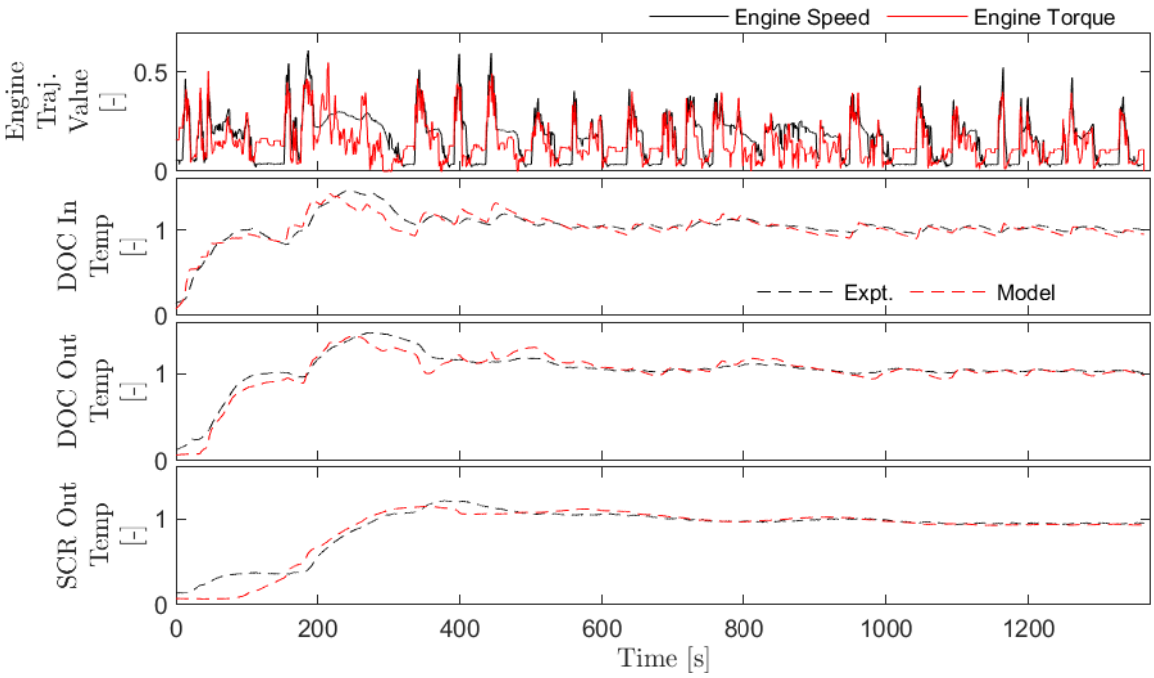


Figure 5.6: Thermal model results over Federal Test Procedure 72 drive cycle (FTP-72). The model tuned using data presented in Figure 5.5 was used directly, without retuning.

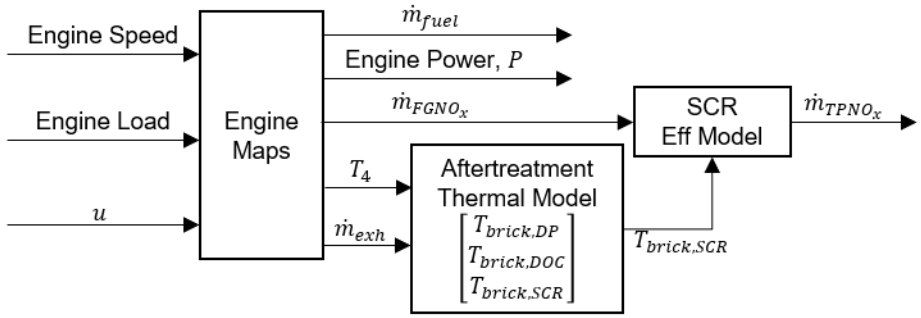


Figure 5.7: Model structure used in MPC and SIL implementations. Note that u denotes the calibration selection variable.

5.3 Optimization Structure

The goal of the work is to use receding horizon model predictive control (MPC) with historical measurement feedback to minimize fuel consumption subject to a tailpipe emissions constraint by selecting one of two engine calibrations for the given control interval. This requires forming an optimization problem, ideally one that can be solved in real-time. Figure 5.2 shows the structure of the proposed MPC architecture, leveraging the previously discussed regression and dynamic models. Table 5.1 contains details on both

Table 5.1: MPC and SIL configuration summary.

MPC Model Integration Method	RK4
MPC Model Time-step	1.0 s
Control Steps in Preview Horizon (N_c)	20
Preview Horizon Duration	600 s
$BSNO_X$ Window Length (N_{BSNO_X})	10 Steps / 300 s
$BSNO_X$ History Window ($N_{BSNO_X,H}$)	4 Steps / 120 s
SIL Model Integration Method	RKDP (ODE45)
SIL Model Time-step	0.1 s

the MPC and Software-in-the-Loop (SIL) settings. One modification to the models used in the optimization function is that the thermal transport delay effects are ignored; instead, the thermal models are re-tuned without the delay to reduce the complexity in the optimization. Note that when performing the simulation experiments, the full model with delays is used in the SIL simulations.

