AN ABSTRACT OF THE THESIS OF

Shane Daly for the degree of Master of Science in Mechanical Engineering presented on September 21st, 2015.

Title: Chemometrics-based Approach for Predicting Low Temperature Combustion Engine Fuel Performance

Abstract approved: _____

Christopher L. Hagen Kyle E. Niemeyer

Low temperature combustion (LTC) engines offer improved efficiency and reduced emissions over traditional gasoline spark-ignited and diesel compression-ignition engines. The ability to produce fuels optimized for LTC engines could help increase efficiency and therefore improve fuel economy, insomuch that Federal fuel economy standards could be met. Currently, a fuel performance rating that can indicate how a fuel will perform in LTC engines has not been well established; this is needed to aid in fuel design/optimization. Furthermore, traditional fuel performance metrics, e.g., Research Octane Number (RON), may not fully explain LTC performance and require expensive, time-consuming experiments. In this work, the LTC index fuel performance metric, which has been recently introduced, was chosen to represent a quantitative measure for LTC fuel performance. This fuel performance metric rates fuels by their potential fuel savings for useful real world operating conditions. In a parallel effort, statistical models based on chemometric techniques were developed to correlate liquid hydrocarbon infrared absorption spectra first to RON for validation and insight into the approach, and finally to the LTC index. Here we show that the developed chemometric models can predict the performance of the FACE gasolines (research grade gasoline-like fuels) for RON ($\pm$10 for all 22 FACE) and the LTC index ($\pm$0.4 for FACE A and C). To the author's

knowledge, these models are novel in that actual gasoline samples are not used to inform the chemometric model; instead, only neat hydrocarbon components are used (individually or up to five component mixtures). The chemometric models could aid researchers and industry alike as an easy, cost-effective, and fast tool to determine LTC performance of gasoline fuels and help in an effort to design next-generation fuels.

Chemometrics-based Approach for Predicting Low Temperature Combustion Engine Fuel Performance

by

Shane Daly

A THESIS

submitted to

Oregon State University

in partial fulfillment of
the requirements for the
degree of

Master of Science

Presented September 21, 2015

Commencement June 2016

Master of Science thesis of Shane Daly presented on September 21, 2015

APPROVED:

Major Professor, representing Mechanical Engineering

Head of the school of Mechanical, Industrial, and Manufacturing Engineering

---

Dean of the Graduate School

I understand that my thesis will become part of the permanent collection of Oregon State University libraries. My signature below authorizes release of my thesis to any reader upon request.

Shane Daly, Author

ACKNOWLEDGEMENTS

My sincere thanks goes to Dr. Chris Hagen, who gave me the opportunity to join his research team, provide fruitful work opportunities—especially this work—as well as access to the laboratory and research facilities. While completing my undergraduate studies I had decided to apply to graduate school a week before the application deadline. Within that week, I inquired about graduate studies to Dr. Hagen via email, soon after met him in person, and received full support under his advisement. I digress with this story because it constantly reminds me how lucky I am to have been provided so many career-building opportunities since advisement under Dr. Hagen, not to mention with somebody who I personally enjoy working for.

Besides my main advisor, I would like to thank my co-advisor Dr. Kyle Niemeyer: his insightful comments, encouragement, and scientific guidance for this project and my thesis writing was invaluable. I would also like to thank Dr. Jeff Gautschi, who kindly gave free-range access to his chemistry lab for use of his equipment and lab supplies. Without Dr. Jeff Gautschi's resources and aid, the experimental results for this work would be much less extensive as it is.

I gratefully thank the Chevron Energy Technology Company for funding this work under research contract number K27995. More specifically, Fuels Research manager Dr. William Cannella provided research fuels, sound project guidance and was, and continues to be, a terrific scientist to work alongside.

Finally yet importantly, I would like to thank my family and extended family, i.e., my best friends—you know who you are. You have all wholly supported my pursuit of higher education and encouraged me during times when I felt inadequate. I feel fortunate to have made it this far in my academic studies, and I want to emphasize it is largely in part from the support I have received over the years. Unfortunately, a separate body of work would be required to include the list of family and friends that have affected my life during this time. In short, thanks to all and I hope in the years to come that our relationships persist.

TABLE OF CONTENTS Page

TABLE OF CONTENTS (Continued) Page

TABLE OF CONTENTS (Continued) Page

Figure LIST OF FIGURES Page

Figure LIST OF FIGURES (Continued) Page

Figure LIST OF FIGURES (Continued) Page

Figure LIST OF FIGURES (Continued) Page

Table LIST OF TABLES Page

LIST OF ABBREVIATIONS

AKI	Anti-Knock Index
ATR	Attenuated Total Reflectance
BMEP	Brake Mean Effective Pressure
CA50	Crank Angle at which 50% of fuel has burned (mass basis)
CAD	Crank Angle Degree
CI	Compression Ignition
CRC	Coordinating Research Council
FACE	Fuels for Advanced Combustion Engines
FTIR	Fourier Transform Infrared
HCCI	Homogeneous Charge Compression Ignition
ICE	Internal combustion Engine
IMEP	Indicated Mean Effective Pressure
IMEP	Indicated Mean Effective Pressure
LTC	Low Temperature Combustion
LTC index	OSU-Chevron LTC driving cycle fuel performance metric
LTHR	Low Temperature Heat Release
MEP	Mean Effective Pressure
MON	Motored Octane Number
NHRR	Net Heat Release Rate
OI	Octane Index
OSU	Oregon State University
PCR	Principal Component Regression
RCCI	Reactivity Controlled Compression Ignition
RON	Research Octane Number
SI	Spark Ignition
SVMR	Support Vector Machine Regression

Chapter 1: Literature Review

1.1 Introduction

Combustion of hydrocarbon fuels will maintain their status as the United States's leading source of energy between now and 2040, supplying 80 percent of the nation's energy needs by 2040 (increasing dependence) [1]. Society continues to rely on internal combustion engines (ICEs) for transportation, commerce, and power generation. As a result, the demands to reduce emissions and improve efficiency in these systems are significant to lessen the growing environmental and economic impacts. Therefore, there is motivation to seek out alternative, eco-friendly fuels and to move away from traditional combustion strategies in favor of advanced, next-generation systems that are cleaner and more efficient.

Since the inception of the ICE, harmful pollutants such as particulate matter, nitrogen oxides, unburned hydrocarbons, and carbon dioxide have been expelled into the atmosphere that can cause serious adverse health effects and negatively impact the environment [2], [3]. This has been managed by exhaust after-treatment strategies and by inherently improving the efficiency of the combustion process within the engine. Fuel efficiency steadily improved over the years due to increasing compression ratios, downsizing, and turbo-charging resulting in increased thermal efficiency [4]; in addition, increased electric hybridization led to reduction in overall fuel use [5]. Recent governmental regulations including the CAFE 2025 standards mandate reaching 54.5 mpg within the next 10 years—a significant engineering effort to accomplish in such a short time frame. Solving this challenge will be aided by continued electric hybridization, but it is projected that by 2025 only 11% of total vehicles in the U.S. will be registered as a hybrid vehicle [5]. Therefore, there is motivation to further improve upon the reciprocating internal combustion engine for public transportation to meet CAFE standards.

Low temperature combustion (LTC) engines, which implement a non-traditional combustion strategy, offer higher efficiencies and lower emissions that could satisfy CAFE standards. This strategy is still of current research interest and requires further development to become a viable option. This work focuses on the fuel design aspect of LTC engine research and development, as the success of this combustion strategy is highly fuel (e.g., gasoline, diesel, ect.) sensitive. Traditional fuel performance metrics, which give insight to how a fuel will perform in ICEs, have had little success in describing how a fuel will perform in a LTC engine. Additionally,

traditional fuel performance metrics require time consuming and expensive physical experiments to determine. In order to design fuels for LTC engines, a fuel performance metric is required and a means to determine the metric rapidly would accelerate the process. This work is a parallel effort to mitigate these two issues. That is to say, this work will review newly proposed fuel metrics, and the metric that best indicates how a fuel will perform in a LTC engine will be selected. With this, a robust tool will be created which can predict this LTC fuel performance metric for fuels; as opposed to the need for expensive, time consuming experiments. The goal of this work is to use chemometrics to predict low-temperature combustion performance of real fuels to aid in the effort for LTC fuel design.

First, a brief overview of the workings of the classic reciprocating ICEs will be introduced, along with a discussion of current areas for improvement and where past and current research efforts bring us. Second, next-generation engine concepts will be introduced with a discussion on the benefits they offer over the classic ICE strategies. Third, fuel performance metrics will be discussed that give useful insight into how particular fuels will perform in their respective ICEs, and will be a primary topic of this work as it is mostly concerned with fuel design. Fourth, with any given LTC engine fuel performance metric, the methodology to statistically correlate the metric to chemical data will be discussed (i.e., chemometrics). The last topic of discussion is utilizing chemometrics to accurately predict fuel properties along with the necessary tools to do so.

Internal combustion engines have two predominant thermodynamic cycles utilizing fundamentally different fuels, the gasoline spark ignition engine (SI) and the diesel compression ignition engine (CI). These two ICE types share many of the same attributes, yet they operate in different manners from a thermodynamic, chemical kinetic, and control systems standpoint. These will now be discussed in the following two sections. Following this, LTC engines will be introduced.

1.1.1 Gasoline Spark Ignition (SI) Engine

The SI engine premixes gasoline fuel with air at stoichiometric proportions (or close to stoichiometric, depending on the platform) in an attempt to create a homogeneous charge of reactants prior to entry in the piston-cylinder chamber. A spark plug is used to ignite the reactant mixture to control favorable initiation of combustion. This cycle is most basically modelled

thermodynamically by an idealized Otto cycle, viz.: isentropic compression, constant volume heat addition, isentropic expansion, and constant volume heat rejection [6 pg. 163]. Although the Otto cycle is useful as a tool to understand the SI process on a fundamental level, the idealizations ignore physics at play that have significant impacts on the efficiency and emissions; these are presented as follows.

Here, a brief overview of the efficiency of the SI engine will be introduced. These efficiencies encompass the combustion, thermodynamic, gas exchange—e.g., fluid and gas dynamics—and mechanical losses, where the product of these losses yield the overall efficiency of the system. Combustion efficiency is the fraction of the fuel energy supplied that is released in the combustion process. Typically, combustion efficiency for SI engines is on the order of 95-98% around stoichiometric equivalence ratios—which is the ratio of the actual fuel-air mixture to the stoichiometric proportion of fuel to air—and decreases as the mixture gets richer in fuel proportions due to insufficient oxygen to complete combustion [6 pg. 82]. Thermodynamic efficiency comprises the heat lost in the cycle that is not utilized; this occurs primarily from the heat lost by the exhaust and that is transferred to the engine coolant through the cylinder walls—typically around 34-45% and 17-26% of the total losses respectively [6 pg. 674]. This efficiency increases when the geometric compression ratio and isentropic exponent (ratio of thermodynamic specific heats γ, or C_p/C_v) of the working fluid of the engine is increased, in addition to when the combustion temperatures decrease for less heat transfer losses [7, 8]. Compression ratios for gasoline-fueled SI engines typically range from 8-12 [9]. Gas exchange efficiency, usually called pumping efficiency, in SI engines is caused by inducting/expelling air/exhaust through restrictions. The intake and exhaust performance are usually defined by the volumetric efficiency, and are governed by the gas dynamics of the process. SI engines have a throttle plate that restricts airflow to manage the load of the engine, and accrues significant pumping losses at lower loads when it is mostly closed. Mechanical efficiency accounts for overcoming resistance of all the moving parts in the engine as well as engine accessories. Pumping and mechanical efficiency are typically lumped into the friction efficiency category, since the work used to overcome this resistance dissipates as heat. Friction efficiency can account for 3-10% of total losses in the system [6 pg. 674]. In total, typically around 25-28% of the total fuel energy supplied will be converted to useful shaft work at high load operation [6 pg. 674]; i.e., all the losses combined take nearly 70% of the fuel energy and reject it as waste heat.

Note that the numbers described here apply to an SI engine operating at maximum power; at lower loads, the overall efficiency is lower due to the higher pumping losses from the throttle plate, reducing it to around 20% or less.

Figure 1 [10] visually depicts the efficiency losses of an SI engine, where the size of the arrows represent the magnitude of the heat lost through inefficiencies. Rather than being represented as efficiency, mean effective pressure (MEP) is used, which is the normalized work that can be achieved (ratio of work to the displacement volume of the engine). As can be seen, the maximum work that can be done is represented initially as the Fuel MEP that comes into the

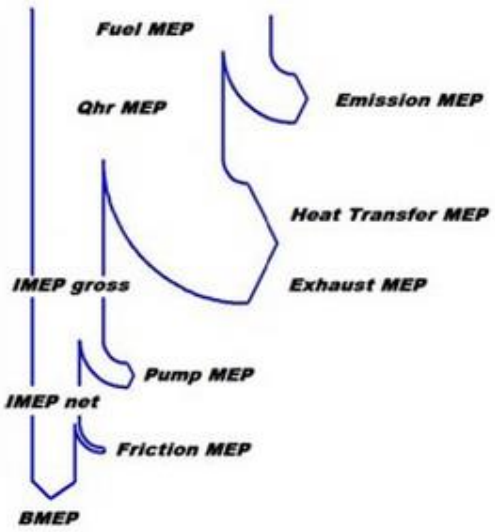

Figure 1: Qualitative fuel availability and pathways of heat losses (inefficiency). [10]

system with the fuel. After all the losses, the final amount of useful work that makes it to the shaft is represented as the brake mean effective pressure (BMEP), and is at best a third of what came originally into the system with the fuel.

Nitrogen oxides (NO_x) are formed in SI engines due to the high combustion temperatures. In addition, carbon monoxide (CO) and unburned hydrocarbons are produced due to incomplete conversion of the fuel, associated with the 95-98% combustion efficiency. This is dealt with by using a three-way catalyst to eliminate these emissions to insignificant proportions by conversion into benign species (e.g., H_2O and N_2). While effective, they are expensive due to their reliance on noble metals such as platinum, palladium, and rhodium for the catalyst.

1.1.2 Diesel Compression Ignition Engine (CI)

CI engines do not premix the fuel and air; instead, only air is drawn into the piston-cylinder assembly. Diesel fuel is then directly injected into the high-temperature, pressurized air wherein spontaneous ignition of the diesel occurs. This cycle is most basically modelled thermodynamically by an idealized Diesel cycle, viz.: isentropic compression, constant pressure heat addition, isentropic expansion, and constant pressure heat rejection [6 pg. 163]. The efficiency and emissions of the CI engine will now be discussed.

Combustion efficiencies of diesel engines are typically higher than SI, but not drastically: about 98% [6 pg. 83]. Thermodynamic efficiency is comparatively higher than the SI engine, due to the higher compression ratios that can be achieved in this cycle—typically 12-24 [9]—along with lower combustion temperatures for less heat loss. The coolant and exhaust losses are typically 16-35% and 22-35%, respectively [6 pg. 674]. Considered together, the losses are comparatively lower than for SI engines. The fluid friction losses associated with the CI engine, around 2-6% [6 pg. 674], are also lower than in SI engines. Unlike SI engines, CI engines require no intake throttling and thus achieve low fluid friction losses at low loads, where at higher loads SI and CI are nearly the same [6 pg. 713]. All these considered, the overall efficiency of the Diesel cycle is around 34-38% [6 pg. 674]. Thus, diesel seems to be the superior option over SI engines due to the higher overall efficiency. However, the diesel fuel itself has a larger molecular weight and results in a more complex combustion process from a chemical kinetics standpoint; in addition, the injection strategy results in regions of extremely rich combustion and thus formation of unburned hydrocarbons and soot that are challenging to counteract, as discussed next.

The combustion of diesel is complex and heterogeneous. When diesel is injected into the engine cylinder assembly, a distribution in air/fuel ratio develops across the space as shown in Figure 2. This distribution causes locally rich and lean mixtures. At the extreme limits, the lean regions can be lean enough to be below the lean combustion limit, which is at about $\phi \approx 0.3$. In contrast, extremely rich zones produce soot, and near-stoichiometric regions produce NO_x due to the associated high combustion temperatures.

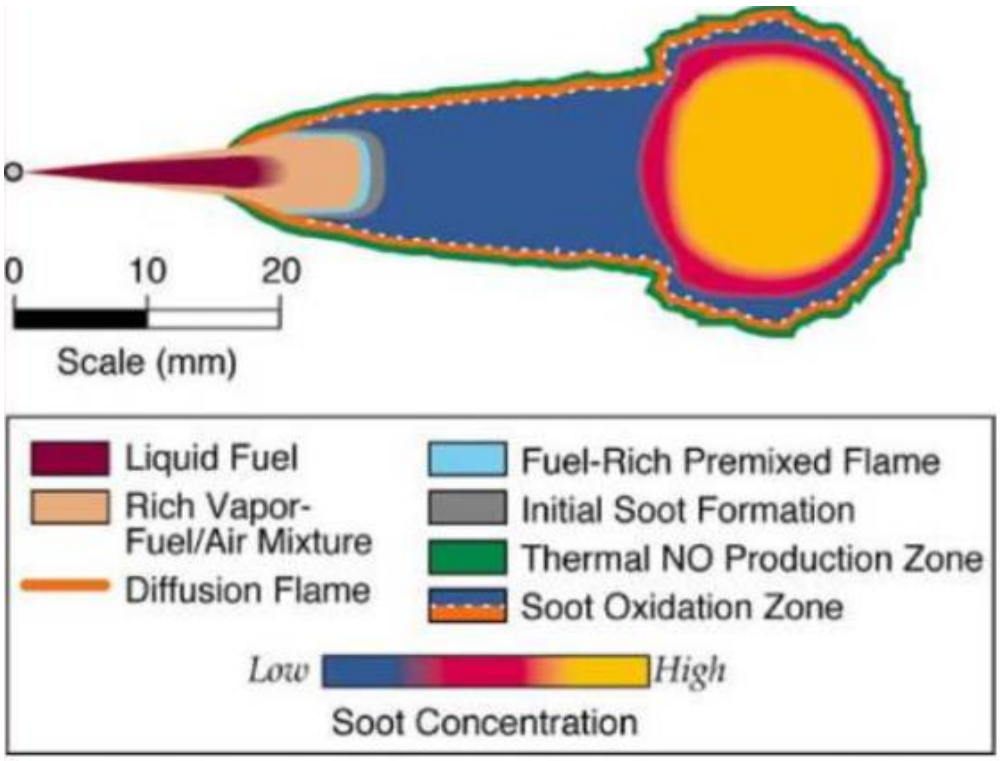

Figure 2: Diesel spray distribution with soot and NO production zones. [11]

At the time of ignition, the locally over-lean mixture reacts slowly or not at all, and results in products of incomplete combustion. Additionally, regions of over-rich—but combustible—mixtures can cool down due to the latent heat of vaporization and do not burn all the way to products; this is known as quenching. As a result, there are incomplete combustion products from the over-lean and rich mixtures; which include unburned hydrocarbons in volatile and solid form (soot). Soot formation can visually be witnessed as a black smoke emitting from the tail pipe. Pyrolysis can also occur, which is another mechanism leading to more unburned hydrocarbons [6 pg. 620].

Particulate matter traps in the exhaust have been designed to substantially reduce soot emissions, but are difficult to implement in all applications. The filters cause problems when clogged and are typically not used in heavy-duty applications, but have been implemented in light-duty automobile engines where the soot formation is smaller. Over time the filters clog up, and as a result the engine injects more fuel to compensate for the loss of power. This increases the exhaust temperature, and in-turn heats up the catalytic material on the soot trap to oxidize the trapped soot: the process was put into effect in the late 1980s. This method proved to reduce

particulate emissions by 70% or more [12]. The amount of soot expelled after exhaust treatment relative to SI is still higher. The efficiency of burning the soot in particulate filters has increased over the years, largely due to the improvement of control systems. Today, the engine computer will run in a manner to increase exhaust temperatures (e.g., variable valve timing, injecting fuel at retarded timing) in conjunction with an extra fuel injector placed in the exhaust. Fuel is injected into the hot exhaust to react with a catalyst element to burn the excess soot. The soot is converted into ash and trapped [13].

High in-cylinder temperatures result in the production of NO_x, due to the dissociation of molecular nitrogen (N_2) and the subsequent reactions of N with O radicals. To mitigate this, exhaust gas recirculation and retarding fuel injection timing are used to lower the temperatures and freeze the chemistry required for NO_x formation [6 pg. 587]. To further reduce NOx—or to replace the aforesaid techniques to reduce NOx—the use of selective catalytic reduction has been recently implemented. This technique regulates the injection of a synthetically produced aqueous solution called urea. When heated, this liquid forms ammonia and reacts with NO_x to produce nitrogen and water. Since diesel engines always operate on the lean side of stoichiometric, CO emissions are low enough to be insignificant [6 pg. 592]. Three-way catalysts cannot be used with diesel engines because they require stoichiometric air-fuel proportions in order for the catalyst to work effectively. Diesel engines run largely diluted mixtures of fuel to air, which causes a lack of UHC and CO relative to NO_x in the exhaust that the catalyst needs in order to effectively reduce all three pollutants to benign species [9].

The current exhaust after-treatment systems (for NO_x and soot management) implemented in Diesel engines are very expensive. In a modern truck, the expense for these systems is roughly the same cost as the IC engine itself, and therefore is a major expense [7]. In the following section, LTC engine concepts will be introduced that offer the potential of reducing or eliminating the need for exhaust after-treatment strategies [14] and offer increased thermodynamic efficiency [10].

1.1.3 Low Temperature Combustion Engines

Diesel engines are capable of operating at higher thermal efficiencies due to the high compression ratios, lean mixtures (therefore higher isentropic exponent), and unthrottled operation (i.e., high volumetric efficiency). However, due to the complex heterogeneous

combustion of the high molecular weight diesel fuel, large amounts of soot form. Therefore, diesel engines are thermodynamically efficient but not considered clean. On the other hand, SI engines combust a homogeneous, stoichiometric charge of reactants and thus produce low amounts of soot. At the same time, the higher combustion temperatures produce NO_x, requiring three-way catalysts to reduce the pollutant emissions. Thus, SI engines are considered clean but not as efficient as diesel engines. Advanced combustion engines are an attempt to operate in a regime that both minimizes emissions and increases efficiency by reducing the combustion temperatures. This idea is called low temperature combustion (LTC), and will be the focus of this section.

LTC engines are an attempt to unify the benefits of SI and CI engines, where a lean fuel-air mixture (to keep temperatures low and isentropic exponent high) is introduced into the piston-cylinder chamber and ignited through compression [7]. Homogeneous charge compression ignition (HCCI) engines implement the LTC concept with a lean, well-mixed (homogeneous) charge of fuel (e..g, gasoline, diesel, or viable hydrocarbons) and air. This cycle is thermodynamically similar to the Otto cycle, but offers higher compression ratios. With this, it is possible to offer the efficiency closer to that of or higher than a diesel engine, but with low emissions of particulates, CO_2, and NO_x without the need for expensive exhaust after-treatment [14]. However, the homogeneous charge coupled with compression ignition leads to control issues (e.g., rapid heat-release, ignition timing) due to the fuel-air mixture igniting all at once, coupled with ignition timing being chemistry controlled and not by a physical process. Other methods, such as reactivity controlled compression ignition (RCCI), leverage multiple fuels and injection strategies to operate in the LTC regime with fewer control issues than HCCI [7]. There are other variants of LTC engines, and interested readers are guided to Yao et al. [15] for a more comprehensive survey.

This work will focus on HCCI operation, as it utilizes the LTC concept and theoretically offers higher efficiency over the other LTC variants due to the homogeneous charge [7], with net indicated efficiencies reaching upwards of 50% [10]. Recall that fuel performance associated with LTC engines is a primary scope for this work. The chosen metric should reflect the fuels ability to operate in the theoretically most efficient LTC mode. As a result, HCCI was chosen as

the cycle to represent LTC operation. Efficiency, emissions and challenges of LTC operation will now be discussed.

From a thermodynamics standpoint, LTC is favorable for two main reasons. The first being a reduction of in-cylinder temperatures, which reduces the heat transfer losses. LTC having temperatures of 1800 K are typical, whereas, for example, diesel CI can exceed 2500 K [16] (see Fig. 4). Assuming similar wall temperatures between LTC and the diesel example (e.g., 450 K), a 5% increase in efficiency can be achieved simply by minimizing the heat transfer effects [17]. Secondly, as we know from analyzing the ideal Otto and Diesel cycles, an increase in the isentropic exponent—the heat capacity ratio, γ—results in more useful work output due to less heat being required to increase the temperature of the gas, and therefore less work input is needed to move the piston. LTC engines achieves higher isentropic exponents by running fuel-lean (see Fig. 3). Furthermore, the isentropic exponent decreases as temperature increases (see Fig. 3). Due to this, reducing the overall combustion temperatures and running fuel-lean is favorable from a work expansion standpoint. Increasing the isentropic exponent from 1.36 to 1.38—considering an ideal Otto cycle where $\phi = 0.5, P_1 = 1\ bar, CR = 14{:}1$, and constant γ—results in an increased gross work output by 55% [8]. LTC operation conveniently provides fuel-lean operation, lower combustion temperatures—as well as diesel CI-like compression ratios—which both provide increased isentropic exponents for an increase in work availability. The overall impacts of these benefits will now be discussed.

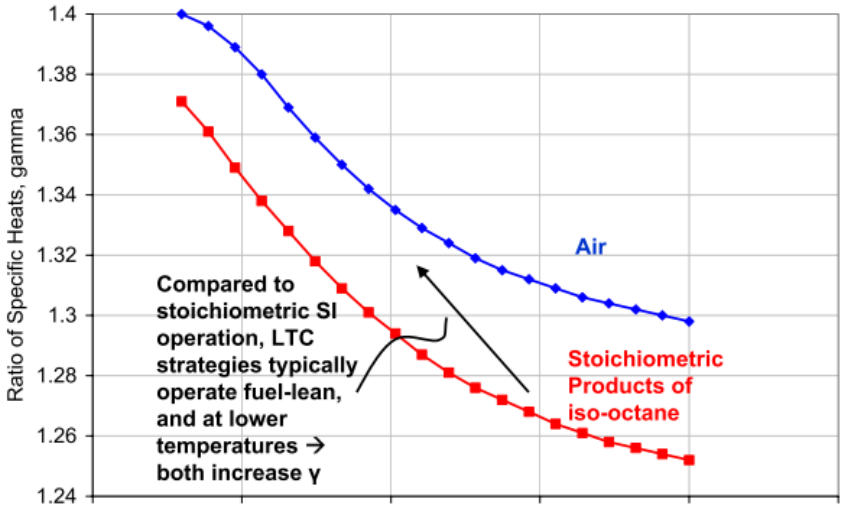

Figure 3: Isentropic exponent (ratio of specific heats C_p/C_v) of air and a stoichiometric mixture of air/iso-octane with temperature dependence. Making the mixture leaner increases the isentropic exponent. [8]

LTC-based engines are potentially the most fuel efficient and clean ICEs as they exploit compression ignition of lean (high-γ) fuel-air mixtures after mostly complete mixing [7]. This sets the stage for lower temperatures and quiescent flows—resulting from the lean mixtures with no flame-front propagation—which decrease the convective heat transfer losses to the cylinder walls to increase thermal efficiency [7]. As a result, fuel efficiency can be up to 20% higher over CI diesel and 40-50% over conventional SI gasoline engines [7]. If LTC engines were adopted in favor of SI engines, a 34% reduction in oil usage could be achieved—i.e., all the current U.S. oil imports from the Persian Gulf [7]. It is also possible to reduce our reliance on fossil fuels with LTC operation, as renewable fuels such as bio-diesel, methanol, or ethanol can be utilized.

As can be seen in the Figure 4 on the following page, LTC engines operate at low temperature and lean fuel-air mixtures to avoid operating regions that are required for NOx and soot to be produced. Note that the y-axis is in equivalence ratio. Thus, an equivalence ratio greater than unity represents a fuel-rich mixture. Temperatures cannot be too low because fuel combustion efficiency is compromised, as seen in Figure 5, which results in the production of CO and UHC. With overall efficiencies being higher, the CO_2 emissions are lower; since LTC operation is clean enough to get rid of expensive exhaust after-treatment systems and inherently does not rely on high pressure injection systems, such engines (e.g., HCCI, RCCI) could be cheaper to produce [14].

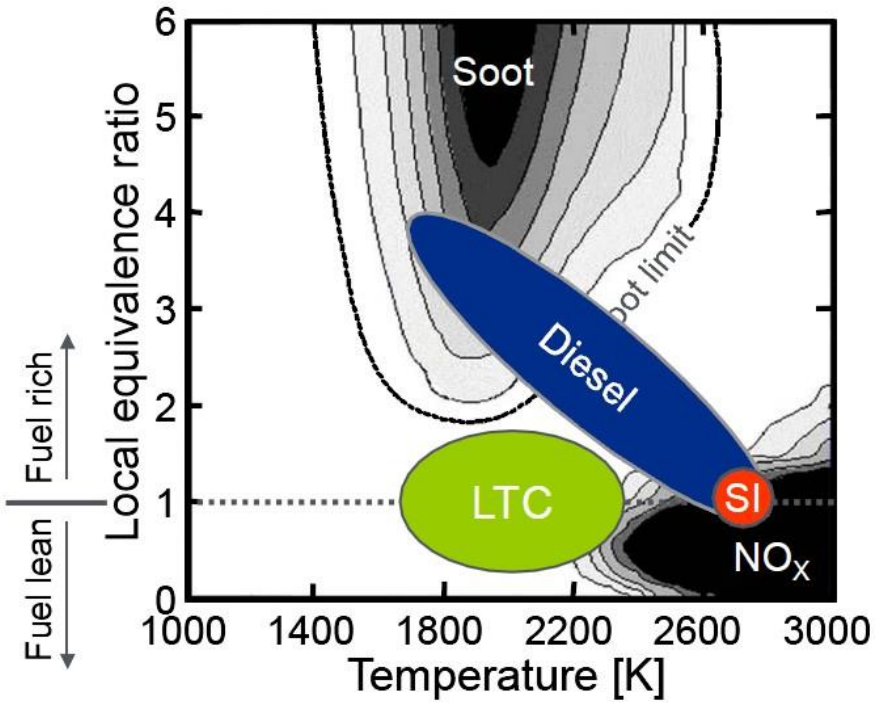

Figure 4: Operating envelopes of diffusion diesel CI, SI, and LTC. [18]

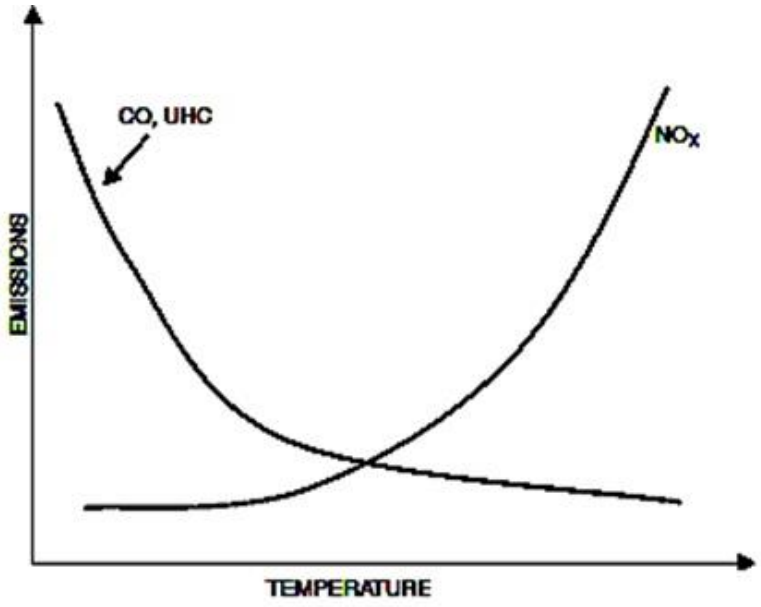

Figure 5: Temperature dependence of complete combustion, and the tradeoff with NOx formation. [19]

There are two major challenges that face advanced LTC combustion engines, including HCCI and variants thereof, and their continued development. The first being that there is no mechanical means to control ignition of the fuel, such as a spark or directly injecting a fuel that reacts nearly spontaneously; instead, the chemical kinetics of the fuel govern when the fuel will ignite during compression and how fast it will burn. The first challenge is then a timing issue that is governed by the fuel chemistry. The effect of ignition timing issues can be seen in Figure 6. Some fuels alone, such as iso-octane in Fig. 6, are not adequate for LTC combustion from a phasing standpoint for most operating conditions. The second issue with HCCI is the low torque per unit displaced volume, i.e. low torque for the size of the engine, caused by the onset of high pressure-rise rates (knock) that are highly audible as a ringing noise, and more importantly highly destructive to the ICE itself [20]. Knock is the result of rapid chemical heat release, and can cause engine damage—this is further discussed in Sec. 1.2.1. Because of the onset of knock, high load limits are difficult to achieve making LTC unfavorable as a power source for transportation means in its current state. Additionally, low-load limits are difficult to achieve due to incomplete combustion of the reactants leaving behind unacceptable UHC and CO emissions. These challenges can be visualized in Figure 7, which shows typical upper and lower limits of HCCI engine operation due to the aforementioned reasons. In order to make contributions to LTC development, one must study engine combustion as a whole; the overall fuel reaction rate is highly dependent on the ICE environment.

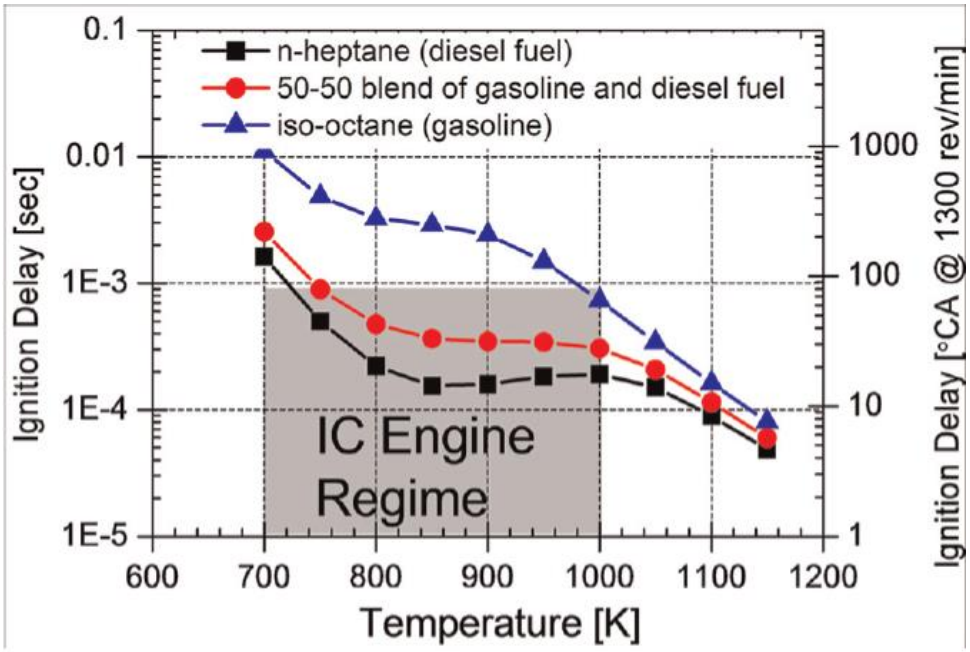

Figure 6: Ignition delays of various fuels and the required values for IC operation. [9]

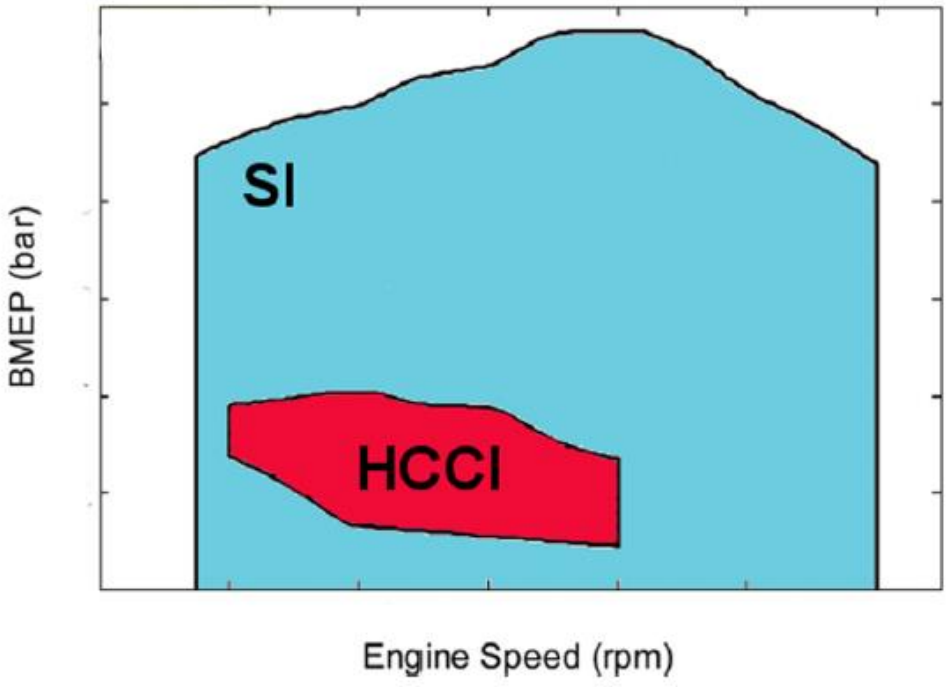

Figure 7: Typical HCCI and SI operating envelopes (specific power vs. engine speed). [21]

Much of the current research and development (R&D) targeted towards LTC involves fuel [22]–[26] and control systems design [8], [9], [27]. Interested readers are directed to Yao et al. [15],

who summarize the inception of HCCI research, key developments, HCCI fundamental theory, control systems, fuel considerations, and recent trends. Control system design encompasses concepts such as reactivity controlled compression ignition (RCCI), where multiple fuels with different reactivity (e.g., gasoline and diesel) are used in varying proportions in an effort to control ignition timing at different engine speeds and loads [9]. Another control system effort is for dual-mode HCCI/SI operation, where SI operation is used for high loads and HCCI is used for low loads to achieve better fuel economy [27]. Alternatively, conventional diesel engines have been converted to run on gasoline in the LTC regime with clever multiple-stage injection strategies utilizing premixing and direct injection [8]. This current research, however, focuses on the fuel design aspect of HCCI R&D, and therefore the remainder of the introduction will focus on fuel performance and metrics, followed by how this work intends to contribute to this area of research.

1.2 Fuel Performance Metrics

Fuel performance has been traditionally measured in two metrics, anti-knock index (AKI) for gasoline—or pump octane number, as it is the number assigned to the different qualities of gasoline at the pump—and Cetane number (CN) for diesel fuels. The first ignition metric discussed will be for gasoline.

1.2.1 Gasoline Spark Ignition (SI) Engines

In ICEs, the higher the geometric compression ratio used, the higher the thermal efficiency achieved for an Otto cycle [6 pg. 170]. The limitation on the compression ratio for SI engines is dictated by the onset of engine knock that is itself influenced by the combination of fuel chemistry and operating conditions—higher compression ratios in general increase the propensity for knock. Knock is the result of rapid chemical heat release, and can cause engine damage. For completeness, knock will be briefly explained here.

With knock, extremely rapid chemical heat release occurs. This additional heat release results in higher pressures and the higher pressure rise rates, potentially causing damage to the engine—this is known as “knock”, as the process creates an audible knocking or ringing sound. Two theories exist for the cause of knock: autoignition or detonation theory. Autoignition theory refers to the spontaneous ignition of the end gas—i.e., the fuel, air, and residual gas mixture ahead of the spark-ignited propagating flame—that occurs due to pre-flame chemistry from

elevated temperature and pressures [6 pg. 457]. The detonation theory suggests that the advancing flame front actually accelerates to supersonic velocities. More recently, Dibble et al. suggest that the causation of knock is from hot spots, caused by non-uniformity in temperature or concentrations, in the end-gas that auto-ignite too quickly for the pressure to equilibrate [28]. This results in pressure waves, which can cause the formation of detonation waves [28]. Simply put, this is a combination of auto-ignition and detonation theory—interested readers are guided to Heywood [6] for a brief overview of the concepts and to Dibble et al. for a more current understanding of the phenomena [28]. Regardless of the causation, the result is potential engine damage and noise pollution.

Some fuels are more resistant to knock than other fuels, and knock constrains engine efficiency. Effectively limiting the temperature and pressure of the end-gas—to combat the occurrence of knock—limits the engine compression ratio [6 pg. 453]. Since particular fuels are more resistant to knock withstanding higher compression ratios can be achieved. It is therefore useful to define the ignition quality of gasoline fuels as the propensity to withstand knock (e.g., octane requirement). The AKI assigns a performance metric to the gasoline auto-ignition quality, and is defined by the arithmetic average of the research octane number (RON) and motored octane number (MON). The two octane numbers are determined in an experimental facility using a Cooperative Fuels Research (CFR) engine via the ASTM-CFR standard testing procedures D2700-14 and D2699-13b [29], [30]. Fuels are combusted in the engine while the compression ratio is increased, till the fuel starts to exhibit significant cyclic variation (instability) or auto-ignition. Each test has specific operating points, the primary difference being air-fuel charge temperature regulation and engine speed [29], [30]. The operating points can be seen in Table 1.

Table 1: Testing conditions for RON and MON tests; from Heywood [6 pg. 473]

Operating Conditions for Research and Motor Methods		
	Research Method (RON)	Motor Method (MON)
Inlet Temperature	$52^o\ C$ (125^oF)	$149^o\ C$ (300^oF)
Inlet Pressure	Atmospheric	
Humidity	0.0036-0.0072 kg/kg dry air	
Coolant Temperature	100^oC (212^oF)	
Engine Speed	600 rev/min	900 rev/min
Spark Advance	13^o BTC (constant)	$19-26^o$ BTC (varies with compression ratio)
Air/fuel ratio	Adjusted for maximum knock	

The octane scale is based on the primary reference fuels (PRFs) iso-octane (2,2,4-trimethylpentane) and n-heptane; the former corresponds to a value of 100, and the latter 0 for both RON and MON. The octane rating of PRF mixtures scales with volume percent of the respective fuels. For example, PRF20, a 20 [vol%] mixture of iso-octane and 80 [vol%] of n-heptane corresponds to an octane rating of 20. The scales are quantitatively determined by the compression ratio required to make the PRFs spontaneously ignite.

In order to attribute an octane rating to a fuel of research interest, the critical compression ratio is measured from the RON and MON tests, and then compared to what mixture of the PRFs is required to yield the same compression ratio results. Then, the octane rating of the fuel in question refers to the volume percentage of iso-octane in this similarly performing PRF mixture.

1.2.2 Diesel Compression Ignition (CI) Engine

Diesel engines are capable of operating at higher thermal efficiencies due to the high compression ratios, lean mixtures (therefore higher isentropic exponent), and unthrottled operation (i.e., high volumetric efficiency). High compression ratios are achievable because the diesel fuel is introduced into the piston-cylinder chamber when it is desirable to ignite the fuel, and not along with the air during the intake stroke as is done with SI engines (with the exception

of direct injected SI—also known as stratified SI, which is not yet widely adopted). When the diesel is injected, there is a finite time delay until auto-ignition occurs. Since knowing the delay time is significant to ensure optimal combustion timing, and therefore efficiency, the cetane number (CN) corresponds to this ignition delay. CN can be thought of as the antithesis of sorts to AKI. That is, resistance to autoignition results in a high AKI, but low CN. A high CN is in most cases desirable for reduced emissions (e.g., fewer unburned hydrocarbons) and lower noise [31].

The scale is defined by Cetane (hexadecane) and alpha-methyl naphthalene (iso-cetane); the former ignites very quickly and is associated a value of 100, and the latter very slowly and assigned as 0. The values are determined in a similar manner to that of RON and MON with a Cooperative Fuels Research (CFR) engine, where instead the compression ratio is increased until a set ignition delay time is achieved. The resulting CN is determined by the resulting mixture of Cetane and iso-Cetane required to achieve the compression ratio attained by the test.

1.2.3 Low Temperature Combustion Engines

LTC engines currently do not have any sort of accepted standard to quantify fuel performance, although a brief overview of some significant efforts focused in this area are presented below. Knock in SI engines is a similar phenomenon to what occurs in HCCI engines; and with this, a question comes to light. Since the octane rating measures the fuel's resistivity to knock, is it therefore possible to utilize known RON and MON to predict HCCI performance? Unfortunately, in SI engines when knock occurs it primarily involves flame propagation in the end-gas at elevated temperatures and pressures. In HCCI engines, however, knock occurs with the entire air-fuel charge at lower temperatures and lean equivalence ratios potentially below the lean flammability limit for flame propagation. Different chemistry dominates these processes, so the octane metric may not pertain to HCCI.

Despite the physical differences of SI and HCCI knock phenomena, modifications to the original AKI were made—in an original attempt to create a better index for SI engines—to characterize fuel performance with HCCI. The octane index (OI), proposed by Kalghagti [32], uses RON and MON along with an engine parameter, K, whose value accounts for the specific engine and operating conditions that RON and MON do not contain alone. Note that if K is set to 0.5, the AKI is returned. The octane index is determined using

$$OI = (1 - K)\, RON + K\, MON = RON - K\, S\ ,$$

where S is the sensitivity of the fuel:

$$S = \ RON - MON\ .$$

The value of K is determined by a best-fit linear regression of OI to the crank angle at which 50% of the combustion has occurred (CA50), as seen below in Figure 8, where K is the slope of the best fit line. Kalghagti discovered that K could be correlated to the temperature in the cylinder when pressure is at 15 [bar] during the compression stroke [32]. The value of K could then be correlated to the compression temperature for a given engine, and provide a useful correlation for OI.

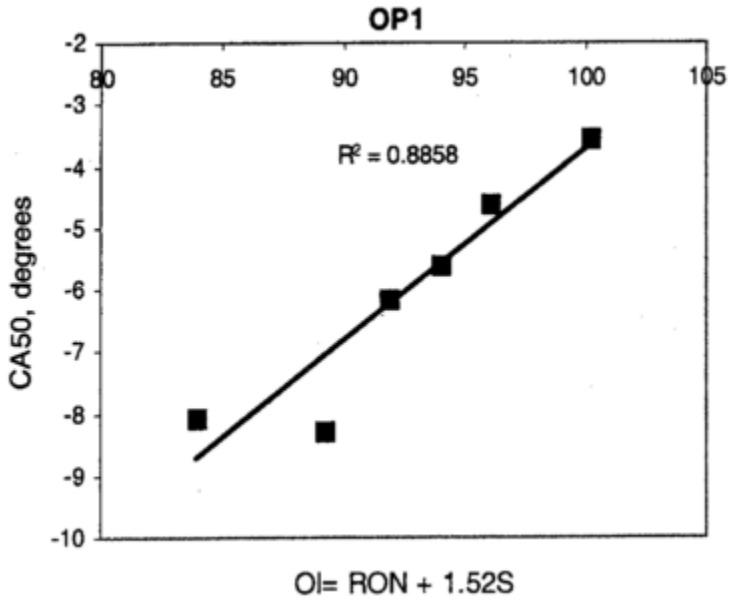

Figure 8: Kalghagti Octane Index correlation to determine K [32].

The work by Kalghagti discovered that HCCI performance does not correlate well with RON or MON taken individually, and instead the two must be considered together. It was proposed that highly sensitive fuels are likely to allow an HCCI engine to operate over a wider load range at a given speed than less sensitive fuels of the same RON [32]. It has been shown by Liu et al. that the Kalghagti OI does not work well to predict the performance oxygenated fuels [33]; most gasolines today contain ethanol. Rapp et al. found that the OI correlates well with auto-ignition of PRF's, but poorly predicts gasoline fuel blends containing naphthenes, aromatics, and ethanol [34].

Kalghatgi's OI reliance on only RON and MON to describe the fuel properties led Shibata and Urshihara to conclude that the OI is insufficient to describe HCCI performance [18–20]. They proposed a new index based on MON; the relative volume percentage of n-alkanes, iso-alkanes, olefins, aromatics, and oxygenates; and temperature-dependent constants. A linear relationship between their index and the crank angle at which 20% of the high temperature heat release occurred was found. The Shibata-Urushihara absolute HCCI index is as follows:

$$\mathrm{S-U\ HCCI\ Index\ (abs)} = \mathrm{m\ MON} + \mathrm{a\ (nP)} + \mathrm{b(iP)} + \mathrm{c(O)} + \mathrm{d(A)} + \mathrm{e(OX)} + \mathrm{Y}$$

where nP, iP, O, A, and OX correspond to volume percent's of n-paraffins, iso-paraffins, olefins, aromatics, and oxygenates, respectively. The remaining symbols (m, a, b, c, d, e, and Y) are constants that depend on temperature and the chemical compound used to represent each chemical class (e.g., the compound n-heptane could represent the class of n-paraffins). A correlation that used RON as opposed to MON was also developed since RON is more widely used as an individual octane descriptor of gasoline fuel performance. Rapp et al. found that the S − U HCCI index correlates well with auto-ignition of PRF's, but poorly predicts gasoline fuel blends containing naphthenes, aromatics, and ethanol [34].