The states for the optimization model include the thermal states of the aftertreatment system ($T_{brick,DP}, T_{brick,DOC}, T_{brick,SCR}$), the fuel consumed ($m_{fuel}$), the amount of tailpipe NO_X emitted (m_{TPNO_x}), and the work produced by the engine (W).

$$x = \left[T_{brick,DP}, T_{brick,DOC}, T_{brick,SCR}, m_{fuel}, m_{TPNO_x}, W\right]^T \tag{5.7}$$

The cost function is the total fuel consumed (fuel consumed at end of horizon, $m_{fuel}(N)$), solved by finding the optimal calibration sequence, $u(k)$:

$$J = \underset{\{u(k)\in\{0,1\}\}_{k=0}^{N}}{\text{Minimize}} \quad m_{fuel}(N) \tag{5.8}$$

Table 5.2: Example of windowed constraint structure using parameters listed in Table 5.1.

Enforcement Range	z	a	b
-120 seconds $\rightarrow$ 180 seconds	1	-4	6
-90 seconds $\rightarrow$ 210 seconds	2	-3	7
...	...	...	...
270 seconds $\rightarrow$ 570 seconds	14	9	19
300 seconds $\rightarrow$ 600 seconds	15	10	20

subject to a sliding time-windowed emissions constraint with window length $N_{\mathrm{BSNO_X}}$, sliding from the start of the history length, $N_{\mathrm{BSNO_{X,hist}}}$, over the entire preview horizon, N_c. Let $m_{TPNO_x,wind}$ and W_{wind} denote the $\mathrm{NO_X}$ and work over a constraint window, which can be calculated by subtracting state values at discrete points in the horizon, as shown in Equations (5.9)-(5.10).

$$m_{TPNO_x,wind}[z] = m_{TPNO_x}[a] - m_{TPNO_x}[b] \tag{5.9}$$

$$W_{wind}[z] = W[a] - W[b] \tag{5.10}$$

$$\tag{5.11}$$

Note that z is an index for referring to individual windows, and a and b represent indices for time in both the historical window and preview horizon; $a,b = 0$ represents the current state, whereas $a,b < 0$ reference points in the $\mathrm{NO_X}$/work history window and $a,b > 0$ reference points in the preview horizon. If only reference points in the preview horizon are used, model or preview uncertainties could result in excess emissions and non-compliance. Including historical data allows the controller to adjust future actions due to historical behavior. Example values for z, a, and b based on the MPC configuration presented in Table 5.1 are shown in Table 5.2. With the windowed $\mathrm{NO_X}$ and work data available, it is possible to calculate the windowed $\mathrm{BSNO_X}$ values for each window z:

$$BSNO_{x,wind}[z] = \frac{m_{TPNO_x,wind}[z]}{W_{wind}[z]} \tag{5.12}$$

and constrain them to be less than some target value times a safety factor (< 1) to provide further insurance against issues poised by noise factors, model errors, or preview errors:

$$\begin{aligned} BSNO_{x,wind}[z] &\leq BSNO_{x,limit} \cdot SF_{BSNO_x} \\ \forall z &\in \left[1, N_{\mathrm{BSNO_X}} + N_{\mathrm{BSNO_X},H}\right] \cap \mathbb{Z}. \end{aligned} \tag{5.13}$$

In practice, the safety factor used could be a function of the legislative conformity factor applied to in-unit testing. Note that $BSNO_{x,wind}[z]$ is a nonlinear function of tailpipe $\mathrm{NO_X}$ and engine work due to the division operation.

The decision variable output and nonlinear constraint means the optimization problem as presented is a Mixed-Integer Nonlinear Programming (MINLP) problem. The problem was coded using CasADi, a modular optimization framework.

5.3.1 Simplification of mixed-integer problem

Mixed-integer optimizations using branch-and-bound techniques are computationally expensive as they involve repeatedly solving relaxations of the original optimization problem while trying to force variables declared discrete to integer values. Attempts to solve the originally formed mixed-integer nonlinear programming problem for long preview horizons using BONMIN, an open-source MINLP solver, yielded solution times in some cases exceeding an hour for a single control step (30 seconds) [57].

To avoid long solution times that would render real-time control infeasible, the problem was instead solved as a constrained continuous optimization problem using IPOPT, where the input interpolates between the results of the two proposed calibrations. The reduction in complexity resulted in solution times that could enable real-time control ($<$ 5 seconds for 30 second control interval). The nonlinear and non-monotonic response behavior of diesel engines to control inputs means that commanding interpolated actuator values between the two calibrations could result in undesirable behavior. To avoid this, the continuous solution (u_{cont}) is converted back to a discrete decision variable (calibration choice) by either rounding the value (to 0 when $u_{cont} < 0.5$ or 1 when $u_{cont} \geq 0.5$) or converting it to a PWM signal (with frequency proportional to control interval time, and duty cycle

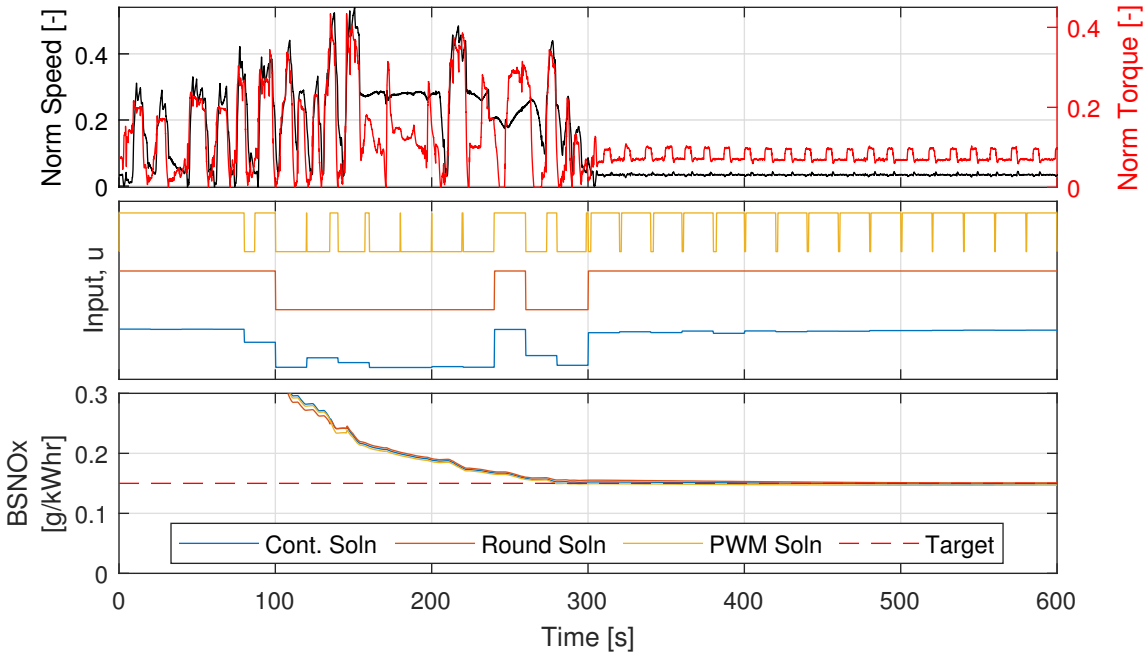


Figure 5.8: Simulation experiment exploring simplification techniques of MINLP. In this figure, the $BSNO_X$ target was 0.150 g/kWh. Note the 2 presented simplification techniques (rounding, PWM) create comparable $BSNO_X$ profiles.

proportional to the continuous fractional command value).

These techniques were evaluated in simulation by performing a terminal $BSNO_X$ trajectory optimization over a 600 second preview trajectory, and then applying the resulting control sequence to the SIL model. In this instance, the optimization was only performed once; in practice, the optimization would be conducted every 30 seconds in a receding horizon fashion. Both techniques recovered similar results compared to the continuous problem in simulation, as shown in Figure 5.8 and Table 5.3. Note that in some instances the PWM and rounding solutions exceeded the $BSNO_X$ limit; when properly integrated in the receding horizon feedback structure, the iterative updates with NO_X/work historical feedback should correct errors induced by the actuator conversion scheme and model uncertainty. Although the PWM scheme results in emission values that are closer to the continuous solution, from a pragmatic perspective it is less desirable because it can command rapid changes in actuator setpoints that could cause other undesirable behaviors; adjustments to the MPC problem formulation would be needed to make the PWM option more attractive. The results in this paper were generated using the rounding scheme, where $u_{cont} \geq 0.5 \rightarrow u = 1$ and $u_{cont} < 0.5 \rightarrow u = 0$.

Table 5.3: MIP Conversion Results

Control Type	$BSNO_X$ Limit [g/kWh]		
	0.150	0.175	0.200
Continuous	0.149	0.174	0.199
Rounding	0.151	0.181	0.202
PWM	0.147	0.173	0.200

5.4 Optimization Results

Experimental vehicle driving data was used to synthesize a drive-cycle for controller evaluation. It consists of a repeated 300 second low-speed, low-load driving maneuver interspersed with idle sequences of increasing length. Nominal simulations of the trajectory, without control, were run and a $BSNO_X$ setpoint for the closed-loop controller was selected based on an achievable target. Closed-loop SIL control simulations were performed while adding noise factors to evaluate robustness. Table 5.4 summarizes the simulation tests conducted.

Figure 5.9 presents the closed-loop SIL data for the described simulation experiments. For the closed-loop control cases, the $BSNO_{x,limit}$ was set to 0.10 g/kWh and the safety

Table 5.4: Numerical experiment description

Legend Name	Description
Nominal	No control
CL-Nominal	Closed-loop control simulation; no noise factors
CL-FGNOx +20%	Closed-loop control simulation; FGNOx increased 20%; model unaware of increase
CL-TempOfset	Closed-loop control simulation; Temperature states offset +50° Celsius