Truedsson et al. defined an HCCI number similar to that of RON and MON [23]. Like RON and MON, the PRFs were used as the scale to define 0-100. Recall for RON and MON, the compression ratio was recorded that caused engine knock. Truedsson et al. defined the HCCI number with the compression ratio to phase combustion at 3^o ATDC (using CA50). This was performed for several intake temperatures at a constant engine speed of 600 RPM, with a constant equivalence ratio of 0.33. A quadratic equation was developed to solve for HCCI number if the critical compression ratio is measured:

$$CR_{crit} = ax^2 + bx + c$$

$$HCCI\ Number = -\frac{b}{2a} + \sqrt{\left(\frac{b}{2a}\right)^2 - \left(\frac{c - CR_{crit}}{a}\right)}$$

where *x* is the volume percent of iso-octane in the PRF used to develop the correlation; and a, b, and c are temperature-dependent constants.

Engine speed was also investigated at three intake temperatures to consider the effect on the HCCI number; this generated new constants for a, b, and c as expected. The HCCI number was found to have almost no correlation to RON and MON, and little correlation to low temperature heat release (LTHR). This HCCI number needs to be further refined by more experiments of the primary reference fuels in the interval, and additionally a method for determining HCCI numbers outside of the 0-100 interval. Also, this work should be investigated in other engines. Lastly, more equivalence ratio's, boosted conditions, and EGR should be investigated. Nonetheless, this HCCI index offers a representation of the required engine geometry, intake temperature, and engine speed required for a particular fuel to operate at optimal combustion timing for LTC operation.

Rapp et al. [34] studied the performance of various fuels in a CFR engine operating in HCCI mode, and discovered that neither octane ratings nor the recently developed HCCI indices of Kalghatgi, or Shibata and Urushihara, reliably predict the autoignition behavior for a wide range of fuels; the Lund-Chevron number was not included in that study. Specifically, gasoline blends with high levels of aromatics, naphthenes, and ethanol are not sufficiently predicted. The HCCI number developed by Truesdsson et al., while a good metric to rate some fuels for successful HCCI operating points, has minimal predictive capability of real fuel HCCI performance due to the limited fuels studied and operating points considered. Also, a new HCCI equation for each equivalence ratio, intake pressure, and engine tested would have to be created; similar to that of the OI, which requires a new K-equation for the aforesaid—having so many relations would be cumbersome.

All prior HCCI performance metrics, while partially successful at describing combustion behavior for gasoline-like fuels, have little relation to realistic engine performance such as fuel economy. It would be helpful if a metric could rank the performance of fuels in such a way that gave insight to potential performance gains from use in LTC engines, in addition to identifying how the fuel will chemically behave in the LTC engine. A LTC fuel performance metric should be able to help designers avoid potentially damaging auto-ignition/knock, but also indicate how efficient the fuel is in HCCI. This metric could then also be used to identify attractive fuels for LTC engine operation based on the potential performance gain associated by using that fuel in addition to containing information on combustion behavior.

A novel LTC index which rates fuels by their potential fuel savings in order to capture useful engine performance from HCCI operation was developed. Niemeyer et al. [35] recently introduced this index (OSU-Chevron LTC index) based on operating envelopes—engine speeds and loads that can be achieved with a given fuel governed by attributes such as auto ignition timing, knock constraints, fuel conversion efficiency, etc.—and comparing the HCCI results to useful operating conditions. The definition of the OSU-Chevron LTC-index is as follows:

$$I_{LTC} = \frac{m_{f,SI\ \epsilon\{HCCI\}}}{m_{f,SI}} * 100\%$$

where m_f represents the mass of fuel used, and subscripts SI and HCCI refer to the IC modes. The numerator represents the mass of fuel that HCCI operation could potentially replace, and the denominator that of the entire driving cycle: the ratio of the two is a mass weighting of sorts that encapsulates potential fuel savings achieved using bi-modal operation of HCCI-SI over its 100% SI counterpart. The (useful) operating envelopes are those required for realistic operating and driving conditions, determined from transient driving cycle simulations. A fuel with poor performance has an operating envelope that is not within realistic conditions. If a fuel cannot achieve realistic operating conditions via HCCI, governed by combustion constraints, then there are no potential fuel savings and therefore a low LTC index assigned. This LTC index is distinct from prior efforts in developing a fuel rating index for HCCI due to the real-world impact on fuel economy that is represented within the index, in addition to the combustion constraints that are held to define viable HCCI operation.

1.3 Chemometrics

1.3.1 Concept

In the last section, it was explained how fuel metrics are determined and used to give an indication of how a fuel will perform in an ICE. This section will discuss how it is possible to predict these metrics with statistical methods. These statistical methods will leverage chemical data to correlate to a metric of choice, such as a fuel performance metric like RON or MON. The purpose of this is creating a model that can accurately predict the performance metric of a fuel sample as to avoid extensive experimental effort. This could be done with a LTC metric to be able to predict LTC fuel performance properties. The model can then aid in next-generation fuel design, as it could predict LTC properties of fuels or fuel blends that have not yet been

targeted in past research efforts. The science to achieve this is known as chemometrics, and will now be introduced.

Chemometrics is the science of extracting information from chemical systems by way of data-driven means, i.e., performing calculations on the data. In this work, infrared (IR) absorption spectra are the chemical data, and the calculations done are statistical correlations relating the IR absorption spectra to chemical properties of interest—the core concept being that infrared absorption spectra provides information on chemical functional groups present in the fuel species, and that these functional groups can be statistically correlated to a chemical property (e.g., index) of interest. More so, the IR absorption spectra require significantly little experimental effort to collect.

A test dataset is used to create the statistical model, which includes set of IR spectra and the chemical property associated with each spectra. This test dataset limits the predictive capabilities of the model, as it is based entirely from the data provided—the larger and more robust the training dataset, typically the better the model. The final result of this is a regression equation, which uses IR absorption spectra as an input and returns a prediction for the chemical property of interest. In other words, chemical properties of interest, such as density, Reid vapor pressure, octane rating, etc., can be retrieved if the IR spectra is known.

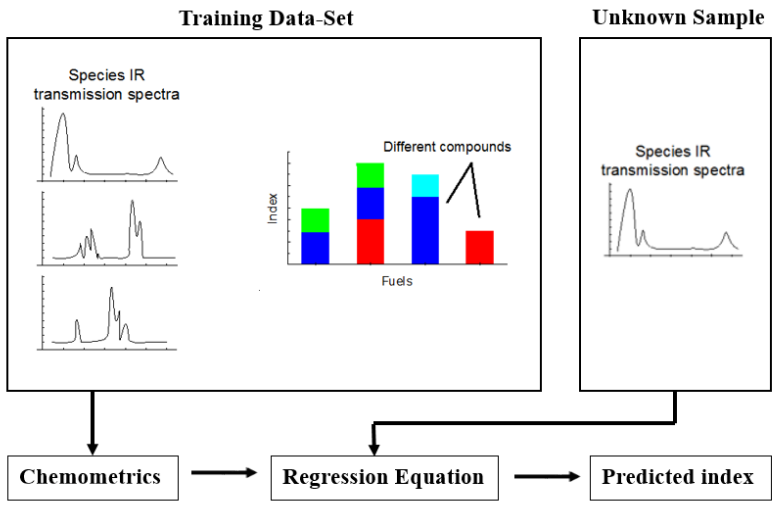

Figure 9: Flow chart of chemometric analysis. A regression model is built from the training data-set (absorption spectra and known indices of each spectra), and the model gives a predicted index based on the input of the unknown samples absorption spectra.

1.3.2 Applications of Chemometrics

Kelly et al. showed that chemometrics can be used to predict the octane number of gasolines to a standard error within 0.4–0.5 [36], which is within the error bounds of the ASTM RON tests itself at $\pm$0.7 [37]. Predicting the octane number of an unknown fuel sample is desirable since physical testing requires time—around 20 min per fuel sample—and is very costly: CFR engines cost around $100,000, require constant maintenance, and each test consumes around a pint of fuel [36]. Collecting the IR spectra of an unknown fuel sample can take as little as a minute—depending on IR collection technique used—and use merely a few milliliters, or even a couple drops, then immediately predict the octane rating using the chemometric model. The significant cost associated with this method is the investment of the IR collection equipment, which can cost half or less than a CFR engine. In addition, collecting IR spectra requires significantly less safety and expert technician/operator considerations and lab modification compared to operating a CFR engine. Kelly et al. demonstrated this with octane number [36]; likewise, the CFR testing procedure for Cetane number could be displaced with this same methodology.

Products developed from industry have leveraged chemometrics to create instruments that analyze gasoline or diesel on-site or in the lab with ease. One of the leading instruments today—the IROX, developed by Garber [38]—uses a robust training dataset of over a thousand gasoline and diesel samples from around the world to create a regression model for properties including octane and cetane number, distillation point, and vapor pressure.

Zeltex, Inc. developed the ZX-101C [39] to predict octane number, which operates on the same fundamental principles as the IROX. However, this version is much simpler and cheaper as it uses 14 selected wavelengths (instead of a broad spectrum) by using infrared-emitting diodes instead of a Fourier transform infrared spectrometer (FTIR,). As a result, the equipment is significantly cheaper than the IROX, but most likely less robust as it uses less data points. Regardless, the accuracy of the instrument proved to predict the octane number of gasoline within $\pm$0.5 [39]. Instruments such as the IROX and ZX-101C, which package the chemometric statistical algorithms and the IR measurement source in a compact manner, demonstrate that chemometrics is a powerful tool for fast, accurate, and easy determination of chemical properties.

1.4 Intent of work

The goal of this work is to use chemometrics to predict low-temperature combustion performance of real fuels. The ability to produce fuels optimized for LTC engines (such as HCCI, as considered in this work) could help reduce fossil fuel consumption, insomuch as a 34% reduction could be achieved in the U.S. [7]. The OSU-Chevron fuel performance index will represent the LTC performance of the fuels: heretofore called the LTC index. Indices for any given fuel of interest can be determined computationally (not to be confused with statistically predicted) as opposed to physical experiments. The LTC index represents potential real-world benefits and encapsulates successful combustion behavior of any fuel in LTC engines. Because of this, it is believed that this index is the best choice as of yet, and compliments the efforts of this work. Investigating alternative LTC performance metrics to that of the LTC index would be disadvantageous, primarily because the LTC index is computationally determined which accelerates this works primary effort. Leveraging indices that rely on experimental studies for the fuels and operating points considered in this work would require vast experimental effort—prior efforts for other LTC performance metrics considered very limited operating conditions [16–20]. By using computational studies to collect LTC performance of fuels rather than time-consuming experiments, the chemometric study can consider a larger data set of fuels. The end goal will be a useful tool that can take an unknown fuel's IR absorption spectra and predict a quantitative value pertaining to LTC performance (LTC index). The model can target hydrocarbon components that result in favorable LTC performance to aid in next-generation advanced fuel design. This tool can therefore point experimentalists in the right direction without the need for time-consuming and costly engine testing.

With the creation of a novel LTC performance index predictive model, instrumentation such as the IROX or ZX-101C fuel analyzer can include this, for researchers and industry alike, as an easy tool to determine LTC fuel performance and help in the effort to design next-generation fuels. It should be noted that the LTC index in itself should not dictate the success of this project; rather, it is a means to prove the concept that chemometrics could predict LTC fuel performance of any kind. That is to say, the success of a chemometric model that can accurately predict the LTC index proves that this is a viable technique—it just so happens that the LTC

index can also be computationally determined, which helps accelerate the process of creating a data set large enough to inform a chemometric model.

1.5 Computational Modeling

The LTC index, as implemented here, requires accurate computational modeling of HCCI engine operation. Single-zone simulations—i.e., assuming spatial homogeneity and no multi-dimensional fluid transport—are computationally efficient and reliable for predicting the ignition quality and knock occurrence of fuels [34, 35]. Multi-zone models exist that divide the piston-cylinder assembly into a few subspaces to consider the boundary layer, core, and crevices, but are computationally more expensive; however, multi-zone models are far less computationally expensive than spatially resolved computational fluid dynamics (CFD) simulations. While these more expensive methods increase the accuracy of heat release, peak pressure, and unburned hydrocarbon emissions predictions to more closely match experimental results, single-zone models suffice for the interest of studying trends and relative fuel performance. Secondly, when considering large chemical kinetic mechanisms that represent fuel chemistry along with a vast parametric study of engine operating points, computationally efficient single-zone models are favored.

For single and multi-zone HCCI simulations, Reaction Design's Chemkin and Chemkin-Pro [41] are widely used due to their built-in models for such. Alternatively, Cantera is an open-source C++ library for chemical kinetics and thermodynamics problems that integrates with Python, MATLAB, C++, and Fortran, and has capabilities to perform stiff chemical kinetic integration and work in conjunction with built-in reactor functions [42]. This approach requires knowledge of programming with the aforesaid programing languages, and requires the user to create the piston-cylinder control volume, heat transfer effects, etc. Cantera does, however, offer more user controllability, and is an attractive approach for researchers interested in developing simulations particular to their problem at hand. Cantera allows complete control and flexibility in terms of both the calculations and the resulting data. As such, Cantera was chosen for the basis of single-zone simulations in the current work.

1.6 Gasoline and the Surrogate Representation

Gasoline fuel is a mixture of hundreds of hydrocarbon species [43] divided into various classes: n-alkanes, iso-alkanes, naphthalene's, aromatics, alkenes, and oxygenates (alkanes and alkenes are also known as paraffins and olefins, respectively). The chemical structure of these various compounds can be visualized in Figure 10.

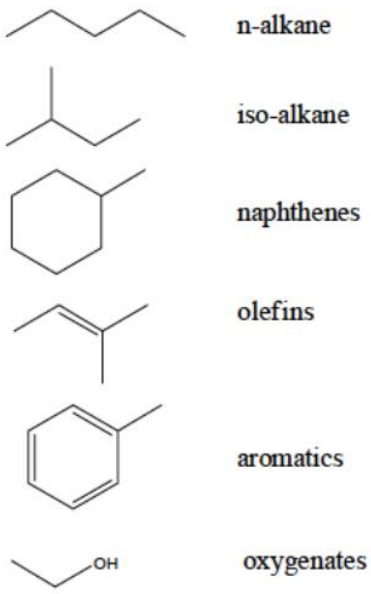

Figure 10: Molecular structures of hydrocarbon classes. [43]

Commercial gasoline consists of all the above compounds in various amounts. The compositions vary widely depending on factors such as the refinery, the source of the crude oil, and additives tailored for optimized drivability depending on the season [44]. Because gasoline is a mixture of different compounds, all of which have different boiling points, gasoline boils, or distills, over a range of temperatures; Figure 11 shows a typical distillation curve for summer and winter gasolines (ASTM D86). The distillation characteristics (volatility) of the gasoline correlate with engine drivability, e.g., how the engine starts, warms up, and runs [44]. For instance, the "front-end" (low temperature distillation characteristics) volatility is known to affect easy cold and hot starting—cold starting being starting the car before it has warmed up, not that it is actually cold outside—freedom from vapor lock (i.e., fuel vaporizing in the fuel line and stalling the fuel

pump), and running-loss emissions (i.e., the fuel evaporates straight from the gas tank to the environment) [44]. Midrange and tail-end volatility are known to effect many other properties; to name a few good short-trip fuel economy and freedom from engine deposits, respectively [44]. Due to all the prior stated reasons, the typical compositions of gasoline in the U.S. vary widely and can be seen in Figure 12.

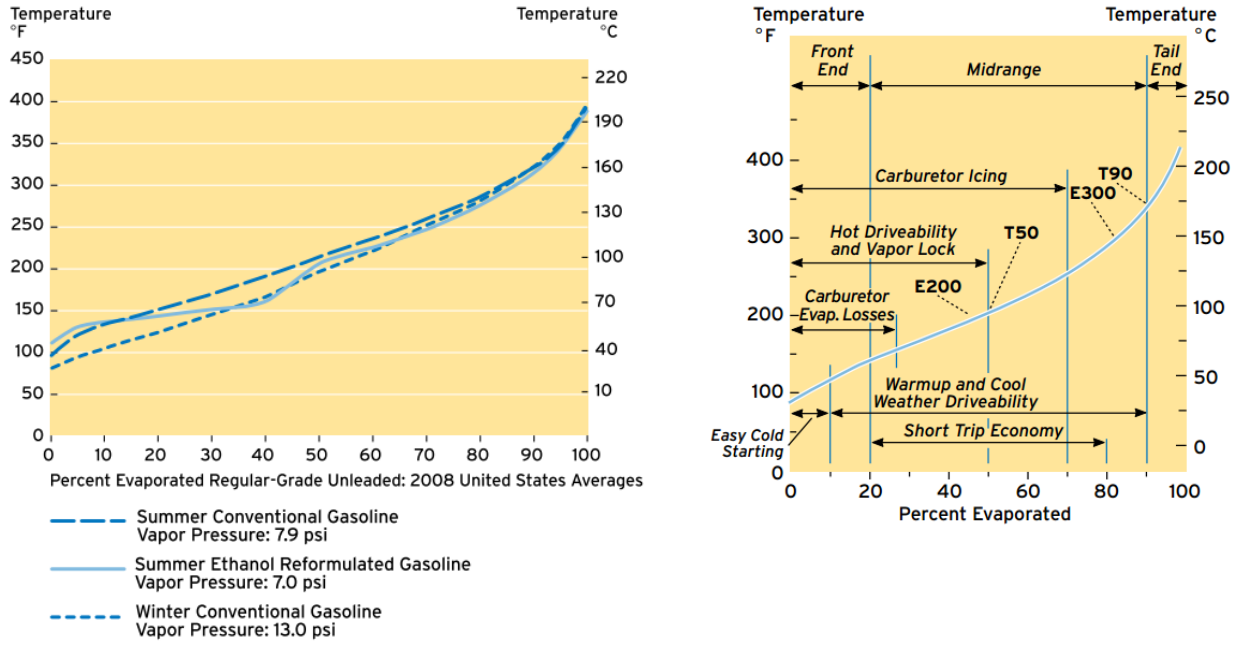

Figure 11: LEFT: Distillation profile for typical commercial gasoline. RIGHT: Distillation profile effects on engine performance. [44]

As discussed, gasoline is a very complex mixture and as a result is difficult to study computationally. In the interest of studying fuel properties, specifically how different hydrocarbon classes and isomers comprising the classes affect fuel properties, gasoline is studied with a surrogate representation. This is done by using a few compounds, typically one of each hydrocarbon class, to create a mixture that approximates the chemical behavior of actual gasolines. This gives the ability for researchers to study the same fuel mixtures, since gasoline compositions widely vary, as well as gain understandings of how individual components interact.

Typical surrogate mixtures consist of primary reference fuels (PRFs) and toluene reference fuels (TRFs). Recall that PRFs are represented by relative amounts of n-heptane (n-alkane) and iso-octane (iso-alkane), and TRFs with the addition of Toluene to the PRF. These are simplest representation for gasoline, but often does not predict the chemical behavior of real gasolines as

many have, olefins, naphthenes, and oxygenates. In an effort to better represent gasoline, the PRFs are often mixed with relative amounts of ethanol and toluene, among other compounds, to represent oxygenates, aromatics, naphthenes, and olefins present in today's gasoline.

Truedsson et al. studied all the aforementioned fuel surrogates for the Lund-Chevron HCCI index [23], and Foong et al. investigated ethanol and toluene for their effects on gasoline RON via ERF and TERF mixtures [45]. Surrogate PRF and TRF mixtures do not include olefin and naphthenes that are present in gasolines. Perez et al. studied auto-ignition behavior for various mixtures of TRFs, with the addition of 1-hexene and methylcyclohexane to represent the olefin and naphthene class, respectively [46]. Kalghatgi et al., while studying OI with relation to HCCI, considered TERF mixtures with the addition of diisobutylene to represent olefins [32].

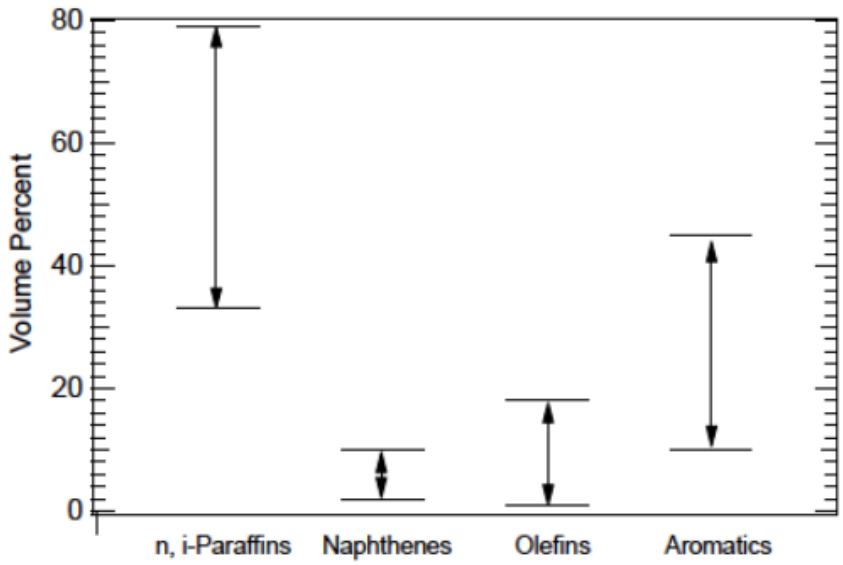

Figure 12: Ranges of hydrocarbon classes found in U.S. gasoline. [47]

1.7 Chemical Mechanisms

Accurate prediction of fuel combustion and pollutant emissions by way of numerical simulation requires detailed knowledge of the molecular reactions that take place during the process, i.e., a chemical mechanism. Chemical mechanisms are the systematic sequence of elementary reactions by which overall fuel oxidation occurs. This is a vital component in reactor simulations; they inform the model that govern combustion processes. In the end, chemical mechanisms provide means to solve species conservation and the first law of thermodynamics, as well as provide data for chemical equilibrium (second law of thermodynamics).

Chemical mechanisms currently encompass pure hydrocarbon species, such as n-heptane, or many pure components and their interactions to represent surrogates for gasoline, jet, or diesel fuels. For TERF mixtures, the Princeton detailed mechanism is an option [48]–[51]. However, TERF mixtures lack a hydrocarbon representing the naphthene class. Methylcyclohexane is frequently suggested as a candidate component of real fuels [52], [53] for the naphthene class, and a mechanism for this component was created by Weber et al. [54]. Many more mechanisms exist for components that comprise gasoline such as n-alkanes, 2-methylalkanes, iso-alkanes, olefins, aromatics, and naphthenes [55]–[65]. Some of these mechanisms are very large, and would require species and reaction reductions if a multitude of studies are needed in a succinct time period. For example, solving a single-zone engine model with the Weber et al. methylcyclohexane mechanism—without a mechanism reduction—on a modest office machine (Intel Core i7-4790K @ 4.4 GHz with 95% CPU utilization) for calculation of the LTC-index would take nearly a month to solve.

Mechanism Automatic Reduction Software (MARS), developed by Niemeyer et al., is capable of automating the process of mechanism reduction given the initial conditions and fuels of interest [66]–[68]. The result of the software yields a reduced mechanism that improves computational efficiency of combustion calculations for the initial parameters selected (pressure, equivalence ratio, hydrocarbon considered), while still retaining the most significant reactions to predict the combustion behavior for the initial conditions considered. Given the size of some of the chemical mechanisms used in this work, this software was utilized to perform mechanism reduction.

1.8 Outline of work

The remainder of this thesis is presented as the methodology (equations, theory, etc.), experimental design, and results for each of the individual topics that comprise this work. The topics include the LTC index, infrared spectroscopy, and chemometrics. Chapter 2 will cover how the LTC-index calculations were performed. Chapter 3 covers Infrared spectroscopy, which compares transmission-based and attenuated total reflectance-based spectroscopy (see Appendix B for in-depth detail on these subjects). Chapter 4 discusses the underlying statistical methods for chemometrics—Principal Component Analysis (PCA) and Support Vector Machine Regression (SVMR). In addition, this section will combine the results of the LTC index and IR absorbance data to create the final LTC index chemometric model.

Chapter 2: LTC Index

2.1 Methodology

In the following sections, firstly the computational approach to calculate the LTC index is explained and secondly the experimental section is presented (i.e., the fuels chosen to be computationally simulated).

2.1.1 HCCI Simulations

Computational simulations were initially performed with Reaction Design's Chemkin software, using a single-zone HCCI model. Later, a Cantera-based model was created in an attempt to better parallelize parametric studies while also having the ability to add more complexity to the model for future research. All documentation on the governing equations, model development, validation, and benefits (over Chemkin) of the Cantera HCCI model can be found in Appendix A. In short, the Cantera-based model created mimics the operation of Chemkin (property evaluation, heat transfer, etc.) and was determined to give virtually identical results to that previously found from Chemkin. With the capability to computationally model the thermodynamics of HCCI operation, the chemical mechanism aspect that is required to model the chemistry of fuels of research interest will be discussed.

HCCI simulations require chemical kinetic mechanisms as an input to model the fuel chemistry. For the fuels of research interest, the chemical mechanisms employed from recent literature include: Chaos et al. for TERF's [48]–[51], Weber et al. for methylcyclohexane [54], Sarathy et al. for 2-methylalkanes and TRF mixtures [55], Sarathy et al. for the FACE A surrogate [69], and lastly Ahmed et al. for the FACE C surrogate [70]. Each mechanism has significant complexity, as particular mechanisms describe chemical behavior of different fuels and have varying levels of species and reactions. If mechanisms found were too large—and therefore too computationally expensive to utilize as-is, in addition to having no other alternative mechanisms to leverage for that given fuel—Niemeyer et al.'s MARS software was used to perform mechanism reduction [66]–[68]. With the appropriate chemical mechanisms at hand and, if needed, reduced to a manageable size to computationally model in a timely manner, the parametric studies to calculate the LTC indices could take place. The parameters studied for the parametric study will now be discussed, including their importance on HCCI engine performance.

The goal of the HCCI simulations is to find all the operation points at which HCCI combustion is viable. To do this, parametric studies of equivalence ratio ($\phi = [0.15: 0.01: 0.45]$), air-fuel charge temperature at BDC ($T_i = [350: 20: 550]$), compression ratio (CR = [9.5, 13]—primarily 13), and engine speed (RPM = [800:100:3000]) were investigated for each fuel; heat release profile, combustion duration, and phasing are dependent on these, which dictate HCCI performance. By simulating through all these conditions, the operational boundaries at which HCCI operation is successful and unsuccessful will be found to adequately determine the operation "map" that any given fuel can accomplish. The reasons that these parameters were selected as they were and their importance will now be explained.

The equivalence ratio array was selected based on typical operating conditions that HCCI engine researchers have successfully accomplished and then extended the range to find unsuccessful/successful operation boundaries [15]; the step-size of 0.01 is chosen to give a fine resolution of the LTC performance, as performance was expected to be highly sensitive to equivalence ratio. Equivalence ratio significantly effects the torque output of the HCCI engine; lower equivalence ratio is less torque, and visa-versa. Due to this, equivalence ratio is used as a means to adjust engine load. Too little fuel (low equivalence ratio) and the fuel/air mixture may not ignite or burn to completion, and too much fuel can cause combustion to occur too rapidly (i.e., knock). The equivalence ratio array is designed to find these two extremes that limit successful HCCI operation, i.e., to find the limits of incomplete combustion and excessive heat release rates that cause engine problems. Equivalence ratio, in other words, affects the amount and duration of heat release of the combustion event. Additionally, equivalence ratio can significantly affect the ignition delay—being the residence time the air/fuel charge is in the engine prior to significantly raising the charge temperature—due to low temperature heat release (LTHR). This is caused by low temperature chemistry that causes an exothermic reaction and acts to raise the mixture temperature and break down fuel components into smaller components early in the compression stroke. This leads to advance the initiation of the more significant, higher temperature heat release (ignition). Figure 13 illustrates the equivalence ratio effects that can be seen for LTC operation.

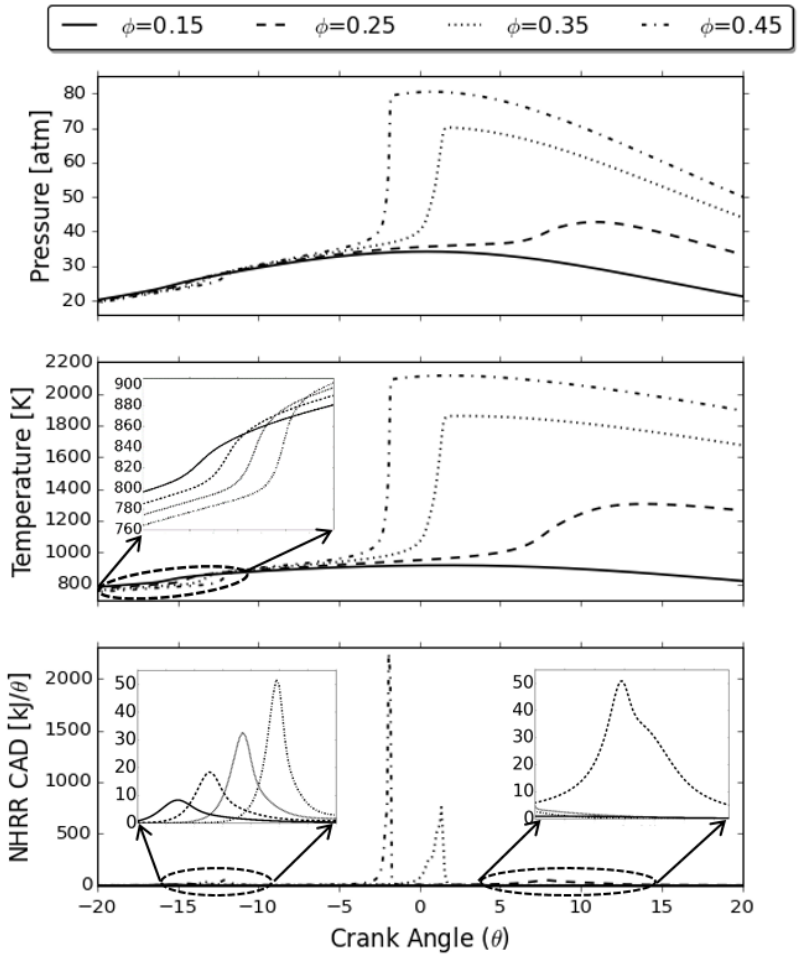

Figure 13: Fuel H60E40 (60 vol% n-Heptane, 40 vol % Ethanol) HCCI combustion at T_i = 350 [K], P_i = 1 [atm], Engine Speed = 800 RPM, CR = 13, ϕ = [0.15: 0.1: 0.45]. Pressure, Temperature, and NHRR per CAD as a function of crank angle. Shown here is the effect of varying equivalence ratio while olding all other parameters constant.

As can be seen for this particular fuel and operating conditions, as equivalence ratio increases the maximum amount of heat release increases, depicted by the increased temperature, pressure and net heat release rate. The duration of the heat release decreases as timing is advanced, due to compression heating. Compression heating accelerates the reaction, due to in-cylinder temperatures rapidly increasing from the heat release occurring during the compression stroke. Furthermore, we observe LTHR from -15 to 10 CAD, which acts to preheat the mixture and advance the ignition of the main, high-temperature heat release. The LTHR increases as the amount of fuel increases, acting to advance timing with the additional preheating. Interestingly, the LTHR is delayed as equivalence ratio increases due to the increase in the specific heat ratio of the air/fuel mixture. Temperature effects will now be further discussed.

The temperature array represents the amount of intake air preheating to the air prior to entry into the piston/cylinder assembly. It is well known that the fuel chemistry is highly dependent on initial temperature, and therefore significantly affects HCCI performance through combustion phasing. Higher intake temperatures, and therefore higher in-cylinder temperatures, aid the initiation of heat release to happen earlier in the compression stroke. This can yield advanced timing beneficial for faster engine speeds, while lower intake temperatures serves to delay timing for slower engine speeds. With the given engine speed range, the temperature range was designed to be large enough for any given fuel to achieve successful HCCI operation for the engine speed limits. Figure 14 portrays the effect of changing intake temperature holding all other parameters constant.

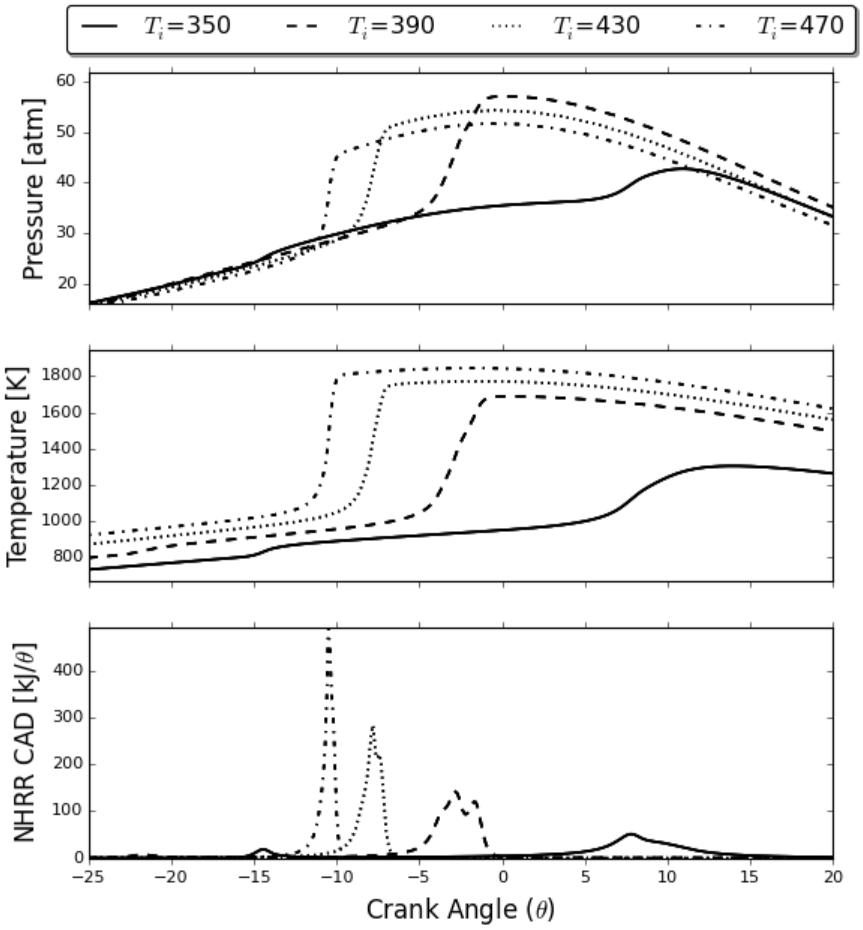

Figure 14: Fuel H60E40 (60 vol% n-Heptane, 40 vol % Ethanol) HCCI combustion at T_i = [350:40:470] [K], P_i = 1 [atm], Engine Speed = 800 RPM, CR = 13, ϕ = 0.25. Pressure, Temperature, and NHRR per CAD as a function of crank angle. Shown here is the effect of varying initial temperature while holding all other parameters constant.

As intake temperature is increased, the phasing of heat release is advanced and the burn duration decreases. In general, since the heat release at higher intake temperatures advanced earlier in the compression stroke, much higher temperature and pressures are seen that lead to more rapid heat release thus decreasing the burn duration.

To summarize equivalence ratio and temperature effects on HCCI operation, they both can affect the ignition delay, duration, and total amount of heat release. Together, the combustion event can be tailored to achieve various torque outputs at different engine speeds depending on the fuel used. If the fuel was found to successfully operate at the equivalence ratio and temperature conditions mentioned above such that the speed and load boundaries were not found, the array bounds were extended to find the point at which LTC operation was unsuccessful.

Recall that the successful HCCI operating points will be compared to real-world engine use to determine potential fuel savings, hence to calculate the LTC fuel index. The operating points that can be accomplished with HCCI will supplement the SI engine that was originally used to accomplish the real-world operating points. As will be explained in detail in the following section, a 2012 Toyota Camry—the best-selling passenger vehicle in North America—is simulated over a driving cycle. The driving cycle represents required vehicle speeds the vehicle must accomplish over a given time period. The vehicles engine then has to match this response with enough torque at any given time to overcome the vehicle weight, road and aerodynamic drag, etc., to achieve the desired speeds. The vehicle simulations with the Camry platform proved that the powertrain is capable of accomplishing the driving cycle in the engine speed range of 800 to 3000 RPM. Therefore, for the HCCI simulations used the same range to directly compare the LTC engine operation to that of the SI engine. Figure 16 shows the effect of varying engine speed holding all else equal.

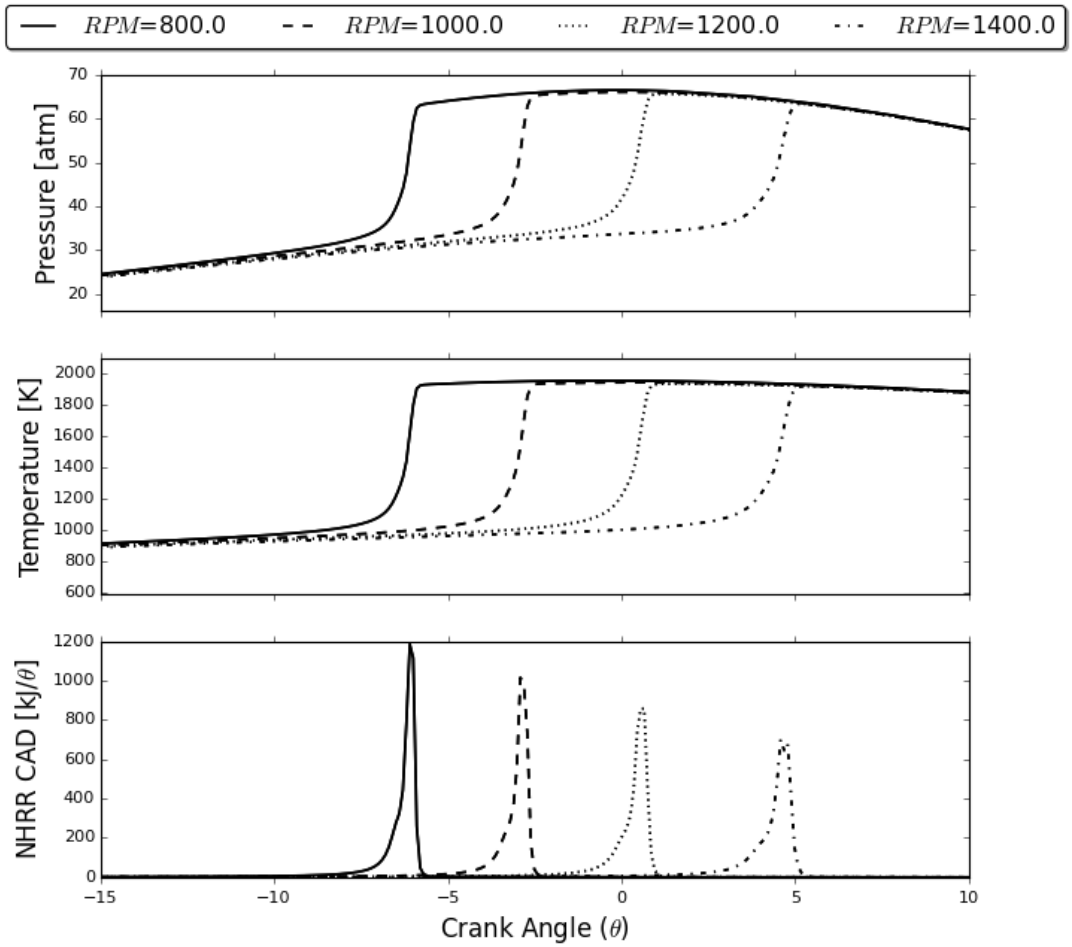

Figure 15: Fuel H60E40 (60 vol % n-Heptane, 40 vol % Ethanol) HCCI combustion at T_i = 390 [K], P_i = 1 [atm], Engine Speed = [800:200:1400] RPM, CR = 13, ϕ = 0.25. Pressure, Temperature, and NHRR per CAD as a function of crank angle. Shown is the effect of varying engine speed while holding all other parameters constant.

Figure 15 illustrates that changing the RPM shifts the phasing of combustion in the engine, and changes the rate of heat release due to a finite time chemical kinetic effect. Up to this point, it is clear that for any given engine speed and torque output desired, a combination of equivalence ratio and intake temperature would have to be well matched to get combustion phasing to a desired value—this will be discussed in the following section. Now, compression ratio effects will be introduced.

Lastly, the compression ratios were chosen to represented typical SI compression ratios used in the majority of older and basic engines today (CR=9.5) [4]. Higher compression ratios (CR=13) were also chosen that are representative of newer engines due to technology advancements such

as variable valve timing or direct fuel injection [4]. The majority of this work is focused on the higher compression ratio of thirteen.

For any given fuel and compression ratio, the entirety of the parametric study equates to typically 7,834 initial conditions for the HCCI simulations. The importance of a compact chemical mechanism paired with a single-zone HCCI model becomes apparent in order to simulate a large number of fuels. With each operating point, crank-angle-resolved parameters were obtained directly from the simulation; such as temperature, pressure, net-heat release, and mass fractions of species; were used to determine engine performance. Post processing the obtained values from the engine simulation yields parameters that characterize engine performance. These parameters include indicated mean effective pressure (IMEP), being the non-dimensional representation of work done by the cycle; CA50, which is the crank angle at which 50% of net-heat release is observed and a descriptor of combustion timing; and maximum pressure rise rate (MPRR)—the last parameter in addition to the amounts of CO and CO_2 generated by the end of the engine cycle being descriptors of engine knock and fuel conversion efficiency, respectively. With these parameters determined from the HCCI simulations, they can be used to determine the viable HCCI operating points. The bounds that dictate successful HCCI operation were briefly introduced qualitatively in this section, and will now be quantitatively introduced.

Successful operation of HCCI combustion is defined by realistic operation limits. These limits for successful operation were bounded by ideal combustion phasing, the maximum pressure rise rate in the combustion chamber, and lastly fuel conversion efficiency. The quantitative limits used for the aforesaid parameters used in literature have been discussed by Niemeyer et al. [35], [71]. A brief overview of the importance of the parameters used to define successful HCCI operation will be discussed. First, the upper and lower constraints that limit successful HCCI operation will now be introduced.

Often, at higher engine loads much of the heat-release occurs rapidly causing in-cylinder pressures to rise too quickly. The pressure rising in this manner can be destructive to the engine and cause unacceptable noise. This is the upper limit of engine operation, and is characterized by the maximum pressure rise rate (MPRR), $dP/d\theta$, and has been set to a limit of 20 bar/(^{o}CA), guided by literature [72], [73].

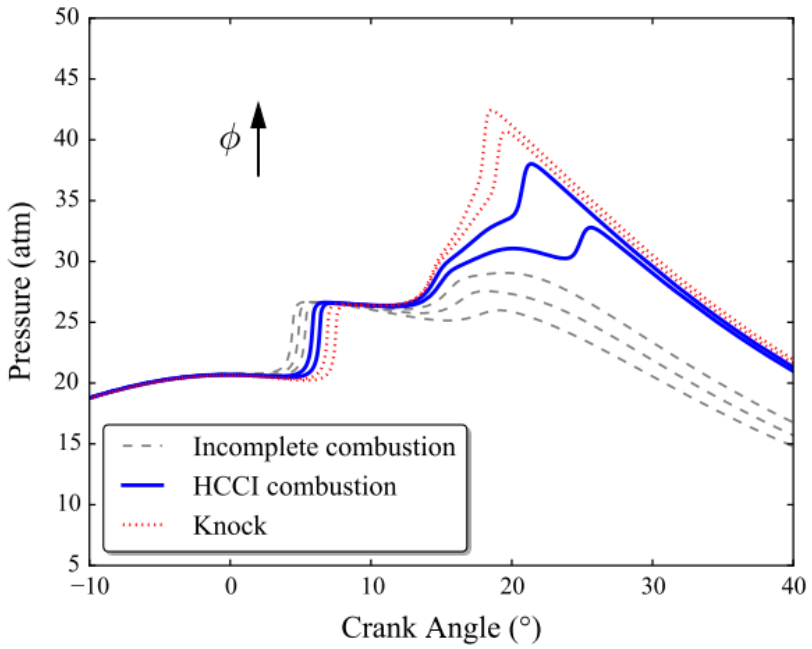

Figure 16: Representative pressure traces for PRF0 (n-heptane) HCCI combustion at CR = 9.5, $T_i = 330$ K, $P_i = 1$ [atm], engine speed = 1400 [rpm], and varying equivalence ratios, demonstrating knocking combustion, normal combustion, and misfire (incomplete combustion). [71]

The lower limit defining successful HCCI operation is represented when misfire occurs; i.e. incomplete combustion. This is determined by the fuel conversion efficiency, defined by the molar conversion of carbon in the fuel to carbon dioxide—the value is set to 90% [74], [75]. These combustion limits can be visualized in Figure 16 for a few operating points. Lastly, the third parameter used to define successful HCCI operation will now be discussed.

As in any reciprocating ICE, the phasing at which combustion occurs is of significant importance to HCCI engines. This timing dictates the overall performance of the engine including emissions and power output. For HCCI engines, CA50 is often used for defining combustion phasing, as it correlates well with engine performance (e.g., efficiency, load, emissions) [76]. Optimal combustion timing for HCCI is determined by selecting the best balance between power, emissions, and acceptable operating conditions. *Higher* torque (power output) typically requires late CA50—more than 10 CAD after top dead center (aTDC)—to reduce the ringing intensity, i.e., MPRR in the engine, for safe operation [77]–[79]. However, *lower* loads favor around 2-6 CAD aTDC for the lowest production of NO_x and unburned hydrocarbons, and ensuring low ringing intensity [77]. *All* loads, however, have the highest combustion efficiency at early CA50

($\leq 5^o$)—but advanced timings are limited to the allowable ringing intensity [77] . In conclusion, CA50 at 2-6 CAD aTDC is a good target for emissions, efficiency, and low ringing intensity for the majority of the load range for HCCI—numerous authors used this range [23], [80]–[84]. Higher loads require boosted intake pressures (turbo/supercharging), which will not be considered for this work, and therefore the required CA50 greater than 6 CAD aTDC will not be considered. Guided by the literature, a CA50 tolerance of 3 ± 2 CAD aTDC will be used to control combustion phasing.

With all three performance limitations defined to ensure realistic HCCI operation, the operating envelopes of LTC operation are then determined. The useful work—dictated by the HCCI operation limit criteria—from the LTC cycle is calculated *a posteriori* from the pressure-volume history and converted to indicated mean effective pressure (IMEP). Recall that IMEP is a normalized representation of work, and is later used to compare to the driving cycle simulations. The relation for IMEP is as follows:

$$imep = \frac{W_c}{V_d}$$

where W_c is the net work from the engine cycle and V_d is the engine displacement volume. An example of successful operating conditions for HCCI operation, with load represented via indicated mean effective pressure, can be seen in Figure 17.

With the successful operating points determined (left) a map is drawn around the points using a convex-hull algorithm (right) [85], [86]. This map of successful HCCI operation will later be compared with that determined from the ADVISOR vehicle driving cycle simulations in order to find the useful fuel economy savings from the LTC cycle: the LTC index. The driving cycle operating envelopes for realistic real-world use, determined from vehicle simulations, will now be discussed.