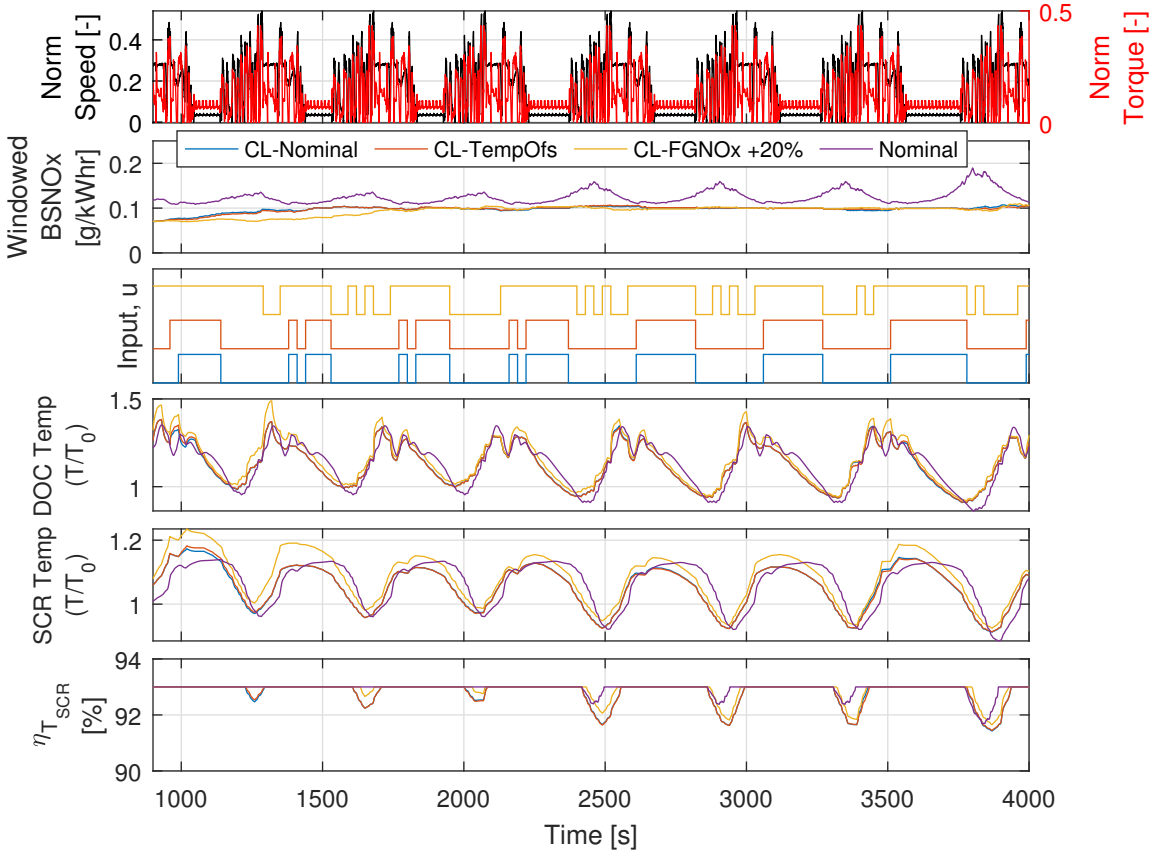


Figure 5.9: Software-in-the-loop MPC controller results for three test cases, which are described in Table 5.4.

factor, SF_{BSNO_x}, was set to 1 to evaluate the tracking performance of the controller

without the safety factor. Note that initially, the nominal trajectory (purple) is very close to the prescribed limit, but grows in magnitude as the idle periods grow larger. This can be observed in the closed-loop trial actuator profiles; the amount of time spent in calibration 1 increases further along in the cycle. The yellow line, trial CL-FGNOx +20%, demonstrates the benefits of the feedback architecture; even with a unmodeled 20% increase in $FGNO_X$ emissions, the historical NO_X data informs the controller of increased NO_X emissions and it responds accordingly. Even with the safety factor set to one, the controller does an excellent job regulating the windowed $BSNO_X$ to the target value in the presence of temperature and emission noise factors. Note that in the presented drive-cycle, aftertreatment temperatures are less of a concern; in fact, exhaust mass flow differences between the calibrations during the idle phase cause the exhaust temperatures for the nominal trajectory to be higher until the engine operates in the portion of the map with the T_4 benefit. Instead, the optimization exploits the dramatic reduction in $FGNO_X$ by running in calibration 1.

5.5 Summary and Conclusions

A novel feedback-oriented model predictive control architecture for diesel engines is presented to address compliance and fuel economy challenges poised by $BSNO_X$ emission regulation changes. In contrast to existing solutions in literature, the MPC model is computationally simpler by using regression maps, first order thermal models, and a single discrete control input corresponding to the desired engine calibration, one which has better fuel economy but increased NO_X emissions, and another which has decreased NO_X emissions, increased fuel consumption, and increased exhaust temperatures.

To avoid a computationally expensive MINLP problem, techniques for converting output from a simpler continuous NLP back to a discrete ("mixed-integer") solution are presented. The simplification adds more error to the system, however issues posed by model error, preview uncertainty, and optimization problem conversion are addressed by using historical $BSNO_X$ feedback as measured by the engine. Software-in-the-Loop (SIL) results show that the controller has the ability to track a windowed emissions target even when using the adapted actuator signal, and appropriately responds to noise factors such as aftertreatment temperatures and emission rates.

CHAPTER 6

Conclusions and Future Work

The work presented in this dissertation focused on using advanced control and feedback strategies to address two challenges facing diesel engine emissions compliance: cold-start emissions and real-world driving emissions.

6.1 Challenge 1: Cold-Start Diesel Emissions

The first challenge was cold-start diesel engine emissions and was approached from the perspective of trying to improve existing engine-based exhaust heating methods that are used to light-off aftertreatment systems. In Chapter 2, literature and experiments highlighted that the magnitude at which these methods can be used are restricted by their potential impacts on combustion variability (NVH consequences). To explore this restriction, an actuator that can be used to increase engine-based exhaust heating, injection timing, was swept at a variety of speed and load conditions to understand the impacts on engine behaviors, including combustion variability, phasing, exhaust temperatures, and emissions. Based on the experimental observations, a novel combustion variance controller using feedback from an in-cylinder pressure sensor was proposed. By controlling combustion variance to a target setpoint, the maximum amount of phasing retard based exhaust heating can be generated for a given acceptable variability level. A portion of the experimental data was used to build a stochastic control-oriented engine model to enable controller design and offline simulation.

Literature was surveyed to identify previous applications of closed-loop variance control to aid in controller design. No such literature was identified, and so detailed controls and statistical analysis were used to develop a theoretical understanding presented in Chapter 3. Windowed estimator analysis was performed under varying assumptions observed during engine operation, including zero-mean processes (steady-state operation) and varying-mean sequence processes (transient operation). Detrending was identified as a

technique which could enable combustion variance estimation, even during intense engine transients. From a controls perspective, the analysis revealed that linear controller design and analysis principles could be used. Numerical experiments were used to verify the theoretical observations and to also understand the implications of closed-loop variance control.

The results from Chapter 2 and Chapter 3 were combined to design feedback combustion variance controllers presented in Chapter 4. Techniques for generating online variance estimates were shown for both a fully-indicated engine, as well as an engine limited to a single cylinder pressure transducer, even during transient operation, critical for real-world applicability. Initial testing focused on steady-state conditions, exploring the impact of feedback window size and controller gain on closed-loop performance. Long-term steady-state testing was also conducted, which demonstrated the controller's ability to adapt to long time-scale combustion variability dynamics. The controller was then tested at transient conditions, where it continued to function and generated the desired exhaust temperature benefit while limiting the variability around the target setpoint.

To ensure real-world functionality, the control setup was adapted to a production test vehicle using production-intent in-cylinder pressure sensors. Vehicle testing demonstrated the controller's general ability to function during real-world driving conditions, as well as the specific ability to adapt to noise factors like fuel type. Test equipment issues limited the ability to get a comprehensive picture of the controller impacts on a FTP-75 emission cycle, but post-processing of the available data showed that the controller could actually reduce hydrocarbon emissions, decrease light-off times, and improve estimated average SCR efficiency. An expected fuel economy penalty was observed, as well as an unexpected feedgas NO_X penalty that could be explained by the testing procedure.

6.1.1 Future Work

The presented work focused on a control architecture only using one actuator for variability control, injection timing retard, where the injections that were adjusted were done so uniformly. Referring back to Figure 2.2, there are many other actuators that could be integrated into the control structure. EGR could be used to enable further feedgas NO_X reductions and potentially reduce pumping work to reduce the fuel economy impact at the cost of a combustion variability increase. Running the controller alongside a cylinder deactivation or early exhaust valve opening strategy could enable taking advantage of the improved exhaust temperature increase opportunities when at higher cylinder loads shown in Figure 2.13. Finally, independent injection control could lead to heat-release shaping,

where the fuel combustion profile could be controlled to maximize efficient exhaust heating while targeting an acceptable variability level.

There are also additional opportunities from a controller tuning/management perspective. A more aggressive controller could drive exhaust temperatures higher and enable faster aftertreatment heating, but this would necessarily drive up heat transfer losses from exhaust components to ambient, harmful from a fuel consumption perspective. A solution may involve using an exhaust plumbing thermal observer, which adjusts the controller usage to constrain exhaust plumbing surface temps to avoid excessive heat losses to ambient air. The observed cycle performance could likely be further improved by properly integrating the controller into the rest of the engine control structure and combining other control techniques. For example, upon confirmation of DOC light-off, post-injections timed to generate exhaust heat could instead be timed to generate feedgas hydrocarbons, which could be oxidized across the DOC for heating of the SCR. By generating the heat at the DOC instead of the engine, un-needed heating of the exhaust plumbing between the turbo and aftertreatment system is avoided.

6.2 Challenge 2: Real-world Driving Diesel Emissions

The second challenge, real-world driving emissions, was approached from a vehicle connectivity and optimization perspective. A review of current state-of-the-art diesel emission MPC approaches revealed that they tend to leverage high-fidelity engine airpath and emission formation models to properly enforce emission constraints. A disadvantage of this approach is the computational complexity, which can limit the maximum preview horizon, as well as relying on output from emission models that could be sensitive to a variety of noise factors. These formulations also tend to focus on instantaneous NO_X emissions rates, rather than the accumulated amounts which are subject to emission standards in the US (NTE) and EU (MAW).