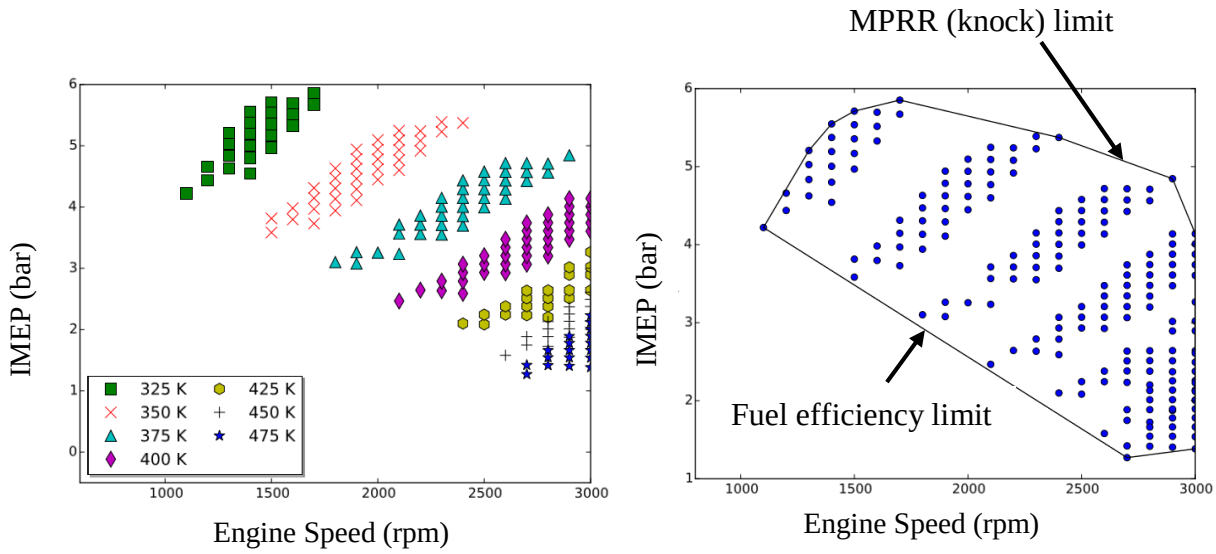

Figure 17: Operating Envelope of PRF 20 (80 vol% n-Heptane, 20 vol% iso-Octane). The left plot depicts temperature and equivalence ratio dependence for successful operation points, i.e., viable HCCI operation at different engine speeds and loads. Note equivalence ratio is not marked, but for any given temperature, load is increased by increasing the amount of fuel. On the right the operation map boundary is depicted, used to overlay with the ADVISOR simulation operating map.

2.1.2 Driving Cycle Operating Envelope Simulations

As mentioned, a light-duty passenger vehicle (2012 Toyota Camry) was simulated in the US EPA FTP-75 driving cycle [87]. Recall that the purpose of this simulation is to determine the engine operating points—in terms of torque/work and engine speed—associated with realistic driving conditions. These points, when found, are compared to the viable HCCI operating points. The FTP-75 driving cycle is a velocity profile over a time period, and can be seen below in Figure 18. This is the standard used to measure the fuel economy and emissions of passenger vehicles in the United States. The cycle itself simulates a short cold start, transient, and a hot start phase. The total distanced travelled is 17.8 km in 1874 s—not including the 10 min break in the cycle—with an average speed of 34.1 km h^{-1}.

The program used to simulate the driving cycle is called Advanced VehIcle SimulatOR (ADVISOR). ADVISOR is a hybrid backward/forward-facing vehicle simulator; where backward indicates the tractive force required to move the vehicle at the directed speed, and forward facing indicating an actual driver directing the engine to meet the necessary speeds.

Combined, an accurate representation of fuel economy and emissions is determined. For more information on these models, please see Wipke et al. [88] or Gao et al. [89]. The vehicle chosen to represent the vehicle is the 2012 Toyota Camry, which is the best-selling passenger car in North America in recent years, as a representative typical light-duty vehicle.

The Toyota Camry is simulated via ADVISOR with the VHE_SMCAR default small passenger car model to determine the operating points to accomplish the driving cycle. This model is based on a 1994 Saturn SL1. The Saturn SL1 model was altered—VEH_SMCAR_CAMRY—and includes the estimated parameters for the 2012 Toyota Camry: coefficient of drag (0.28), frontal area (2.28 m^2), fraction of vehicle weight on front axle (0.54), center of gravity (0.53 m), wheelbase (2.775 m), and curb weight (1447 kg). With this vehicle model the FC_SI95 engine model is used as a representative engine, which is based on the Saturn 1.9L dual overhead cam SI engine [108]. This engine model was modified by scaling the engine default maximum power of 95 kW (Saturn SL1) to 133 kW (Toyota Camry) to more accurately represent the Camry's engine.

ADVISOR yields the engine speed and torque required to achieve the vehicle speed from the driving cycle in one second intervals. The torque value is normalized to indicated mean effective pressure to compare to the HCCI simulation results. The basic relation for a four-stroke engine to calculate this is as follows:

$$imep(t) = \frac{4\,\pi\,T(t)}{V_d}$$

Where $T(t)$ is the torque required to accomplish the necessary vehicle speed as a function of time. The final results of the ADVISOR driving cycle simulation for the 2012 Toyota Camry can be seen in Figure 19.

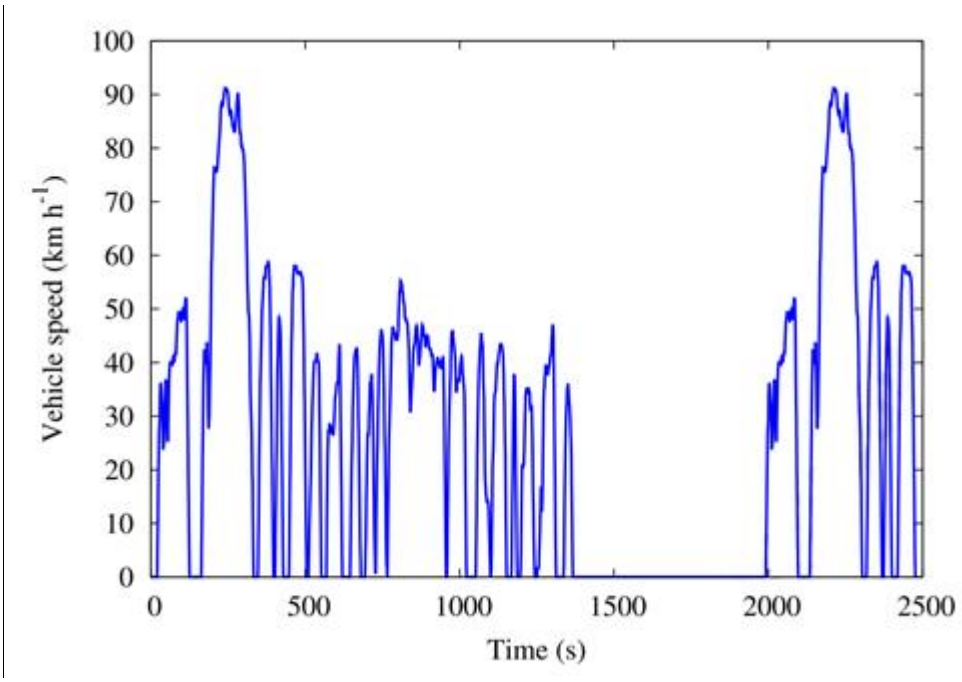

Figure 18: US EPA FTP-75 driving cycle: Required vehicle speed as a function of time, representing typical real-world use. This driving cycle is the U.S. standard for determining fuel economy. [35]

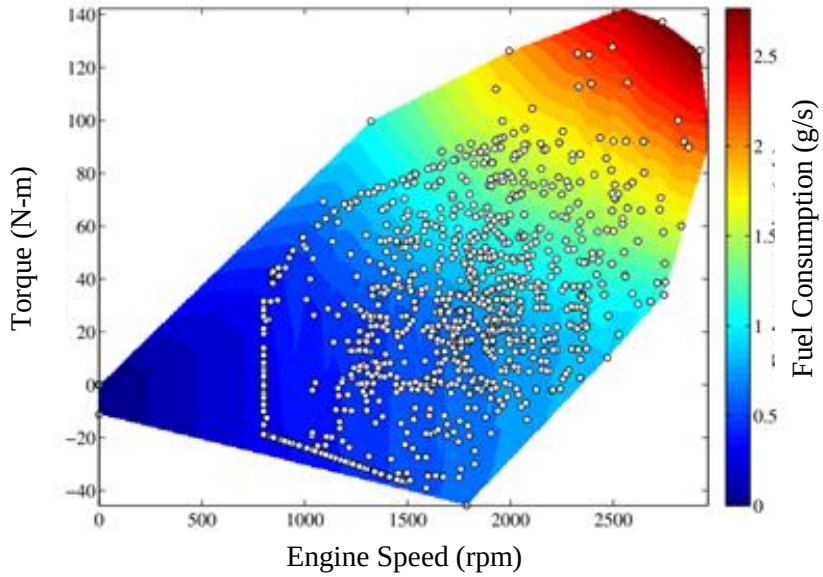

Figure 19: Engine Operating points for the 2012 Toyota Camry that are required to accomplish the driving cycle. White dots represent the torque, or work, required in 1 sec. intervals to accelerate the vehicle to the driving cycle vehicle speeds. The color indicates the amount of fuel required to achieve the engine loads (torque). [35]

2.1.3 LTC Index Calculation

As covered in chapter 1, we found that the LTC index is calculated based on the potential fuel savings offered by LTC operation over its SI counterpart. The fuel savings is defined as the mass of fuel used with the SI cycle within viable LTC operation to that of a pure SI cycle. Again, as shown in Chapter 1, the LTC index can be calculated using

$$LTC_{index} = \frac{m_{f,SI\{LTC\}}}{m_{f,SI}} * 100\%$$

where subscripts SI{LTC} indicate SI cycle within the LTC operating map, and SI only that of SI operation. This calculation can be visualized by the overlay of the HCCI operation envelope with that of the real world requirement from the driving cycle simulation, as seen in Figure 20. With the IMEP map determined for successful HCCI operation, the overlay shows the capable IMEP points that can supplement real world use. The ratio of the mass of fuel used in the operating map to that of all the operating points combined gives the potential fuel savings, as defined by the LTC index. That is, the index is a mass-weighted fraction of operating points. Figure 20 illustrates an LTC index of 34.2.

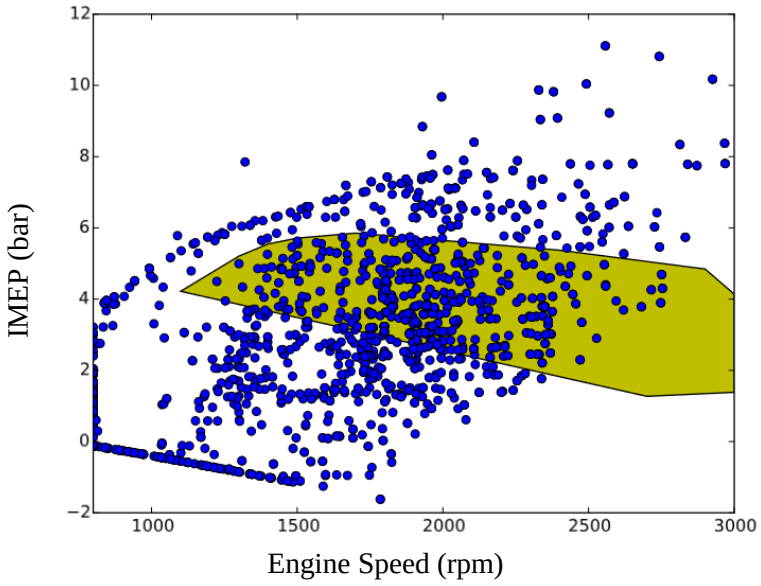

Figure 20: Viable HCCI operation map of PRF 20 at CR=13 (yellow) overlaid with required operation points for real-world operation (blue). The potential fuel savings here (LTC index) were calculated to be 34.2%

2.2 Computational Experiment Design

This section explains the fuels considered to computationally simulated for the LTC index, that is, fuels of importance to utilize for the chemometric model. Ideally, experimental simulations of gasoline's would take place. However, as mentioned, gasoline consists of hundreds of components so researchers begin studies with pure components and mixtures of pure components that are better characterized. In this particular study, the only option to represent gasoline-like fuels are surrogate models that are either neat components or mixtures of neat components. In the end, the LTC indices of the pure hydrocarbon components and mixtures of such will be used to inform the statistical model in combination with their absorbance spectra, and the resulting model used to predict the LTC index of real gasoline fuels.

In the search for fuels for advanced engines which provide fuel efficiency benefits, fuels of research interest to be computationally studied were guided by literature [23], [32], [45], [46], as well as findings from this work. Primarily, the fuels studied were that considered in the work of Truedsson et al. [23], who focused on binary, ternary, and quaternary mixtures of 2,2,4-trimethylpentane (iso-octane), n-heptane, toluene, and ethanol. Other fuel mixtures studied considered constituents such as methylcyclohexane guided by Perez et al. [46]. Later, from findings in this work (see Sec. 4.2.3), it was seen that xylene(s), 2-methybutane and 2-methylhexane could be of significant importance in describing real gasoline fuel behavior, and as a result these were also studied.

Many of the relative mixtures of the pure components studied to represent real fuel qualities or to determine blending effects had already been presented in past literature. However, novel relative mixtures needed to be selected for the new fuels of interest guided by this work. These mixtures were guided by a few motives; the first being to mix the new constituents with the original fuels—iso-octane, n-heptane, etc.—to capture how these fuels interact with each other in LTC mode, and the second being to capture these effects with as little blends as possible to reduce computational efforts. Ultimately, since the chemical mechanisms describing these additional components were relatively large—even after mechanism reduction—the computational effort was still immense. As a result, a smaller number of mixtures in comparison to the original intent were considered due to time limitations. To put the computational effort in perspective for the more complex studies, a single LTC performance index calculation would take around 168 hours

to complete on a server utilizing 36 quad-core CPUs @ 2.9 GHz. The extensive list of fuels computationally studied can be seen in the Appendix D. Mixtures of n-heptane, iso-octane, toluene, ethanol, methylcyclohexane, 2-methylbutane (iso-pentane), 2-methylhexane, and m-xylene were conducted.

In addition to simulating pure components and mixtures of pure components, surrogate representations of specific research-grade gasoline fuels were also studied. Recall that the goal of this work is to be able to predict the LTC performance of real fuels, and therefore the LTC index of real gasolines would have to be known to be able to validate the final chemometric model. Research grade gasoline blends—known as the Fuels for Advanced Combustion Engines (FACE) gasolines—designed by the Coordinating Research Council (CRC, a consortium of fuel and automobile manufacturers conducting environmental research [90]), were used as the gasoline test set. These ten gasoline fuels were statistically designed in an effort to target attractive gasoline blends for next-generation advanced IC engines. Surrogate chemical mechanisms exist for a few of these specific research fuels (FACE A and C) [69], [70], and were computationally studied to determine the LTC indices of these real fuels. The FACE fuels that are able to be computationally simulated will be predicted with the chemometric model, and validated against the computed LTC index. Therefore, FACE fuels “A” and “C” will be computationally studied as means to validate the final chemometric model for its ability to predict the LTC index of real fuels.

2.3 Results and Analysis

2.3.1 HCCI Operating Envelope Examples

Operating envelopes for successful HCCI combustion are shown in Figs. 21–23 for selected fuels; the entirety of the results can be found in the Appendix C. The figures show the temperature dependence to achieve the operating points, as well as the results overlaid onto the Toyota Camry driving cycle operating points. The mass-weighted fuel savings are indicated in each figure via the LTC index, as well.

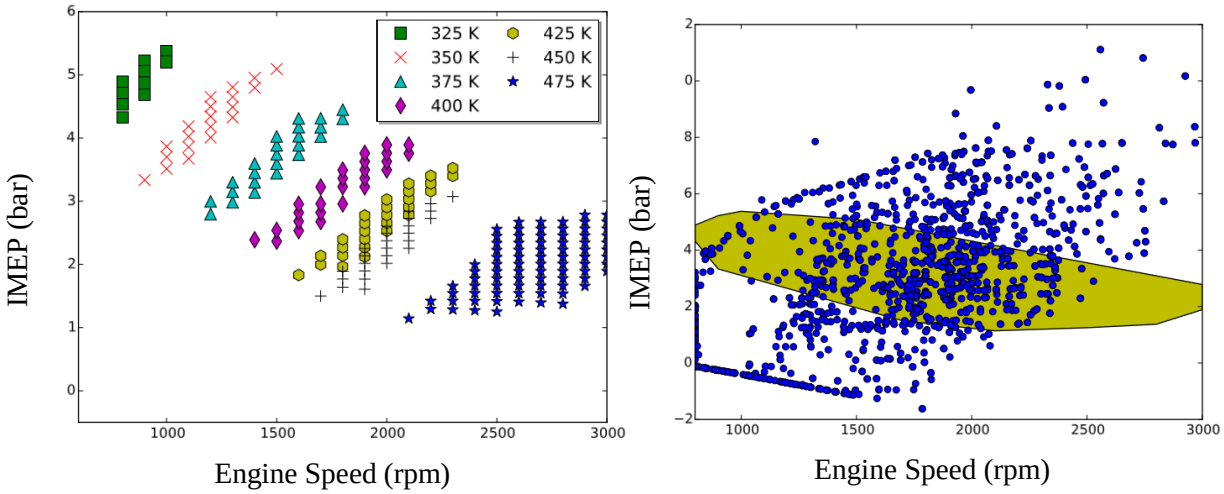

Figure 21: PRF 40 Operating Envelope. RON/MON (40/40), LTC index (32.75%)

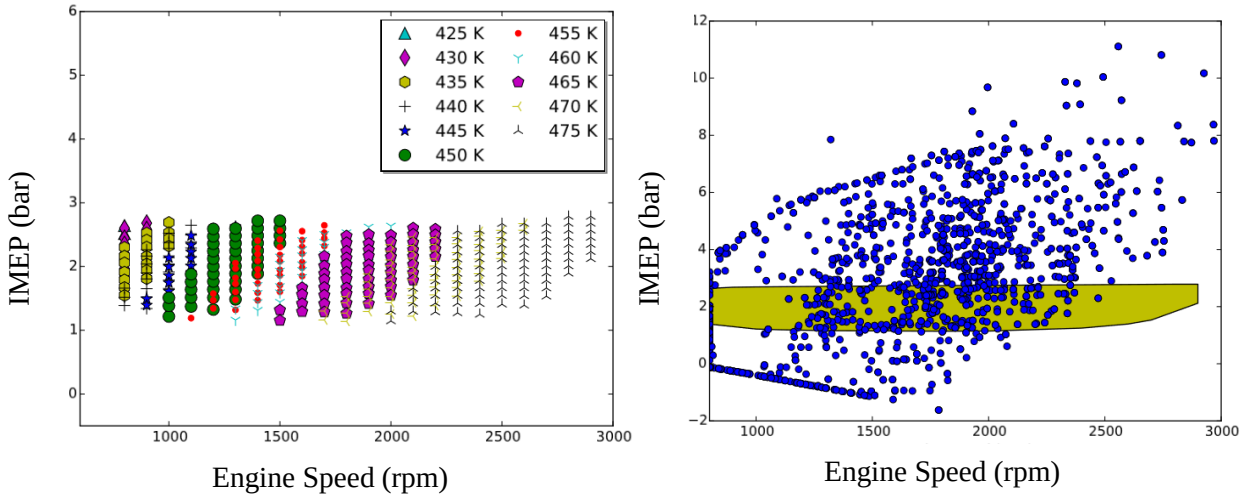

Figure 22: PRF 85 Operating Envelope RON/MON (85/85), LTC index (23.3%)

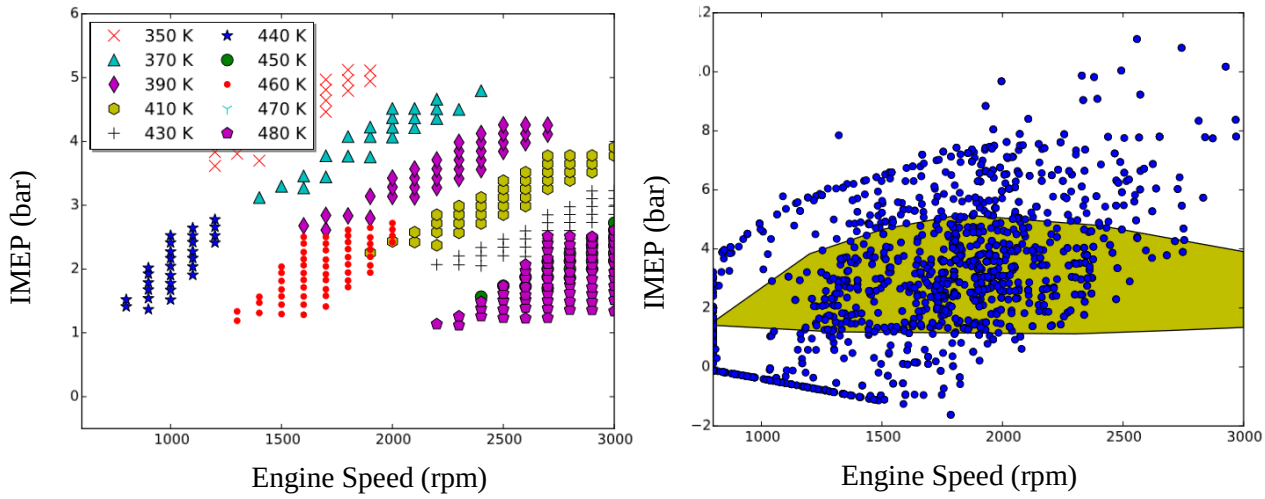

Figure 23: H80T20 Operating Envelope RON/MON (27.7/24.8), LTC index (41.58%)

2.3.2 LTC indices

A table summarizing the calculated LTC indices, as well as known RON/MON values, are found in the Appendix C. A total of 102 fuels were considered consisting of binary, ternary, and quaternary mixtures including PRFs, TRFs, ERFs, TERFs, and TRFs mixed independently with m-xylene and iso-pentane (X and S fuels, respectively). In addition, certain pure components were considered including iso-pentane, methylcyclohexane, xylene, and 2-methylhexane. Lastly, surrogate fuel mixtures that mimicked the behavior of FACE fuels "A" and "C" were simulated. The higher LTC index scoring fuels for the compression ratio of 13 are presented here.

Table 2: High scoring LTC index fuels for CR=13

Fuel Name	RON	MON	LTC index
Two Component Mixtures			
PRF20	20	20	34.2
PRF40	40	40	32.75
H90T10	14.1	12.6	37.72
H80T20	27.7	24.8	41.58
H60E40	71.4	65	35.57
Three Component Mixtures			
TRF1	39	37	35.2
TRF4	82.8	80.7	34.1

2.3.3 LTC index vs. other Fuel Performance Metrics

Traditional performance metrics such as RON and MON, fuel sensitivity (RON-MON), or anti knock index (AKI) were found to have no correlation to the LTC index. It was hypothesized that if the LTC index is a good descriptor of LTC performance, then it should be consistent with similar findings in literature that support the notion that RON/MON do not have significant correlation to LTC performance [23], [34]. The notion that RON and/or MON are not good descriptors of LTC performance alone was shown to be further supported with the LTC index.

As can be seen in Figures 24 and 25, for the fuels studied, the LTC index does not correlate to RON, MON, or octane sensitivity. No linear trend is found, e.g., increasing RON/MON resulting in smaller LTC indices. Fuels with widely varying RON/MON values, due to significantly different fuel compositions, did not necessarily result in different LTC indices: which resolves the possibility of a non-linear trend to correlate the data. Instead, fuels with low RON/MON values proved to be capable of having similar LTC performance to that of high RON/MON. Furthermore, fuels with similar RON, MON, or octane sensitivity showed to have a wide spread of LTC index values. These results support the findings, e.g., of Rapp et al., in that traditional performance metrics do not correlate to those that indicate LTC performance.

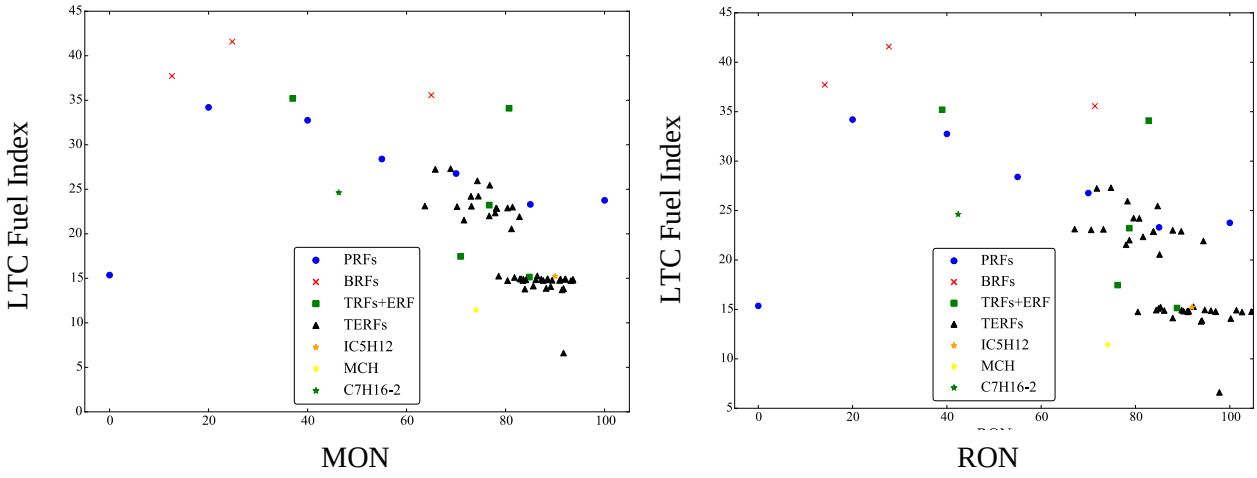

Figure 24: LTC index vs. MON (left) and RON (right)

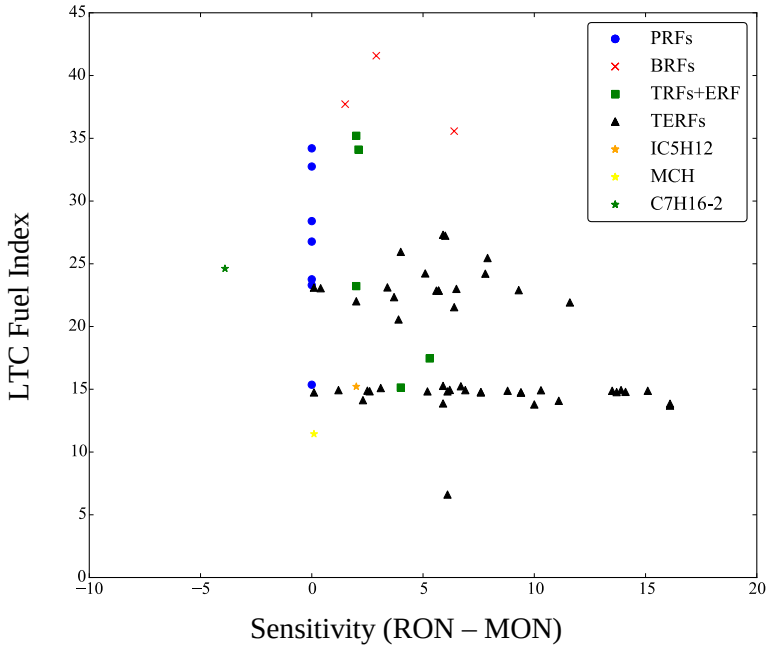

Figure 25: LTC index vs. Sensitivity (RON-MON)

2.3.3 Attractive Next-generation LTC Fuels

It was found that pure components alone do not exhibit the highest or lowest LTC performance. Instead, synergistic blending was seen with mixtures to yield higher or lower LTC performance than any of the respective components alone. For example, pure n-heptane (PRF0) and iso-octane (PRF100) had indices of 15.36% and 23.76%, respectively. However, when PRF20 is considered, the performance increased to a value of 34.2%. Synergistic effects showed that relative mixtures of the components had a strong effect on LTC performance.

In the pursuit of identifying attractive fuels for LTC performance, it is beneficial to have the highest LTC index. In addition to a high LTC index, secondarily it is also desirable to have a high RON/MON in order to be able to operate in SI mode (to resist autoignition) for a dual-mode HCCI/SI control scheme. It was found that fuels with high LTC indices in addition to high RON/MON values exhibited significant equivalence ratio sensitivity to ignition delay, which extended the operating range of these fuels significantly. This trend was exclusively seen for the *higher* RON/MON fuels with high LTC index (e.g., H60E40, TRF4); this was also seen for

lower RON/MON fuels with high LTC index (e.g., H90T10, H80T20, TRF1) with the exception of high LTC index PRF fuels. The equivalence ratio sensitivity is illustrated in the following figure, where a constant-volume homogeneous reactor was simulated to show the trends that explain the increased operating envelopes for these fuels.

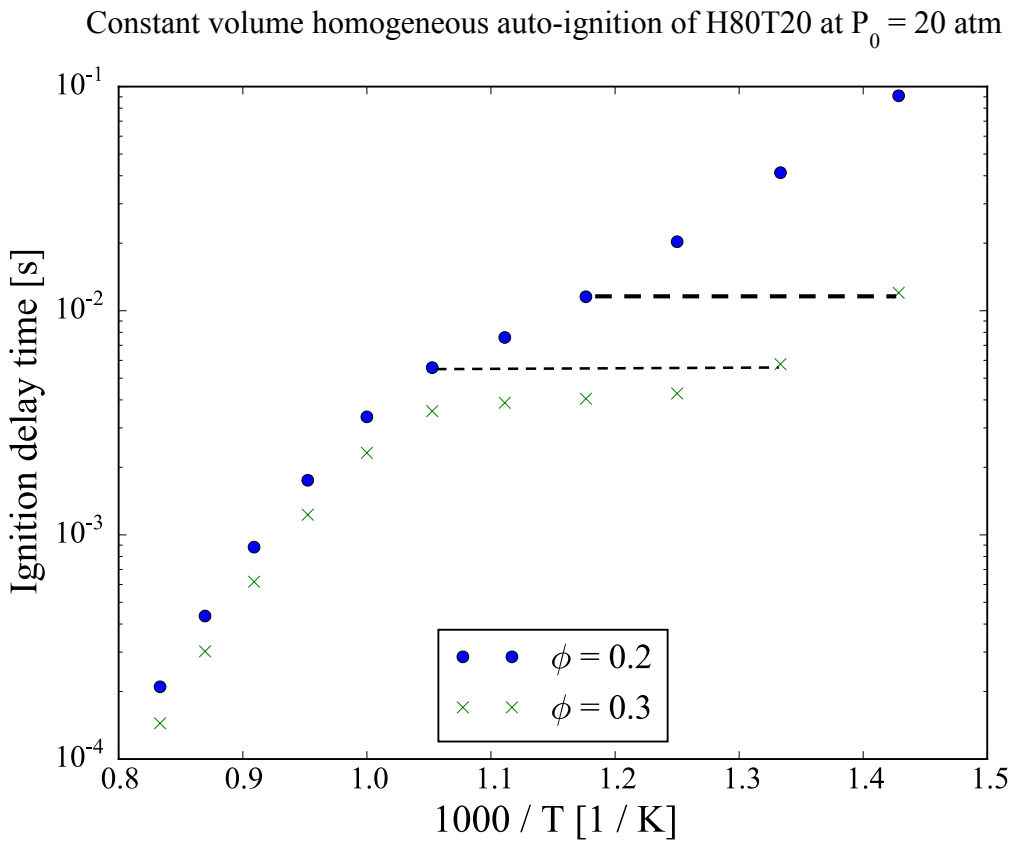

Figure 26: Ignition delay of H80T20 as a function of equivalence ratio and temperature. This trend shown allows for the torque output of the engine to be controlled over a wide range (as equivalence ratio changes) while avoiding negative impacts on combustion timing. Note that temperature is decreasing as the x-axis increases (T^{-1}).

As can be seen in Figure 26, for a constant initial temperature the ignition delay decreases significantly as the equivalence ratio increases from 0.2 to 0.3. Alternatively, ignition delay—closely connected to combustion phasing (CA50) in an HCCI engine—can be held roughly constant as equivalence ratio is increased and temperature decreased (or vice versa) for this fuel. Equivalence ratio scales the engine load (i.e., more fuel results in more torque/power), and ignition timing across these equivalence ratios can be held constant as intake temperature is

appropriately varied. Therefore, the engine load can be varied by way of altering the equivalence ratio and the combustion phasing controlled by way of altering the intake temperature.

The advantage of equivalence ratio sensitivity, i.e., a large allowable fueling range while temperature is varied, is illustrated in Figure 23, which shows the viable operating map for HCCI combustion for fuel H80T20. Comparing this fuel map with the other two PRF fuels illustrated in this section, Figures 21 and 22, it can be seen that fuel H80T20 is nearly a combination of the two PRF operating envelopes. Both of the PRF fuels illustrated did not have the beneficial equivalence ratio sensitivity, and therefore the allowable fueling range is smaller as intake temperature changes. PRF 40 demonstrated higher loads, and PRF 100 lower loads. H80T20 was, however, capable of achieving both load regimes due to the equivalence ratio sensitivity. Because varying equivalence ratio across a wide range of intake temperatures did not affect ignition timing, the amount of fueling could be adjusted to change the engine load without negatively effecting ignition timing that in turn leads to non-viable HCCI combustion. It is now clear that the equivalence ratio sensitivity allows for a wider range of loads to be accomplished as the intake temperature is appropriately varied, whereas fuels that do not have this trend are limited in viable load ranges.

Most of the fuels studied did not exhibit equivalence ratio sensitivity; as a result, they had smaller operating envelopes and therefore lower LTC indices (e.g., PRF 85, see Fig. 22). The fuels that did display this trend (e.g., H80T20, see Fig 23) had the highest LTC indices and were also shown to have a wide range of RON/MON values. In the end, the high LTC index fuels with high RON/MON values pertained to that of fuels H60E40 and TRF4. These fuels should be considered for further research as candidates for next-generation LTC fuels. Designing fuels that can further capitalize on equivalence ratio sensitivity benefits may be the next step to making LTC strategies viable for light-duty automotive applications.

Chapter 3: Infrared Spectroscopy

3.1 Methodology

Infrared spectroscopy operates on the fundamental principle that passing radiation at specific wavelengths into a material excites particular vibration modes acting in response to absorb radiation to some degree. Absorption occurs at varying degrees with frequency dependence on the identity, state, and concentration of a molecule. Simple molecules can experience symmetrical or asymmetrical stretching, scissoring, rocking, twisting, or wagging—each of which correspond to the absorption of a particular amount of radiation at some frequency. In general, the number of degrees of freedom increases as the size of the molecule increases and gets more complex.

Imagining more complex molecules, it is possible to classify these based on the multiple vibrational modes they will uniquely exhibit due to the increased degrees of freedom. Slight changes in chemical structures result in the absorbance to be altered due to the change in vibrational characteristics. Each molecule exhibits unique absorbing frequencies as well as the amount it absorbs.

In Appendix B, an overview of the working principles of Fourier transform infrared spectroscopy (FTIR), data processing techniques, and the two IR collection methods utilized in this current work are discussed. First, Beer's law is derived and the important physical parameters that comprise the final relation are explained; this law is used to convert raw collected light intensity measurements to the more intuitive absorption representation. Second, the working principles of FTIR spectroscopy equipment are briefly introduced. Following this, the two IR collection techniques used in this work—transmission and attenuated total reflectance (ATR)-based spectroscopy—are explained. In these two sections, the theory and relations used to process absorbance data specific to highly absorbing media, e.g., liquids, will be presented in detail. The work of contributing researchers in this field [91]–[103] has been presented in the author's own words. The section is designed to give the reader a sufficient understanding to the extent that they could conduct their own experiments, and provides all the necessary literature resources.

In the following section, the pros and cons of transmission and ATR-based methods are discussed. These two methods were investigated in this work as the IR absorbance spectra with transmission-based data are quantitative, whereas with ATR—to the author's knowledge—is qualitative. Recall the purpose of collecting IR absorbance data is to correlate it to fuel performance ratings by way of chemometrics. As discussed in detail in Chapter 4, determining if one IR collection method works better than the other with chemometrics was an important question to address. Here, the two IR collection techniques are compared and contrasted from a spectrophotometry point of view.

3.1.1 Discussion of ATR vs. Transmission

Typically, the path length is on the order of 10 [μm] for ATR measurements. This is a needed feature for measuring samples with high molar absorption coefficients, such as liquids. Solid absorption spectra can also be measured with ATR, since the beam is reflected. For transmission experiments, measurements at 10 [μm] or less are difficult to achieve since spacers of that thickness are difficult to make and use effectively. As a result, liquids have to be pressed between the transmission slides with no spacer, where the path length is determined from linear extrapolation from known, longer path length data. Alternatively, transmission slides that have a significantly different refractive index than that of the sample can be used to count the interference fringes with no spacer inserted, but this creates a sinusoidal baseline in the data and so this method is typically avoided: see Lacey et al. for this alternate method [91]. Moreover, the film thickness of liquids typically do not permit being pressed together in a repeatable fashion, usually due to variability in the size and number of drops placed on the slides and viscosity differences between samples. In conclusion, transmission-based experiments at low path lengths introduce experimental uncertainty in the final results since the path length has to be linearly extrapolated, otherwise sinusoidal baseline uncertainties are introduced. Even though ATR inherently has a path length with an uncertain definite value—due to the reliance on theory to determine the path length—short path length measurements with transmission are either not feasible or have their own uncertainties from the aforesaid reasons. Therefore, ATR can be a favorable alternative to transmission-based experiments to measure substances with high absorption coefficients, especially if it is a solid.

In addition to the technical feasibility of measuring the IR spectra of liquids and solids using ATR, it is also easy to do in a timely manner. A drop of liquid on top of the ATR crystal, or by pressing a solid sample to ensure complete contact, is all that is needed for sample preparation. This takes little time, and absorptivity experiments can be completed in as little as a few minutes per sample. However, from the authors experience with transmission-based experiments, a single sample can take multiple hours to complete, and slides diminish in quality over time and have to be polished by expert hands to restore them. With this said, it is clear that if a large effort toward measuring many samples are required, ATR is then desired as a complete study could be done in as little as a day, whereas a large transmission-based study may take weeks or months.

3.2 Experimental Design

3.2.1 Transmission

A commercial FTIR, ThermoFisher Nicolet iS10, was used for the transmission-based experiments (450 to 4500 cm^{-1} at 2 cm^{-1} resolution) with a variable path length KBr optical cell.

These transmission-based experiments consisted of absorptivity measurements at multiple path lengths; path length determination will be discussed first. The variable path length transmission apparatus has Teflon spacers of varying thicknesses to insert between the KBr slides—see Figure 27. The apparatus is assembled with no liquid sample inside, followed by measuring the IR spectrum for that of air and then the transmission slides alone. These results displayed in percent transmittance shows the interference fringe pattern. The interference fringe pattern and the associated wavelength values were used to calculate the actual path length.

With the path length determined, the liquid sample is injected into the container using a push-pull method. Two Norm-ject 1 [ml] latex-free (VWR-53548-001) disposable syringes, one filled with the sample and the other empty, could fill the transmission slide apparatus through the access ports. Following this, the absorbency measurements can commence. Prior to each liquid absorbency measurement, the background spectrum was first captured by measuring the air only; the transmission apparatus is then put into place for liquid IR absorbency measurements to follow. Single beam data are saved in comma-separated format (.csv) to later be post-processed.

The apparatus is disassembled after each trial, cleaned with cotton and methanol, and another path length set-up is constructed with a new spacer. The path length studies were done in descending order (e.g., 500, 50, 15, and $\leq$15 μm). Interested readers are directed to Figure 53 in Appendix B to see the effect of path length. Measurements less than 15 μm, if needed for resolving highly absorbing frequencies, were accomplished by using no Teflon spacer—no spacers were available for $\leq$15 μm—with a few droplets of the sample pressed between the slides. After setting up the apparatus with a new path length, the process was repeated.

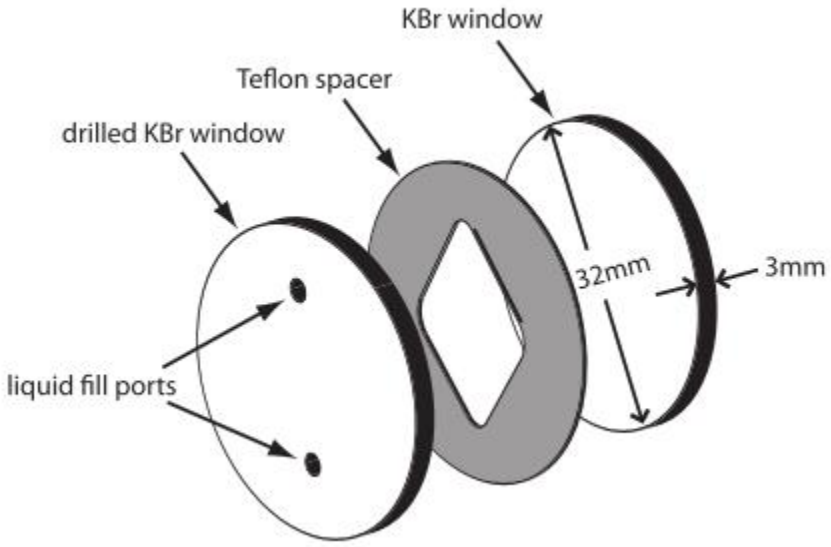

Figure 27: Optical cells used in liquid transmission experiments. Teflon spacers used had thicknesses 15, 50, and 500 μm. [100]

3.2.2 Attenuated Total Reflectance

Here, the same Nicolet FTIR used for transmission-based experiments was also used, but the Nicolet Smart ATR accessory is implemented to replace the transmission-based slide holder. There is a solid sample press (the blue handle), which can be used to ensure proper contact when measuring solid samples—this was rotated to make way for liquid samples. To prepare the fuels that are mixtures of pure components, a Hamilton Gastight syringe was used.

It was detailed in section 3.1.1 that experimental effort by way of ATR is straightforward. For all the fuels measured, Norm-ject 1 [ml] latex-free (VWR-53548-001) disposable syringe transferred a few drops of the liquid sample directly onto the ATR crystal. The liquid was free to

spread and evaporate as the measurement took place. It was found that constructing a barrier with a lid to hold 1 [ml] of liquid, to ensure no evaporation occurred and to prevent the liquid from spreading too thin, yielded consistent absorbency results. The latter method was avoided because it required more of the liquid sample and effort to clean between fuel sample experiments. Additionally, there was risk of the liquid seeping into the crystal and contaminating the ATR accessory. Prior to sample collection, the FTIR was purged with nitrogen to remove water vapor contamination in the data. With the easier "droplet" method utilized, the sample surface was wiped clean with methanol and cotton to prepare for the next fuel sample.

3.3 Results and Analysis

3.3.1 Transmission

The measured optical constants of liquid toluene were compared to the previously published data of Bertie et al. [111] (Figs. 28 and 29) to validate the results. The reference data were measured independently by six spectroscopists in four different laboratories, and showed that the reproducibility of the optical constant technique (the IR data post-processing procedure) for the molar absorption coefficient is within 2.5 % [98]. The measurements made in this work were in good agreement with the published data, though some slight discrepancies were found. The possible cause of the discrepancies will be discussed here.

On average, molar absorption coefficients across the spectrum were measured about 10% higher than in the reference data. The path lengths used in this work for absorption measurements, in addition to transmission slides used, are likely the cause for the variation. Bertie et al. measured absorptivity at more path-lengths (finer resolution) in addition to path lengths greater than 500 μm where appropriate [98] . Additionally, Bertie et al. performed the long path length measurements with NaCl, CaF2, and KBr slides to more accurately determine linear absorption coefficients at the anchor points [98] . In addition, Bertie et al. measured at a resolution of 1 cm^{-1} whereas this work is at 2 cm^{-1} [98] . The combination of reproducibility of the optical constant method (2.5%), path lengths used, and multiple transmission slides considered for long path lengths could account for the overall discrepancy found between this work and that of Bertie et al. [98].

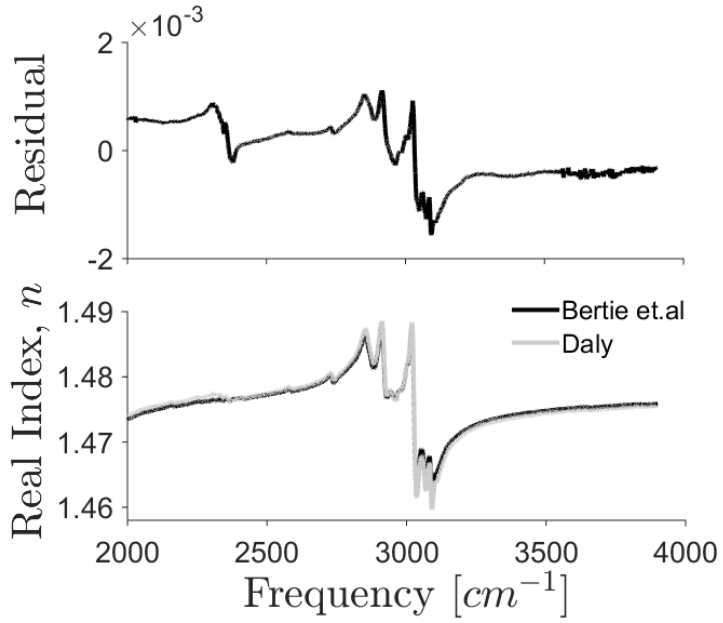

Figure 28: Measured real refractive index of liquid toluene compared with published data [98]. The residual is defined as $n_{Daly} - n_{Bertie}$

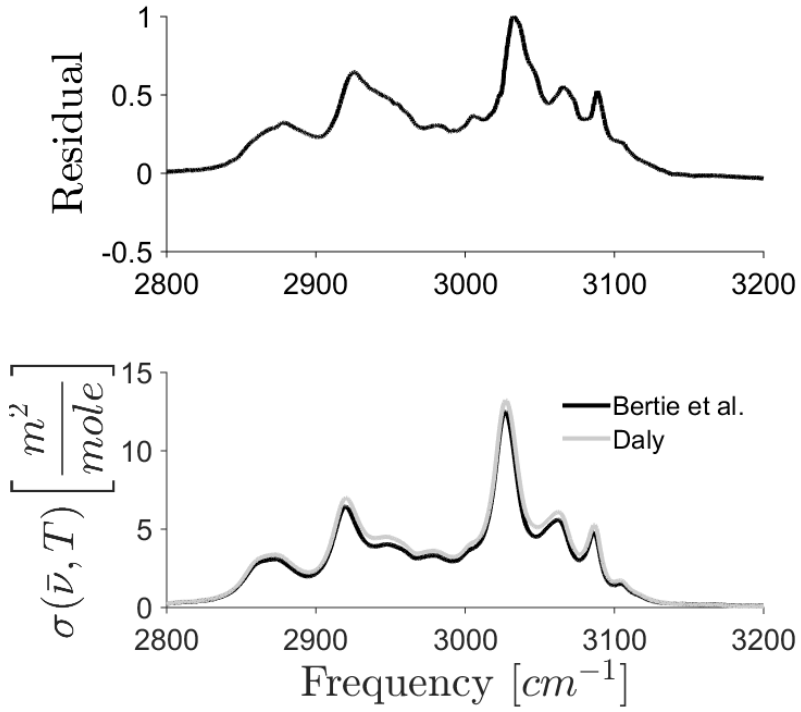

Figure 29: Molar Absorption Coefficient of liquid toluene compared with published data [98]. The residual is defined as $\sigma_{Daly} - \sigma_{Bertie}$

IR spectra of neat components and the FACE gasolines, collected via transmission-based methods, are presented below. The entirety of the results are found in Appendix E. Analyzing Figure 30, the intrinsic absorbing capabilities of the neat components measured can be realized. Now considering FACE A, depicted in Figure 31, interesting insight into blending effects of ethanol are seen. An approximate representation of the absorption spectrum for FACE A with no ethanol can be achieved by a blend of iso-octane and n-heptane; these components are the majority that constitutes FACE A [104]. Introducing ethanol in FACE A should result in absorbing frequencies intrinsic to ethanol to be exhibited in the mixture. This does occur, as seen in the major absorbing frequencies around 1100 and 3300 [cm^{-1}], although at some frequencies interesting synergistic blending effects occur. At around 2950 [cm^{-1}], with a small introduction of ethanol (e.g., 10% and 15%), the absorptivity increases. The magnitude of this absorptivity is higher than iso-octane, n-heptane, or ethanol alone. Therefore, with the introduction of ethanol, there is no linear combination of the three individual spectra that results in what was observed—this is atypical of the usual linear blending of hydrocarbon absorbance. Non-linear blending effects of alcohols with hydrocarbons have been observed and explained in literature [105]–[107]. In general, it is suggested that alcohols interact with hydrocarbons in various ways by means of van der Walls forces, and the molecular structures formed (e.g., double-bonded dimer, linear polymer, water-like structure, etc.) are a function of the alcohol concentration. These various interactions thereby alter the original molecular structure of the hydrocarbons and therefore their absorbing characteristics. This may explain the synergistic blending effects, i.e., the combined effects being greater than the sum of their separate effects, seen with the FACE A mixture, as well as other FACE gasolines that exhibit similar trends with ethanol introduction, since ethanol effects are concentration-dependent. These spectroscopic trends should be investigated further in future work.