Chapter 5 presented a novel feedback-oriented model predictive control architecture for diesel engines to address compliance and fuel economy challenges poised by the evolving $BSNO_X$ emission regulation changes involving moving emission windows. The controller combines route preview information, which could be sourced from a vehicle's navigation system, with feedback from an onboard NO_X sensor to optimize fuel economy subject to windowed emission constraints. To enable longer preview times, critical when managing aftertreatment thermal dynamics, the MPC model is computationally simpler by using regression maps, first order thermal models, and a single discrete control input corresponding to the desired engine calibration, one which has better fuel economy but

increased NO_X emissions, and another which has decreased NO_X emissions, increased fuel consumption, and increased exhaust temperatures. The use of feedback was intended to correct errors created by the modeling simplifications. Software-in-the-Loop (SIL) experimental results show that the controller has the ability to track a windowed emissions target, and appropriately respond to noise factors such as aftertreatment temperatures and emission rate errors.

6.2.1 Future Work

The presented work was limited to SIL testing - immediate future work should include experimental testing, either using a hardware-in-the-loop setup or vehicle implementation. Additional work using the current framework could include revisiting engine calibration design guidelines to take best advantage of the optimization structure.

Changes to the optimization structure could include adding additional calibration options; for example, a calibration that has a feedgas NO_X reduction without an emphasis on additional exhaust heat. Such a calibration would be useful when the aftertreatment system is below ideal temperatures, but the preview trajectory includes higher speeds and loads where exhaust temperature increases are easier to generate.

The simplicity of the presented MPC architecture lends itself to a cloud-based implementation to avoid increases in onboard vehicle processing power that limit the feasibility of other techniques. Figure 6.1 shows a proposed cloud-based implementation; the vehicle would only need to send the vehicle states and destination to the cloud, and in return receive the simple decision variable at the desired control interval times. Avoiding the use of time-resolved actuator profiles or calibration setpoints reduces the bandwidth needed between the vehicle and cloud, making production implementation more feasible so preview and other forms of data can be used to further drive real-world emission

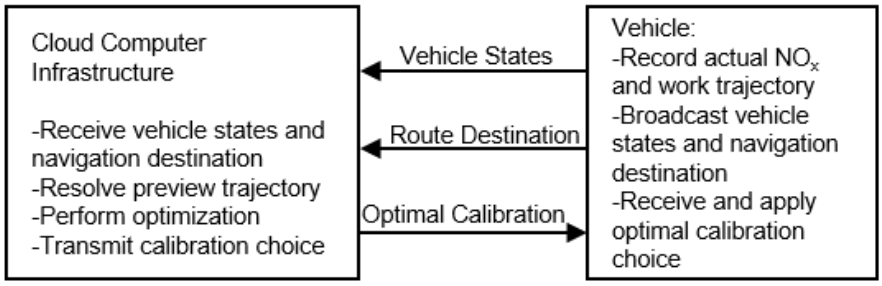


Figure 6.1: Potential real-world implementation to leverage cloud resources and minimize onboard vehicle processing power.

improvements. Such an implementation could also be used to enforce geo-specific $BSNO_X$ restrictions ("Green Zones"), where the controller $BSNO_X$ limit is adjusted as the vehicle travels between urban areas (where NO_X reductions are critical due to high concentration of vehicles) and more rural areas where it will diffuse into the environment with reduced risk of smog formation.

BIBLIOGRAPHY

[1] Mitchell Bieniek, Anna Stefanopoulou, John Hoard, Brien Fulton, and Michiel Van Nieuwstadt. Combustion Variability Model for Control of Injection Timing for Diesel Exhaust Heating, nov 2018.

[2] Eric Kurtz and Christopher J Polonowski. The Influence of Fuel Cetane Number on Catalyst Light-Off Operation in a Modern Diesel Engine. *SAE Int. J. Fuels Lubr.*, 10, 2017.

[3] Jason Miwa, Darius Mehta, and Chad Koci. Evaluation of Cold Start Technologies on a 3L Diesel Engine. In *SAE 2016 World Congress and Exhibition*. SAE International, apr 2016.

[4] Christopher A. Sharp, Cynthia C. Webb, Gary D. Neely, and Ian Smith. Evaluating Technologies and Methods to Lower Nitrogen Oxide Emissions from Heavy-Duty Vehicles, 2017.

[5] Aswin Karthik Ramesh, Dheeraj Bharadwaj Gosala, Cody Allen, Mrunal Joshi, James McCarthy, Lisa Farrell, Edward D. Koeberlein, and Gregory Shaver. Cylinder Deactivation for Increased Engine Efficiency and Aftertreatment Thermal Management in Diesel Engines. In *SAE Technical Papers*, volume 2018-April, 2018.

[6] Dheeraj B Gosala, Aswin K Ramesh, Cody M Allen, Mrunal C Joshi, Alexander H Taylor, Matthew Van Voorhis, Gregory M Shaver, Lisa Farrell, Edward Koeberlein, James McCarthy, and Dale Stretch. Diesel engine aftertreatment warm-up through early exhaust valve opening and internal exhaust gas recirculation during idle operation. *International Journal of Engine Research*, 2017.

[7] J B Heywood. *Internal Combustion Engine Fundamentals*. Automotive technology series. McGraw-Hill, 1988.

[8] Tayyar Ozel, Matthew J. Hall, and Ron Matthews. Increasing Exhaust Temperature of an Idling Light-Duty Diesel Engine through Post-Injection and Intake Throttling. In *SAE Technical Papers*, volume 2018-April, 2018.

[9] C S Draper and Y T Li. *Principles of optimalizing control systems and an application to the internal combustion engine*. American Society of Mechanical Engineers, New York, 1951.