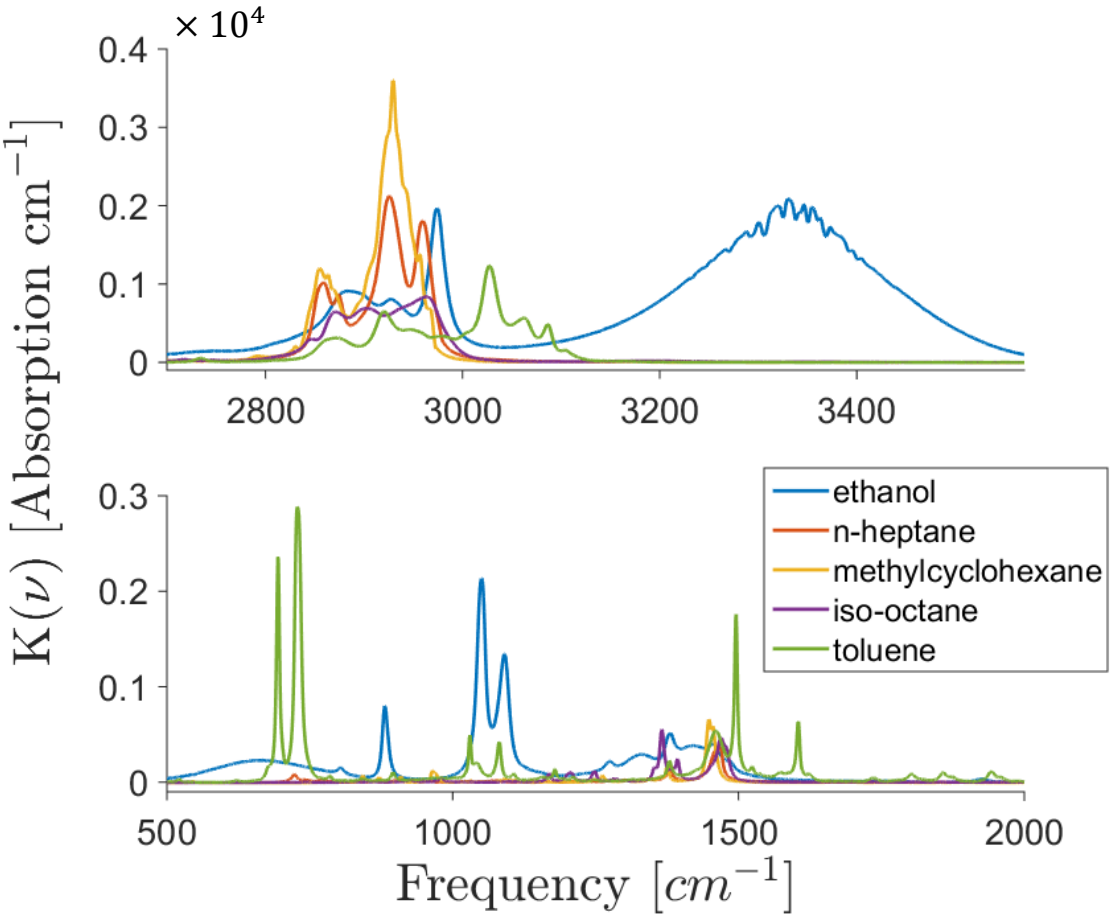

Figure 30: Linear Absorption Coefficient Spectra of neat component fuels. Note both y-axes are multiplied by 10^4 and x-axes are continuations of each other.

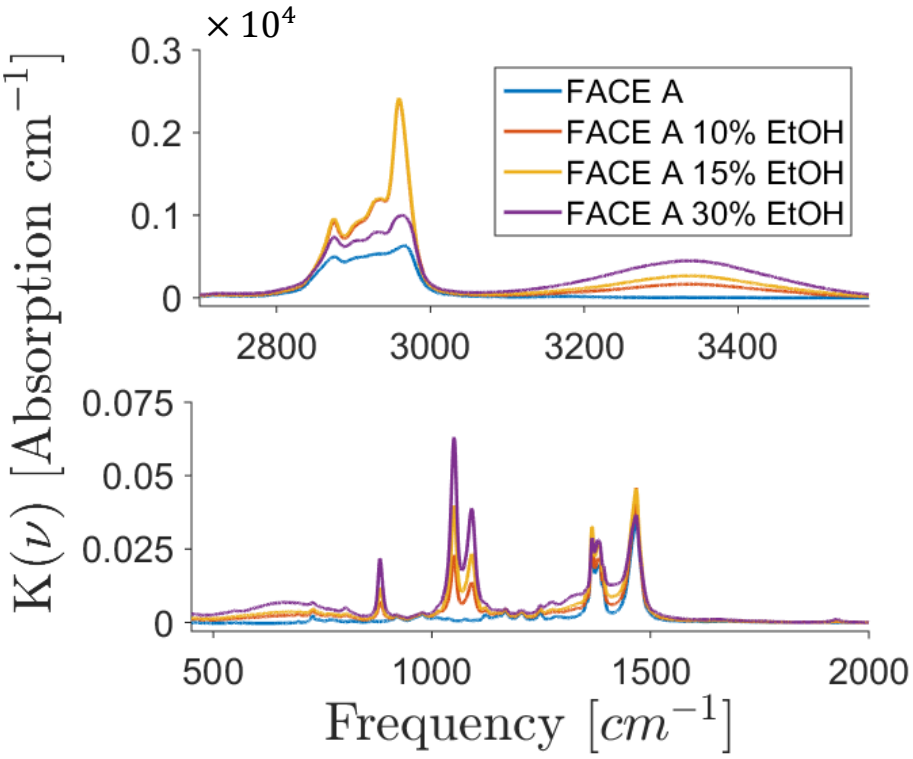

Figure 31: Linear Absorption Coefficient Spectra of FACE A and various blends with ethanol. Atypical synergistic blending effects caused by ethanol introduction are best seen around 2900 cm^{-1} for ethanol concentrations of 10 and 15%.

3.3.2 Attenuated Total Reflectance

Over 200 fuels were procured by the author and measured on the ATR equipment, 34 of which were pure hydrocarbon components. The extensive lists of fuels studied can be found in Appendix D. Figure 32 shows some of the pure hydrocarbon absorptivity spectra collected in this work. All the hydrocarbons shown here represent the aromatic class, and shown here to only illustrate a sub-set of the hydrocarbons IR absorption spectra measured.

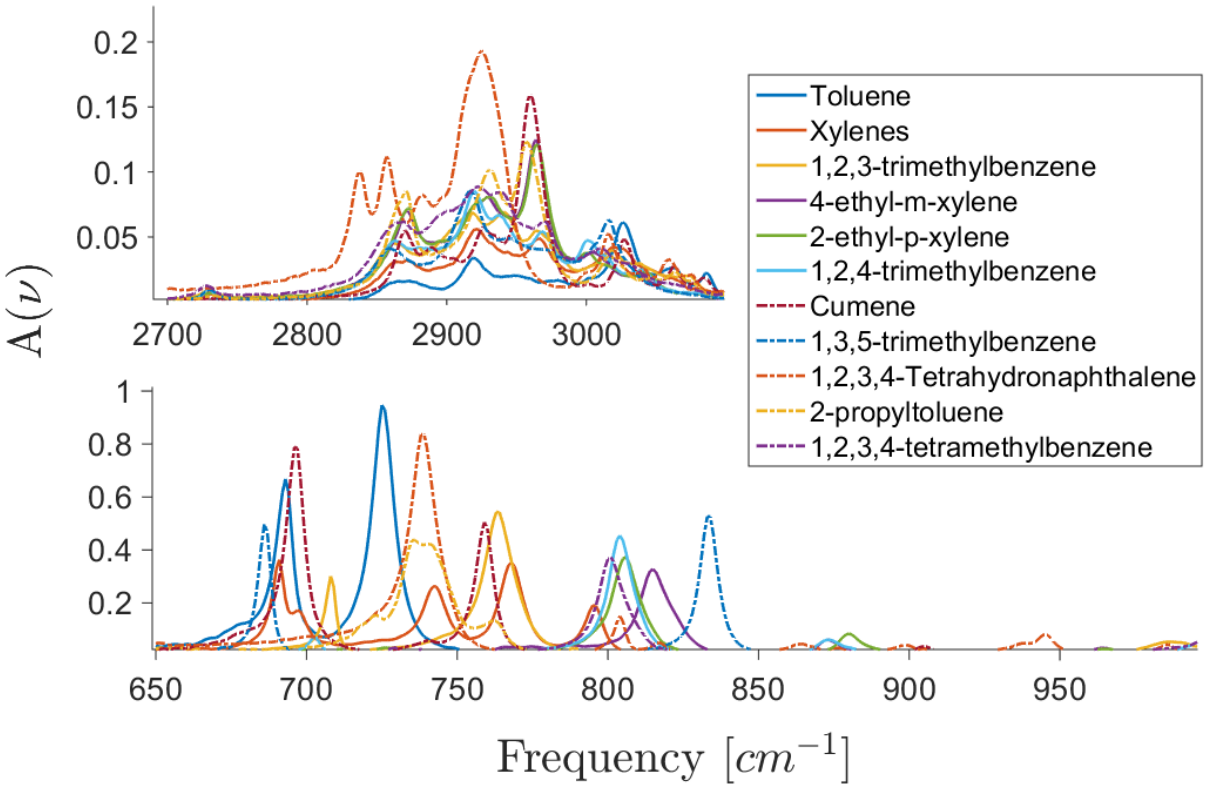

Figure 32: Example sub-set of absorption spectra measured by the author. Aromatic hydrocarbons considered in this work are shown here. Note that all ATR results shown are uncorrected for any effects, e.g., variable path length and dispersion.

Chapter 4: Chemometrics

The goal of this work is to collect IR absorption spectra for fuels of research interest, calculate the LTC index for each fuel, and use chemometric analysis to create a model. This model could then take an unknown fuel's IR absorption spectrum as an input and predict the LTC performance of that fuel.

First, the basic working principles of two multivariate statistical techniques are explained (Sec. 4.1). In this work, these methods correlate IR absorbance spectra to the fuel performance ratings RON and LTC index. Secondly, the experimental section (sec. 4.2) investigates how the statistical models are affected by various parameters: multiple studies were conducted to inform this. The studies conducted are explained in detail, including how the statistical models were developed, and the findings of the studies are discussed. Lastly and most importantly, the LTC index correlation is created and discussed.

4.1 Methodology

The reader may recall from the earlier section 1.3.2 the discussion of the industry products IROX [38] and Zeltex ZX-101C [39]. These devices are capable of measuring a fuel's IR absorption spectrum and predicting various fuel-manufacturing specifications such as the octane rating, density, Reid vapor pressure, etc. These products do so with a built-in FTIR and chemometric model developed for these particular applications. A short introduction to the statistical theory that allows chemical data to be correlated to a desired variable of choice will be provided. This overall process is known as chemometrics [108].

Chemometrics is the science of extracting information from chemical systems by data-driven means. Statistical methods, such as principle component regression (PCR) and support vector machine regression (SVMR), extract patterns from the chemical data and attempt to create a correlation to an explanatory variable of choice. In this work, IR absorption spectra are the chemical data, and the explanatory variable are the fuel performance ratings RON and LTC index. This section will cover the theory of the two statistical methods investigated: Principal Component Regression (PCR) and Support Vector Machine Regression (SVMR).

4.1.1 Principal Component Regression (PCR)

Principal component regression, or PCR, is a linear method that identifies patterns in multi-dimensional data and correlates these to an explanatory variable, i.e., a variable that may explain the patterns found in the chemical data. In short summary, the fundamental mathematics of this technique is eigenvector decomposition of the datasets' covariance matrix. Readers that are unfamiliar with multi-dimensional statistics or linear algebra techniques are directed towards the tutorial of PCR by Smith [109], which offers an introductory explanation of statistical covariance and eigenvectors; a more rigorous mathematical explanation of the methodology can be found in the book by Anton et al. [110]. MATLAB software was used for PCR. A tutorial on how to use MATLAB's PCR tools for chemometric applications can be found online [111]. Since PCR is relatively easy to understand, a brief summary of the concepts is given in the order of operations that take place for the method.

Covariance is a measure of the standard deviation between two arrays, and can be seen as

$$cov(X,Y) = \frac{\sum_{i=1}^{n}(X_i - X_{mean})(Y_i - Y_{mean})}{n-1} \tag{4.1}$$

where *n* is the number of variables within the dataset represented by arrays X and Y. This does not yield a meaningful number; instead, the sign is only of importance: a positive sign indicates that *X* and *Y* increase together, a negative sign indicates that one dimension increases while the other decreases, and a value of zero represents no correlation between the two dimensions.

For high-dimensional datasets, the covariance values are organized in a symmetrical matrix about the diagonal. This is mostly for mathematical convenience for the following operations. An example is visualized here with a three-dimensional dataset, represented by *x*, *y*, and *z*.

$$C = \begin{pmatrix} cov(x,x) & cov(x,y) & cov(x,z) \\ cov(y,x) & cov(y,y) & cov(y,z) \\ cov(z,x) & cov(y,z) & cov(z,z) \end{pmatrix} \tag{4.2}$$

With the covariance matrix, eigenvalue decomposition can be performed on the covariance matrix to find the eigenvectors, i.e., the principal components.

Eigenvalue decomposition is laborious for matrix dimensions greater than three. The methodology for this is beyond the scope of this work; interested readers are directed to Anton

[110] for decomposition techniques. The result of this process will be illustrated with a two-dimensional dataset. Figure 33 shows a random dataset plotted after subtracting the mean value. The eigenvectors of the covariance matrix are overlaid with the data, and oriented in the directions of maximum covariance. As can be seen, one eigenvector is aligned with the primary linear relationship that the data exhibits, whereas the other is aligned with the small spread in the data. Each eigenvector has an associated eigenvalue, where the magnitude of the eigenvalue represents the amount of variance of the data in the direction of each eigenvector. The eigenvector aligned with the primary linear relationship has a large eigenvalue, as it explains the majority of the trend the data exhibits, while the other eigenvector has a smaller eigenvalue.

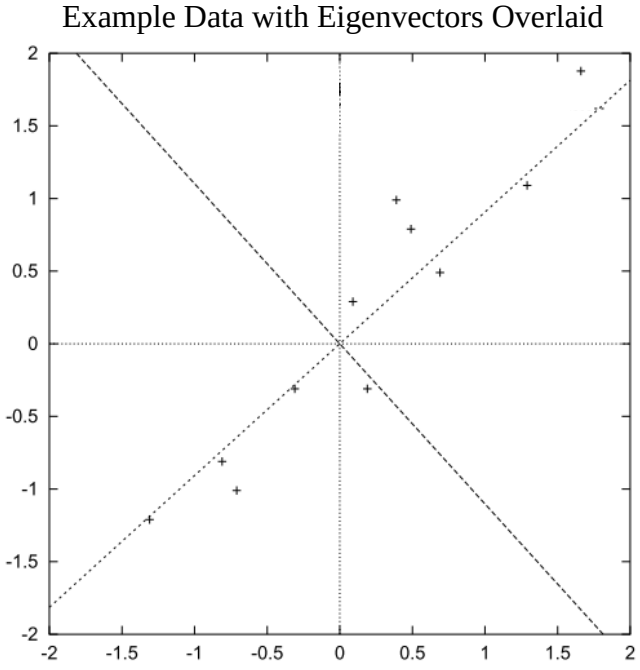

Figure 33: Example data set that has been normalized (mean subtracted) with eigenvectors overlaid to visually show the maximum covariance of the data. [109]**.**

N-dimensional datasets have N eigenvectors and corresponding eigenvalues. The eigenvalues for each eigenvector are a weighting of sorts that explain the variance in the data, where the eigenvalues represent the spatial direction of the variance. This information can then be used with a least-squares regression method to correlate the data to an explanatory variable (see Canale et al. [112] regarding least-squares regression).

4.1.2 Kernel-Based Support Vector Machine Regression (SVMR)

Support Vector Machines (SVM) are learning algorithms capable of modeling complex non-linear datasets by way of user-defined kernel functions. This statistical learning algorithm was invented by Vapnik in 1979, though has not received attention until recently [113]. According to Devos et al., this technique has been recently introduced to chemometric analysis and has proven to be powerful in IR spectra classification problems [114].

To summarize the governing theory for this complex statistical technique would require an exhaustive amount of effort, even if the goal were to be a mere basic introduction. The reader is guided to the useful work of Burges [113], for an in-depth understanding of the theory and supporting mathematics. The basic idea and systematic approach are presented here, instead, guided by the helpful work of Devos et al. [114] and Luts et al. [115]. These sources provide useful graphics to illustrate the principles, and apply the methodology to multiple types of problems. For example, the robustness of this technique is demonstrated in brain tumor classification, near-IR absorbance spectrum classification, and more. First, the basics of SVMs is introduced following by the kernel methods that make nonlinear analysis possible.

In order to classify patterns in data sets, hyperplanes are mathematically constructed to define decision boundaries. That is to say, the decision boundary separates data that fall into two defined classes: one class being data that has a positive influence on the descriptor variable, and the other proving to have a negative influence on the descriptor variable. For example, if certain data prove to have a higher "performance value", they will be separated from data that has "low performance". The idea here is to map the original data points from the original dimensions, called input space, into a high-dimensional—or even infinite-dimensional—feature space to make the problem feasible. Take, for example, the data set provided in Figure 34. As can be seen, the two concentric circles are not linearly separable; i.e., the positive and negative influential data points cannot have a linear decision boundary drawn between them in any fashion. However, mapping these data into a higher-dimensional feature space by way of a mathematical transformation allows a clear decision boundary, or hyperplane, to be drawn. With

the data "stretched" into a higher-dimensional form, the next step in the SVM process is to determine a hyperplane that can now separate the data, as seen in Figure 35 (from [116]).

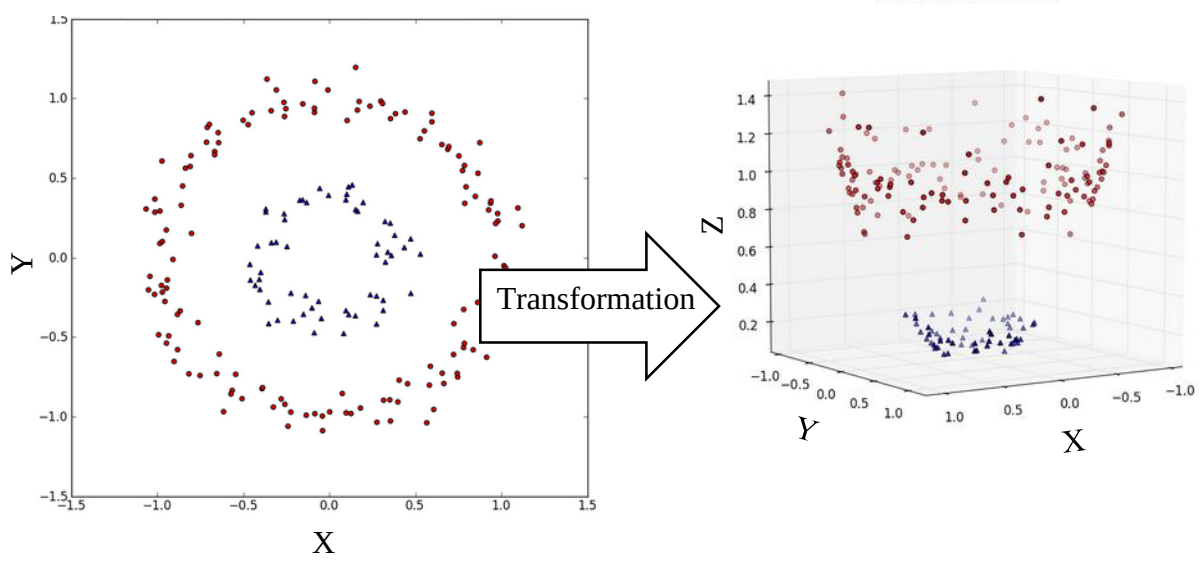

Figure 34: Non-linear data set (left) and transformed data set (right) into a higher dimension. [116]

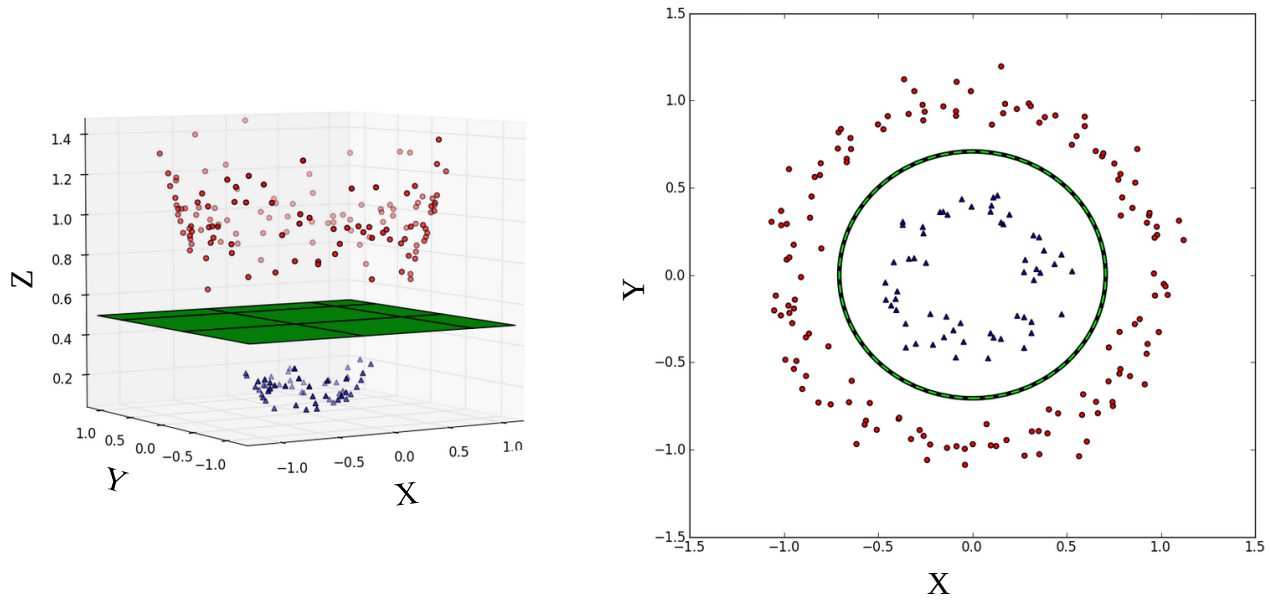

Figure 35: Transformed data set and the hyperplane constructed via SVM to separate the data set (left) and shown after transforming back to the original dimensional space (right). [116]

Kernel functions provide a means to transform data into a higher dimension prior to the SVM algorithm that determines the hyperplane decision boundary. In the previous section, Figure 34 illustrated that the concentric circles were not linearly separable. However, after the mathematical transformation, i.e., the kernel function, a hyperplane (via. SVM) could then be constructed that separates the classes in data (Fig. 37). The mathematical form of the kernel function can be linear or highly nonlinear. There are many types of kernel functions used in literature [113]–[115], some of which are

- Linear: $K(x,y) = x^T y$ (4.6)
- Polynomial: $K(x,y) = (\tau + x^T y)^d, \tau \geq 0$ (4.7)
- Gaussian: $K(x,y) = \exp\left(-\frac{||x-y||^2}{2\sigma^2}\right)$ (4.8)
- Square Root Gaussian: $K(x,y) = \exp\left(\frac{\sqrt{(x-y)(x-y)^T}}{2\sigma^2}\right)$ (4.9)
- Sigmoid: $K(x,y) = \tanh(ax^T y + b)$ (4.10)

where *x* is the data, *y* is the explanatory variable we wish to correlate to, and the other constants τ, σ, a, b, d, are user-defined kernel parameters. With this brief introduction to kernel functions complete, the methodology to determine what kernel function and kernel parameters to use, in addition to a meta-parameter required for the SVM hyperplane decision boundary, will be discussed. This leads us to SVMR parameter optimization.

Multiple parameters, known as kernel and meta parameters, dictate the hyperplane boundary complexity and therefore the performance of the statistical model. Kernel parameters vary the way in which the data is transformed into a higher-dimensional space, and the meta parameter affects the error tolerance used for the hyperplane decision boundary by the SVM algorithm. Guided by the work of Devos et al. [114] and Luts et al. [115], a cross-validation study, i.e., a

parametric study of kernel and meta-parameters, is performed. Figure 36 visualizes an example of such a cross-validation study from Devos et al. [114].

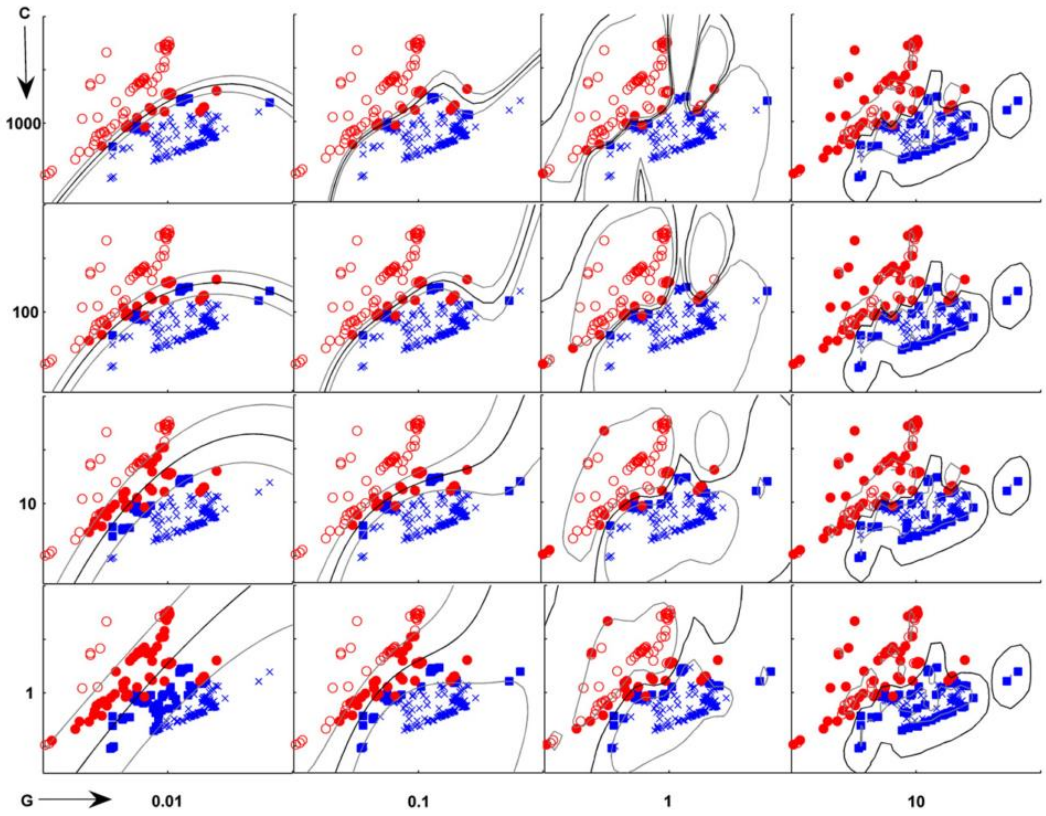

Figure 36: Cross-validation study, from Devos et al. [114]**, showing the effect of the hyperplane decision boundary (black line) as a function of meta (C) and kernel-parameter (G). In this work, the kernel used is the Gausian function and the kernel parameter G is $1/\sigma^2$ in relation to eq. 4.8. The light black lines surrounding the darker black line (decision boundary) indicate the error tolerance dictated by the meta parameter, C.**

As can be seen in Figure 36, the tunable parameters affect the decision boundary. With the cross-validation methodology, Devos et al. [114] and Luts et al. [115] then performed a leave-one-out validation study, where a random subset of the data is removed prior to building the statistical model. The statistical model created utilized only a fraction of the original data, where subsequently it was applied to predict the performance of the data that were left out. When the minimum error is determined from this study, those kernel and meta-parameters are selected. This process is repeated for several kernel functions until the optimal kernel function, kernel parameters, and meta parameter are determined for the dataset at hand. At this point, the kernel-based SVMR is complete.

4.2 Experimental Design

Recall the goal of this work is to collect a training data set, i.e., collect IR absorption spectra for and calculate the LTC index fuels of research interest, and use chemometric analysis to create a model that can predict LTC index of an unknown sample. Three problems needed to be addressed in order to accomplish this goal:

1. Which spectroscopic method is better for the statistical model—transmission or ATR? Recall ATR yields qualitative results and could potentially not inform the chemometric models well enough compared to the quantitative transmission-based method.
   Addressed in section 4.2.1
2. Can the IR spectra of a blended fuel, e.g., a mixture of n-heptane and iso-octane, be represented by the linear combination of the pure component fuel spectra? In other words, is it required to measure the IR absorbance spectrums of fuels that are simple mixtures of pure components, or can the linear combination of spectra be used?
   Addressed in section 4.2.2
3. Which and how many fuels are required to build the chemometric model? Each fuel's IR absorption spectrum and LTC index, which can require a significant effort to procure, are necessary to inform the chemometric model.
   Addressed in section 4.2.3

Guided by the findings of the above questions, a final chemometric model predicted the LTC index of the FACE fuels (Sec. 4.3.1).

4.2.1 Transmission vs. ATR-based models

In order to answer these questions, the initial effort of this work was to determine if ATR methods suffice to inform the chemometric model. This was of interest as collecting fuel samples in this manner is easy, takes little time, and only a few drops of the fuel sample are required. First, a few fuel absorption spectra were measured using transmission and ATR methods. Fuels initially considered were n-heptane, iso-octane, toluene, and ethanol. These pure component fuels were simulated computationally to find the LTC performance index, as discussed in Chapter 2. With the absorption spectra and performance variables determined, some initial chemometric analysis could take place. This initial analysis used PCR as the

statistical method, as it was easily understood and required no parameter optimization in contrast to SVMR. However, even though PCR is in essence a simpler technique, it should still inform whether transmission or ATR-based data may be preferred in future chemometric analysis.

In order to investigate the effects of ATR or transmission-based data on the chemometric models, a statistical model was created with both methods. The statistical models were created with pure components only—i.e., informing the model with four absorption spectra and corresponding LTC indices—and using this model to predict the LTC performance of blends of these pure components. The blends of pure components were computationally simulated for the LTC index, so that the actual LTC performance was known. Absorption spectra used were linear combinations of the pure components, where the relative proportions of each spectrum known by the mole fractions of the fuel mixture. ATR-based spectra were not post-processed for any dispersion or path-length effects; the data used with the chemometric model was therefore qualitative and in terms of absorbance. Transmission spectra were post-processed for path length and dispersion effects; these data should be quantitative and are in terms of linear absorption coefficient. The results of this study are found here.

Building a chemometric model for the LTC-index (CR=9.5) with ATR and transmission-based absorbance spectra gave the same coefficient of determination (R^2) and standard deviation (σ) to within seven significant figures; hence no discernable difference between ATR and transmission-based results. Since ATR methods gave the same statistical performance as transmission-based, the remainder of the work utilized ATR methods as it is simpler to implement. This finding should hold true if and only if the same ATR equipment is used throughout the study. Each instrument is different for one way or the other, and therefore measured absorbance spectra will vary.

Guided by the findings of this study, question #2 was addressed and is presented in the following section.

4.2.2 Linear Combination Spectra

Guided by literature, it was suspected that ethanol would cause non-linear effects on the absorption characteristics when blended with hydrocarbons. This was also discussed in section 3.3.1, where this phenomenon was experimentally observed with the FACE gasolines containing ethanol (FACE fuel spectra are found in Appendix E). Therefore, a linear combination spectrum for simple fuel blends with ethanol could most likely not be used. This was first investigated by measuring the FACE gasoline IR absorption spectra via transmission-based methods and analyzing the effects of ethanol introduction—to the author's knowledge, the data itself is of research interest as quantitative spectra has not yet been published for these research fuels. As can be seen with the FACE fuels, to spectroscopically represent fuels with ethanol the fuel's actual absorption spectra needs to be measured.

Since the chemometric models to follow in this work were to be built using ATR-based data, dispersion effects, i.e., variable index of refraction as a function of light frequency, in the ATR data were of concern (see Appendix B2.2.1). Pure components have different refractive indices and therefore different path lengths as a function of light frequency. Blends of these components now have a "blended" refractive index that could cause minor phase shifts in the data; therefore, assuming a linearly blended spectra would not capture these effects. To investigate this, ATR absorbance spectra of mixtures of the pure component spectra were measured. Two chemometric models, powered by PCR, were both built with pure components *and* pure component mixtures. One model used linearly blended absorbance spectra for the fuel blends, and the other model used the actual fuel IR spectrum for each fuel blend. The results of this study are preseted here.

Figures 37 and 38 on the following page show that regardless of the fuel type the actual spectra are required to inform the model due to non-ideal ATR effects and non-linear blending phenomena. For the PCR model, adding linearly combined spectra for the blended absorbance spectra did not add any intrinsic information to the model. However, when adding the actual blended fuel spectrum to the model, the fuel performance metrics were perfectly predicted, as expected. As a result, actual fuel IR spectra of blended mixtures were used. With question #2 addressed, question #3 was investigated and is covered in the following section.

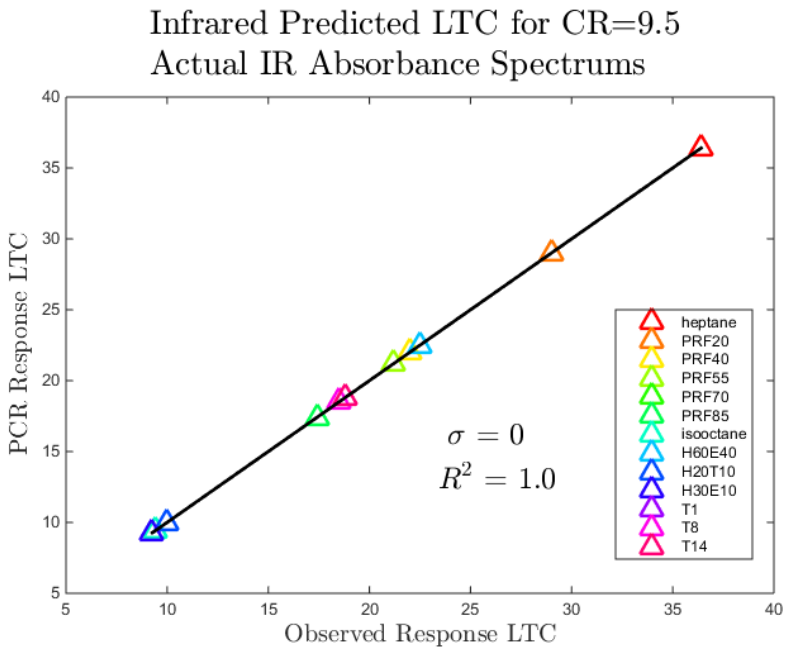

Figure 37: LTC Index (CR=9.5) prediction using actual IR absorbance spectra to build the chemometric model.

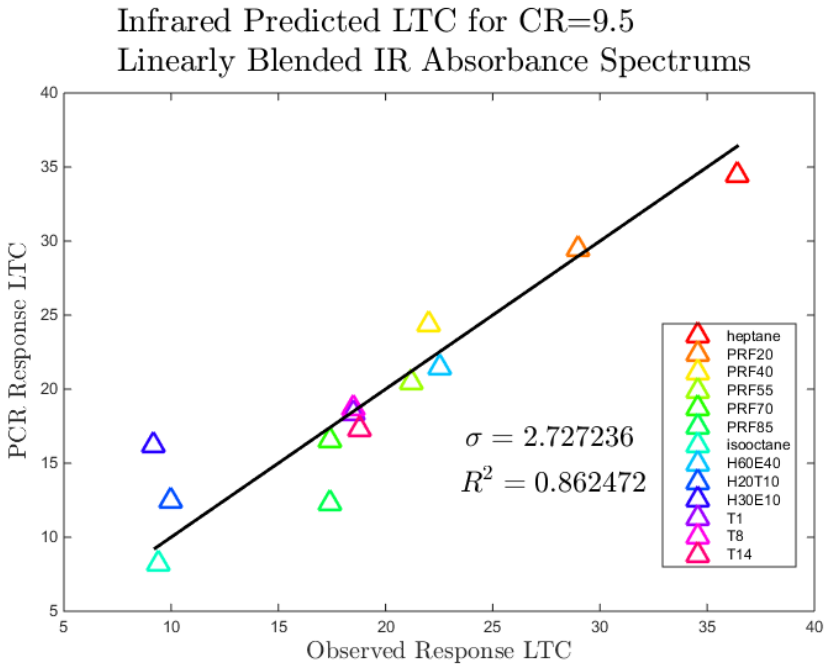

Figure 38: LTC Index (CR=9.5) prediction using using linearly blended IR absorbance spectra to build the chemometric model.

4.2.3 RON Predictions for CRC FACE Gasolines

In an effort to determine which fuels were needed to inform the chemometric model, i.e., what functional groups are required to accurately predict LTC performance, a chemometric study was done with research octane number (RON) as the fuel performance rating. RON is known for many pure component hydrocarbons and the FACE gasolines. Calculating the LTC index as the fuel performance rating, which requires time-consuming computational studies, can be avoided in favor of using RON to determine the necessary components. In addition, at this time there are chemical mechanisms available for only two FACE gasolines and selective pure component, as discussed in Section 1.7, which limits the types of fuels that can be investigated. With the availability of RON for the fuel performance metric, only the absorbance spectra have to be determined from FTIR measurements, and therefore a robust dataset can be analyzed. With this study, many fuel's IR spectra can be determined quickly with ATR to build a chemometric model. Following that, the performance of the FACE gasolines can be predicted with this model and validated against the true RON values.

For this work, a novel chemometric model was created wherein only pure component fuels informed the statistical model to predict real gasoline fuel RON. As discussed in section 1.3.2, Kelly et al. [36] created a RON model with great success, but informed the model using a large dataset of real gasoline fuels; this yielded great statistical performance. However, in relation to LTC performance (i.e., LTC index), at this time only pure components and mixtures of pure components are capable of being computationally simulated for the LTC index. Therefore, real gasoline data sets cannot inform the model with the exception of the few FACE gasolines with surrogate mechanisms available. As a result, the effect of only using pure components to predict RON of the FACE fuels was investigated to determine the feasibility of this technique. From this, pure component fuels that increase the accuracy of the RON prediction for the FACE fuels can be targeted for the LTC index model.

The major constituents found within the FACE gasolines drove fuels targeted for this study. The CRC report [104] has a succinct summary of these primary compounds found in the FACE gasolines. Nearly the entirety of these fuels were procured and the IR absorbency spectra were measured with the ATR equipment.

Many chemometric models were created by way of PCR and SVMR, that is, the fuels used to inform the chemometric model were selectively added to determine the relative impacts each fuel has on the predictive capabilities. This was accomplished by first creating a baseline statistical model that incorporated no pure component spectra. Spectra were collected for n-heptane, iso-octane, toluene, ethanol, methylcyclohexane, and 1-hexene, both individually and in blends based on mixtures given by Truedsson [23], Foong et al. [45], and Perez et al. [46]. The "baseline" fuel list primarily consists of every single fuel mixture (binary and above) from the three previously stated sources, among a few other mixtures: the list is found in Appendix D. This initial model was used to then predict RON values for the FACE gasolines. Following that, the statistical model would then be "re-built" with the addition of one new pure component. The FACE gasolines are then predicted again, and the new results compared to the baseline results—this was repeated for all pure components. The results are discussed next.

The results are depicted in Figure 39. The y-axes represents the difference between the baseline model and that with a pure hydrocarbon component added. Mixtures of five pure components represent the baseline, and as a pure component is selectively added (moving along x-axis) the chemometric model is then informed by the baseline fuels in addition that single pure component. Noting the first x-axes point, the "baseline" y-axes value is zero as the baseline chemometric predictive values has been subtracted out of the data. This way, when a pure component is selectively added to the baseline model the *sensitivity* of adding the pure component is represented in the y-axes. For example, looking at the third x-axis value (2-methylbutane), the model is informed by the baseline fuels and *only* that hydrocarbon component (i.e., the previous fuel at the second x-axis value (n-heptane) is not included). A large y-axis value indicates that the chemometric model prediction is significantly effected when a hydrocarbon component is added to inform the chemometric model. A small y-axes value indicates that adding a hydrocarbon component has no effect on the chemometric model prediction. The y-axes has no indication as to the accuracy of the model, only if the predictions significantly changed. The legend labels "FACE_X_%", which indicates the "X" FACE gasoline with "%" volume percent ethanol blended into that fuel.

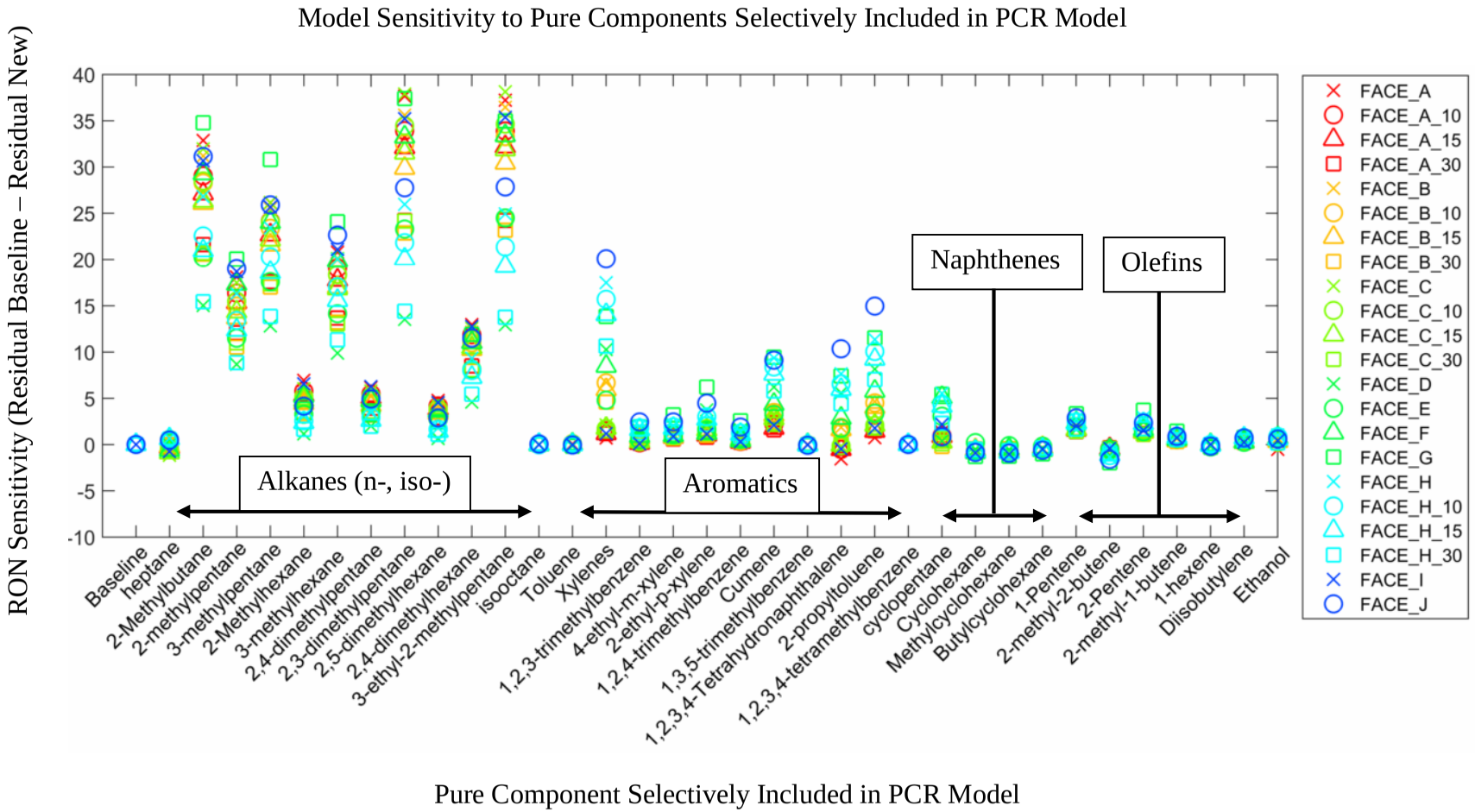

Figure 39: RON sensitivity plot versus pure components selectively added. Y-axes represents the difference in the residual from the baseline. The residual is defined as the actual RON less the predicted RON value. Therefore, the y-axes is the residual of the baseline less the residual of the baseline with the new pure component added to the model ($RON\ Residual_{baseline} - RON\ Residual_{baseline+pure}$). The larger the y-axes value, the larger the effect the pure component has on the RON prediction for FACE gasolines. X-axes represents the pure component used with the baseline. FACE_X_# indicates FACE gasoline X blended with # vol% ethanol.

The results indicate that the alkane and aromatic classes are not represented adequately by iso-octane and toluene, respectively. It appears that including the information of nearly any of the other alkanes considered here should improve the accuracy of the model, and as for the aromatic class xylene and 2-propyltoluene should be considered.

At this point, another question was raised: *will adding additional pure components to the model add redundant information*? For example, if two important pure components representing the alkane class are added to the model, are there diminishing returns for predictive performance? That is to say, could we add one pure component of the two and still get the same predictive performance of having both? Figure 40 illustrates this effect. Instead of adding only one pure component to the baseline and analyzing the sensitivity of the model to that component, the pure components here are added to the baseline additively as we move along the x-axis, as seen on the following page.

Figure 40's x-axes represents adding the hydrocarbon components additively, e.g., the third x-axes point (xylene) represents that the chemometric model is built with the baseline, the second x-axes point (2,3-dimethylpentane) and xylene. The y-axes represents the accuracy of the model, where the magnitude is the actual RON value of the FACE gasolines less the predicted RON value. Unlike the previous study, the y-axis here represents the model performance. Refer to Table 3 for a summary of the predicted and actual RON values.

Clearly, it is possible to reach predictive performance of ± 10 RON by adding two additional pure components that represent the alkane and aromatic class, instead of requiring all of the pure components to follow. Here, 2,3-dimethylpentane represented the alkane class and xylene(s) that of the aromatic class. Adding additional pure components after these first two—including pure components which show a high RON sensitivity effect (Fig. 39)—demonstrates that the predictive capabilities of the model has diminishing returns. This may indicate that it is instead feasible to select a few components from each class to get adequate predictive performance. There also appears to be some leeway as to which pure components to selectively use—this would like to be investigated in future work. For example, instead of using 2,3-dimethylpentane, to which the model showed a large sensitivity (Fig. 39), it could be replaced by a few 2-methylalkanes for the same effect. Figure 41 illustrates this principle.

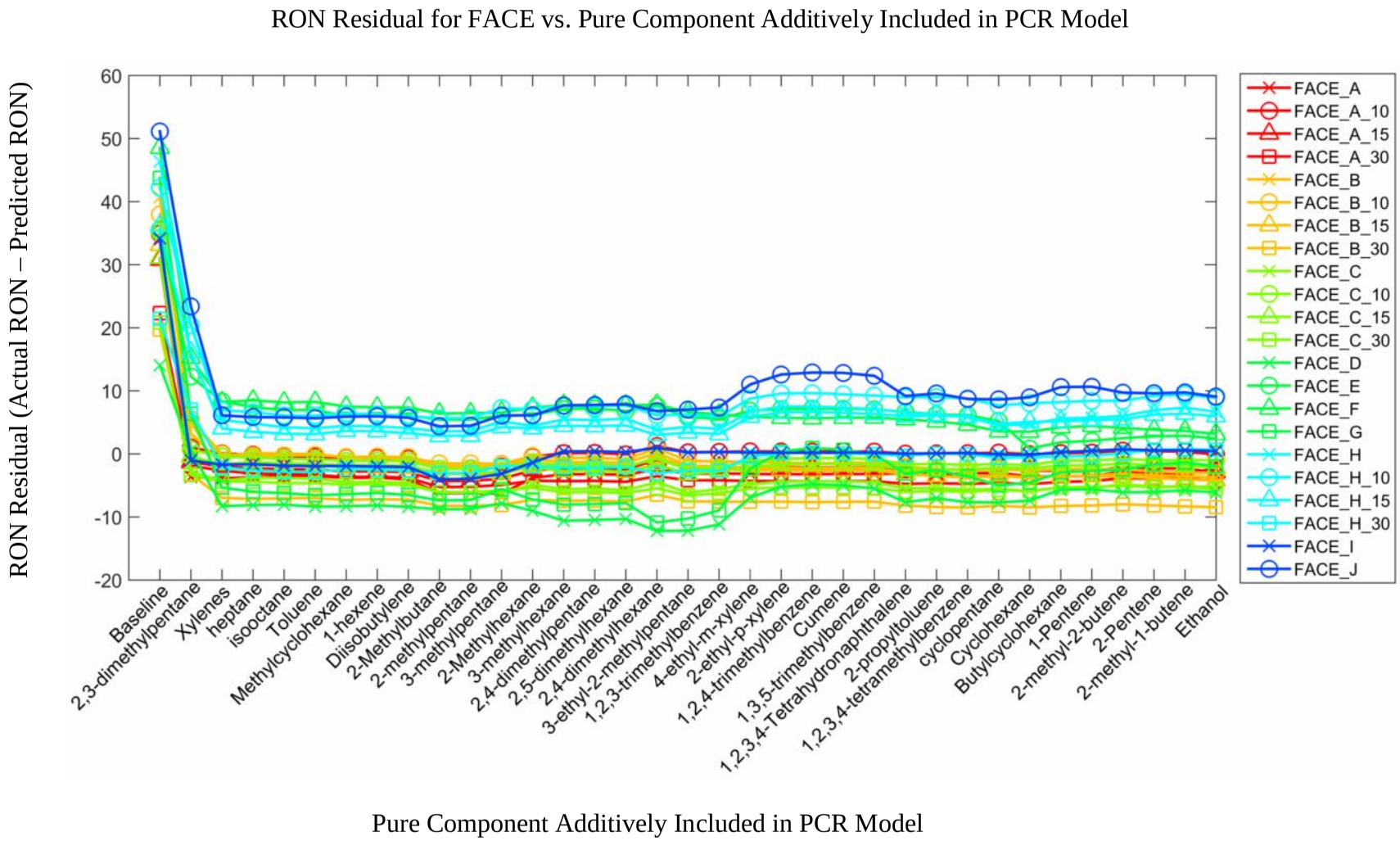

Figure 40: FACE Fuel RON residual versus pure components in the model. For example, at the x-axis point "xylene", the chemometric model has been built with baseline + 2,3-dimethylpentane + xylene. This shows that with addition of a few pure components that are not represented in the baseline yields in significantly better RON predictive performance for real gasolines.

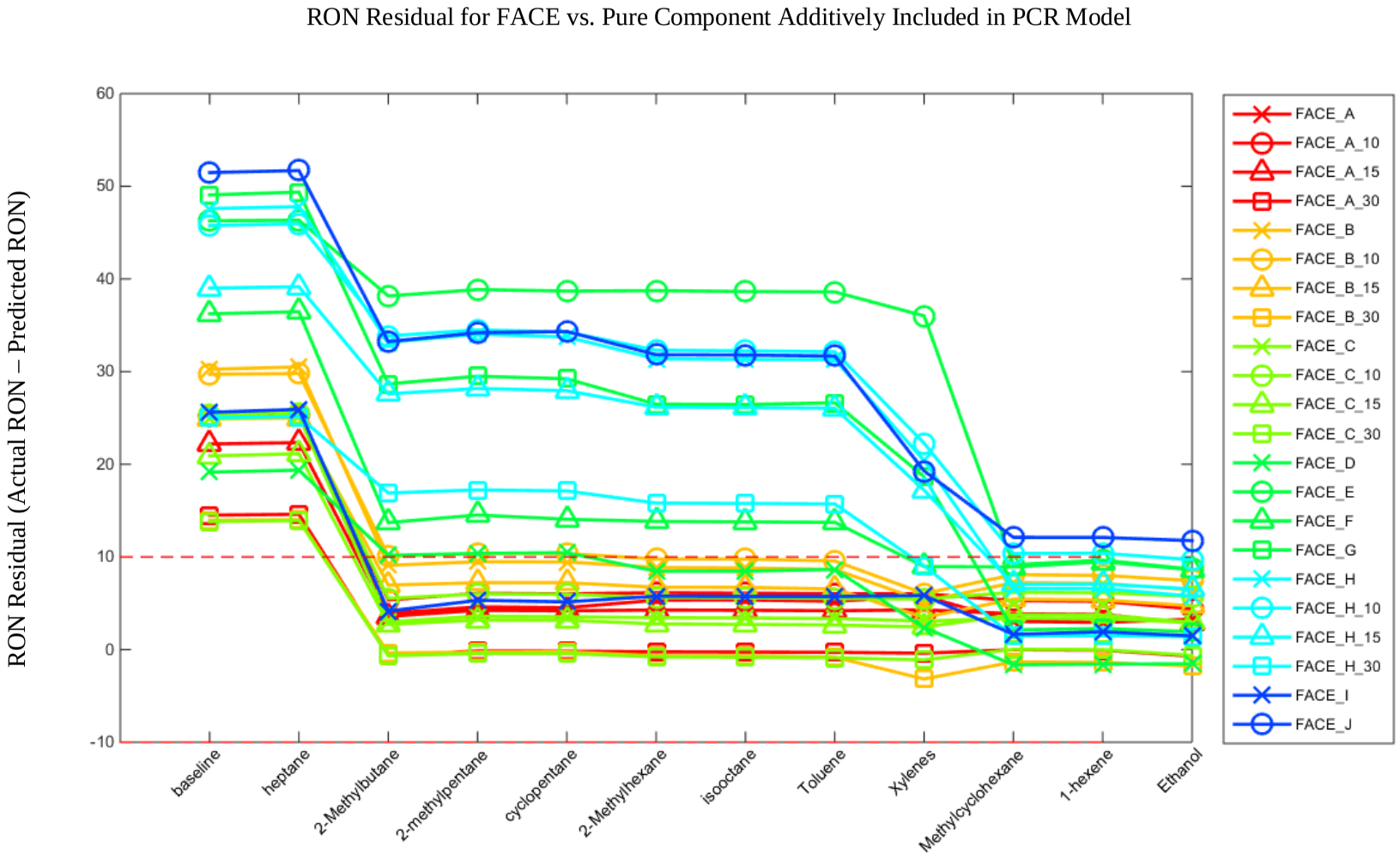

Figure 41: FACE gasoline RON residual versus pure components added into the model. This demonstrates the importance of the order in which pure components are added into the model, and shows there is leeway as to which pure components can be used to improve the RON model.

In this additive RON study, 2-methylalkanes replaced 2,3-dimenthylpentane. Some other pure components were also randomly distributed. It was found that 2-methylbutane or 2-methylhexane could be interchanged for the same effect, i.e., if 2-methylhexane was first added in lieu of 2-methylbutane it would have the same initial effect, and when 2-methylbutane was later added it had no effect. Interestingly, when these initial pure components are used, methylcyclohexane now shows a strong effect on the RON prediction. The earlier sensitivity studies indicated that methylcyclohexane, and the naphthenic class in general, had little effect on the predictive capabilities for RON. Furthermore, in the previous additive study methylcyclohexane had no effect when it was added after the other pure components. This may indicate that adding pure components from one class can significantly affect the predictive capabilities of hydrocarbons from other classes. This could be due to hydrocarbon IR absorbance spectra in fundamental vibrating frequencies, i.e., the C-H stretching frequency, from one class overlapping with hydrocarbons from other classes, albeit in varying magnitudes. Spectroscopically, this results in redundant information being added to the model, and may explain why some hydrocarbons can inform the chemometric model in the same way as other hydrocarbons, even if it represents a completely different class. As a result, it is difficult to determine which components can be selectively used, as adding one can affect the other regardless of hydrocarbon classes considered and vice versa.

The chemometrics relates the chemical functional groups, i.e., the IR absorbance spectra, to the RON performance of real gasolines. These sensitivity RON studies targeted which functional groups could be most important for predicting real gasoline performance. In a way, we are developing spectroscopic surrogates for real gasolines. While this study was originally intended to help target fuels to simulate for the LTC index, the studies may also have another broad impact. This pure component chemometric predictive approach, not necessarily with RON, could help aid in targeting hydrocarbon species that explain the majority of a combustion phenomenon for gasolines. Therefore, these studies show that there could be additional effort towards creating spectroscopic surrogates to represent real gasolines. Future work includes developing pure component spectroscopic surrogates, informed by chemometrics, to represent real gasolines.

The results of the selective and additive RON studies, along with available chemical mechanism considerations, informed the decision to target methylcyclohexane, 2-methylbutane, 2-methylhexane, and xylene(s) for the LTC index model.

In addition to investigating the effects of individual fuels, the relative predictive performance of PCR and SVMR were also analyzed. This was done by using all the fuel blends and pure components to build the model, i.e., all of the fuels considered in this study. It was found for the RON studies that SVMR yields the same results as PCR when a linear kernel is used; non-linear kernel choices for SVMR proved to give detrimental performance over the linear kernel.

Table 3: Actual and predicted RON values for CRC FACE gasolines. Chemometric model (PCR) informed with baseline and all pure components tested in this work.

CRC FACE Gasolines			
Fuel	**Actual RON [104]**	**Predicted RON**	**% Error**
FACE_A	83.9	86.2	2.7521
FACE_A_10	92.0	92.0	0.0985
FACE_A_15	94.8	97.7	3.1000
FACE_A_30	102.3	106.6	4.2294
FACE_B	95.8	99.5	3.8875
FACE_B_10	101.1	104.1	3.0089
FACE_B_15	103.0	107.2	4.0938
FACE_B_30	106.0	114.4	7.9669
FACE_C	84.3	89.5	6.1757
FACE_C_10	91.7	92.6	1.0738
FACE_C_15	94.8	97.0	2.3458
FACE_C_30	101.5	107.0	5.4878
FACE_D	93.2	99.3	6.5785
FACE_E	87.4	84.0	2.8021
FACE_F	94.0	90.8	3.3785
FACE_G	96.5	98.4	2.0141
FACE_H	86.9	80.2	7.6807
FACE_H_10	92.1	83.2	9.6402
FACE_H_15	94.1	88.1	6.3486
FACE_H_30	99.4	98.7	0.6844
FACE_I	70.0	69.5	0.6421
FACE_J	73.8	64.7	12.3038

4.2.4 LTC Index Chemometric Model

The previous sections addressed three important questions prior to the arrival at creating a chemometric model for the LTC index—recall that creating this model is the primary motivation of this work. We previously found that actual IR absorbance spectra were required to be used for the chemometric models (see section 4.3.1), i.e., it is not advised to use linearly blended spectra due to unideal ATR effects and non-linear blending effects. Additionally, ATR proved to give the same results as transmission despite no corrections performed on the ATR data to account for non-ideal effects (see section 4.3.2). Lastly, based on the pure component RON study, pure components that drastically affected the statistical prediction of the FACE gasolines' RON values were determined. Based on these results, in addition to considering available chemical mechanisms to simulate these pure components, additional pure components—2-methylbutane, 2-methylhexane and xylene blended with TRFs, and pure methylcyclohexane—were computationally simulated for the LTC index. In future work, additional pure components should be simulated to enhance the statistical model, in addition to more FACE gasolines to validate the model.

A chemometric model for predicting LTC performance (LTC index), built from all the considered pure components and mixtures thereof, was used to predict the performance of the FACE gasolines. At this time, the model can be validated against FACE gasolines A and C, as chemical mechanisms were readily available to simulate these fuels. The meta and kernel parameters for SVMR-based models were chosen to minimize the coefficient of determination for predicting FACEs A and C, that is, the indices of FACEs A and C were the two data points used as the data subset for model creation (as explained in section 4.1.2). The final model yields promising results, as seen in the results and analysis section.

4.3 Results and Analysis

4.3.1 Predicted LTC index

The final chemometric model was created and used to predict the LTC index of all the FACE fuels. The model was validated against the known values for FACEs A and C. All the fuels simulated (for LTC index) to create this model can be found in Appendix C. The final results are depicted here in Figure 47. The x-axes represents the FACE gasolines being predicted by the

chemometric model, and the y-axis is the predicted LTC index ("x" markers). Also shown are the true values of FACE gasolines A and C ("triangle" markers), which serve to validate the model.

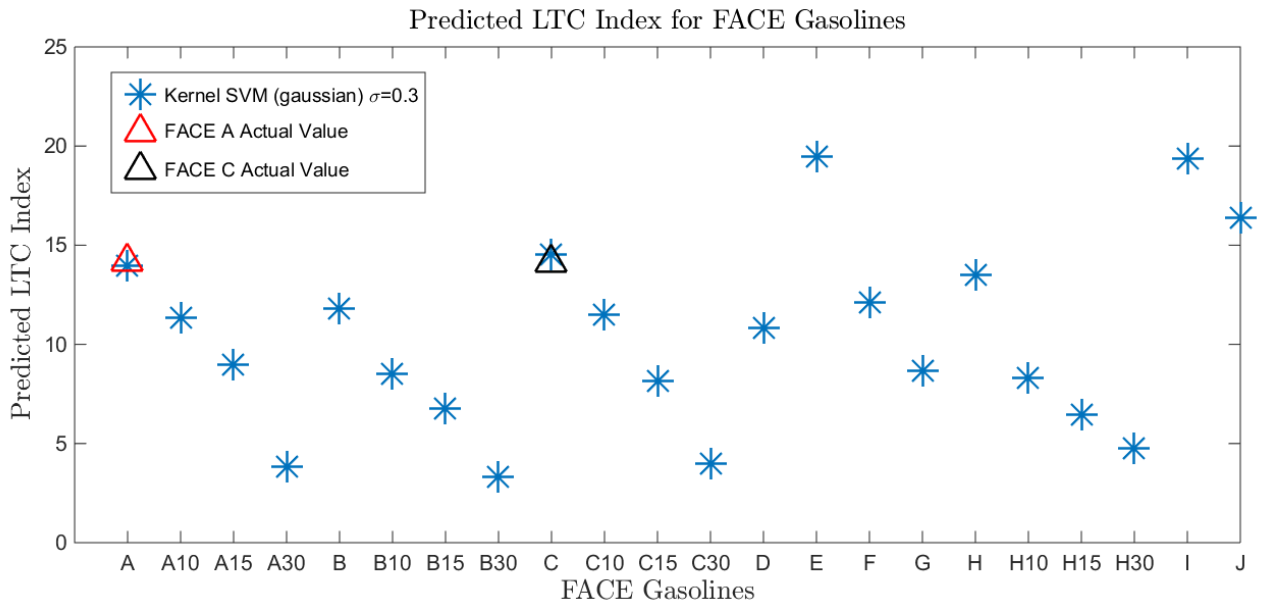

Figure 42: Final model for LTC index. This chemometric model utilizes SVMR with a non-linear kernel. The kernel used is a modified Gaussian profile (see equation 4.9), where the meta and kernel parameters were chosen to minimize the co ient of determination for predicting FACE fuels A and C—kernel parameter, σ, is 0.273 and SVM meta parameter, c, is 1000. The true values of FACEs A and C are indicated by the triangle markers.

The results in Figure 42 show a close LTC index prediction of FACE gasolines A and C: both within $\pm$ 0.4. In order to further prove the accuracy of this model, the other FACE fuels should be simulated to check against. Due to the close prediction of FACEs A and C, this work helps motivate further surrogate mechanism development for the remaining FACE fuels.

Figure 43 shows other chemometric predictive model results overlaid with the previously presented results, i.e., the models utilized different statistical methods than the prior. The legend in the figure indicates whether PCR, linear-based SVMR, and non-linear-based SVMR were used (and the Kernel type and parameter values).

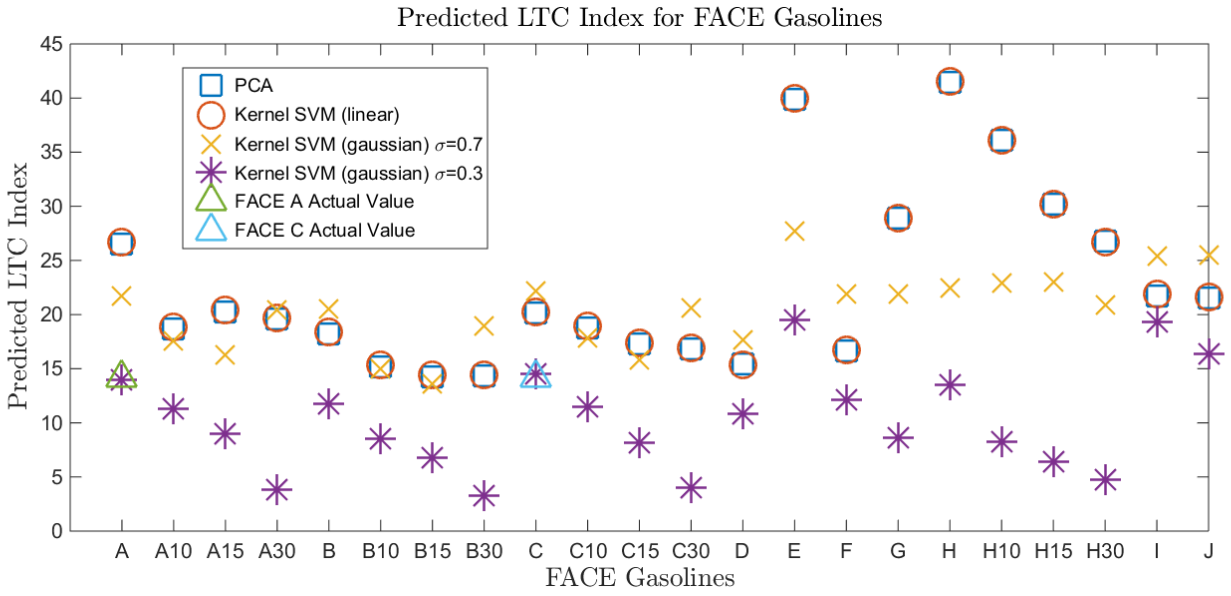

Figure 43: Multiple chemometric models for LTC index overlaid. It is seen that PCR and SVMR with a linear kernel yield the same results (red circles and blue squares, respectively). The final SVMR regression is shown here (purple "X"), as well as another SVMR model with a larger kernel parameter (yellow "X"). The larger kernel parameter value (0.7 versus 0.273) gives significantly different results.

Figure 43 indicates that non-linear SVMR methods proved to be more accurate in predicting FACE fuels A and C over the linear PCR and linear SVMR approaches (given the current training data set). Interestingly, recall that RON proved to be better predicted with linear methods. The IR absorbance spectra are non-linearly related to LTC index, yet linearly related to RON according to the performance of the chemometric models. This idea needs to be investigated further, however, due to the limited FACE fuel data points to validate the LTC index chemometric model.

Chapter 5: Conclusion

5.1 Overview

Low temperature combustion (LTC) engines are an attempt to unify the benefits of traditional gasoline spark-ignited and diesel compression-ignition engines, where a lean fuel-air mixture is introduced into the piston-cylinder chamber and ignited through compression [7]. With this, it is possible to offer efficiencies closer to diesel engines, but with low emissions of particulates and NO_x without the need for expensive exhaust after-treatment. The ability to produce fuels optimized for LTC engines could help reduce fossil fuel consumption, insomuch as a 34% reduction could be achieved in the U.S. [7] . It is desirable to create fuels that can operate in the LTC regime in order to promote advanced, next-generation engines as a viable alternative. The scope of this work is creating a tool that can predict the LTC performance of gasoline fuels, as well as target attractive fuel components to aid in designing fuels for LTC operation. This tool will avoid the need for expensive and time-consuming physical testing as required with traditional fuel performance metrics such as octane rating.

5.2 Methodology

This body of work investigated chemometric modeling as a means to predict next generation engine performance of real gasoline fuels. Chemometrics extracts information from chemical systems from data-driven means by performing calculations on the data. In this work, infrared (IR) absorption spectra are the chemical data, and the calculations done are statistical correlations relating the IR absorption spectra to chemical properties of interest.

The chemical property of interest in this work is a fuel performance rating specific to next generation engines operating on the LTC concept. This is defined by the LTC index, which rates fuels by their potential fuel savings in order to capture useful engine performance from LTC operation. Niemeyer et al. [35] recently introduced this index based on operating envelopes—engine speeds and loads that can be achieved with a given fuel—and comparing the LTC results to useful operating conditions.

Over 200 infrared absorption spectra comprising pure hydrocarbon components, mixtures of pure components, and research grade gasolines (FACE) were collected to provide the chemical data. The IR absorbance spectra were collected via transmission-based and attenuated total reflectance

(ATR)-based spectroscopy—the effects of using either approach with chemometric modeling were investigated.

Two novel chemometric models were created, one for predicting Research Octane Number (RON)—as a means to validate the chemometric approaches and to investigate attractive fuels for the LTC model—and the other model to predict the LTC index. The models predicted the performance of FACE gasolines and blends containing ethanol. This served to bound the task of creating a model to predict real gasoline fuels that contain hundreds of various hydrocarbons. The FACE gasolines are 22 fuels that were statistically designed by experts in the field to represent the chemical and ASTM performance variations found in gasolines. The important findings of this work will now be discussed.

5.3 Summary of Results

Traditional performance metrics such as RON, Motored Octane Number (MON), or Octane Sensitivity (RON-MON) were found to have little correlation to the LTC index. This was found to be true for over 100 fuels that were considered consisting of pure components, mixtures of pure components, and gasoline-like fuels (FACEs A and C). Fuels with widely varying RON/MON values, due to significantly different fuel compositions, did not necessarily result in different LTC indices. Instead, similar LTC performance was observed with fuels having low or high RON/MON. Furthermore, fuels with similar RON, MON, or octane sensitivity showed wide spreads of LTC index values. These results support the findings that the traditional fuel performance metrics do not fully explain LTC performance, and there is motivation to adopt a new industry standard for fuel performance rating for advanced LTC engines.

Attractive fuels that had high LTC performance, in addition to favorable high RON/MON values, exhibited significant equivalence ratio sensitivity for ignition delay, which extended the operating range of these fuels significantly. Designing fuels that can further capitalize on the benefits from equivalence ratio sensitivity may be the next step to making LTC strategies viable for light-duty automotive applications.

The two infrared absorption collection techniques, quantitative transmission and the qualitative attenuated total reflectance (ATR), yielded the same statistical performance with the chemometric models—so long as all the ATR data used is from the same equipment. This is a

significant finding as ATR-based spectroscopy is easy, fast, requires a few droplets of sample to test, and does not need extensive data post-processing—further highlighting the merit of this work as it proves to reduce physical experimental effort over traditional fuel performance testing (as well as significantly reduce cost).

Chemometrics were verified to be capable of predicting RON for pure hydrocarbon components and mixtures of such. More importantly, it was also shown that chemometric modeling—informed by a limited dataset of pure hydrocarbon components—could predict RON for all 22 gasoline-like fuels (FACE gasolines) considered in this work to within ± 10 RON. These models, to the author's knowledge, are novel in that only pure components are used to inform the chemometric model (as opposed to real gasoline samples) to predict the performance of real gasoline fuels. This initial study helped verify that particular chemical functional groups, represented by infrared absorbance spectra of pure hydrocarbon components, could explain gasoline performance.

Informed by the RON modeling results, a novel chemometric model was created to predict the LTC index. This model predicted the LTC index of FACE gasolines A and C within 3% (± 0.4) of the true value. Due to the high accuracy of the LTC chemometric model for these gasoline samples, this work merits further attention. This model could aid researchers and industry alike, as an easy tool to determine LTC performance of gasoline fuels and help in the effort to design next-generation fuels. This tool not only predicts the LTC index of an unknown fuel sample, but also indicates which functional groups influence the LTC index. That is to say, the weighting coefficients in the final regression equation indicate which wavenumbers in the absorbance spectra attribute positive or negative LTC index. With this, we can target beneficial functional groups to create fuels with high LTC index.

5.4 Future Work

Significant non-linear blending effects with the IR absorption spectra were observed with mixtures containing ethanol. Will non-linear effects significantly affect this work, e.g., with predicting the LTC index of gasoline fuels containing ethanol? The causes and implications of this phenomena should be investigated in future efforts.

A chemometric sensitivity study with RON was conducted, wherein the predictive capabilities of the model for the FACE gasolines were determined as a function of the pure components. This technique demonstrated that particular hydrocarbon components that prove beneficial to the predictive performance of the chemometric model could be targeted. This methodology is of sorts a means to create surrogate representations of infrared absorption spectra of real gasoline fuels; that is to say, pinpointing the most important chemical functional groups to represent gasoline. However, it was found that there was leeway as to which pure components could be used to the same effect—this needs to be further investigated. Potentially, if the previous question is addressed, this methodology can help inform new surrogates that emulate gasoline fuels. This pure component chemometric predictive approach, not necessarily with RON, could help aid in targeting hydrocarbon species that explain the majority of combustion phenomena for gasolines. Using this methodology to aid in the effort for LTC engine fuel design should also be investigated as future work.

The LTC index model requires further validation. Accomplishing this is possible when surrogate chemical mechanisms for the remaining FACE gasolines (or any research grade gasoline) are developed; this would provide the additional data needed to better validate the LTC chemometric model. The results of this work, until then, are only promising.

References

[1] "Hydrocarbons America's Leading Source of Energy Consumption" [Online]. Available: http://instituteforenergyresearch.org/topics/encyclopedia/fossil-fuels/.

[2] "United States Environmental Protection Agency - Particulate Matter." [Online]. Available: http://www.epa.gov/pm/health.html.

[3] U. S. E. P. Agency, "How nitrogen oxides affect the way we live and breathe," no. September, 1998.

[4] J. B. Heywood and O. Z. Welling, "Trends in Performance Characteristics of Modern Automobile SI and Diesel Engines," vol. 2, no. 1, pp. 1650–1662, 2009.

[5] "US diesel car registrations increased by 24% from 2010-2012, hybrids up 33%," 2013. [Online]. Available: http://www.greencarcongress.com/2013/04/diesel-20130426.html.

[6] John B. Heywood, *Internal Combustion Engine Fundamentals*, McGraw-Hill, New York, 1988.

[7] R. D. Reitz, "Directions in internal combustion engine research," *Combust. Flame*, vol. 160, no. 1, pp. 1–8, 2013.

[8] C. P. Loeper, "Experimental Investigation of Gasoline Compression Ignition Combustion in a Light-Duty," 2013.

[9] S. L. Kokjohn, R. M. Hanson, D. a. Splitter, and R. D. Reitz, "Fuel reactivity controlled compression ignition (RCCI): a pathway to controlled high-efficiency clean combustion," *Int. J. Engine Res.*, vol. 12, no. x, pp. 209–226, 2011.

[10] D. B. Johansson, "High Efficiency Combustion Engines– What is the Limit?," *Lund Univ.*

[11] Dec, J., "A Conceptual Model of DI Diesel Combustion Based on Laser-Sheet Imaging", SAE Technical Paper 970873, 1997, doi:10.4271/970873.

[12] G. Abthoff, J., Schuster, H., Langer, H., and Loose, "The Regenerable Trap Oxidizer –An emission Control Technique for Diesel Engines," *SAE*, no. 850015, 1985.

[13] J. H. Kim, M. Y. Kim, and H. G. Kim, "-Assisted Soot Regeneration Behavior in a Diesel Particulate Filter with Heavy-Duty Diesel Exhaust Gases," *Numer. Heat Transf. Part A Appl.*, vol. 58, no. 2, pp. 725–739, 2010.

[14] H. Yun, N. Wermuth, and P. Najt, "Extending the High Load Operating Limit of a Naturally-Aspirated Gasoline HCCI Combustion Engine," vol. 3, no. 1, pp. 681–699, Apr. 2010.

[15] M. Yao, Z. Zheng, and H. Liu, “Progress and recent trends in homogeneous charge compression ignition (HCCI) engines,” *Prog. Energy Combust. Sci.*, vol. 35, no. 5, pp. 398–437, Oct. 2009.

[16] J. et al. Neely, G., Sasaki, S., Huang, Y., Leet, “New Diesel Emission Control Strategy to Meet US Tier 2 Emissions Regulations,” *SAE*, 2005.

[17] J. A. Caton, “A Review of Investigations Using the Second Law of Thermodynamics to Study Internal-Combustion Engines,” *SAE*, no. 2000–01–1081, 200AD.

[18] “No Title.” [Online]. Available: http://www.princeton.edu/puceg/perspective/combustion/fig1.jpg.

[19] Boyce, P. M, 2002, ‘Gas turbine engineering handbook’, Gulf Professional Publishing, 2nd edition, Huston, Texas, USA.

[20] P. E. Yelvington, M. B. I. Rallo, S. Liput, J. W. Tester, W. H. Green, and J. Yang, “Prediction of Performance Maps for Homogeneous-Charge Compression-Ignition Engines,” *Combust. Sci. Technol.*, vol. 176, no. 8, pp. 1243–1282, Aug. 2004.

[21] Lacey, Joshua S., et al. "HCCI operability limits: the impact of refinery stream gasoline property variation." *Journal of Engineering for Gas Turbines and Power* 135.8 (2013): 081505.

[22] Kalghatgi, G. T. "Fuel anti-knock quality-Part I. Engine studies." No. 2001-01-3584. SAE Technical Paper, 2001.

[23] Truedsson, Ida. "The HCCI Fuel Number-Measuring and Describing Auto-ignition for HCCI Combustion Engines." PhD diss., Lund University, 2014.

[24] T. Shibata, G., Urushihara, “Auto-Ignition Characteristics of Hydrocarbons and Development of HCCI Fuel Index,” *SAE*, vol. 2007–01–02, 2007.

[25] K. Shibata, G., Oyama, “Correlation of Low Temperature Heat Release With Fuel Composition and HCCI Engine Combustion,” *SAE*, vol. 2005–01–01, 2005.

[26] T. Shibata, G., Oyama, K., Urushihara, T., and Nakano, “The Effect of Fuel Properties on Low and High Temperature Heat Release and Resulting Performance of an HCCI Engine,” *SAE Tech. Pap.*, vol. 2004–01–05, 2004.

[27] X. Yang and G. Zhu, “SI and HCCI combustion mode transition control of an HCCI capable SI engine,” *IEEE Trans. Control Syst. Technol.*, vol. 21, no. 5, pp. 1558–1569, 2013.

[28] J. Warnatz, U. Maas, and R. W. Dibble, *Combustion*, ISBN-10 3-540-25992-9 4th ed. Springer Berlin Heidelberg New York, 2006.

[29] ASTM D2700-14, *Standard Test Method for Motor Octane Number of Spark-Ignition Engine Fuel*. West Conshohocken, PA: ASTM International, 2014.

[30] A. D2699-13b, *Standard Test Method for Research Octane Number of Spark-Ignition Engine Fuel*. West Conshohocken, PA: ASTM International, 2014.

[31] J. C. Guibet, *Fuels and Engines - Technology, energy, environment*, Éditions T. Paris, 1999.

[32] G. T. Kalghatgi, "A Method of Defining Ignition Quality of Fuels in HCCI Engines," *SAE*, 2014.

[33] Z. Liu, H., Yao, M., Zhang, B., Zheng, "Influence of Fuel and Operating Conditions on Combustion Characteristics of a Homogenous Charge Compression Ignition Engine," *Energy & Fuels*, no. 23, pp. 1422–1430, 2009.

[34] V. H. Rapp, W. J. Cannella, J.-Y. Chen, and R. W. Dibble, "Predicting Fuel Performance for Future HCCI Engines," *Combust. Sci. Technol.*, vol. 185, no. 5, pp. 735–748, May 2013.

[35] Niemeyer, K. E., Daly, S. R., Cannella, W. J., & Hagen, C. L. (2015). A Novel Fuel Performance Index for Low-Temperature Combustion Engines Based on Operating Envelopes in Light-Duty Driving Cycle Simulations. Journal of Engineering for Gas Turbines and Power, 137(10), 101601. doi:10.1115/1.4029948

[36] J. J. Kelly, C. H. Barlow, T. M. Jinguji, and J. B. Callis, "Prediction of Gasoline Octane Numbers from Near-Infrared Spectral Features in the Range 660-1215 nm," Analytical Chemistry, vol. 61, no. 4, pp. 313–320, 1989.

[37] A. B. of A. Standards, "ASTM Procedures D 2699 and D 2700," vol. 5.02, 1995.

[38] http://www.grabner-instruments.com/products/fuelanalysis/index.aspx.

[39] http://www.zeltex.com/eval.html.

[40] J.-Y. C. G. Chin, "Modeling of emissions from HCCI engines using a consistent 3-zone model with applications to validation of reduced chemistry," *Proc. Combust. Inst.*, no. 33, pp. 3073–3079, 2011.

[41] R. Design, "Reaction Workbench 15131." San Diego, 2013.

[42] Dave Goodwin, Nicholas Malaya, Harry Moffat, and Raymond Speth, "Cantera: An object-oriented software toolkit for chemical kinetics, thermodynamics, and transport processes. Version 2.1a1." url: http://www.cantera.org/

[43] H. Pitz, W.J., Cernansky, N.P., Dryer, F.L., Egolfopoulos, F.N., Farrell, J.T., Friend, D.G., Pitsch, "Development of an experimental Database and Chemical Kinetic Models for Surrogate Gasoline Fuels," *SAE 2007-01-0175*, 2007.

[44] C. Company, "Motor Gasolines Technical Review," vol. FTR-1, pp. 1–69, 1996.

[45] T. M. Foong, K. J. Morganti, M. J. Brear, G. Da Silva, Y. Yang, and F. L. Dryer, "The octane numbers of ethanol blended with gasoline and its surrogates," *Fuel*, vol. 115, pp. 727–739, 2014.

[46] P. L. Perez and A. L. Boehman, "Experimental Investigation of the Autoignition Behavior of Surrogate Gasoline Fuels in a Constant-Volume Combustion Bomb Apparatus and Its Relevance to HCCI Combustion," *Energy & Fuels*, vol. 26, no. 10, pp. 6106–6117, Oct. 2012.

[47] N. and Grumman, "Northrop Grumman Motor Gasolines, Winter 2004-05," no. Report No. NGMS-240 PPS, 2005.

[48] M. Chaos, A. Kazakov, Z. Zhao, "A high-temperature chemical kinetic model for primary reference fuels," *Int. J. Chem. Kinet*, no. 39, pp. 399–414, 2007.

[49] M. Chaos, Z. Zhao, A. Kazakov, P. Gokulakrishnan, M. Angioletti, "A PRF+toluene surrogate fuel model for simulating gasoline kinetics, in: Fifth Joint Meeting of the US Sections of the Combustion Institute," *Pap. E26*, 2007.

[50] J. Li, A. Kazakov, M. Chaos, "Chemical kinetics of ethanol oxidation, in: Fifth Joint Meeting of the US Sections of the Combustion Institute," *Pap. C28*, 2007.

[51] F. M. Haas, M. Chaos, F. L. Dryer, "Low and intermediate temperature oxidation of ethanol and ethanol–PRF blends: An experimental and modeling study", Combust. Flame 156 2346–2350, 2009.

[52] T. Bieleveld, A. Frassoldati, A. Cuoci, T. Faravelli, E. Ranzi, U. Niemann, "Experimetnal and kinetic modeling study of combustion of gasoline, its surrogates and components in laminar non-premixed flames." *Proc. Combust. Inst*, no. 32, pp. 493–500, 2009.

[53] C.V. Naik, W.J. Pitz, C.K. Westbrook, M. Sjöberg, J.E. Dec, J.P. Orme, H.J. Curran, "Detailed Chemical Kinetic Modeling of Surrogate Fuels for Gasoline and Application to an HCCI Engine," *SAE, Int.*, no. Report No. 2005–01–3741, 2005.

[54] B. W. Weber, W. J. Pitz, M. Mehl, E. J. Silke, A. C. Davis, and C. J. Sung, "Experiments and modeling of the autoignition of methylcyclohexane at high pressure," *Combust. Flame*, vol. 161, no. 8, pp. 1972–1983, 2014.

[55] S. M. Sarathy, C. K. Westbrook, M. Mehl, W. J. Pitz, C. Togbe, P. Dagaut, H. Wang, M. a. Oehlschlaeger, U. Niemann, K. Seshadri, P. S. Veloo, C. Ji, F. N. Egolfopoulos, and T. Lu, "Comprehensive chemical kinetic modeling of the oxidation of 2-methylalkanes from C7 to C20," *Combust. Flame*, vol. 158, pp. 2338–2357, 2011.

[56] Curran, H. J., P. Gaffuri, William J. Pitz, and Charles K. Westbrook. "A comprehensive modeling study of n-heptane oxidation." *Combustion and flame* 114, no. 1 (1998): 149-177.

[57] Westbrook, Charles K., William J. Pitz, Olivier Herbinet, Henry J. Curran, and Emma J. Silke. "A comprehensive detailed chemical kinetic reaction mechanism for combustion of n-alkane hydrocarbons from n-octane to n-hexadecane." *Combustion and Flame* 156, no. 1 (2009): 181-199.

[58] Curran, Henry J., P. Gaffuri, W. J. Pitz, and C. K. Westbrook. "A comprehensive modeling study of iso-octane oxidation." *Combustion and flame* 129, no. 3 (2002): 253-280.

[59] Oehlschlaeger, Matthew A., Justin Steinberg, Charles K. Westbrook, and William J. Pitz. "The autoignition of iso-cetane at high to moderate temperatures and elevated pressures: Shock tube experiments and kinetic modeling." *Combustion and flame* 156, no. 11 (2009): 2165-2172.

[60] Westbrook, C. K., W. J. Pitz, J. E. Boercker, H. J. Curran, J. F. Griffiths, C. Mohamed, and M. Ribaucour. "Detailed chemical kinetic reaction mechanisms for autoignition of isomers of heptane under rapid compression." *Proceedings of the combustion institute* 29, no. 1 (2002): 1311-1318.

[61] Westbrook, Charles K., William J. Pitz, Henry C. Curran, Janice Boercker, and Eric Kunrath. "Chemical kinetic modeling study of shock tube ignition of heptane isomers." *International Journal of Chemical Kinetics* 33, no. 12 (2001): 868-877.

[62] Mehl, Marco, Guillaume Vanhove, William J. Pitz, and Eliseo Ranzi. "Oxidation and combustion of the n-hexene isomers: A wide range kinetic modeling study." *Combustion and Flame* 155, no. 4 (2008): 756-772.

[63] Mehl, Marco, William J. Pitz, Charles K. Westbrook, Kenji Yasunaga, Christine Conroy, and Henry J. Curran. "Autoignition behavior of unsaturated hydrocarbons in the low and high temperature regions." *Proceedings of the Combustion Institute* 33, no. 1 (2011): 201-208.

[64] Mehl, Marco, William J. Pitz, Charles K. Westbrook, and Henry J. Curran. "Kinetic modeling of gasoline surrogate components and mixtures under engine conditions." *Proceedings of the Combustion Institute* 33, no. 1 (2011): 193-200.

[65] Silke, Emma J., William J. Pitz, Charles K. Westbrook, and Marc Ribaucour. "Detailed chemical kinetic modeling of cyclohexane oxidation." *The Journal of Physical Chemistry A* 111, no. 19 (2007): 3761-3775.

[66] K. E. Niemeyer, C.-J. Sung, and M. P. Raju, "Skeletal mechanism generation for surrogate fuels using directed relation graph with error propagation and sensitivity analysis," *Combust. Flame*, vol. 157, no. 9, pp. 1760–1770, Sep. 2010.

[67] Niemeyer, Kyle E., and Chih-Jen Sung. "Mechanism reduction strategies for multicomponent gasoline surrogate fuels." In *7th US National Combustion Meeting of the Combustion Institute*. 2011.

[68] Niemeyer, Kyle E., and Chih-Jen Sung. "Mechanism reduction for multicomponent surrogates: A case study using toluene reference fuels." *Combustion and Flame* 161, no. 11 (2014): 2752-2764.

[69] Sarathy SM, Kukkadapu G, Mehl M, Wang W, Javed T, Park S, et al., "Ignition of alkane-rich FACE gasoline fuels and their surrogate mixtures". Proc Combust Inst 2015;35:249–57. doi:10.1016/j.proci.2014.05.122.

[70] Ahmed A, Goteng G, Shankar VSB, Al-Qurashi K, Roberts WL, Sarathy SM. "A computational methodology for formulating gasoline surrogate fuels with accurate physical and chemical kinetic properties". Fuel 2015;143:290–300. doi:10.1016/j.fuel.2014.11.022.

[71] K. E. Niemeyer, S. R. Daly, W. J. Cannella, and C. L. Hagen, "Investigation of the LTC fuel performance index for oxygenated reference fuel blends," *Fuel*, vol. 155, no. x, pp. 14–24, 2015.

[72] A. Aldawood, S. Mosbach, M. Kraft, "HCCI combustion control using dual-fuel approach: Experimental and modeling investigations", SAE Technical Paper 2012-01-1117, 2012. doi:10.4271/2012-01-1117.

[73] T. Aroonsrisopon, V. Sohm, P. Werner, D. E. Foster, T. Morikawa, M. Iida, "An investigation into the effect of fuel composition on HCCI combustion characteristics", SAE Technical Paper 2002-01-2830, 2002. doi:10.4271/2002-01-2830.

[74] A. Oakley, H. Zhao, N. Ladommatos, T. Ma, "Experimental studies on controlled auto-ignition (CAI) combustion of gasoline in a 4-stroke engine", SAE Technical Paper 2001-01-1030, 2001. doi:10.4271/ 2001-01-1030.

[75] P. Yelvington, W. Green, "Prediction of the knock limit and viable operating range for a homogeneous-charge compression-ignition (HCCI) engine", SAE Technical Paper 2003-01-1092, 2003. doi:10.4271/ 2003-01-1092.

[76] Jacob R. Zuehl, Jaal Ghandhi, Christopher L. Hagen, William Canella., "Fuel Effects on HCCI Combustion Using Negative Valve Overlap". SAE Technical Paper on 2010-12-04, doi:10.4271/2010-01-0161.

[77] S. Saxena, J.-Y. Chen, and R. Dibble, “Maximizing Power Output in an Automotive Scale Multi-Cylinder Homogeneous Charge Compression Ignition (HCCI) Engine,” no. x, Apr. 2011.

[78] S. Saxena, N. Shah, I. Bedoya, and A. Phadke, “Understanding optimal engine operating strategies for gasoline-fueled HCCI engines using crank-angle resolved exergy analysis,” *Appl. Energy*, vol. 114, pp. 155–163, Feb. 2014.

[79] Eng et al. US patent US20070119417.

[80] I. Truedsson, M. Tuner, B. Johansson, W. Cannella, "Pressure sensitivity of HCCI auto-ignition temperature for primary reference fuels", SAE Int. J. Engines 5 1089–1108., 2012.

[81] I. Truedsson, M. Tuner, B. Johansson, W. Cannella, Emission formation study of HCCI combustion with gasoline surrogate fuels, SAE Technical Paper 2013-01-2626, 2013. doi:10.4271/2013-01-2626.

[82] I. Truedsson, M. Tuner, B. Johansson, W. Cannella, Pressure sensitivity of HCCI auto-ignition temperature for oxygenated reference fuels, J. Eng. Gas. Turb. Power 135 072801, 2013.

[83] Hyvönen, Jari, Göran Haraldsson, and Bengt Johansson. *Balancing cylinder-to-cylinder variations in a multi-cylinder VCR-HCCI engine*. No. 2004-01-1897. SAE Technical Paper, 2004.

[84] R.H. Butt et al., Improving ion current of sparkplug ion sensors in HCCI combustion using sodium, potassium, and cesium acetates: Experimental and numerical modeling, http://dx.doi.org/10.1016/j.proci.2014.06.084.

[85] E. Jones, T. Oliphant, P. Peterson, et al., SciPy: Open source scientific tools for Python, http://www.scipy.org/, 2001–. [Online; accessed 2014-09-03].

[86] C. B. Barber, D. P. Dobkin, H. Huhdanpaa, The quickhull algorithm for convex hulls, ACM Trans. Math. Software 22 (1996) 469–483.

[87] EPA Office of Transportation and Air Quality, 2013. Dynamometer drive schedules. http://www.epa.gov/ nvfel/testing/dynamometer.htm.

[88] Wipke, K. B., Cuddy, M. R., and Burch, S. D., 1999. ‘ADVISOR 2.1: A user-friendly advanced powertrain simulation using a combined backward/forward approach’. IEEE Trans. Vehicular Technol., 48(6), pp. 1751–1761.

[89] Gao, D. W., Mi, C., and Emadi, A., 2007. 'Modeling and simulation of electric and hybrid vehicles'. Proc. IEEE, 95(4), pp. 729–745.

[90] http://www.crcao.org/workshops/LCA%20October%202013/Opening/Bailey,%20Brent.pdf.

[91] L. a. Averett, P. R. Griffiths, and K. Nishikida, "Effective path length in attenuated total reflection spectroscopy," *Anal. Chem.*, vol. 80, no. 8, pp. 3045–3049, 2008.

[92] Harrick, N. J.Internal Reflection Spectroscopy; Interscience Publishers: New York, 1967.

[93] N. J. Harrick and F. K. du Pré, "Effective thickness of bulk materials and of thin films for internal reflection spectroscopy.," *Appl. Opt.*, vol. 5, no. 11, pp. 1739–1743, 1966.

[94] Harrick, N. J. J. Opt. Soc. Am. 55, 851-857, 1965.

[95] Robert Siegel and John R. Howell. Thermal radiation heat transfer. Taylor & Francis, New York, 2002.

[96] Craig Bohren and Donald Huffman. Absorption and Scattering of Light by Small Particles. Wiley., 1998.

[97] David W. Ball. The basics of spectroscopy. Tutorial texts in optical engineering, v. TT 49. SPIE- The International Society for Optical Engineering, Bellingham, WA, 2001.

[98] Bertie, J., Infrared Intensities of Liquids XIII: Accurate Optical Constants and Molar Absorption Coefficients Between 6500 and 435 cm-1 of Toluene at 25 C, from Spectra Recorded in Several Laboratories., Applied Spectroscopy Vol:48 No:1," 1994.

[99] Jason Morgan Porter, "Laser-Based Diagnostics for Hydrocarbon Fuels in the Liquid and Vapor Phases", PhD. Thesis, Stanford University, 2009.

[100] J. M. Porter, J. B. Jeffries, and R. K. Hanson, "Mid-infrared absorption measurements of liquid hydrocarbon fuels near," *J. Quant. Spectrosc. Radiat. Transf.*, vol. 110, no. 18, pp. 2135–2147, Dec. 2009.

[101] C. Dale Keefe, "Computer programs for the determination of optical constants from transmission spectra and the study of absolute absorption intensities," *J. Mol. Struct.*, vol. 641, no. 2–3, pp. 165–173, Nov. 2002.

[102] Bertie, John E., C. Dale Keefe, and R. Norman Jones. "Infrared intensities of liquids VIII. Accurate baseline correction of transmission spectra of liquids for computation of absolute intensities, and the 1036 cm-1 band of benzene as a potential intensity standard." *Canadian journal of chemistry* 69, no. 11 (1991): 1609-1618.

[103] Mirabella, Francis M. *Internal reflection spectroscopy: Theory and applications*. Vol. 15. CRC Press, pg. 20, 1992.

[104] W. Cannella et al., "Face Gasolines and Blends with Ethanol : Detailed Characterization of Physical and Chemical Properties", CRC Report No . AVFL-24, July 2014.

[105] S. Corsetti, F. M. Zehentbauer, D. McGloin, and J. Kiefer, "Characterization of gasoline/ethanol blends by infrared and excess infrared spectroscopy," *Fuel*, vol. 141, pp. 136–142, Feb. 2015.

[106] J. T. Reilly, a. Thomas, a. R. Gibson, C. Y. Luebehusen, and M. D. Donohue, "Analysis of the Self-Association of Aliphatic Alcohols Using Fourier Transform Infrared (FT-IR) Spectroscopy," *Ind. Eng. Chem. Res.*, vol. 52, pp. 14456–14462, 2013.

[107] H. C. Van Ness, J. Van Winkle, H. H. Richtol, and H. B. Hollinger, "Infrared spectra and the thermodynamics of alcohol-hydrocarbon systems," *J. Phys. Chem.*, vol. 71, no. 6, pp. 1483–1494, 1967.

[108] Barry K. Lavine and Jerome Workman, "Chemometrics," *Anal. Chem.*, vol. 74, no. 12, pp. 2763–2770, 2002.

[109] L. I. Smith, "A tutorial on Principal Components Analysis Introduction," 2002.

[110] *"Elementary Linear Algebra 5e" by Howard Anton, Publisher JohnWiley & Sons Inc, ISBN 0-471-85223-6. .*

[111] http://www.mathworks.com/help/stats/partial-least-squares-regression-and-principal-components-regression.html.

[112] Chapra, Steven C., and Raymond P. Canale. *Numerical methods for engineers*. Vol. 2. McGraw-Hill, 2012.

[113] C. J. C. Burges, "A Tutorial on Support Vector Machines for Pattern Recognition," *Data Min. Knowl. Discov.*, vol. 2, pp. 121–167, 1998.

[114] O. Devos, C. Ruckebusch, A. Durand, L. Duponchel, and J. P. Huvenne, "Support vector machines (SVM) in near infrared (NIR) spectroscopy: Focus on parameters optimization and model interpretation," *Chemom. Intell. Lab. Syst.*, vol. 96, no. 1, pp. 27–33, 2009.

[115] J. Luts, F. Ojeda, R. Van de Plas Raf, B. De Moor, S. Van Huffel, and J. a K. Suykens, "A tutorial on support vector machine-based methods for classification problems in chemometrics," *Anal. Chim. Acta*, vol. 665, no. 2, pp. 129–145, 2010.

[116] E. Kim, "Everything You Wanted to Know about the kernel Trick." [Online]. Available: http://www.eric-kim.net/eric-kim-net/posts/1/kernel_trick.html.

[117] “FTIR graphic” [Online]. Available: https://en.wikipedia.org/wiki/Fourier_transform_infrared_spectroscopy#/media/File:FTIR_Interferometer.png.