[10] R J Hosey and J D Powell. Closed Loop, Knock Adaptive Spark Timing Control Based on Cylinder Pressure. *Journal of Dynamic Systems, Measurement, and Control*, 101(1):64–69, mar 1979.

[11] Jacob Klimstra. The Optimum Combustion Phasing Angle—A Convenient Engine Tuning Criterion, 1985.

[12] Jaesung Chung, Kyunghan Min, Seungsuk Oh, and Myoungho Sunwoo. In-cylinder pressure based real-time combustion control for reduction of combustion dispersions in light-duty diesel engines. *Applied Thermal Engineering*, 99:1183–1189, 2016.

[13] BorgWarner. Borgwarner Equips New Generation Of Diesel Engines From Volkswagen With Pressure Sensor Glow Plugs.

[14] Frank Willems. Is Cylinder Pressure-Based Control Required to Meet Future HD Legislation? *IFAC-PapersOnLine*, 51(31):111–118, 2018.

[15] Roberto Saracino, Maria Rosaria Gaballo, Soenke Mannal, Stefan Motz, Antonio Carlucci, and Marco Benegiamo. Cylinder Pressure-Based Closed Loop Combustion Control: A Valid Support to Fulfill Current and Future Requirements of Diesel Powertrain Systems, 2015.

[16] US EPA. Learn About Volkswagen Violations, 2017.

[17] Peter Mock. Real-Driving Emissions Test Procedure for Exhaust Gas Pollutant Emissions of Cars and Light Commercial Vehicles in Europe. *Icct*, (January):1–10, 2017.

[18] Code of Federal Regulations. CFR 40 Part 86 Subpart N.

[19] CARB MSCD/ECCD. Heavy-Duty Low NOx Program: Proposed Heavy-Duty In-Use Compliance Testing Public Workshop. In *CARB Public Workshop*, Diamond Bar, CA, 2019.

[20] European Union. Regulation (EU) No 582/2011, 2019.

[21] Southwest Research Institute. SwRI'S New Low-Load Cycle Targets Heavy-Duty Diesel Engine Emissions, 2020.

[22] Chunan Huang, Rasoul Salehi, Tulga Ersal, and Anna G. Stefanopoulou. An energy and emission conscious adaptive cruise controller for a connected automated diesel truck. *Vehicle System Dynamics*, 58(5), 2020.

[23] Di Chen, Youngki Kim, and Anna Stefanopoulou. Predictive Equivalent Consumption Minimization Strategy with Segmented Traffic Information. *IEEE Transactions on Vehicular Technology*, 9545(c):1–1, 2020.

[24] Esteban R. Gelso and Johan Dahl. Diesel Engine Control with Exhaust Aftertreatment Constraints. *IFAC-PapersOnLine*, 50(1), 2017.

[25] Dominic Liao-McPherson, Mike Huang, Shinhoon Kim, Masanori Shimada, Ken Butts, and Ilya Kolmanovsky. Model predictive emissions control of a diesel engine airpath: Design and experimental evaluation. *International Journal of Robust and Nonlinear Control*, 30(17), 2020.

[26] Saravanan Duraiarasan, Rasoul Salehi, Fucong Wang, Anna Stefanopoulou, Marc Allain, and Siddharth Mahesh. Real-Time Embedded Models for Simulation and Control of Clean and Fuel-Efficient Heavy-Duty Diesel Engines. In *SAE Technical Papers*, volume 2020-April, 2020.

[27] Bryan P. Maldonado, Mitchell Bieniek, John Hoard, Anna G. Stefanopoulou, Brien Fulton, and Michiel Van Nieuwstadt. Modelling and estimation of combustion variability for fast light-off of diesel aftertreatment. *International Journal of Powertrains*, 9(1-2), 2020.

[28] Mitchell Bieniek, Bryan Maldonado, Anna G. Stefanopoulou, and John Hoard. Online Control of Process Variance Using Feedback. In *Proceedings of the American Control Conference*, volume 2020-July, 2020.

[29] Mitchell Bieniek, Bryan P Maldonado, Anna G Stefanopoulou, Brien Fulton, John Hoard, and Michiel Van Nieuwstadt. Combustion Variance Controller for Fast Diesel Aftertreatment Light-Off. *Submitted to IEEE Transactions on Control Systems Technology*, 2020.

[30] Mitchell Bieniek, Anna Stefanopoulou, John Hoard, Bryan Maldonado, Brien Fulton, and Michiel Van Nieuwstadt. Retard to the Limit: Closed-Loop COVIMEP Control for Aggressive Exhaust Heating. *IFAC-PapersOnLine*, 52(5):624–629, 2019.

[31] Brien Fulton, Michiel Van Nieuwstadt, Michael Hopka, and Mitchell Bieniek. Patent Application: System and Method for Previewing Vehicle Emissions, 2020.

[32] Timothy Johnson and Ameya Joshi. Review of Vehicle Engine Efficiency and Emissions. In *SAE Technical Paper*. SAE International, 2017.

[33] John B.Heywood. *Internal Combustion Engine Fundamentals Second Edition*, volume 7. 2015.