[118] “Nicolet FTIR commercial graphic” [Online]. Available: http://www.thermoscientific.com/content/tfs/en/product/omnic-series-software.html.

[119] “Hawranek,J.P., The control of errors in i.r. spectrophotometry—III. Transmission measurements using thin cells, Spectrochimica acta. Part A, Molecular spectroscopy, Vol:32, is: 1,” 1976.

[120] J. E. Bertie and S. L. Zhang, “Infrared intensities of liquids. IX. The Kramers–Kronig transform, and its approximation by the finite Hilbert transform via fast Fourier transforms,” *Can. J. Chem.*, vol. 70, no. 2, pp. 520–531, Feb. 1992.

[121] D. C. Keefe, "Computer Programs for Optical Constant Method" [Online] Available: http://faculty. uccb.ns.ca/dkeefe/ programs.

[122] Scherzer, Julius. "Octane-enhancing, zeolitic FCC catalysts: scientific and technical aspects." *Catalysis Reviews—Science and Engineering* 31, no. 3 (1989): 215-354.

[123] Colorado school of mines, “Hydrocarbon Database”.

[124] Francis A. Carey, Richard J. Sundberg, "Advanced Organic Chemistry: Part A: Structure and Mechanisms", Springer Science & Business Media, (2007), ISBN: 0387448977, 9780387448978.

[125] K. Owen, T. Coley: Automotive Fuels Handbook, Society of Automotive Engineers,Warrendale, PA, 1990, pp.564–565.

[126] API Research Project 45: "Knocking characteristics of pure hydrocarbons". 1958, ASTM.

[127] Perez, P.L. and A.L. Boehman, "Low-Temperature Oxidation of Hydrocarbon Fuels in a Motored Engine", In ACS Fuel Chemistry Division Preprints, 55(1), 566-567 (2010).

[128] T. a. Albahri, “Structural Group Contribution Method for Predicting the Octane Number of Pure Hydrocarbon Liquids,” *Ind. Eng. Chem. Res.*, vol. 42, pp. 657–662, 2003.

[129] M.J. Pilling, *Low-temperature Combustion and Autoignition*. Amsterdam: Elsvier Science B.V., 1997.

[130] Hunwartzen I., ‘Modifications of CFR test engine unit to determine octane numbers of pure alcohols and gasoline-alcohol blends’. SAE Technical Paper 820002; 1982.

[131] Anderson JE, Leone TG, Shelby MH, Wallington TJ, Bizub JJ, Foster M, et al., "Octane numbers of ethanol–gasoline blends: measurements and novel estimation method from molar composition". SAE Technical Paper 2012-01- 1274; 2012. doi: http://dx.doi.org/10.42.

Appendices

Appendix A: Development of Python/ Cantera Model

A1 Engine Geometry

Table 4: Engine Geometry/ Parameters for HCCI Simulations

Engine Geometry/ Parameters for CRC Research Engine	
Parameter	*Value / Equation*
Compression Ratio, rc	13
Cylinder Bore [m], B	0.828
Displacement Volume [m^3], V_d	616 E-06
Ratio of Stroke to Crank Radius, c	4.44
Stroke [m], s	$4V_d/(\pi B^2)$
Maximum Cylinder Volume [m^3]	$4V_d r_c/(r_c - 1)$
Piston Area [m^2]	$\pi B^2/4$
Rotation Rate [rad/s], $\dot{w}$	$RPM\pi/30$

A2 Control Volume Relations

Table 5: Piston Speed and Control Volume Relations for HCCI Simulations

Piston Speed/ Control Volume Relations	
Parameter	*Equation*
Piston Speed(t) [m/s]	$\dot{w}\frac{s}{2}\sin(t\,\dot{w})\left(1-\frac{\cos(t\,\dot{w})}{\sqrt{c^2-\sin(t\,\dot{w})^2}}\right)$
Volume(t) [m^3], V(t)	$\frac{Vd}{r_c-1}+\frac{Vd}{2}*\left(1+c-\cos(t\,\dot{w}-\pi)-\sqrt{c^2-\sin(t\,\dot{w}-\pi)^2}\right)$

A3 Heat Transfer Relations

Table 6: Woschni Heat Transfer Correlation for HCCI Simulations

Heat Transfer Correlation	
Parameter	*Equation*
Viscosity, μ	$0.1 * C1_{\mu} \, (T)^{C2_{\mu}}$
Thermal Conductivity [W/m-K], k_g	$10^{-5} * C1_k (T)^{C2_k}$
Average Piston Speed [m/s], Sm	2*s*RPM/60
Woschni average cylinder gas velocity, $\overline{w}$	C11*Sm
Reynolds Number, Re	$B\overline{w}\rho/\mu$
Nusselt Number, Nu	$aRe^{b}Pr^{c}$
Convective Heat Transfer Coefficient [W/m^2-K], h	$Nu * k_g/B$
Heat Transfer Flux [W/m^2], q''	$\dfrac{\left(h * (T_w - T) * \dfrac{(2 * \pi * B^3 + 16 * V(t))}{4 * B}\right)}{Ap}$

A4 Thermo-Physical Property Evaluation

During the evaluation of the Cantera model, it was found that results considering heat transfer were not consistent with Chemkin results. The issue was narrowed down to the thermo-physical property evaluation used in the heat transfer model. The form for the property evaluation within Chemkin are as follows:

$$\mu = 0.1 * C1_{\mu} \, (T)^{C2_{\mu}}$$

$$k_g = 10^{-5} * C1_k (T)^{C2_k}$$

Reaction Design kindly provided the constants (C1 and C2) to evaluate the above expressions. These properties are used to determine the convective heat transfer coefficient for the control surface boundary, via the Woschnii heat transfer correlation. Using the same heat transfer correlation and thermo-physical constants as that of Chemkin within Cantera, the heat transfer results were still found to be significantly different. A small study was done to investigate this discrepancy.

Chemkin outputs crank resolved heat losses through the control surface, piston-cylinder heat transfer area, and temperature—it is therefore possible to back-solve for the thermo-physical properties with the heat transfer correlation. Through this study, it was found that the constants Reaction Design provided were "half correct". If the constants were not user-defined and left blank within the Chemkin menu, it would be expected that the default values would be referenced. This is not the case, and in fact the default values provided by reaction design were not actually being used for this scenario. Albeit, if one or more of the constants were user-defined and a few left blank, the default constants were being used for the constants not user-defined (left blank). Regardless, it was of importance for this study to match Chemkin since prior fuel studies had been done with Chemkin, so with the transition to Cantera it was then desired to have comparable results. With the previous Chemkin studies all constants were left undefined, so the constants *actually* being used by Chemkin were left to be determined.

The thermo-physical properties from Chemkin—back-solved from the heat transfer correlations—were plotted versus temperature. Seen in Figure 44, it is observed that after TDC new constants are used. It is inferred that Chemkin is taking into account that air is primarily

present BTDC, while ATDC it is exhaust products. Since the variation is very slight, a best fit line was determined and the constants yielding this were recorded:

$C1_{\mu}$= 1.1258e-05

$C2_{\mu}$= .485

$C1_{k}$= 132.7906

$C2_{k}$= .5

When these above constants are used within the Cantera model, excellent agreement between heat transfer results was achieved to that of past Chemkin results. The LTC index for CR9.5 with pure n-Heptane, which was determined with Chemkin, was found to be the same value with Cantera-based results.

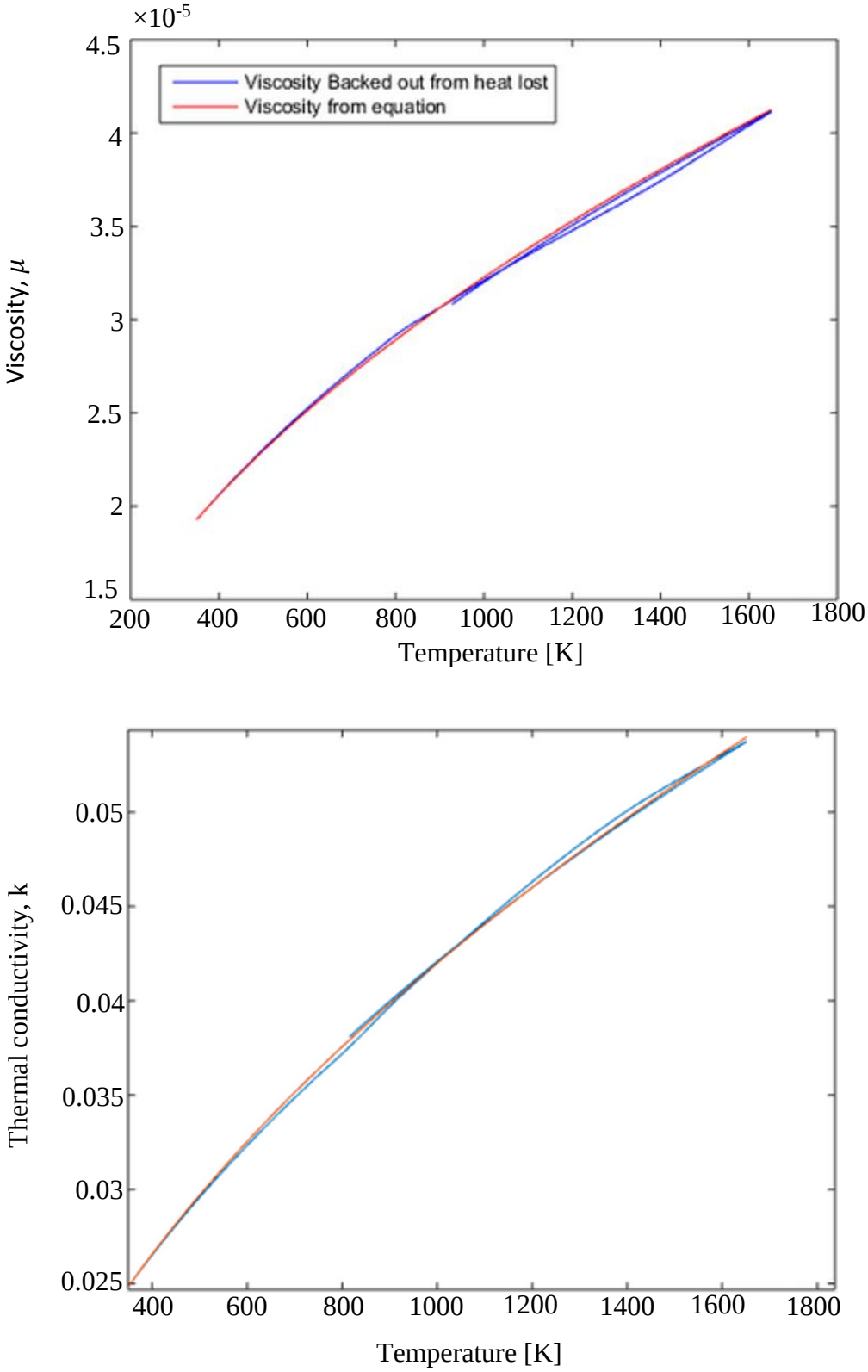

Figure 44: Chemkin viscosity (top) and thermal conductivity (bottom) back-solved from heat transfer correlations. A best-fit line is fit to solve for the constants used within Chemkin.

A5 Chemkin vs. Python Comparison Studies

A5.1 Adiabatic compression

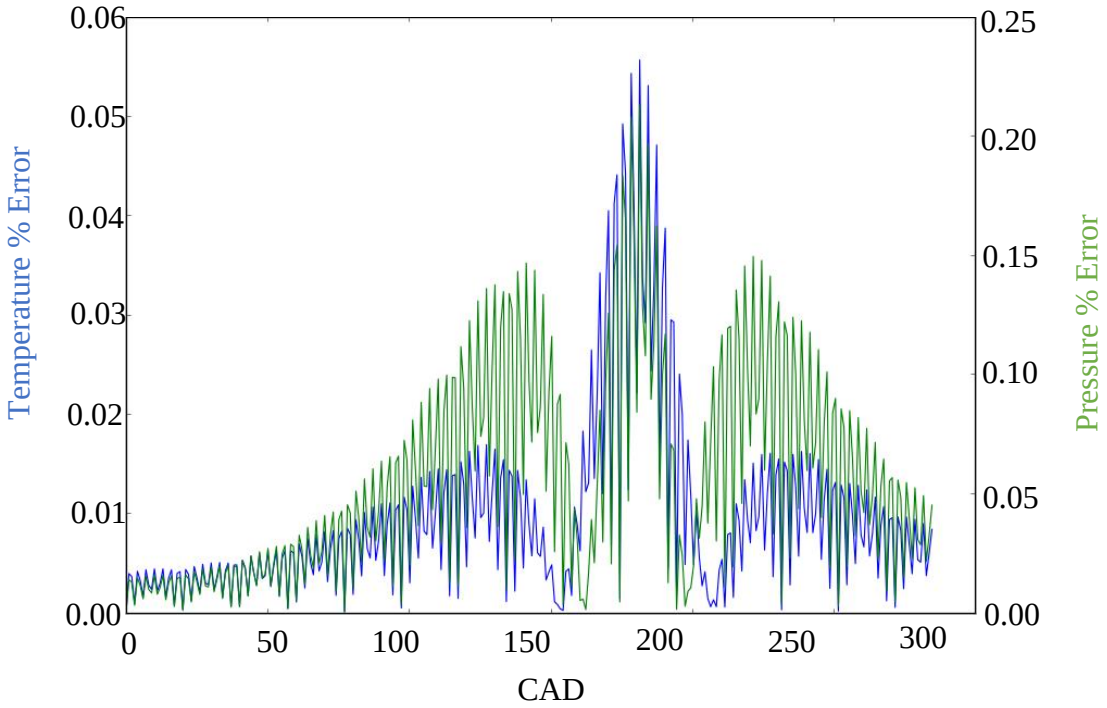

Figure 45: Adiabatic compression results of Python/Cantera model vs. Chemkin. % Error of temperature and pressure of Python/Cantera vs. Chemkin for Adiabatic compression results

A5.2 Adiabatic combustion

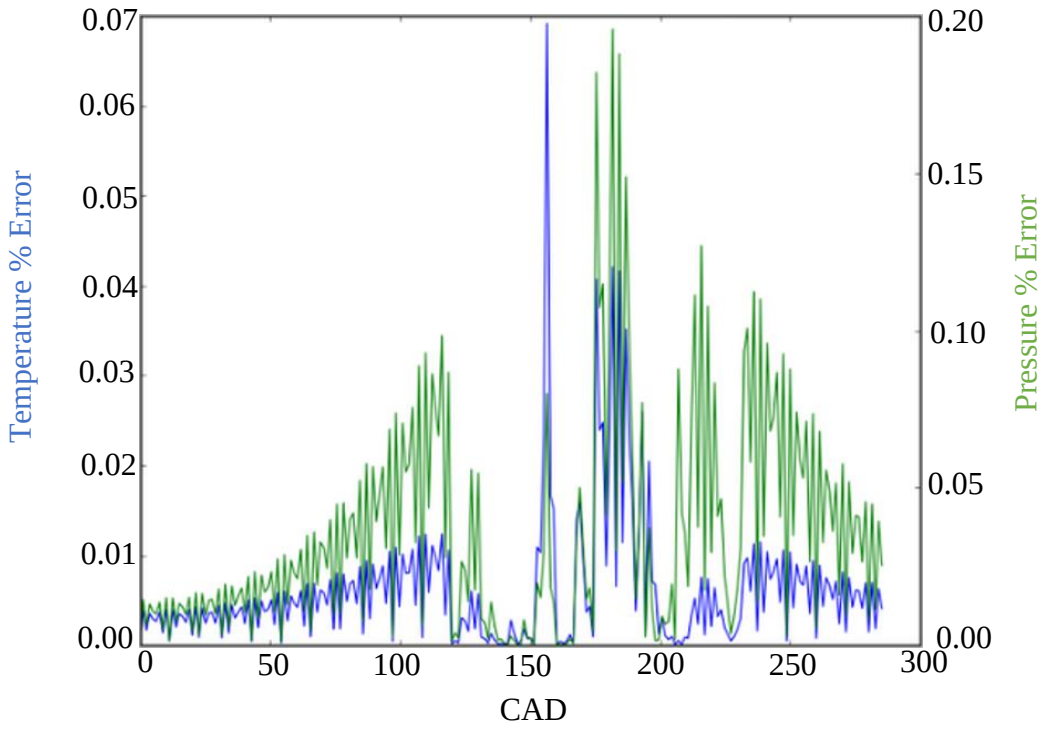

Figure 46: Adiabatic combustion of pure n-heptane results of Python/Cantera model vs. Chemkin. % Error of temperature and pressure of Python/Cantera vs. Chemkin for Adiabatic combustion of pure n-heptane results.

A5.3 Combustion with Heat Loss

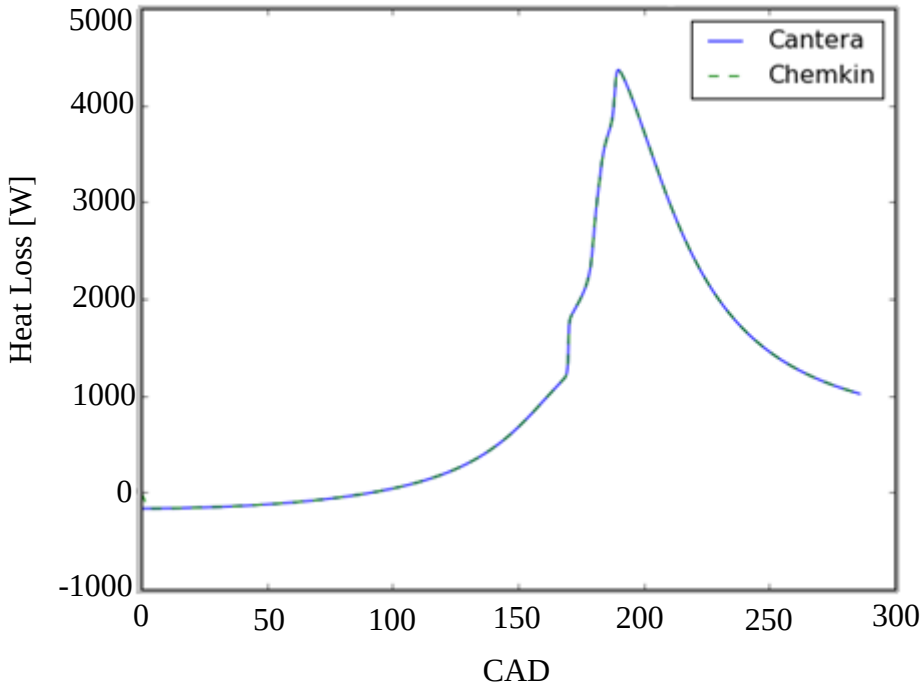

Figure 47: Combustion with heat loss of pure n-heptane: heat loss results of Python/Cantera model vs. Chemkin.

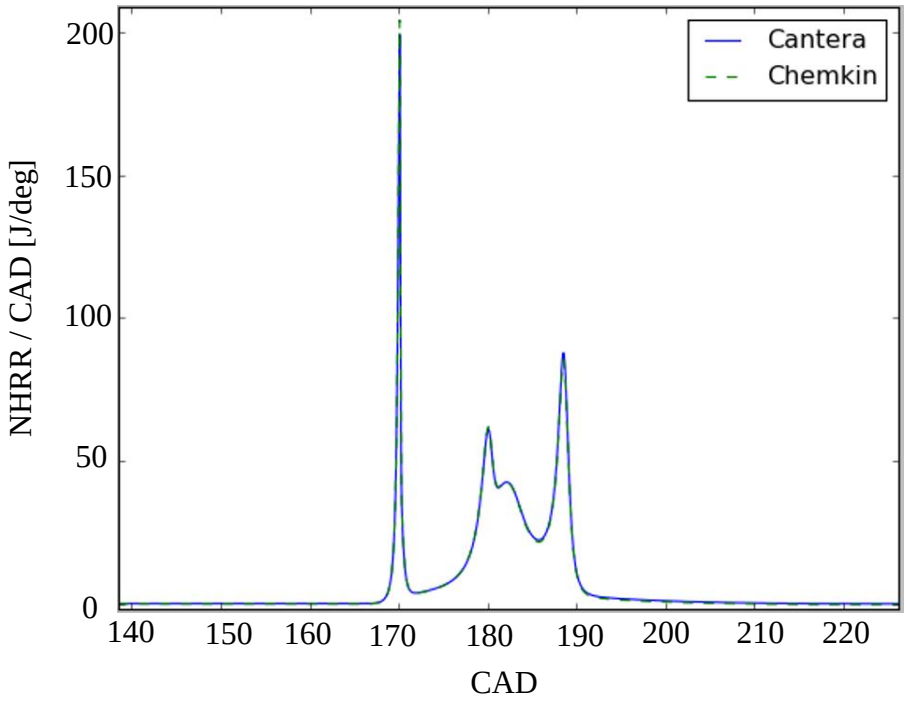

Figure 48: Combustion with heat loss of pure n-heptane: net heat release per crank angle results of Python/Cantera model vs. Chemkin.

Appendix B: Infrared Spectroscopy Theory / Post-processing

B1 Beer's Law

Beer's law describes how the attenuation of light through materials depends on the sample thickness and concentration of the material. The thickness, i.e., path length, dependence was discovered independently by Pierre Bouguer and Johann Lambert in 1729 and mid-1750s, respectively [97]. Following this, August Beer contributed the concentration dependence in 1852 [97] . These concepts combined together create the Beer-Lambert-Bouguer law, or simply Beer's Law. The mathematical derivation of the law can be found in many texts (e.g., [95], [96]), but will briefly be presented here.

Take a beam of coherent light at some wavelength and intensity, I_ν, incident upon a material of some thickness (visualized below). The incoming radiation will pass through the material and be absorbed to some degree, indicated by the change of a finite amount of intensity dI_ν.

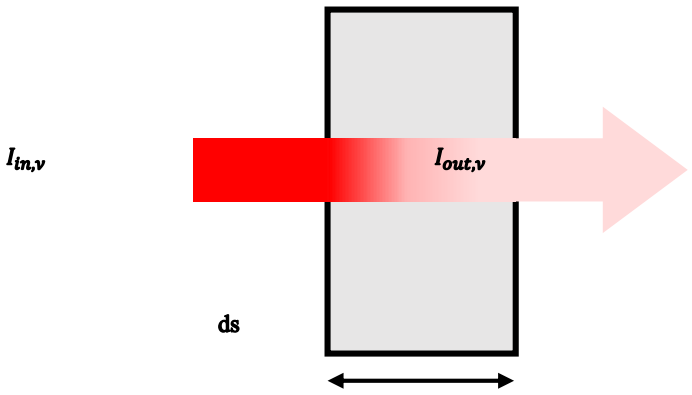

Figure 49: Illustration of attenuation of light through an absorbing medium for Beer's law

An energy balance can be performed that equates the energy in, out, and absorbed within the control volume:

$$dI_\nu = I_{in,\nu} - I_{out,\nu}$$

The addition of the attenuation coefficient, β, will be introduced to represent the scalar amount that could be multiplied by the incident radiation to retrieve the exiting radiation after absorption

occurs at some differential thickness, *ds*. Note that it has been assumed that the material under consideration has no fluorescent qualities and Mie scattering can be ignored.

$$dI_\nu = -\beta_{e,\nu,T}\, I_{in,\nu} ds$$

Rearranging I_{in} to the left side of the equation and integrating results in,

$$\ln\left(\frac{I}{I_{in}}\right)_\nu = \ln(T) = -\beta_{e,\nu,T} s$$

where transmittance has been defined as the ratio of I to I_{in}. Now, defining the absorbance, A, and separating the attenuation coefficient out to account for concentration and molecular size effects:

$$-\ln(T) = A = \sigma_{\nu,T}\, c\, s \qquad (B.1)$$

The absorbance represents the amount of light absorbed by the material, and is linearly proportional to the path length and attenuation coefficient. The attenuation coefficient represents the product of the concentration of the sample, *c*, and the effective cross sectional area of the absorbing material, $\sigma_{\nu,T}$.

B2 Fourier Transform Infrared Spectroscopy (FTIR)

The working principle of FTIR for collecting IR data is based on the Michelson Interferometer. In a Michaelson Interferometer, light is split into two beams at a beam splitter, where one beam is reflected back from a stationary mirror and the other from a moving mirror. The moving mirror acts to delay the beam's arrival to create destructive and constructive interference of the light wave depending on wavelength and the mirror position. This makes it possible to take the polychromatic light source and modulate many different wavelengths. With the moving mirror's known position as a function of time, the Fourier transform can be taken of the interferogram—the light intensity as a function of mirror position—to determine the frequency of the generated light.

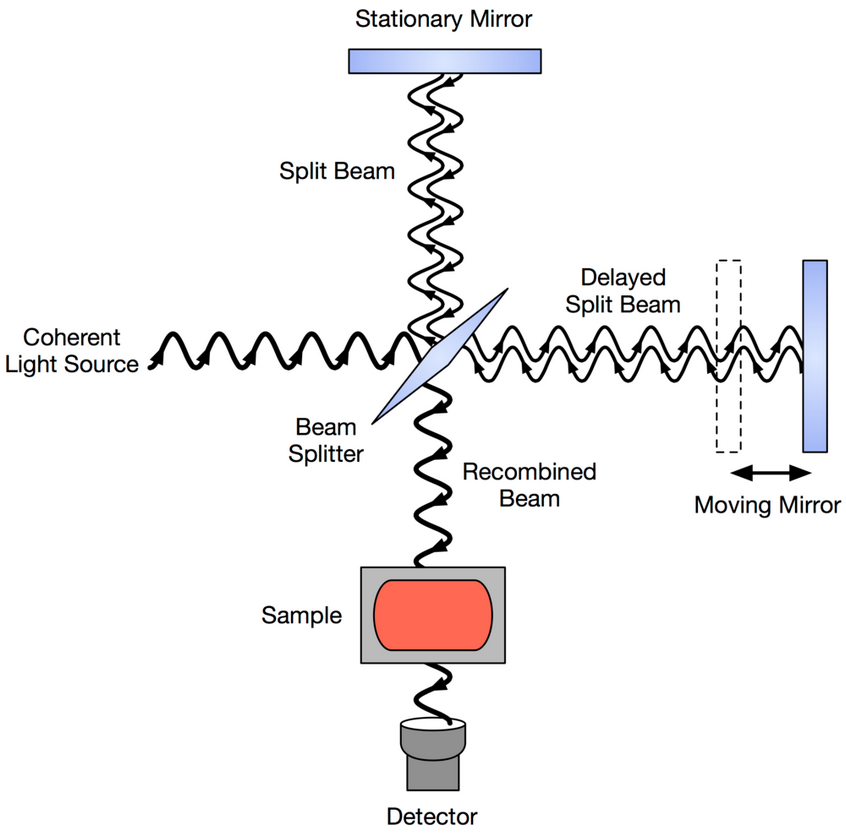

Figure 50: working principle of Michelson Interferometer. [117]

A scan of mirror positions and light intensity are taken without a sample (Io) in the FTIR beam path and with a sample in the beam path (I). These two signals are ratioed as described in Equation B.1 and used to determine wavelength specific absorbance. Two common approaches for collecting sample measurements with and FTIR are explained next.

B2.1 Transmission

Transmission-based FTIR (hereafter referred to as transmission) is a *quantitative* IR collection technique, in which the light physically passes through a sample. With transmission, it is possible to know with certainty the important physical parameters that are present in Beer's law. These properties are the path length—controlled by the geometry of the sample container—and the concentration(s) (i.e., density) of the sample of interest. With these parameters, as well as other parameters and relations to account for reflection losses, it is possible to determine the effective absorbing cross-sectional area, i.e., molar absorption coefficient, which is a quantitative measure of the optical absorbing capability of that medium. Since it is possible to determine a

quantitative measure of the optical property of the material, transmission is a quantitative spectroscopic technique. The post-processing theory for transmission data will now be discussed.

B2.1.1 Post Processing Transmission Data

In liquid transmission experiments the sample is confined between two optically transparent windows, such as potassium bromide (KBr), at short path lengths. First, determining the path length of the sample container will be discussed, followed by how to account for ideal and non-ideal reflection losses caused by the cell windows.

The path length is determined by counting the interference fringes caused by internal reflections between the two cell windows (i.e., an etalon). To do this, light (I_o) is passed through the atmosphere, and then again (I) through the cell windows with no sample inside. Light reflects at the boundaries of disparate media, which causes destructive interference between the two windows in addition to the interference within the cell windows itself. However, the windows are sufficiently thick and cause a large overtone that is not captured in the spectral range under consideration. The path length is determined using

$$L = \frac{M}{2(\nu_1 - \nu_2)}, \tag{B.2}$$

where *M* is the number of intensity peaks between optical frequencies ν_1 and ν_2. Because of the interference pattern created (see Figure 51 for an example), it is possible to determine the path length. These interference fringes also produces noise in the absorbance data if the baseline measurement includes this.

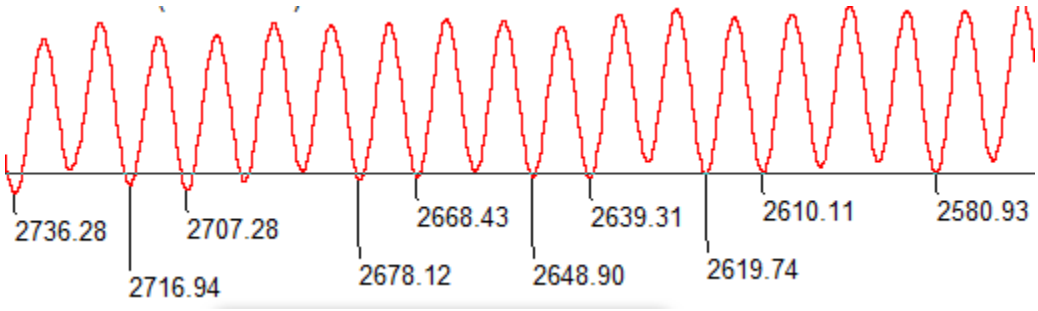

Figure 51: 0.515 [mm] path length interference fringe pattern. Peak values, labeled with associated frequencies, were determined with OMNIC™ series software [118].

Therefore, the baseline used for these experiments (I_o) is that of only the atmosphere in the beam's path with no cells, and the sample measurement (I) is with the windows and sample. This is done because the liquid samples of interest have an index of refraction close to that of the potassium bromide cells, and thus the interference fringes are negligible. When the liquid sample is inside the transmission windows no fringes are created. As a result, only the atmosphere for the baseline measurement is used when measuring liquid hydrocarbons. Otherwise, when the baseline (I_o) and measurement (I) are ratioed the interference fringe pattern would appear in the post-processed data.

Path length measurements of less than $15\ [\mu m]$ were performed for a few highly absorbent samples. In this case, the liquid is pressed between the two transmission slides directly (i.e., with no spacer), and as a result there is no way to determine the path length by way of counting the interference fringes. Instead, the path length is linearly extrapolated based on the absorbance measurements from the longer path length experiments at non-saturated absorbing frequencies. For these shorter path length cases, the value of path length was calculated using

$$L_{short}[cm] = \frac{A_{short}(\nu)-b}{A_{long}(\nu)-c} L_{long} \tag{B.3}$$

where the constants *b* and *c* represent values that correct non-absorbing frequencies to zero absorbance; this serves as a first-order approximation to correct for the reflection losses from the lenses. This is only done for path length measurements less than 15 microns. The process for longer path lengths will be discussed next.

With the path length determined experimentally (or extrapolated for small path lengths), the required parameters for Beer's law are all known. However, the absorbance hides an additional term: the amount of light reflected by the lenses themselves. Accounting for these reflection losses is important as they inflate the absorbance attributed to the sample, since in reality a portion of the measured absorbance is reflected away from the light detector and not in fact absorbed by the material at all. These windows reflect some of the incident radiation due to the difference between the refractive indices when mediums change (i.e., air to slides, slides to liquid sample). Therefore, the light captured by the detector has been attenuated by the medium *and*

the light lost due to reflection—and therefore the measured absorbance is larger than the true value for the absorbing medium. The work of Bertie et al. (see Bertie et al [102], and Porter et al. [99], [100]) presents the theory, solution procedure, and developed software to account for ideal and non-ideal losses due to reflection for this application. A complete summarization of the referenced work [99], [100], [102] is presented here.

In order to account for light lost due to reflection in the transmission slides, it is important to leverage other developments from both the particle and wavelike behaviors of light. The remainder of this section will introduce the theory that makes it possible to account for these reflection losses: guided by the useful summarization by Porter et al. [99] of the collective efforts of Bertie, Hawranek, Keefe, and many others towards liquid transmission I.R. spectroscopy. First, an introduction of the real and imaginary optical refractive index will be introduced. Second, the imaginary refractive index will be related to the molar absorption coefficient present in Beer's law. Third, Kramers–Kronig integration will be established as a means to relate the real and imaginary optical constants. Lastly, the step-by-step process of how all these relations allow us to determine ideal and non-ideal reflective losses in the transmission lenses will be laid out in full detail—the entirety of this process to be discussed is known as the *optical constant* method.

The frequency dependent optical constants for materials are referred to as the real and imaginary refractive indices, which sum together to form the complex refractive index [119].

$$m(\nu) = n(\nu) + ik(\nu) \tag{B.4}$$

The real refractive index, $n(\nu)$, is understood by the physical distortion of light at the interface of disparate media, e.g., when light "bends" as it transitions from air to water. The imaginary refractive index, i.e., absorption index $k(\nu)$, is related to the molar absorption coefficient and measures the attenuation of electromagnetic radiation propagating through an absorbing medium. The above optical constant relation may be better understood by comparing it with the solution of harmonic oscillations for a damped-mass-spring system; which has real (e.g., physical position, much like the physical distortion of light) and imaginary portions (e.g., energy decay through absorbance by the dampener, much like the energy decay of light by being absorbed by the medium). The imaginary refractive index, the value attributed to energy being absorbed by the medium, will now be related to the molar absorption coefficient.

It is possible to derive Beer's law from Maxwell's equations, which also gives the useful expression relating the imaginary refractive index to the molar absorption coefficient. Moreover, it illustrates an intuitive definition of the refractive index(s), and shows that real and imaginary refractive index are dependent on each other—interested readers are directed to the thesis of Porter et al. [99] for a full, succinct derivation. The end result is summarized as follows:

$$\sigma(\nu) = \frac{4\pi\nu k(\nu)}{n_a} \tag{B.5}$$

where σ is the effective molar cross sectional area in cm^2/mole (e.g., molar absorption coefficient), and n_a is the concentration in mole/cm^3. The final physical relation required for the optical constant method will now be introduced.

With the above expression (eq. B.5) derived from classical mechanical arguments, another useful relation—derived with little physical intuition—known as the Kramers–Kronig relations are presented. The relation states that an imaginary solution for a given system can be retrieved from the real solution, and vice versa. Again, for illustrative purposes, imagining a damped-mass-spring system, if one knows the position as a function of time the imaginary portion (e.g., energy dissipated in the dampener) can be deduced. This process can be done analytically with the Kramers–Kronig relations—also known as dispersion relations specific to the application of optics—and is accomplished through integral equations. The form of the following relation is specific to optical constants, though is not too far off from the general relation:

$$n(\nu_i) - n(\infty) = \frac{2}{\pi}\, P \int_0^\infty \frac{\nu k(\nu)}{\nu^2 - \nu_i^2}\, d\nu \tag{B.6}$$

where *P* indicates taking the principal value of the improper integral, due to the singularity at $\nu = \nu_i$ [120].

Recall that the purpose of the optical constant method is a means to correct the absorbance data to account for reflection losses in the transmission lenses. The reasons for this process will be discussed due to unknown variables and underlying assumptions; first, the unknown variables and the procedure to solve for them are to be outlined.

The Fresnel relations determine the *ideal* loss of transmission of light from the lenses—not defined in the *Optical Constant Method –Equations* section. These relations, and the specific application towards thin film spectrophotometry, can be found in Hawranek [119]. The Fresnel relations solve for the fraction of electromagnetic radiation that is lost, i.e., refracted away from its original path at the interface of disparate media due to the real index of refraction differences. Excluding some variables, the relations depend on the real index of refraction of both substances (the lenses and the sample). The real index of refraction as a function of wavenumber, i.e., n-spectrum, for the lenses are typically known and can be found readily, while that of the samples are typically unknown—especially for liquid hydrocarbons which have not extensively been studied quantitatively [99], [100]. Since the real index of refraction for the sample is unknown, it has to be determined from the series of equations developed a priori leveraging the actual absorption measurements of the sample.

Since the real refractive index is unknown for the sample, a first approximation is found from the uncorrected absorbance data by utilizing Maxwells' relation and performing Kramers–Kronig integration (eq. B.5 and B.6, respectively). This gives an approximate n-spectrum as it still represents non-ideal and ideal losses present from the uncorrected absorbance data. Nonetheless, this n-spectrum can be used in the Fresnel relations (see Hawranek [119]) to find the approximate ideal reflection losses. These losses are subtracted from the data and referred to as a baseline correction. An example n-spectrum can be seen in Figure 52. In addition to ideal losses caused by reflection, non-ideal losses will be explained and a solution process to account for this will now be introduced. This process first requires an explanation of the importance of using multiple path length measurements.

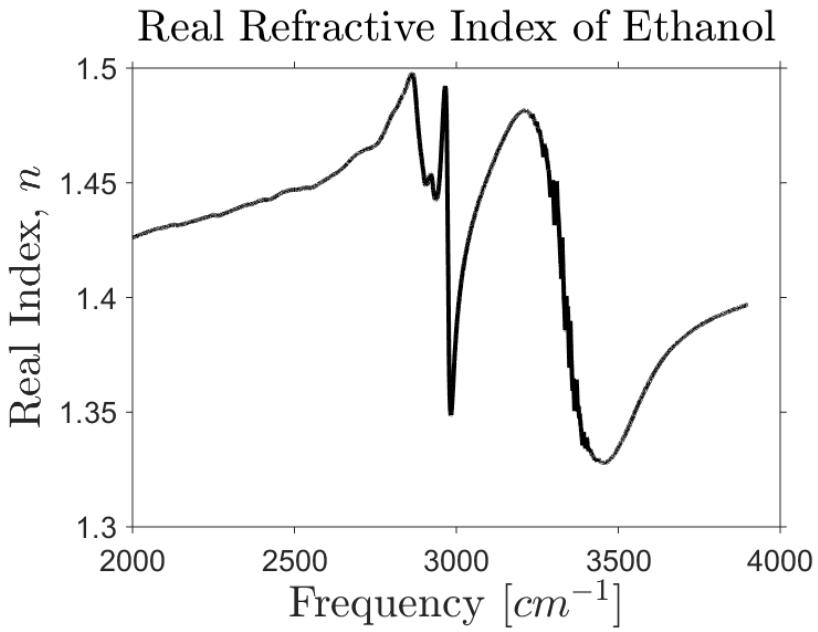

Figure 52: n-spectrum of Ethanol near 3000 [cm^{-1}] from Daly

Multiple path length absorption measurements and why such a process is used to create a composite spectrum, in addition to how it provides a means to solve for non-ideal losses will now be discussed. Liquid hydrocarbons have high molar absorption coefficients at absorbing frequencies, and therefore short path lengths are required for measurements to avoid spectral saturation—i.e., absorbance values greater than three. Guided by literature [99], [100], an absorbance value greater than three is taken to be saturated data, which corresponds to 0.1% of the incident radiation transmitting through the sample. It was found from experiments in this work that our instrument followed this "rule of thumb", as anything higher than an absorbance of three was beyond the instrument resolution; e.g., see Figure 53 and notice the longer path lengths do not follow the trends of the unsaturated shorter path length samples. Therefore, long path length measurements give better resolution at lower-absorbing frequencies, and the short path lengths capture the higher absorbing frequencies without spectral saturation. In the end, a composite spectrum is constructed that combines the multiple spectra of each path length. Note that in Figure 53, the ethanol absorbance at the longer path lengths at highly absorbing frequencies cannot be used as the data is completely saturated, whereas as the path length decreases the saturation is eliminated. However, at the smaller path lengths some of the moderately absorbing peaks are barely captured. The process by which the longer path length samples are used to correct for non-ideal losses will now be presented.

Longer path lengths are sought after, as they give better spectral resolution (i.e., signal to noise ratio) at lower-absorbing frequencies that cannot be achieved with short path length measurements. More importantly, non-ideal losses make little-to-no contribution to the overall losses (e.g., due to cell positioning, which enhance non-ideal losses as the path length decreases) in long path length measurements [102]. It is assumed that for long path lengths the light absorbed is the sum of the ideal transmission losses through the cell and that by the sample, since

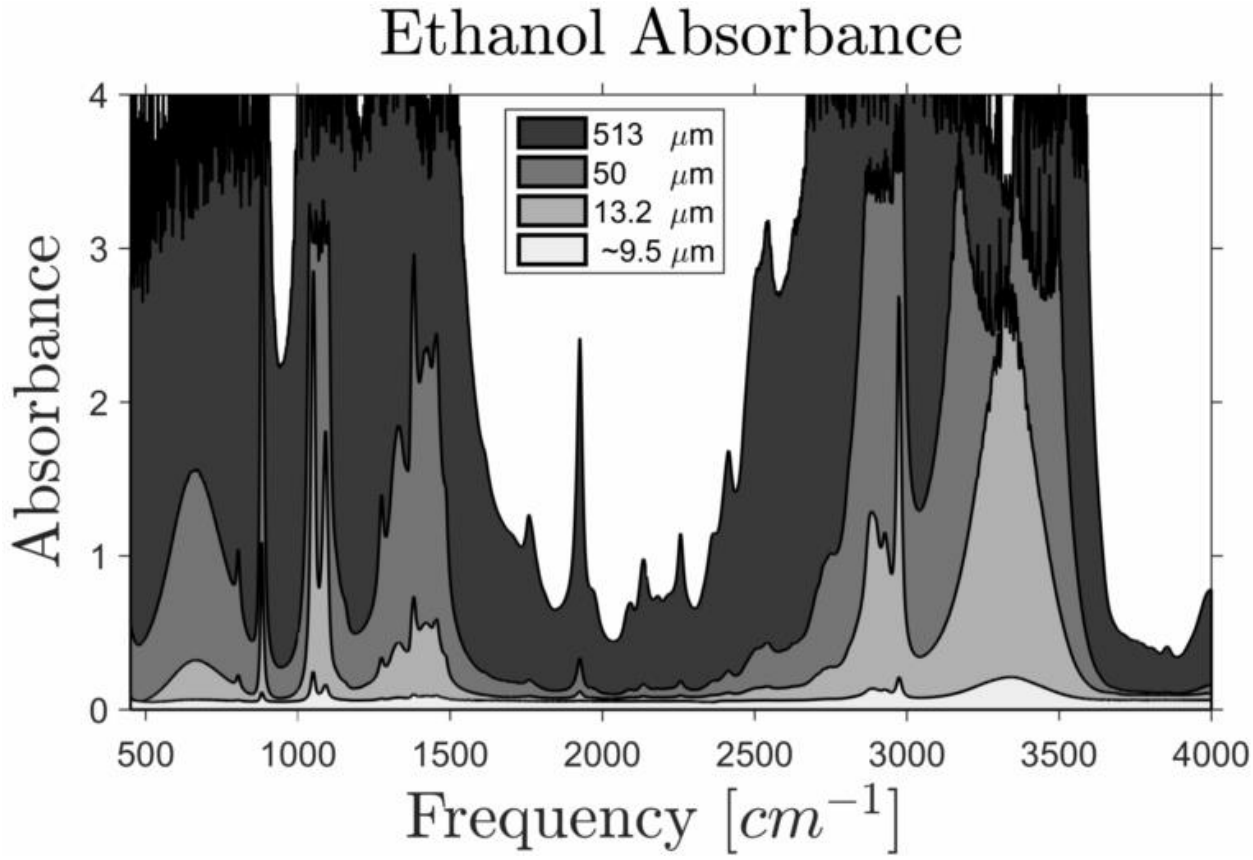

Figure 53: Ethanol IR Absorbance Spectrum at various path lengths from Daly

Understanding the importance of multiple path length measurements brings us to the idea of Anchor point correction. Invented by Bertie et al. [102], this method leverages the longer path length samples (that lack non-ideal losses) to determine the non-ideal losses in the shorter path length samples that we require to avoid saturated data. Before continuing, the linear absorption coefficient is introduced:

$$K(\nu) = \frac{A(\nu)}{s} \tag{B.7}$$

The linear absorption coefficient is the absorption (from Beer's law in Equation B.1) normalized by the path length used at that measurement. Thus, if one were to compare the linear absorption coefficient of various path lengths for a sample—for a constant concentration throughout—the coefficients should be equal. Linear absorption coefficients are determined at all the path length studies at *anchor points*—user-determined points where low-absorbing frequencies exist, as they are not saturated when long path lengths are used. Take, for example, ethanol, which is illustrated in Figure 55. A good "anchor point" for this example is seen to be around 2450 $[cm^{-1}]$, as the long path length data is not saturated at this point. With the linear absorption coefficients (K(ν)) at the chosen anchor point for all path lengths, the K(ν) values at long path lengths can be compared to the shorter path length samples. The difference between the "long" and "short" linear absorption coefficients gives the non-ideal losses present in the "short" samples at the anchor points. As a result, in addition to accounting for the ideal reflection losses via the Fresnel relations, it is now also possible to subtract out non-ideal losses present in short path length data. However, the non-ideal losses are only known at the anchor points, as at other frequencies the data is too saturated and therefore the linear absorption coefficient cannot be determined. Bertie et al. [102] uses a linear fit of the non-ideal losses between user-defined anchor points across the spectral breadth considered. With an understanding of how to account for ideal and non-ideal losses the solution procedure will now be laid out.

With the physical assumptions and mathematical relations laid out, a map of the solution process from the original, uncorrected data to the corrected absorbance data will be discussed. Guided by the original work of Bertie et al. [102], and the useful summary of that work from Porter et al. [99], the previously stated methodology can be implemented with updated software from Keefe et al. [101]. The software used (multiple windows executable files) is introduced as the solution process is explained, all of which can be found on the World Wide Web [121]. The solution process will now be laid out, and a flow chart illustrating the process is seen in Figure 54.

Using the developed in-house (OSU) Python script, the raw single beam spectra are first taken and converted to transmittance. Following this, the transmittance spectra are converted to absorbance for convenience: recall Beer's law relates absorbance to concentration, path length, and the molar absorption coefficient. These absorbance spectra are now ready to be corrected to account for the ideal and non-ideal reflection losses from the lenses. Before continuing, it will

be noted that the programs to be introduced are windows executable files (.exe) that read and write Meta files and access these when executed. To automate using these programs, another in-house Python script was created to efficiently read and write the Meta files and run the windows executable's—previously, the output of one program, being a Meta file, had to be manually entered into another Meta file repeatedly. The created Python program can be found in Appendix F.

Since the n-spectrum of the fuel sample is unknown, an approximation of the ideal losses cannot be determined via the Fresnel relations. To mitigate this, the uncorrected absorbance spectrum is converted to the imaginary index, $k(\nu)$, via Beers's law (eq. B.1) and Maxwell's relation (eq. B.5). The k-spectrum is then used with the Kramers–Kronig dispersion integral equation (eq. B.6) to calculate the n-spectrum—all these operations are internally done with EXPABS2K. These initial operations can be visualized as step #1 in Figure 54. With the approximate n-spectrum—now moving to step #2 in Figure 54—the Fresnel relations can be utilized to solve for the ideal losses. Subsequently the linear absorption coefficients, $K(\nu)$, in the long path length sample at the anchor points are determined: accomplished with BASELINE. At this point, the long path length sample have been corrected as the non-ideal losses are insignificant for this sample [102]. The non-ideal losses for the short path length samples can be found with anchor point correction, which will now be explained.

EXPABS2K accepts the linear absorption coefficients at the user-defined anchor points from program BASELINE. Recall the difference between the linear absorption coefficient of the long path length samples and the short path length samples yields the non-ideal losses [102]. The non-ideal losses are linearly interpolated between the anchor points and are subtracted from the data at each respective path length throughout the spectrum. Additionally, EXPABS2K performs a baseline correction for the short path length samples that are being corrected in this step. All absorbance spectrums for each path length have now been approximately corrected for ideal and non-ideal losses.

With the absorbance spectrums approximately corrected for reflection losses, the aforesaid process (step #2) is repeated. At this point, a more accurate n-spectrum and absorbance spectrum will be used to repeat step #2: indicated by the dashed lines seen in Figure 54. This will serve to better correct all path length samples. At convergence, which usually occurs in

about three to four iterations, the ideal and non-ideal losses have been adequately accounted for—convergence is defined when the residual of $K(\nu)$ for all anchor points between iterations are within 1E-4. At this point, absorbance spectrums of all path lengths are combined to create a composite spectrum, as illustrated in step # 3. This is done by taking the molar absorption coefficient spectra (or k-spectrum since they are related through Maxwell's relation) and filtering the data. Longer path length samples are preferred unless the data is spectrally saturated, i.e., the absorbance value is greater than three. If this is the case, the next shortest path length data at that frequency is selected, and so on. This corrected composite spectrum of all path lengths is the final result of the optical constant correction method for the k-spectrum. The final step is using Kramers-Kronig integration with the composite k-spectrum to determine the final n-spectrum: done with KKTRANS, a stripped down version of EXPABS2K, which accepts the k-spectrum as an input.