[34] Zhiping Han, Naeim Henein, and Bogdan Nitu. Diesel Engine Cold Start Combustion Instability and Control Strategy. (724), 2001.

[35] Nicolo Cavina, Giorgio Mancini, Enrico Corti, Davide Moro, Matteo De Cesare, and Federico Stola. Thermal Management Strategies for SCR After Treatment Systems. In *SAE Technical Paper*. SAE International, 2013.

[36] Chang Hwan Kim, Michael Paratore, Eugene Gonze, Charles Solbrig, and Stuart Smith. Electrically heated catalysts for cold-start emissions in diesel aftertreatment. *SAE Technical Papers*, (x):2–11, 2012.

[37] R H Middleton and G C Goodwin. *Digital Control and Estimation: A Unified Approach*. Prentice-Hall International Series in Computer Science. Prentice Hall, 1990.

[38] Anthony Hotz and Robert E Skelton. Covariance control theory. *International Journal of Control*, 46(1):13–32, jul 1987.

[39] Michael K. Sain. *On Minimal-Variance Control of Linear Systems with Quadratic Loss*. PhD thesis, University of Illinois - Urbana, 1965.

[40] Chang H. Won, Michael K. Sain, and Stanley R. Liberty. Full-state-feedback minimal cost variance control on an infinite time horizon: The risk-sensitive approach. In *Proceedings of the IEEE Conference on Decision and Control*, volume 1, 2001.

[41] S M Ross. *A First Course in Probability*. Pearson Prentice Hall, 2010.

[42] Jordanka Angelova. On moments of sample mean and variance. *International Journal of Pure and Applied Mathematics*, 79, jan 2012.

[43] Bryan P Maldonado and Anna G Stefanopoulou. Linear Stochastic Modeling and Control of Diluted Combustion for SI Engines. *IFAC-PapersOnLine*, 51(31):99–104, 2018.

[44] Yi Cao and Diane Rossiter. An input pre-screening technique for control structure selection. *Computers and Chemical Engineering*, 21(6):563–569, 1997.

[45] Jin Young Choi, Jietae Lee, Jae Hak Jung, Moonyong Lee, and Chonghun Han. Sequential loop closing identification of multivariable process models. *Computers and Chemical Engineering*, 24(2-7):809–814, 2000.

[46] J. Wahlström and L. Eriksson. Modelling diesel engines with a variable-geometry turbocharger and exhaust gas recirculation by optimization of model parameters for capturing non-linear system dynamics. *Proceedings of the Institution of Mechanical Engineers, Part D: Journal of Automobile Engineering*, 225(7):960–986, 2011.

[47] J. P. Jensen, A. F. Kristensen, S. C. Sorenson, N. Houbak, and E. Hendricks. Mean value modeling of a small turbocharged diesel engine. *SAE Technical Papers*, 1991.

[48] Giovanni Cavataio, Hung Wen Jen, James R. Warner, James W. Girard, Jeong Y. Kim, and Christine K. Lambert. Enhanced durability of a Cu/Zeolite based SCR catalyst. *SAE International Journal of Fuels and Lubricants*, 1(1):477–487, 2009.

[49] Timothy C. Watling, Mehrdad Ahmadinejad, Monica Ţuţuianu, Åsa Johansson, and Michael A.J. Paterson. Development and Validation of a Pt-Pd Diesel Oxidation Catalyst Model. *SAE International Journal of Engines*, 5(3):1420–1442, 2012.

[50] Pingen Chen and Junmin Wang. Oxygen concentration dynamic model and observer-based estimation through a diesel engine aftertreatment system. *Journal of Dynamic Systems, Measurement and Control, Transactions of the ASME*, 134(3), 2012.

[51] Environmental Protection Agency. PR DOJ, EPA ANNOUNCE ONE BILLION DOLLAR SETTLEMENT WITH DIESEL ENGINE INDUSTRY FOR CLEAN AIR VIOLATIONS, oct 1998.

[52] Christopher A. Laroo. MODERNIZING HD IN-USE TEST PROCEDURES TO ACHIEVE REAL WORLD EMISSION REDUCTIONS. In *SAE High Efficiency IC Engine Symposium*, Detroit, 2019.

[53] Chaozhe R. He, Helmut Maurer, and Gábor Orosz. Fuel Consumption Optimization of Heavy-Duty Vehicles with Grade, Wind, and Traffic Information. *Journal of Computational and Nonlinear Dynamics*, 11(6):1–12, 2016.

[54] Chunan Huang, Rasoul Salehi, and Anna G. Stefanopoulou. Intelligent Cruise Control of Diesel Powered Vehicles Addressing the Fuel Consumption Versus Emissions Trade-off. *Proceedings of the American Control Conference*, 2018-June:840–845, 2018.

[55] Olivier Lepreux, Yann Creff, and Nicolas Petit. Motion planning for a diesel oxidation catalyst outlet temperature. *Proceedings of the American Control Conference*, (1):2092–2098, 2008.

[56] Varun Pandey, Bruno Jeanneret, Sylvain Gillet, Alan Keromnes, and Luis Le Moyne. A simplified thermal model for the three way catalytic converter. *21st International Transport and Air Pollution Conference*, 2016.

[57] Pierre Bonami and Jon Lee. Bonmin users' manual. *Retrieved November*, (July):1–24, 2007.