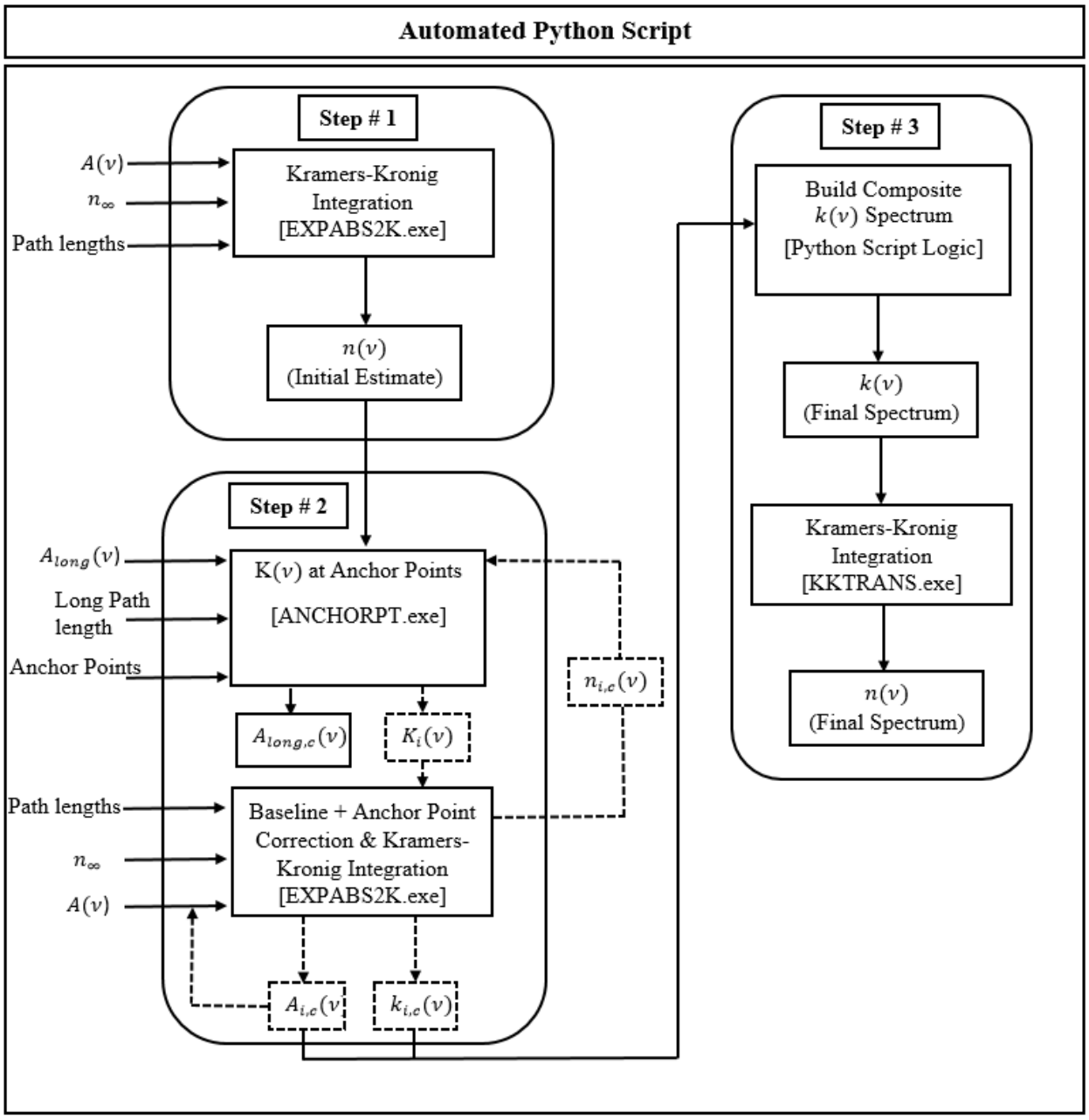

Figure 54: Flow chart of post-processing transmission-based IR absorbance data. This particular strategy is known as the optical constant method with anchor point correction. Dashed lines indicate process paths undergoing iteration. Subscripts "i" indicating the iteration count, and "c" indicating the spectrum has been corrected. After the convergence criteria is reached ($K_{i+1} - K_i \leq 1E\text{-}4$ for all anchor points), $k_{i,c}(\nu)$ and $A_{i,c}(\nu)$ spectra are sent to step # 3 where the final composite spectrum is created. Note: all programs based on log base 10 absorbance. EXPABS2K.exe has built in real refractive indices for various transmission slides, including KBr.

B2.2 Attenuated Total Reflectance (ATR)

In contrast to transmission spectroscopy, attenuated total reflectance (ATR) spectroscopy is a *qualitative* technique to gather IR absorption spectra. Instead of passing the light through the sample, the ATR process merely reflects light against a sample. This is again done with a Michelson interferometer, where the recombined light beam is passed at an angle into a crystal with a high real refractive index. As governed by the Fresenel relations, light can totally reflect at the interface of disparate media, i.e., where the crystal meets the sample, if the index of refraction difference between the crystal and the sample is large enough—hence the high index of refraction of ATR crystals. The critical angle θ_c necessary to ensure total reflectance is described with

$$\theta_c = \sin^{-1}\left(\frac{n_2}{n_1}\right) \qquad (B.8)$$

where n_2 and n_1 are the real refractive indices of the two media, where n_1 represents the media the light is first traveling in. The light comes into contact with the interface of the ATR crystal and the sample, which totally internally reflects back toward the light collector when the angle of incidence is greater than the critical angle θ_c. At this area of reflection an evanescent wave is produced, as predicted by the Maxwell equations. This evanescent wave, literally meaning a wave that tends to vanish, protrudes into the sample at an exponential decay with distance. The sample absorbs the energy of this wave, and as a result the intensity of the beam reflecting back towards the light detector decreases. Due to the nature of this evanescence wave there is uncertainty in the beam's path length, and thus ultimately the path length needs to be calculated from equations based on electromagnetic theory as it cannot be measured directly—therefore, ATR is primarily a qualitative technique. That said, there are multiple methods for determining the evanescent-wave-effective path length, as will be introduced in the following section.

B2.2.1 Post Processing ATR Data

The profile of evanescence waves decay exponentially into the sample. The depth of penetration can be estimated as the point at which the energy depletes to e^{-1} of its value at the surface of the wave [103]. This equation assumes that the efficiency of the polarizations for the instrumentation used are equivalent. The relation is as follows:

$$d_p = \frac{\lambda}{2\pi n_1 \sqrt{\sin^2\theta - n^2{}_{21}}} \qquad (B.9)$$

where λ is the frequency of light in cm, θ is the incident angle, n_1 is the real index of refraction of the sample, and n_{21} is the ratio of the sample and ATR crystal real refractive index.

In an attempt to better compare ATR and transmission absorptivity results, the "effective" path length was defined. This was defined as the equivalent path length in a transmission measurement that would lead to the same absorbency with ATR. The effective path length has parallel and perpendicular polarization contributions, developed by Harrick [92]–[94], which penetrate to different depths as determined by

$$d_{e\perp} = \frac{n_1^2 n_2 \cos\theta}{(n_1^2 - n_2^2)} \frac{\lambda}{\pi\sqrt{n_1^2 \sin^2\theta - n_2^2}} \qquad (B.10)$$

$$d_{e\parallel} = \frac{n_1^2 n_2 \cos\theta}{(n_1^2 - n_2^2)} \frac{\lambda}{\pi\sqrt{n_1^2 \sin^2\theta - n_2^2}} \frac{2n_1^2 \sin^2\theta - n_2^2}{(n_1^2 - n_2^2)\sin^2\theta - n_2^2} \qquad (B.11)$$

The effective penetration for an unpolarized IR beam is then the average of the parallel and perpendicular polarizations:

$$d_e = \frac{(d_{e\perp} + d_{e\parallel})}{2} \qquad (B.12)$$

Harrick assumed [92]–[94] that dispersion effects are small, i.e., the real refractive index of the absorbing sample is constant throughout the measurement range, and therefore it is implied that the absorbance is small—which is not the case for liquid measurements.

Lacey et al. recently applied the Maxwell relation along with Kramers–Kronig integration to ATR absorbance data [91], similar to the methodology described in the earlier transmission-based spectroscopy section in order to account for dispersion effects (i.e., the variation in n_2). The algorithm is proprietary, as it was developed for Thermo Fisher Scientific and is integrated into their spectroscopy software. Based on this, it seems as though ATR could be extended to be a more quantitative spectroscopic technique; this author would like to investigate this further for future work. Nonetheless, it is speculated that the solution procedure is similar to that outlined

for transmission-based data wherein the path length would also be iteratively determined along with the real refractive index until convergence.

Path length considerations for ATR are complicated and rely exclusively on theoretical equations. As a result, to the author's current knowledge, ATR is still primarily qualitative and as a result cannot be compared directly to transmission-based spectra—as mentioned, this would like to be investigated further as future work. Therefore, IR spectra intended for use by any kind of modeling need to be collected exclusively with either ATR or transmission.

Appendix C: LTC Indices / Operating Envelopes

Table 7: LTC indices of fuels considered in this work

Fuel Name	RON	MON	LTC index
Two Component Mixtures			
PRF0	0	0	15.36
PRF20	20	20	34.2
PRF40	40	40	32.75
PRF55	55	55	28.4
PRF70	70	70	26.77
PRF85	85	85	23.3
PRF100	100	100	23.76
H90T10	14.1	12.6	37.72
H80T20	27.7	24.8	41.58
H60E40	71.4	65	35.57
Three Component Mixtures			
TRF1	39	37	35.2
TRF2	88.8	84.8	15.14
TRF3	76.2	70.9	17.46
TRF4	82.8	80.7	34.1
ERF1	78.7	76.7	23.22
H20T20	84.9	81.8	15.1
H20T40	89.8	82.9	14.93
H20T60	93.9	83.9	13.79
H20E1	80.5	80.4	14.75
H20E5	84.4	83.2	14.93
H20E10	87.9	85.6	14.13
H20E20	94.1	88.2	13.87
H30E1	70.6	70.2	23.05
H30E5	73.2	73.1	23.1
H30E10	78.7	76.7	22.01
H30E20	85.1	81.2	20.56
H40E60	94.4	82.8	21.92
H45E55	89.7	80.4	22.9
H50E50	84.7	76.8	25.46
H55E45	78	71.6	21.54
Four Component Mixtures			
TERF1	67.1	63.7	23.12
TERF2	94.7	88.5	14.95
TERF3	97	87.6	14.76
TERF4	80.8	73	24.21
TERF5	79.6	74.5	24.23

TERF6	85.3	78.6	15.24
TERF7	83.8	78.1	22.86
TERF8	71.8	65.8	27.24
TERF9	81.6	77.9	22.34
TERF10	90.2	84.1	14.83
TERF11	83.8	78.2	22.86
TERF12	78.3	74.3	25.95
TERF13	86.1	83.6	14.89
TERF14	74.8	68.9	27.31
TERF15	92.3	86.4	15.27
TERF16	87.9	81.4	23
TERF17	83.8	78.1	22.86
TERF91-15	91	88.4	14.84
TERF91-15-10	97.8	91.7	6.61
TERF91-15-20	102.6	93.2	14.73
TERF91-15-40	107.1	93.6	14.88
TERF91-15-60	107.7	93.6	14.79
TERF91-15-80	1078	91.7	13.85
TERF91-30	91.3	86.1	14.83
TERF91-30-10	97	89.4	14.78
TERF91-30-20	101.4	91.1	14.92
TERF91-30-40	106	91.2	14.93
TERF91-30-60	107.1	92	14.88
TERF91-80	107.5	91.4	13.7
TERF91-40	91.1	86.5	14.75
TERF91-40-15	96	87.2	14.88
TERF91-40-20	100.2	89.1	14.08
TERF91-45-40	104.6	90.9	14.77
	Neat Fuels		
IC5H12 (iso-pentane)	92	90	15.22
methylcyclohexane	74.1	74	11.44
C7H16-2 (2-	42.4	46.3	24.62
	TRF Mixtures with iso-pentane		
S1	?	?	22.38
S4	?	?	28.61
S6	?	?	25.68
S7	?	?	22.08
	TRF Mixtures with m-xylene		
X1	?	?	21.14
X2	?	?	22.34
X3	?	?	14.31
X4	?	?	19.18
X5	?	?	21.31
X6	?	?	22
X7	?	?	22.44

X8	?	?	21.5
X9	?	?	18.67
X10	?	?	9.65
X11	?	?	23.43
X12	?	?	26.99
X13	?	?	28.44
X14	?	?	21.56
X15	?	?	25.45
X16	?	?	26.03
X17	?	?	20.76
X18	?	?	21.8
X19	?	?	26.89
X20	?	?	25.53
X21	?	?	25.55
X22	?	?	29.36
X23	?	?	26.53
X24	?	?	27.05
X25	?	?	27.08
X26	?	?	28.97
X27	?	?	26.92
X28	?	?	26.99
X29	?	?	26.32
X30	?	?	24.06
	FACE Gasolines		
FACE A	84	84	14.19
FACE C	84	84	14.13

Appendix D: Fuel List

Table 8: Pure components procured in this work. Infrared absorption spectra of these were acquired via. ATR spectroscopy. Note (*) indicates multiple sources were averaged.

Pure Components

Fuel name	RON	Carbon Number	Class
heptane	0	7	Alkane
2-methylbutane	92[122]	5	Alkane
2-methylpentane	73.4[123]	6	Alkane
3-methylpentane	74.5[123]	6	Alkane
2-methylhexane	42[123], [124]*	7	Alkane
3-methylhexane	52 [123], [124]*	7	Alkane
2,4-dimethylpentane	83.1 [123], [124]*	7	Alkane
2,3-dimethylpentane	91.1 [123], [124]*	7	Alkane
2,5-dimethylhexane	55.3 [123], [124]*	8	Alkane
2,4-dimethylhexane	65.2 [123], [124]*	8	Alkane
3-ethyl-2-methylpentane	87.3 [123], [124]*	8	Alkane
isooctane	100	8	Alkane
toluene	113[23]	7	Aromatic
xylenes	114[125]*	8	Aromatic
1,2,3-trimethylbenzene	100.5[123]	9	Aromatic
4-ethyl-m-xylene	100.6[123]	9	Aromatic
2-ethyl-p-xylene	100.6[123]	9	Aromatic
1,2,4-trimethylbenzene	101.4[123]	9	Aromatic
cumene	102.1[123]	9	Aromatic
1,3,5-trimethylbenzene	106[123]	9	Aromatic
1,2,3,4-tetrahydronaphthalene	96.4[123]	10	Aromatic
2-propyltoluene	100.3[123]	10	Aromatic
1,2,3,4-tetramethylbenzene	100.5[123]	10	Aromatic
cyclopentane	100.1[123]	5	Naphthene
Cyclohexane	80.7[126]	6	Naphthene
methylcyclohexane	74.1[127]	7	Naphthene
butylcyclohexane	63.8[128]	9	Naphthene
1-Pentene	90[125]	5	Olefin
2-methyl-2-butene	97.3[123]	5	Olefin
2-pentene	98[122]	5	Olefin
2-methyl-1-butene	100.2[123]	5	Olefin
1-hexene	74.9[127]	6	Olefin
diisobutylene	103.8[129]	8	Olefin
ethanol	108.5 [45], [130], [131]*	2	Alcohol

Table 9: Fuel mixtures considered in this work created by the author for LTC index to inform chemometric model.

Fuel Name	n-heptane	isooctane	toluene	ethanol	2-methylbutane	Xylene(s)
S1	50	0	0	0	50	0
S2	30	0	20	0	50	0
S4	30	20	0	0	50	0
S6	70	0	10	0	20	0
S7	50	20	20	0	10	0
X1	30	20	0	0	0	50
X2	30	20	5	0	0	45
X3	30	20	10	0	0	40
X4	30	0	20	0	0	50
X5	30	5	20	0	0	45
X6	30	10	20	0	0	40
X7	30	20	20	0	0	30
X8	30	10	10	0	0	50
X9	30	5	5	0	0	60
X10	30	0	0	0	0	70
X11	50	0	0	0	0	50
X12	50	20	0	0	0	30
X13	50	20	5	0	0	25
X14	50	20	10	0	0	20
X15	50	0	20	0	0	30
X16	50	5	20	0	0	25
X17	50	10	20	0	0	20
X18	50	20	20	0	0	10
X19	50	10	10	0	0	30
X20	50	5	5	0	0	40
X21	70	0	0	0	0	30
X22	70	20	0	0	0	10
X23	70	0	20	0	0	10
X24	70	10	0	0	0	20
X25	70	0	10	0	0	20
X26	70	10	10	0	0	10
X27	70	5	5	0	0	20
X28	70	5	0	0	0	25
X29	70	0	5	0	0	25
X30	90	0	0	0	0	10

Table 10: "Baseline" fuel list for RON study. Hydrocarbon component compositions for the fuel mixtures can be found in the literature. "IBX" fuels indicate iso-octane/diisobutylene mixtures, where the number indicates the diisobutylene composition in vol%. "Perez" fuels 1, 9, 17, and 21 purposefully left out (pure components).

Perez et al. [46]	Truedsson [23]		Foong et al. [45]					Others
Perez 2	H20T10	TERF1	T-E10	TERF91-30 -80	PRF10-E60	PRF40-E60	PRF80-E40	IB20
Perez 3	H20T20	TERF2	T-E20	TERF91-45 -10	PRF20-E20	PRF40-E80	PRF80-E60	IB40
Perez 4	H20T40	TERF3	T-E40	TERF91-45 -20	PRF20-E30	PRF50-E10	PRF80-E80	IB60
Perez 5	H20T60	TERF4	T-E60	TERF91-45 -40	PRF20-E40	PRF50-E20	PRF90-E5	IB80
Perez 6	H20E1	TERF5	T-E80	TERF91-45 -60	PRF20-E50	PRF50-E30	PRF90-E10	
Perez 7	H20E5	TERF6	TERF91-15	TERF91-45 -80	PRF20-E60	PRF50-E40	PRF90-E20	
Perez 8	H20E10	TERF7	TERF91-30	PRF0-E30	PRF20-E80	PRF60-E10	PRF100-E10	
Perez 10	H20E20	TERF8	TERF91-45	PRF0-E40	PRF30-E10	PRF60-E20	PRF100-E20	
Perez 11	H30E1	TERF9	TERF91-15 -10	PRF0-E50	PRF30-E20	PRF60-E30	PRF100-E40	
Perez 13	H30E5	TERF10	TERF91-15 -20	PRF0-E60	PRF30-E30	PRF60-E40	PRF100-E60	
Perez 14	H30E10	TERF11	TERF91-15 -40	PRF0-E70	PRF30-E40	PRF60-E60	PRF100-E80	
Perez 15	H30E20	TERF12	TERF91-15 -60	PRF0-E80	PRF30-E50	PRF60-E80		
Perez 16	H40E60	TERF13	TERF91-15 -80	PRF0-E90	PRF40-E10	PRF70-E10		
Perez 18	H45E55	TERF14	TERF91-30 -10	PRF10-E20	PRF40-E20	PRF70-E20		
Perez 19	H50E50	TERF15	TERF91-30 -20	PRF10-E30	PRF40-E30	PRF70-E30		
Perez 20	H55E45	TERF16	TERF91-30 -40	PRF10-E40	PRF40-E40	PRF80-E10		
		TERF17	TERF91-30 -60	PRF10-E50	PRF40-E50	PRF80-E20		

Appendix E: IR Absorbance Spectra

E1 Transmission-based

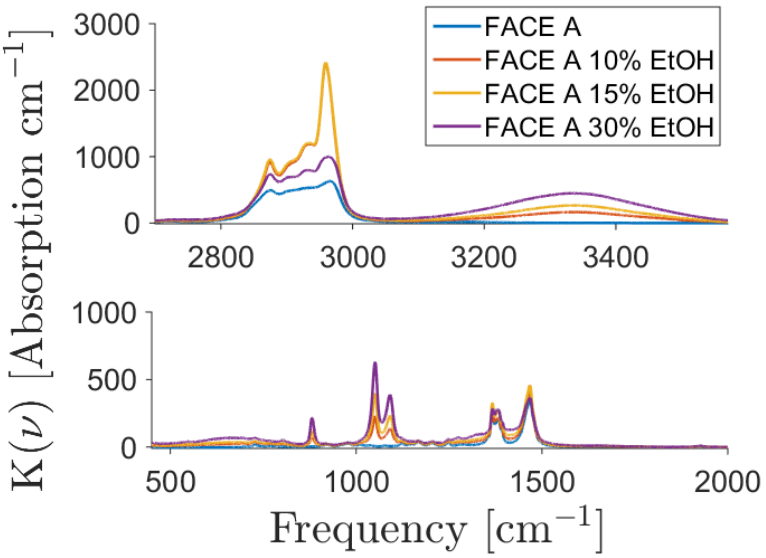

Figure 55: Linear Absorption Coefficient Spectrum of FACE A and various blends with ethanol. Non-linear blending effects are easily seen in vicinity of 2900 cm^{-1}

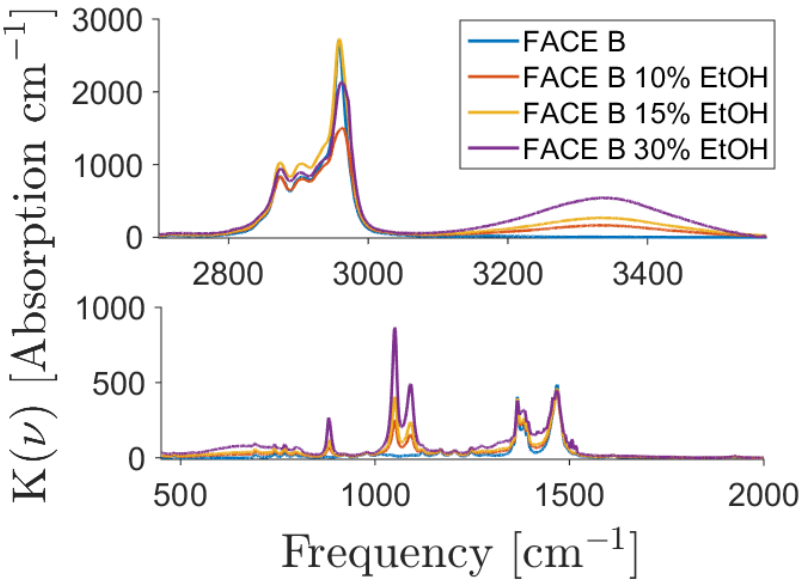

Figure 56: Linear Absorption Coefficient Spectrum of FACE B and various blends with ethanol. Non-linear blending effects are easily seen in vicinity of 2900 cm^{-1}

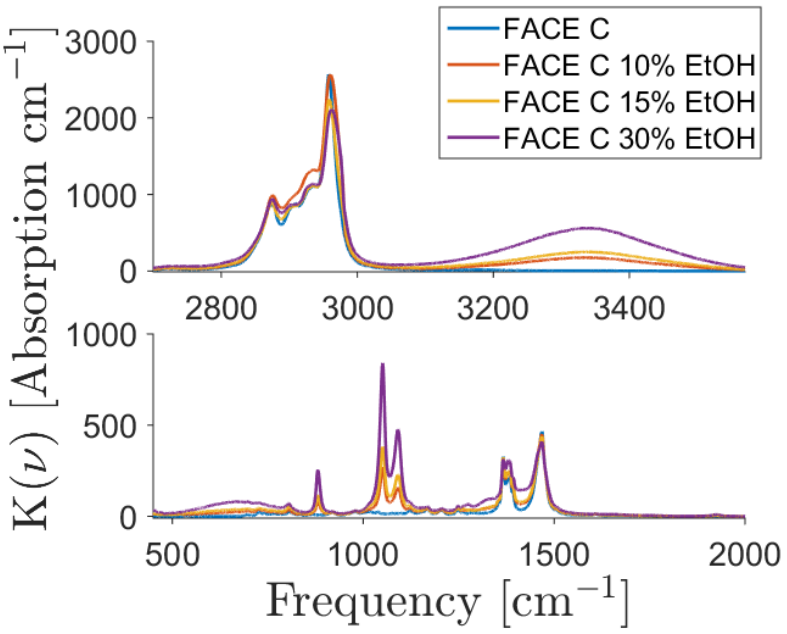

Figure 57: Linear Absorption Coefficient Spectrum of FACE C and various blends with ethanol. Non-linear blending effects are easily seen in vicinity of 2900 cm^{-1}

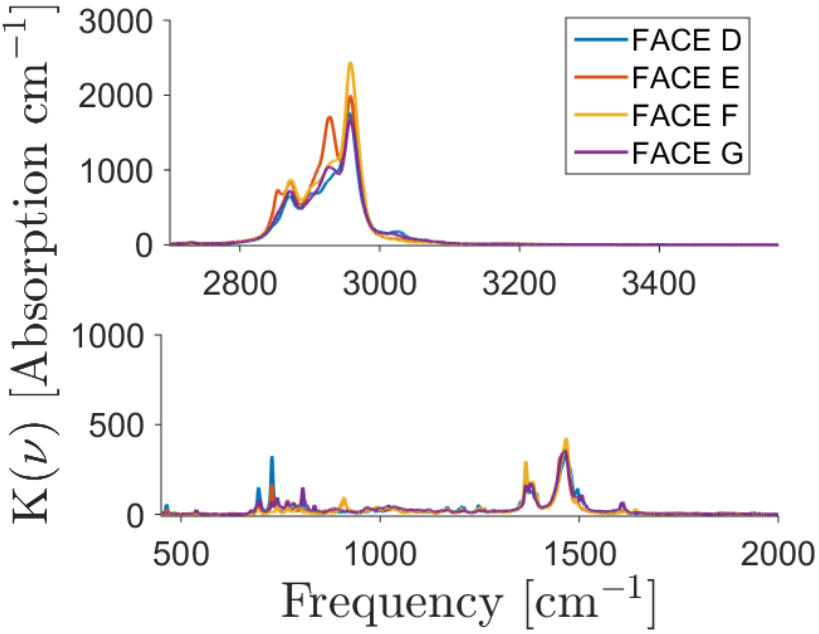

Figure 58: Linear Absorption Coefficient Spectrum of FACE D, E, F and G.

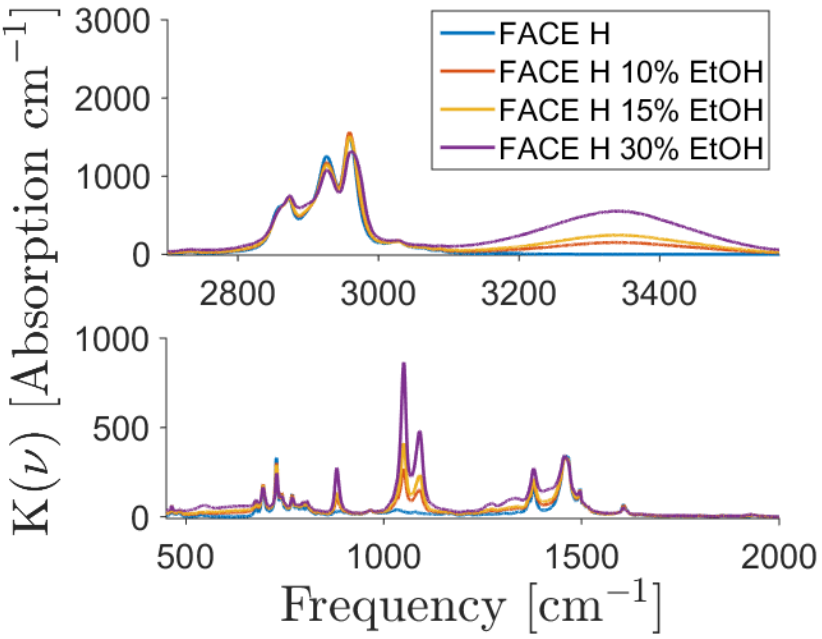

Figure 59: Linear Absorption Coefficient Spectrum of FACE H and various blends with ethanol.

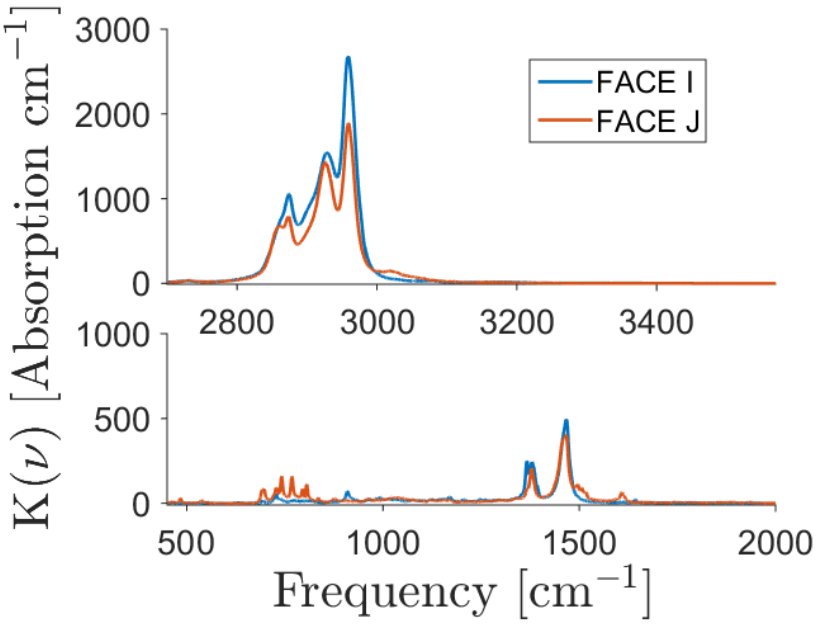

Figure 60: Linear Absorption Coefficient Spectrum of FACE I and J.

E2 Attenuated total reflectance-based

E2.1 Pure Components

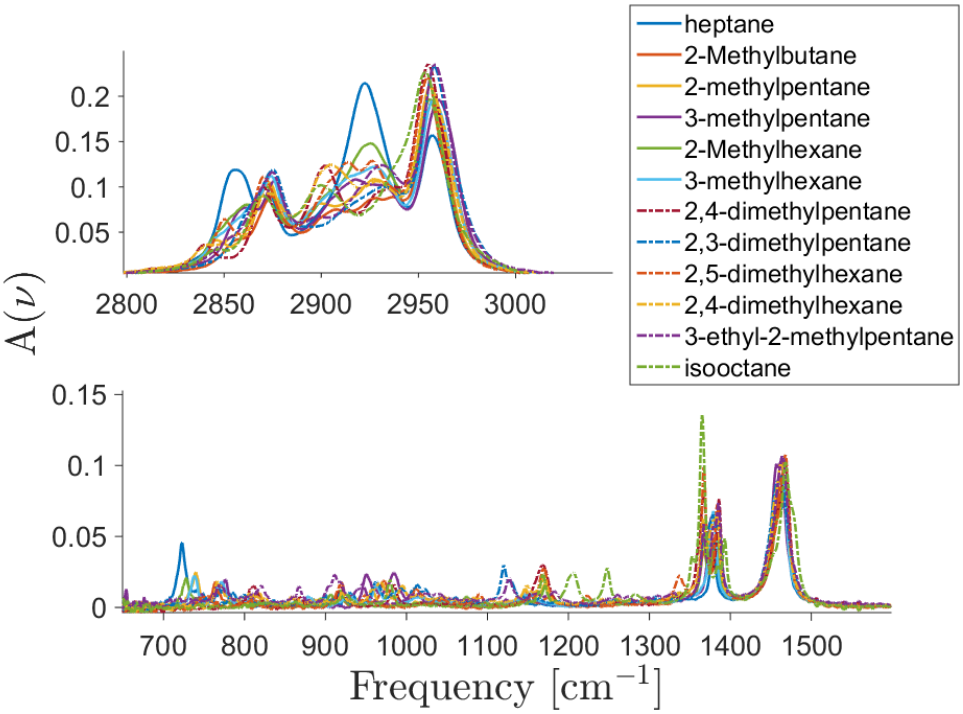

Figure 61: Alkane hydrocarbon absorption spectra

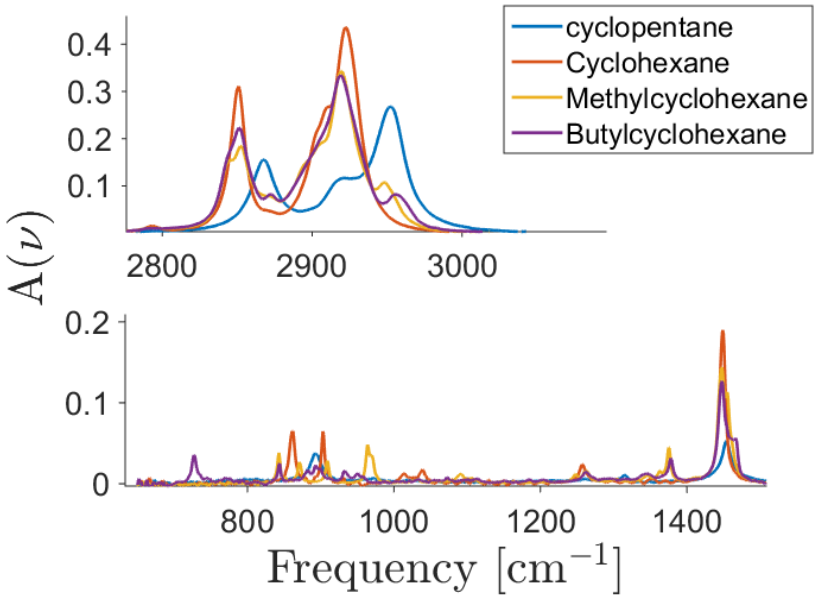

Figure 62: Naphthenic hydrocarbon absorption spectra

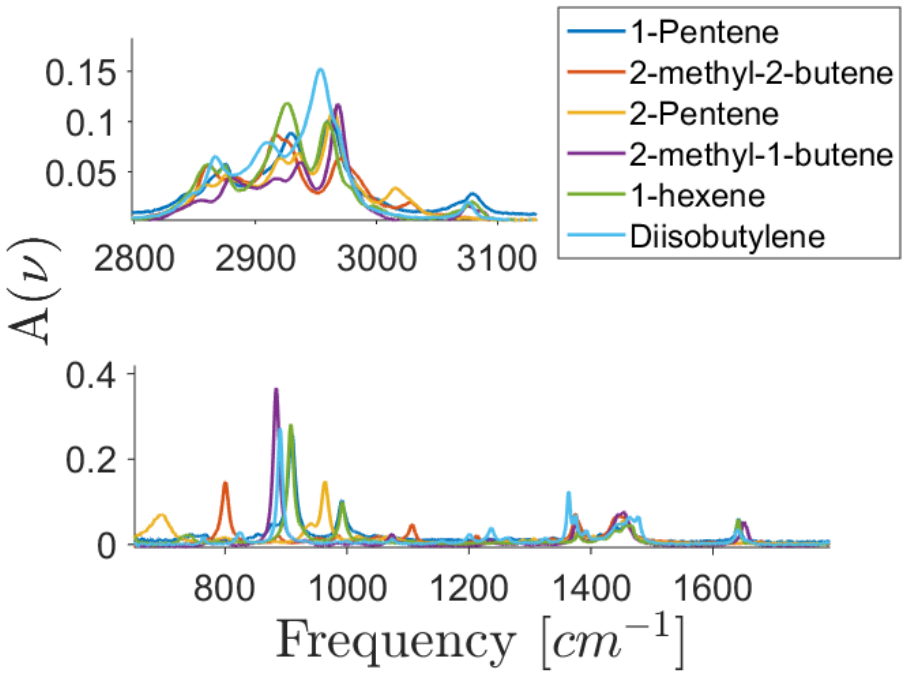

Figure 63: Olefin hydrocarbon absorption spectra

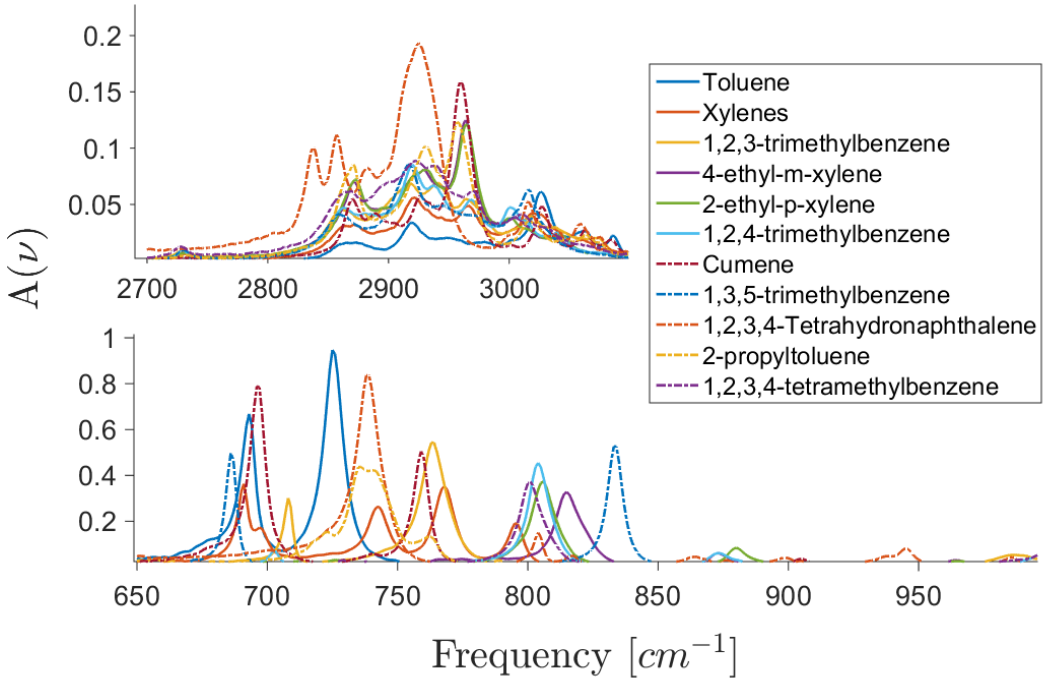

Figure 64: Aromatic hydrocarbon absorption spectra

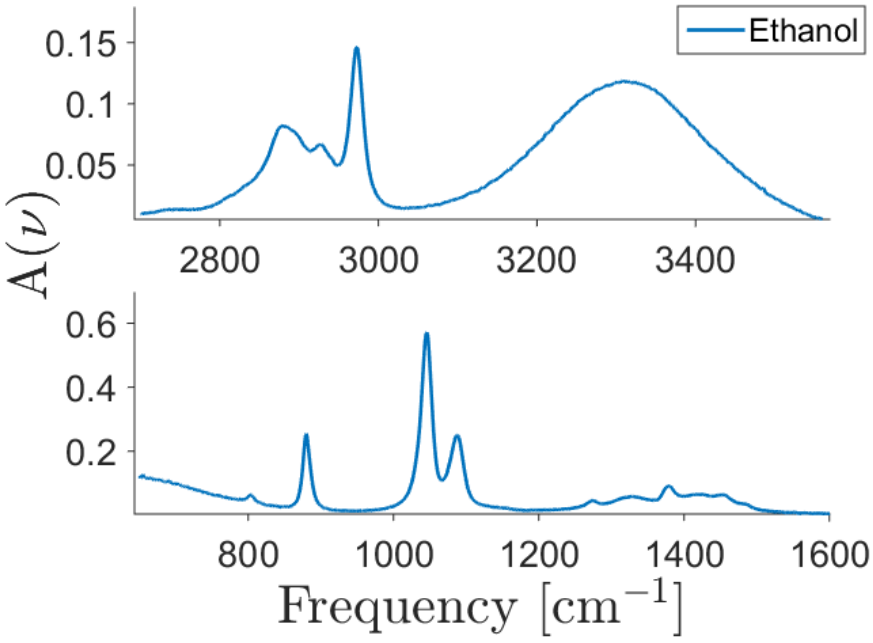

Figure 65: Oxygenated hydrocarbon absorption spectra

E2.2 Pure Component Mixtures

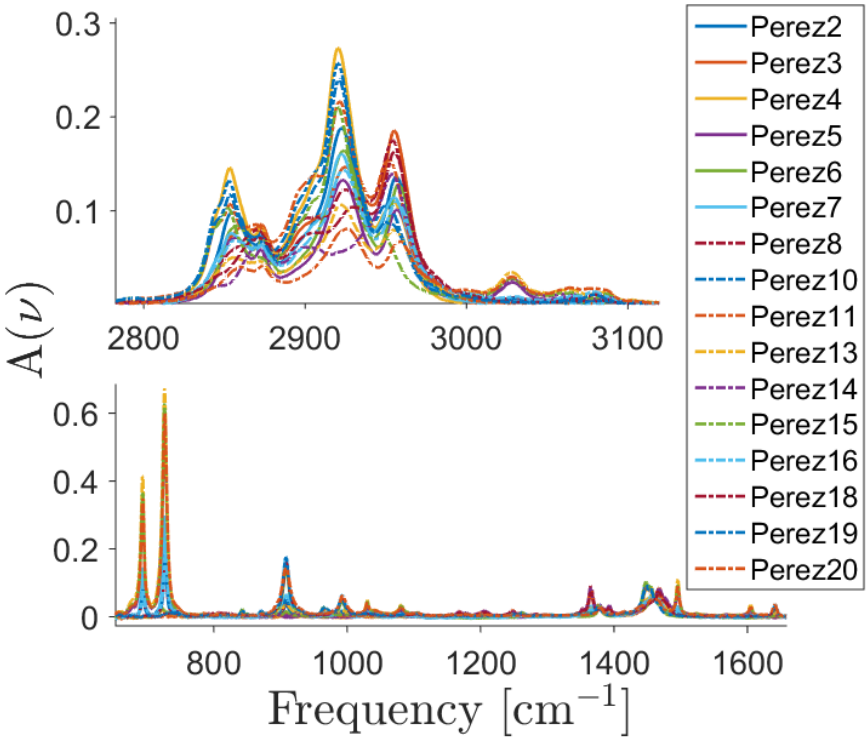

Figure 66: Perez et al. [46] **fuels**

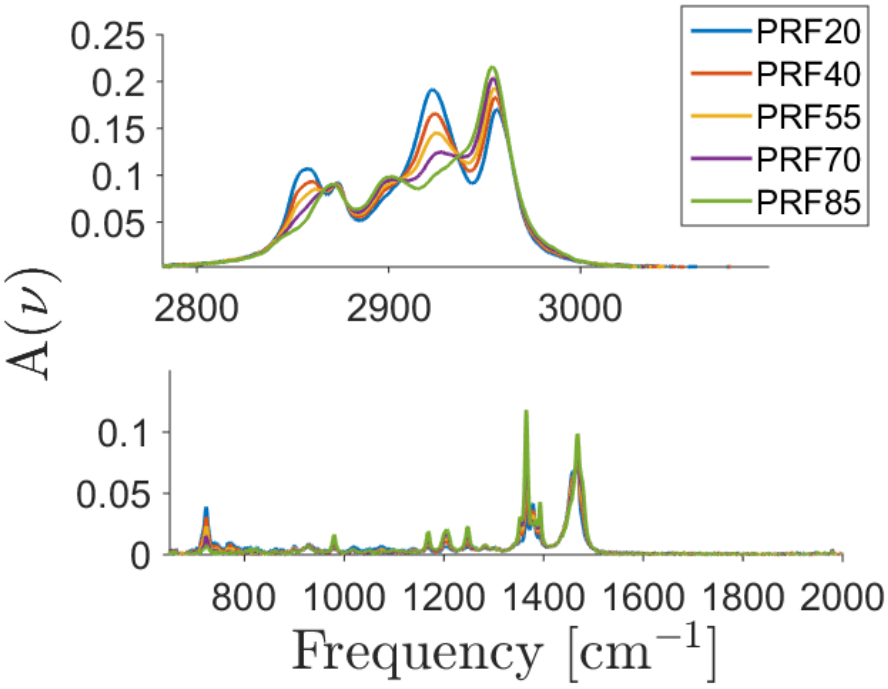

Figure 67: Primary Reference Fuels (PRF)

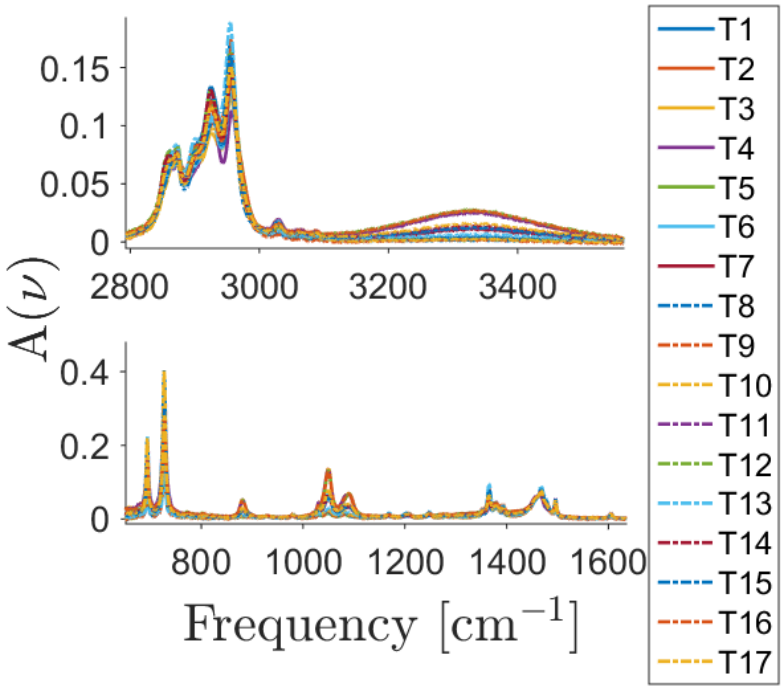

Figure 68: Toluene-Ethanol Reference Fuels (TERF) fuels from Truedsson [23]

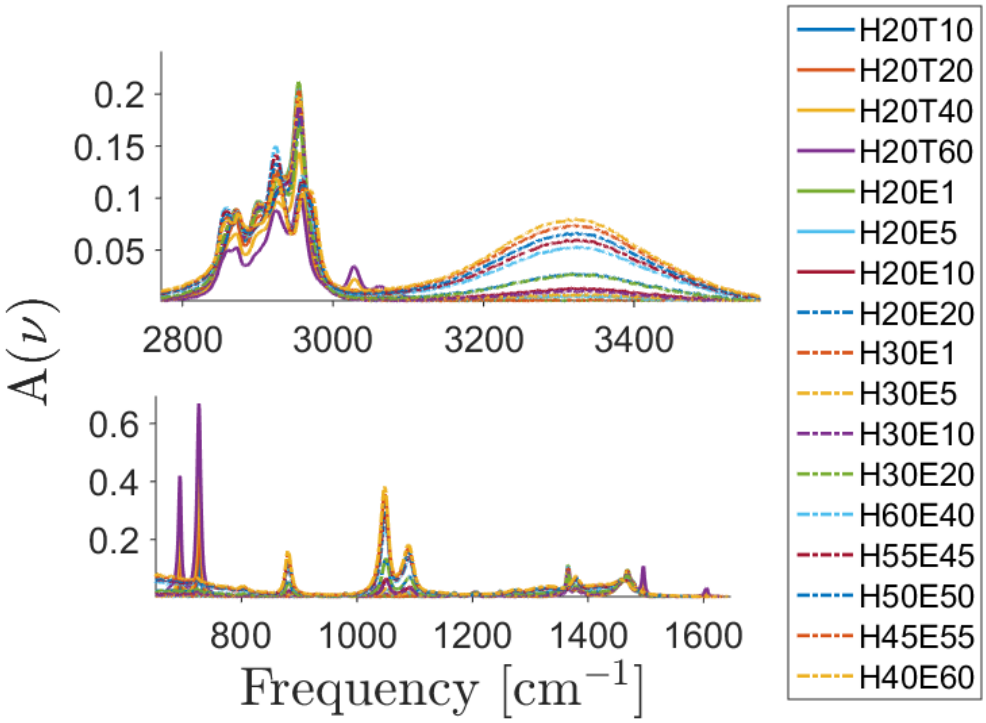

Figure 69: n-heptane/toluene and n-heptane/ethanol mixture fuels from Truedsson [23]

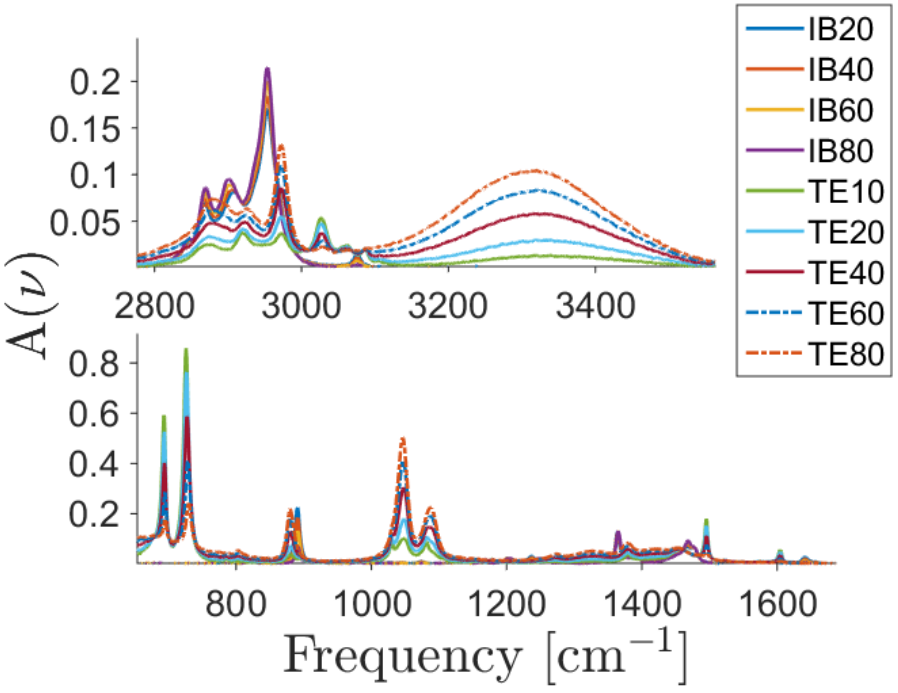

Figure 70: iso-octane and diisobutylene (IBX) mixtures and fuels from Foong et al. [45]

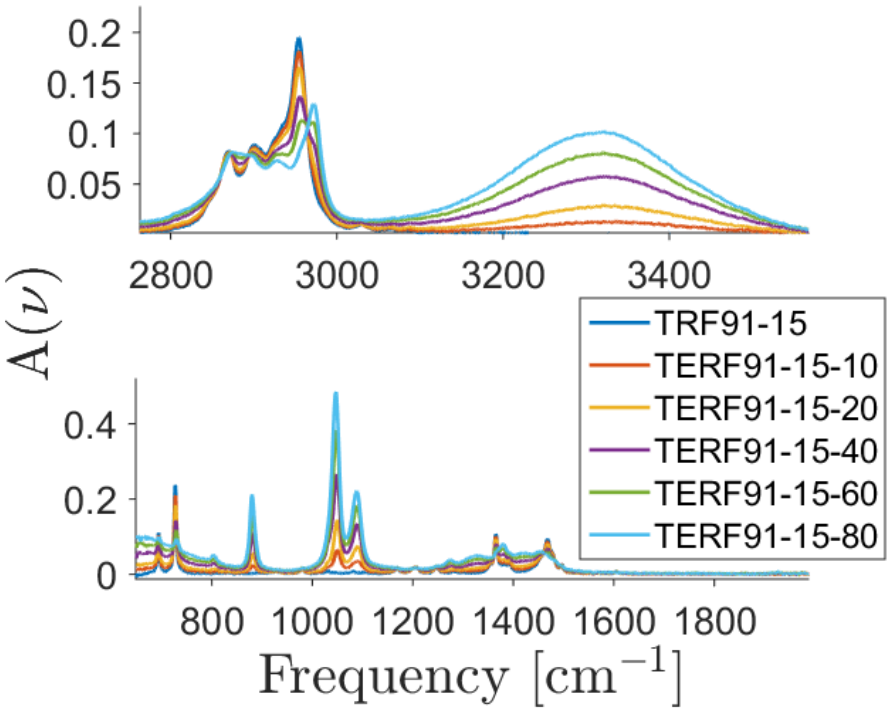

Figure 71: Foong et al. [45] **TERF91-15-EX fuel mixtures**

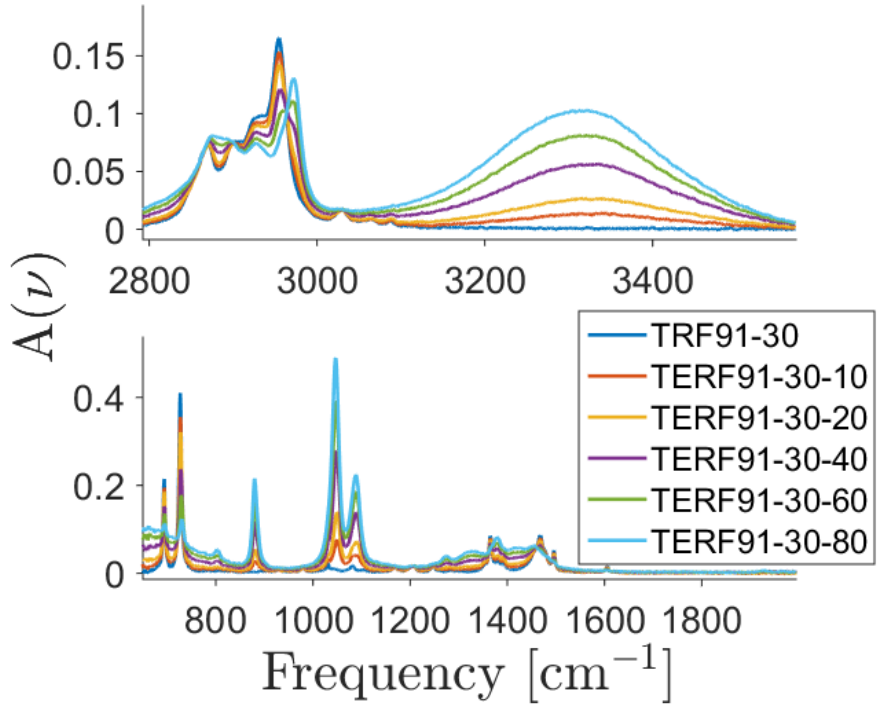

Figure 72: Foong et al. [45] **TERF-30-EX fuel mixtures**

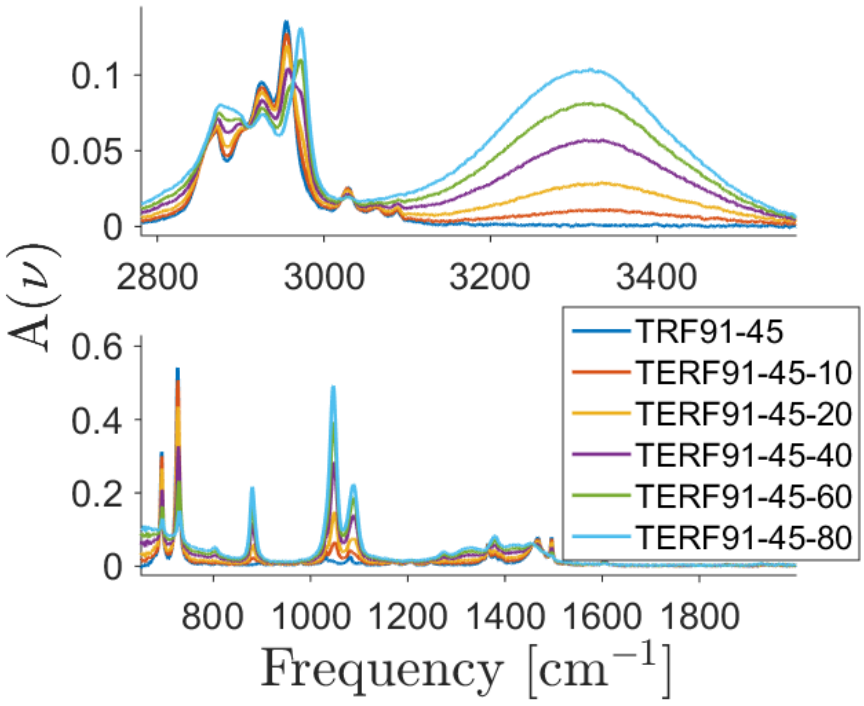

Figure 73: Foong et al. [45] **TERF-45-EX fuel mixtures**

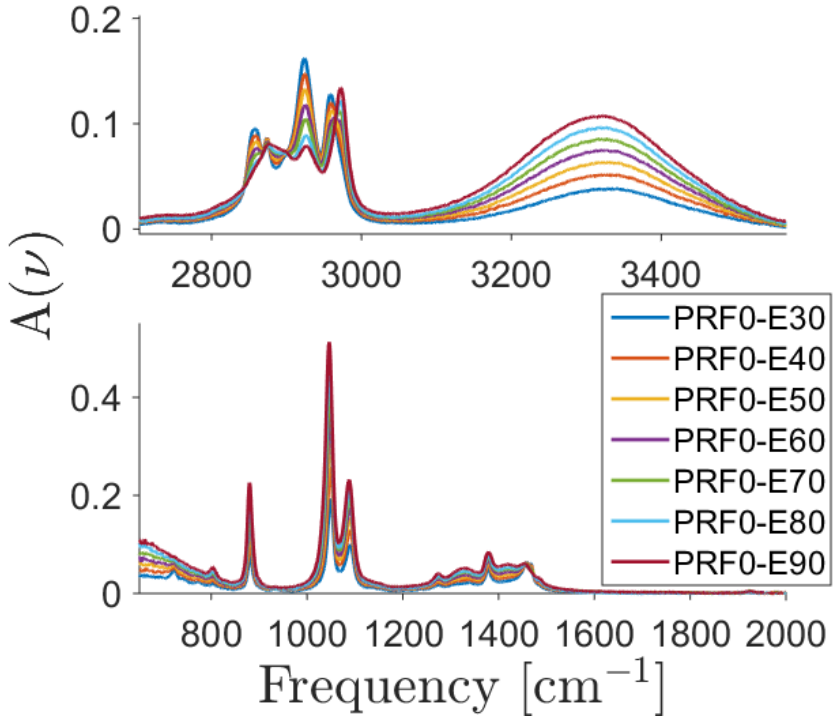

Figure 74: Foong et al. [45] **PRF0-EX fuel mixtures**

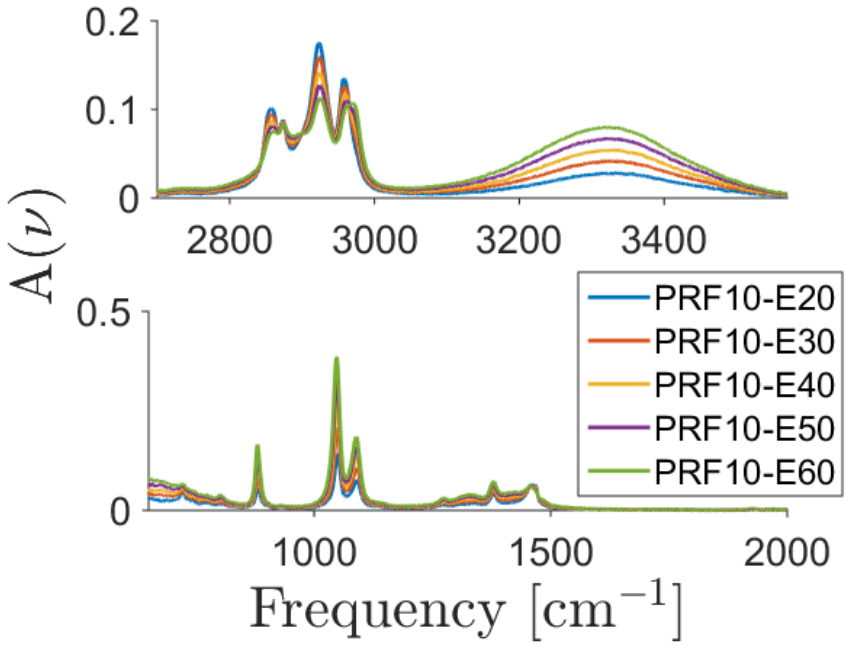

Figure 75: Foong et al. [45] **PRF10-EX fuel mixtures**

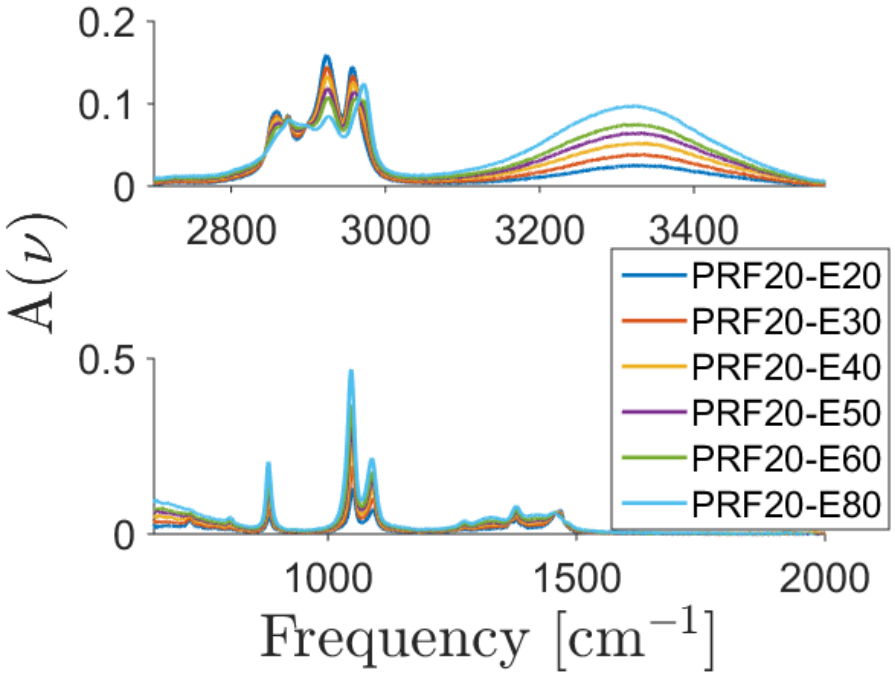

Figure 76: Foong et al. [45] **PRF20-EX fuel mixtures**

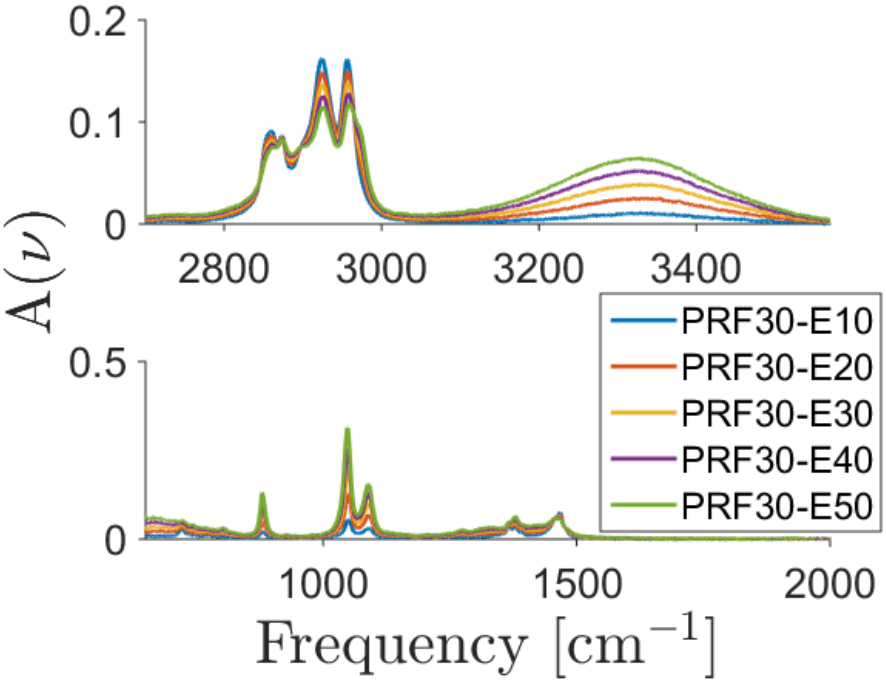

Figure 77: Foong et al. [45] **PRF30-EX fuel mixtures**

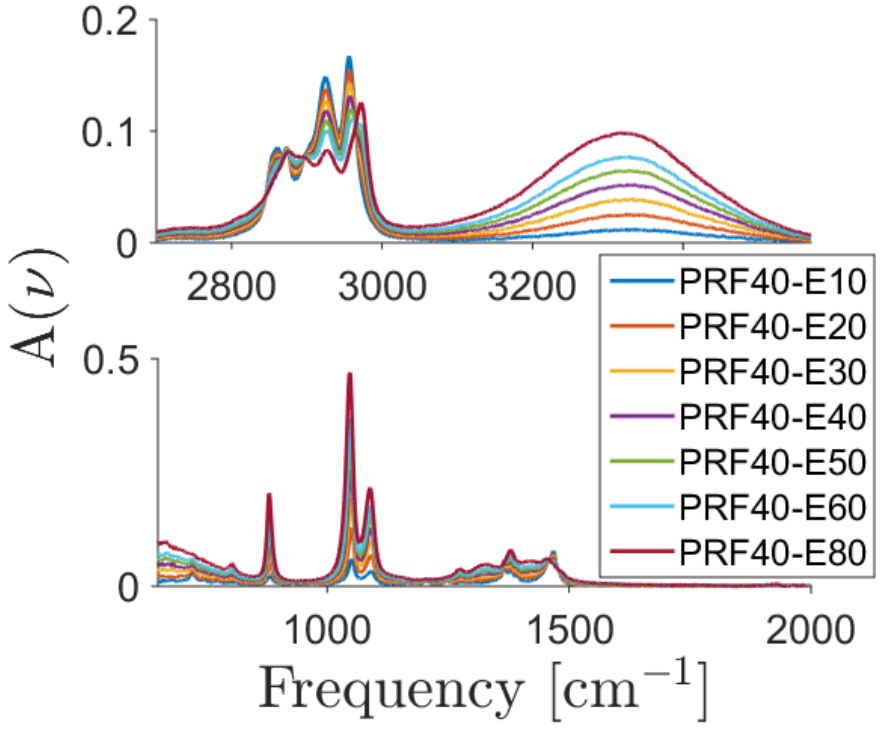

Figure 78: Foong et al. [45] **PRF40-EX fuel mixtures**

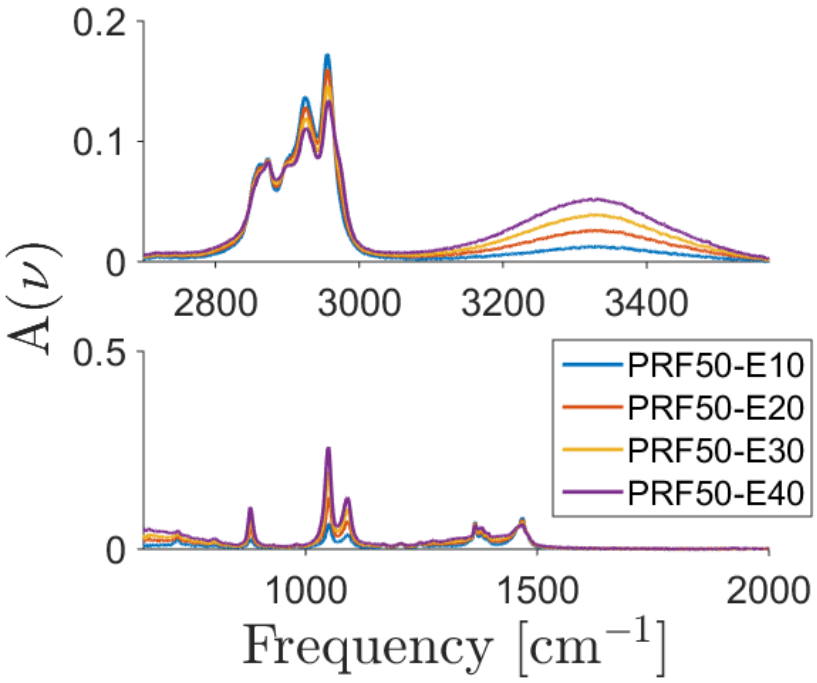

Figure 79: **Foong et al.** [45] **PRF50-EX fuel mixtures**

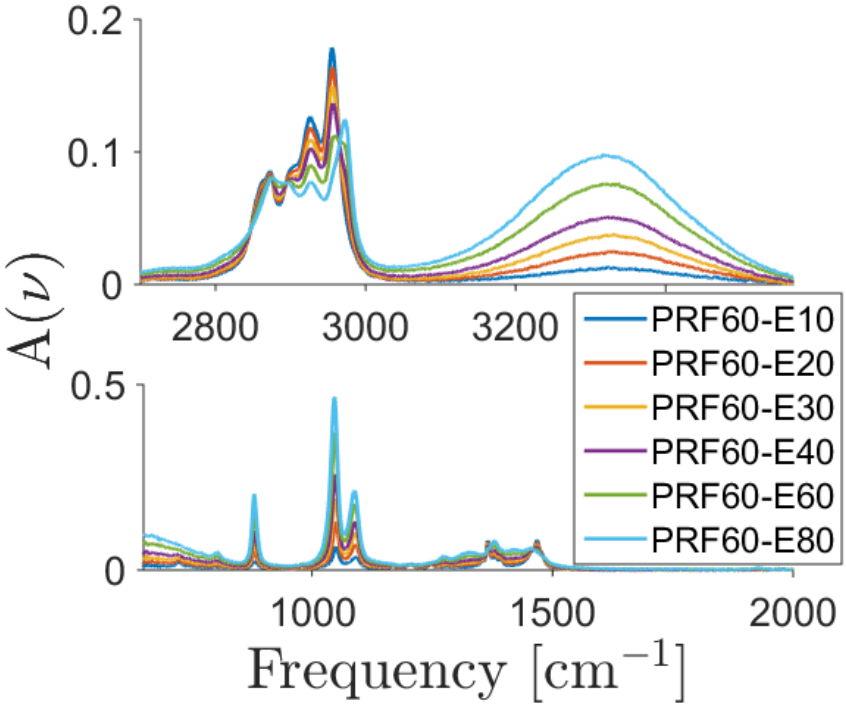

Figure 80: **Foong et al.** [45] **PRF60-EX fuel mixtures**

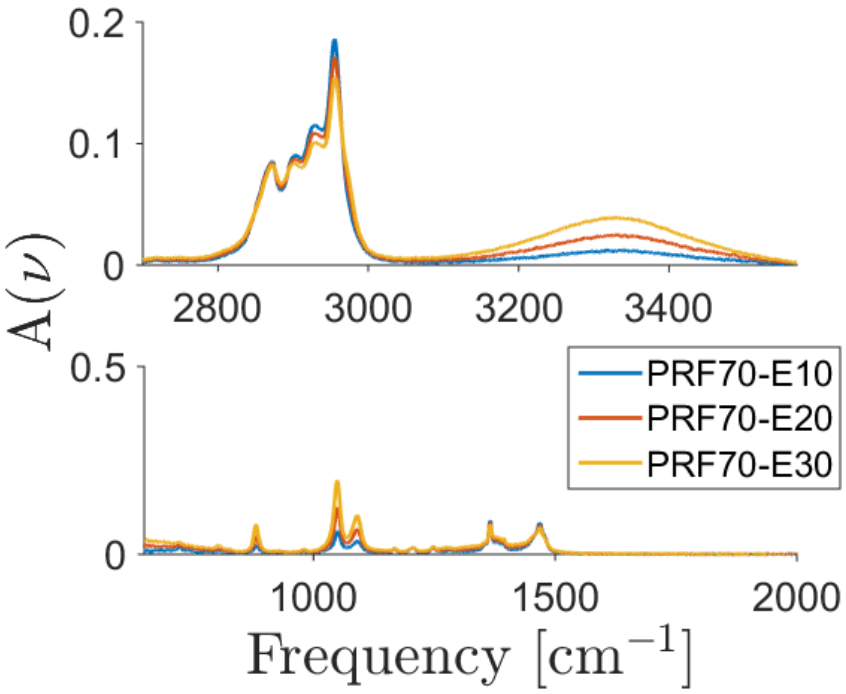

Figure 81: Foong et al. [45] **PRF70-EX fuel mixtures**

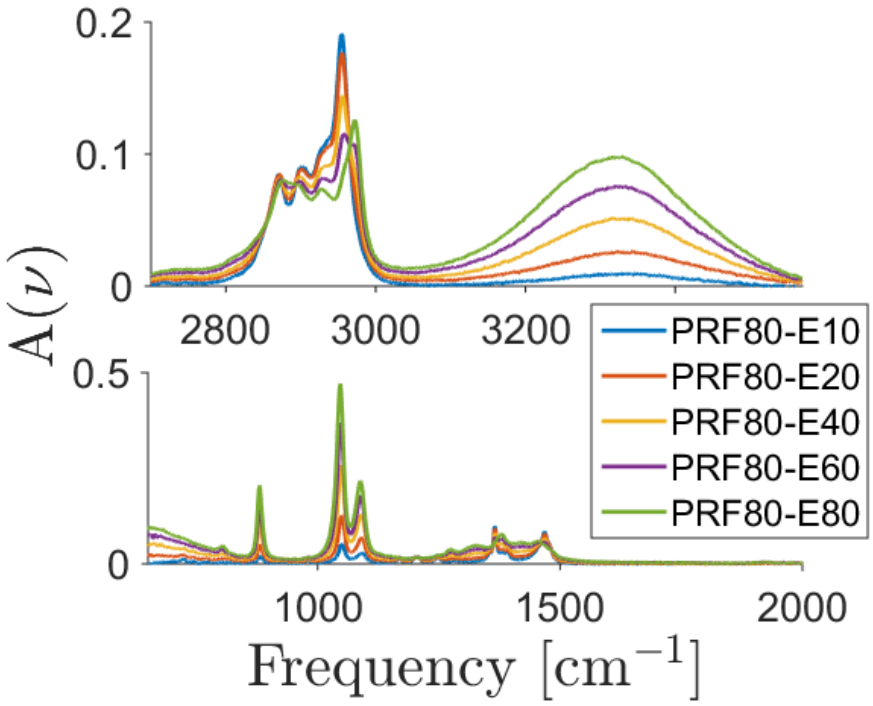

Figure 82: Foong et al. [45] **PRF80-EX fuel mixtures**

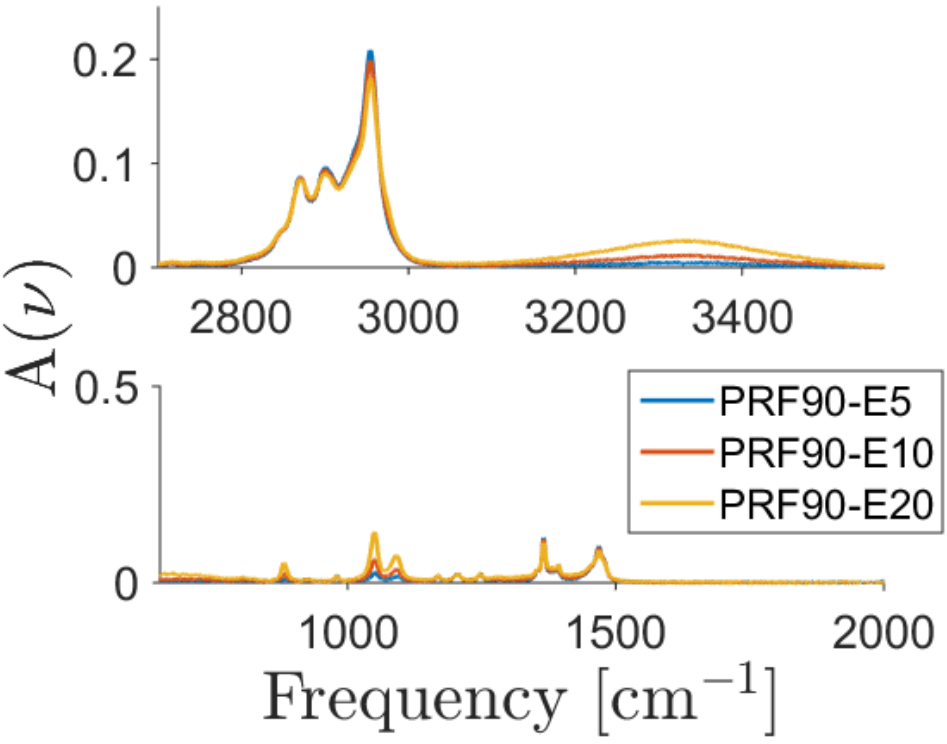

Figure 83: Foong et al. [45] **PRF90-EX fuel mixtures**

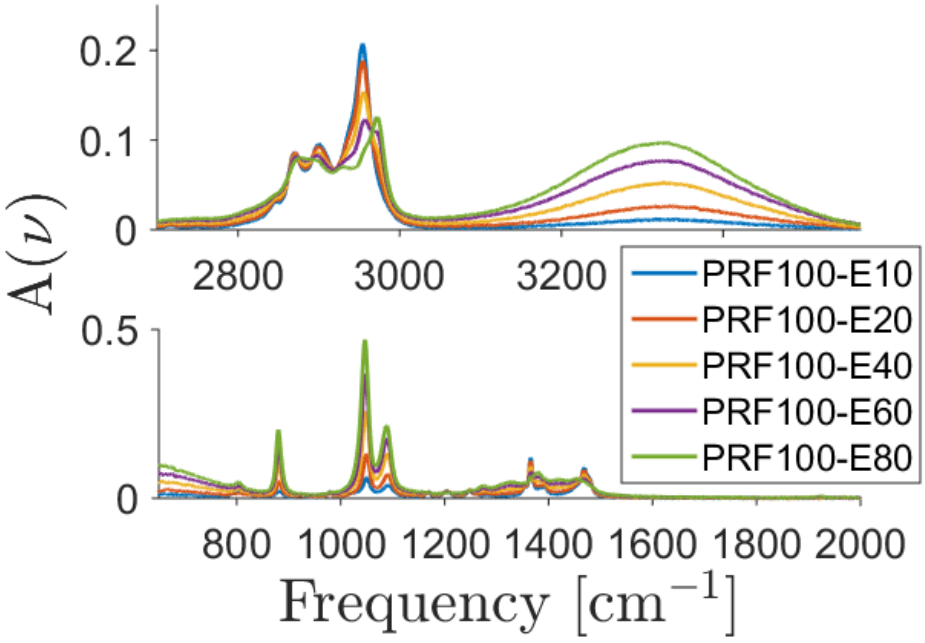

Figure 84: Foong et al. [45] **PRF100-EX fuel mixtures**

Appendix F: In-House Developed Code

F1 HCCI Simulation Code (Python/Cantera)

```
# -*- coding: utf-8 -*-
"""
@author: Shane
Contact: dalys@onid.oregonstate.edu
0-D Homogeneous Charge Compression Ignition (HCCI) Engine Simulation
Features Woschnii heat transfer (variable heat transfer coefficient)
"""
from __future__ import division       #Python version 3 features
import numpy                          #Mathematical operations
import cantera                        #Cantera library
from joblib import Parallel, delayed  #Parallel processsing function import
import h5py                           #Data storage in HDF5 format
import os                             #Operating system operations (to create working directories)
import time                           #Allows us to track processor time
"--->Fuel Specification<---
"
fuel_names =['T1_kal','T2_kal','T3_kal','T4_kal','T5_kal','T6_kal','T7_kal','T8_kal','T9_kal','T10_kal','T12_kal','T13_kal','T14_kal','T15_kal','T16_kal'] #'H20T10'
num_fuels = len(fuel_names) #How many fuel mixtures are in parametric study
fuel_species =[['nc7h16','ic8h18','c6h5ch3'],['nc7h16','ic8h18','c6h5ch3'],['nc7h16','ic8h18','c6h5ch3'],['nc7h16','ic8h18','c6h5ch3'],['nc7h16','ic8h18','c6h5ch3'],['nc7h16','ic8h18','c6h5ch3'],['nc7h16','ic8h18','c6h5ch3'],['nc7h16','ic8h18','c6h5ch3'],['nc7h16','ic8h18','c6h5ch3'],['nc7h16','ic8h18','c6h5ch3'],['nc7h16','ic8h18','c6h5ch3'],['nc7h16','ic8h18','c6h5ch3'],['nc7h16','ic8h18','c6h5ch3']]
X_fuel=[[ 0.3746,    0.3836,    0.2419],[
    0.1147,    0.4441,    0.4411],[
    0.1287,    0.2029,    0.6684],[
    0.2804,    0.1651,    0.5545],[
    0.2639,    0.3734,    0.3628],[
    0.2219,    0.3016,    0.4765],[
    0.2264,    0.3681,    0.4055],[
    0.3528,    0.2206,    0.4266],[
    0.2274,    0.5093,    0.2633],[
    0.1627,    0.4033,    0.4340],[
    0.2655,    0.4499,    0.2846],[
    0.1656,    0.6564,    0.1780],[
    0.3216,    0.2588,    0.4196],[
    0.1375,    0.4427,    0.4198],[
    0.1920,    0.3456,    0.4624]]

"--->File Specifications<---"
working_dir = "C:\Users\Shane\Documents\Python Scripts\\"
"--->Number of Processing Cores<---"
num_cores=6 #cores used in parallel processing
"--->Parametric Study Values<---"
dRPM=100.0                            #RPM stepsize
RPM=numpy.arange(800.0,3100.0,dRPM)   #Revolutions per minute
#RPM=[800.0]
Pi=100000                             #Initial Pressure (BDC)
dphi=0.01                             #equivalence ratio stepsize
phi=numpy.arange(0.15,0.45,dphi)      #Equivalence Ratio
phi=-numpy.sort(-phi)
#phi=[.33]
dT=20                                #Temperature stepsize
```

```
Ti=numpy.arange(350,550+dT,dT)       #Initial Temperature (BDC)
#Ti=-numpy.sort(-
Ti)                 #Sorting Temperature in decending order (higher temperature takes l
onger to simulate, helps with "time-remaining of simulation" accuracy)
#Ti = [350]
"--->Engine Parameters<---"
rc=13                               #Compression Ratio [Vmax/Vmin]
B=.0828                             #Cylinder Bore
Vd=616*10**(-6)                     #Displacement Volume of Engine [m^3]
c=4.44                              #ratio of connecting rod to crank radius
"--->Woschni heat transfer constants<---"
aa=.035                            #Nusselt Number Correlation Constant
bb=.8                              #Nusselt Number Correlation Re power constant
C11=2.28                           #from Chemkin
Tw=430                             #Wall (Bore) Temperature
Pr=0.7
cc=0
"--->Chemical Mechanism<---"
while True: #This simply suppresses the Exception from mechanism import (Happens with p
arallel processing forks trying to access file at same time), and allows a re-
try without spamming alot of useless information in print window.
    try:
        mechanism=cantera.Solution('Princeton_10atm.cti') #mechanism import  #Princeton
_10atm Sarathy_Mech_Reduced_TRF_isopentane.cti Andrae_Xylene_mech FACE_C_mech_re MCH_me
ch_skeletal
    except:
        continue
    break
"--->Recorded Time-step and Crank Angle Duration<---"
RPM_range=286                       #Crank Angle Degree simulation finishes (CAD aBDC)

#timesteps=286/.05                  #Recoreded timesteps for crank angle range
#--------------------->HCCI Simulation Function Code<--------------------------------
def HCCI(RPM,Ti,phi,fuel_species,X_fuel,FuelFolder,FilePrefix):
    #Specifying air for the ideal gas reactor (required for engine simulation) and prop
erty evaluations for viscosity, density, and thermal conductivity
    air=cantera.Solution('air.xml')        #air (the piston model is air separated by a
 wall(piston) to the reactor)
    io2 =mechanism.species_index('o2')     #Finding species indices
    in2 = mechanism.species_index('n2')    #Finding species indices
    ico2 =mechanism.species_index('co2')   #Finding species indices
    ico =mechanism.species_index('co')     #Finding species indices
    ih2o = mechanism.species_index('h2o')  #Finding species indices
    #Engine and Heat Transfer Parameters
    s=4*Vd/(numpy.pi*B**2)                 #Stroke [m]
    Vmax=Vd*rc/(rc-1)                      #Maximum cylinder Volume [m^3]
    Ap=(numpy.pi*B**2)/4                   #Piston Area
    wdot=RPM*numpy.pi/30                   #rotation rate in rad/s
    #Simulation Resolution
    t_end=RPM_range/(6*RPM) #Simulation time domain of crank angle's to be resolved
    #theta=t*wdot*180/numpy.pi                        #Simulation crank angle (Just a re
corded value, not used in any calculations)
    "slider crank rule for piston velocity" #Cantera calls for piston(reactor wall) vel
ocity, not Volume(t) or dV/dt (for the energy equation)
    def piston_speed(t):
        return (wdot*s/2*numpy.sin(t*wdot))*(1-(numpy.cos(t*wdot))/(numpy.sqrt(c**2-
numpy.sin(t*wdot)**2)))
    "Piston Volume Function" #(used for heat transfer Area calculation)
    def Volume(t):
        return Vd/(rc-1)+Vd/2*(1+c-numpy.cos(t*wdot-numpy.pi)-numpy.sqrt((c)**2-
numpy.sin(t*wdot-numpy.pi)**2))
```

```
97.
98.     "Heat Transfer Function"
99.     #Chemkin Property Evaluation constants
100.            C1_mu= 1.1258e-05
101.            C2_mu= .485
102.            C1_k= 132.7906
103.            C2_k= .5
104.            def q(t):
105.                rho = r2.kinetics.density_mass
106.                mu=(C1_mu*(r2.T)**(C2_mu))*0.1       #Viscosity of air
107.                k_g=(C1_k*(r2.T)**(C2_k))*10**-(5)   #Thermal Conductivity of air [W/m-
    K]
108.                Sm=2.0*s*RPM/60                 #Average Piston Speed
109.                w_bar=C11*Sm                    #Woschni average cylinder gas velocity
110.                Re=B*w_bar*rho/mu               #Reynolds Number
111.                Nu=aa*Re**bb*Pr**cc             #Nusselt Number
112.                h=Nu*k_g/(B)                    #convective heat transfer coefficient
113.                return (h*(Tw-r2.T)*(2*numpy.pi*B**3+16*Volume(t))/(4*B))/Ap
114.
115.
116.            def net_heat_release_per_deg(t):
117.                return (-
    Volume(t)*numpy.pi*sum(r2.kinetics.partial_molar_enthalpies*r2.kinetics.net_production_
    rates)+q(t)*Ap)/(wdot*180)
118.
119.            "Fuels and composition Parameters" #(Solves for Stoich F/A micture)
120.            length_f=len(fuel_species)                  #Defining this value for loops b
    elow
121.            ifuel=numpy.zeros(length_f)                 #pre-
    allocating an array for fuel indices
122.            stoich_o2_fuel=numpy.zeros(length_f)        #pre-
    allocating an array for stoich O2 amounts for each pure component in the fuel
123.            x = numpy.zeros(len(mechanism.X))           #pre-
    allocating an array for mole fractions to later define the relative mole fractions for
    the chemical mechanism
124.            output_variables=[]                         #pre-
    defining an empty list for the string list of output variables
125.            for i in range(length_f):
126.                "Solving for stoich O2 mixture"
127.                ifuel[i] =mechanism.species_index(fuel_species[i])  #capturing the prope
    r indiced in the chemical mechanism for the fuel component
128.                stoich_o2_fuel[i]=X_fuel[i]*(mechanism.n_atoms(fuel_species[i],'C')+mech
    anism.n_atoms(fuel_species[i],'H')/4-
    mechanism.n_atoms(fuel_species[i],'O')/2) #solving for stoich amount of oxygen for each
     fuel pure component
129.                "Creating fuel composition array for mechanism input"
130.                x[ifuel[i]] = X_fuel[i]*phi                             #relative amoun
    t of fuel to air
131.                "Creating first performance parameters as fuel mole fraction"
132.                output_variables.append('X'+fuel_species[i])            #the fuel mole
    fractions "output variable names" are being added to the list
133.            "Stoich O2 Mixture"
134.            stoich_o2=numpy.sum(stoich_o2_fuel)                         #The above solv
    ed for the stoich mixture of the fuels components separately. The final answer is the s
    um of the oxygen amounts
135.            x[io2] = stoich_o2                                          #stoichiometric
     A/F amount of Oxygen
136.            x[in2] = stoich_o2*(0.79/0.21)                              #stoichiometric
     A/F amount of Nitrogen
137.            "Setting Fuel and air mixture composition in mechanism"
```

```
            mechanism.TPX=Ti,Pi,x[:]                                        #The chemical mechanism is set to have the user defined initial conditions
            r2 = cantera.IdealGasReactor(mechanism,  name='HCCI Reactor') #specifying that r2 is an ideal gas reactor (varying temperature and pressure)
            r2.volume=Vmax                                                  #Specifying initial volume of ideal gas reactor (BDC volume)


            "Populating output variables -
            - columns are each performance variable and rows indicate crank angle"
            output_variables = output_variables + ['Xo2','Xco','Xco2','Xn2' ,'Xh2o', 'pres','temp','volume','crank_rotation_angle','net_heat_release_rate_per_CA','mass','molecular_weight','mass_density','specific_heat','mixture_enthalpy','time','chem_heat_release_per_deg','net_heat_release_per_deg','heat_loss','rho']
            Performance_Parameters=numpy.zeros([1,len(output_variables)]) #The matrix of numerical values corresponding to the user selected reactor parameters
            Performance_Parameters_add=numpy.zeros([1,len(output_variables)])
            "------>Cantera HCCI Simulation<----------"
            env = cantera.Reservoir(air, name='Environment (Air at the specified heat transfer correlation temperature')
            cantera.Wall(env, r2,velocity=piston_speed,A=Ap,Q=q)   #This is the "piston" in the simulation. The "wall" is moving and has heat transfer through it
            sim = cantera.ReactorNet([r2])  #this is the completed cantera reactor that will simulate HCCI
            time = sim.time                 #starting simulation time (default 0.0s for sim.time)
            n=0
            end_time = t_end           #Simulation End-time
            sim.rtol=1e-
            9                   #Sets maximum integration relative tolerance
            sim.atol=1e-
            20                  #Sets maximum integration absolute tolerance
            count=0
            heat_release = []
            while time < end_time:          #Cantera Simulation integraiton loop
                heat_release.append(net_heat_release_per_deg(time))
                if count ==1:
                    Performance_Parameters = numpy.concatenate((Performance_Parameters, Performance_Parameters_add), axis=0)

                "Reactor Parameters to get"
                for ff in range(length_f):
                    if ff<=length_f:
                        Performance_Parameters[n,ff] = r2.kinetics.X[ifuel[ff]]
                Performance_Parameters[n,(ff+1):] = r2.kinetics.X[io2],r2.kinetics.X[ico],r2.kinetics.X[ico2],r2.kinetics.X[in2],r2.kinetics.X[ih2o],r2.thermo.P,r2.thermo.T,r2.volume,time*wdot*180/numpy.pi,r2.kinetics.enthalpy_mass,r2.mass,r2.kinetics.mean_molecular_weight,r2.kinetics.density,r2.kinetics.cp,r2.kinetics.enthalpy_mass,time,-r2.volume*numpy.pi*sum(r2.kinetics.partial_molar_enthalpies*r2.kinetics.net_production_rates)/(wdot*180),net_heat_release_per_deg(time),q(time)*Ap,r2.kinetics.density_mass


                "Advancing Simulation"
                sim.step(end_time)
                time = sim.time

                count = 1
                n=n+1

            "net_heat_release analysis to determine frequency of data collection (finding appropriate d_theta)"
```

```
            #If net heat release is greater than some threshhold, record at 0.1 CAD, otherwise collect at 2 CAD.


            for xx in range(len(output_variables)):
                if output_variables[xx] == 'crank_rotation_angle':
                    THETA_INDICE = xx
            dtheta=[]
            ind=[]
            p1=0
            for p in range(len(heat_release)-1):
                dtheta.append(abs(Performance_Parameters[p1,THETA_INDICE] - Performance_Parameters[p,THETA_INDICE]))
                if heat_release[p] >= 0.2: #Record at 0.1 CAD
                    #Get every 0.1 indice
                    #If dtheta > 0.1 grab indice and reset beginning d_theta ind.

                    if dtheta[p] >= 0.1:
                        p1=p #reset the initial d_theta_indice
                        ind.append(p)    #grab indice of simulation at which this happened
                else: #Record at 2 CAD
                    #Get every 2 CAD Indice
                    if dtheta[p] >= 2:
                        p1=p #reset the initial d_theta_indice
                        ind.append(p)    #grab indice of simulation at which this happened

            "File Specification"
            FileName = "%s_RPM=%s_Ti=%s_phi=%s.hdf5" %(FilePrefix,RPM,Ti,phi)
               #Generating a filename based on fuel name and initial parametric study values
            f=h5py.File(FuelFolder + FileName,"w")
               #Creating file for each parametric study point #f.name is 'u/'
            "Performance Paramater Saving"
            for v in range(len(output_variables)):
                f.create_dataset("%s" %(output_variables[v]),data=Performance_Parameters[ind,v])        #creating hdf5 datasets for each performance parameter inside the file created above
            f.close()
               #close the file we just created
            return

        "Parallel Process job Executor for HCCI function"
        if __name__ == '__main__':                                                                    #No idea what this does, but Windows machines need it to "protect" the main loop (whatever that is...) for parallel processing
            while True:
                try:
                    start_time = time.clock()
        #capturing time for the start of the simulation
                    for q in range(num_fuels):
        #Fuel Loop
                        FuelFolder= working_dir + "CR_" + str(rc) + "\\" + fuel_names[q]
                      #fuel folder to save data in
                        FilePrefix="\\"+fuel_names[q]+"_CR=%s_Pi=%.1f" %(rc,Pi*10**(-5))     #prefix to name of saved file within fuel folder
                        start_time2 = time.clock()
        #Starting processor time of simulation
```

```
                        if not os.path.exists(FuelFolder):
        #If the fuel working directory already exists, don't try and overwrite the folder
                            os.makedirs(FuelFolder)
        #Creating working directory
                        for z in range(len(Ti)):
                            new_time1= time.clock()
                            for k in range(len(phi)):
                                Parallel(n_jobs=num_cores)(delayed(HCCI)(RPM[o],Ti[z],phi[k],fuel_species[q],X_fuel[q],FuelFolder,FilePrefix) for o in range(len(RPM))) #parallel process executor
                                print "Finished: Ti=%s, phi=%s, all RPM" %(Ti[z],phi[k])  #Delta t
                                "Keyboard interrupt to pause simulation"

                            new_time2=time.clock()
        #This, and new_time1 allows calculation of how long it takes to do a complete temperature study with all phi and RPM
                            "Elapsed time, estimated time until completion, and parametric studies being finished"
                            print "Elapsed time:  %.0f hour(s) and %.1f minute(s)" %((time.clock() - start_time2)/3600,((time.clock() - start_time2)/3600-numpy.fix((time.clock() - start_time2)/3600))*60) #Elapsed t
                            print fuel_names[q] + " " + "is estimated to take a remaining %.0f hour(s) and %.1f minute(s)" %((new_time2 - new_time1)/3600*len(Ti[(z+1):]) , (((new_time2 - new_time1)/3600)*len(Ti[(z+1):])-numpy.fix(len(Ti[(z+1):])*(new_time2 - new_time1)/3600))*60)

                        print "The simulation for"+" "+"[" +FilePrefix+"]"+" "+ "took %.0f hour(s) and %.1f minute(s)" %((time.clock() - start_time2)/3600,((time.clock() - start_time2)/3600-numpy.fix((time.clock() - start_time2)/3600))*60) #Total Elapsed t
                        print "The remaining fuel studies are estimated to take a remaining %.0f hour(s) and %.1f minute(s)" %((new_time2-start_time2)/3600*(num_fuels-(q+1)) , (((new_time2-start_time2)/3600)*(num_fuels-(q+1))-numpy.fix((num_fuels-(q+1))*(new_time2-start_time2)/3600))*60)
                    print "The simulation took a total of %.0f hour(s) and %.1f minute(s)" %((time.clock() - start_time)/3600,((time.clock() - start_time)/3600-numpy.fix((time.clock() - start_time)/3600))*60) #Total Elapsed t
                    if q == max(q):
                        break
                except KeyboardInterrupt:
                    print '\nPausing...  (Hit ENTER to continue, type quit to exit.)'
                    try:
                        response = raw_input()
                        if response.lower() == 'quit':
                            break
                        print 'Quitting...'
                    except KeyboardInterrupt:
                        print 'Resuming...'
                        continue
```

F2 Transmission Spectra Optical Constant Method Correction (Python)

```
#Automates ANCHORPT.exe EXPABS2K.exe and updates .ASC files for iterative process
import numpy as py              #"Numpy" for Mathematical operations
from subprocess import Popen
from time import sleep


wdir=["C:\Users\Shane\Documents\IRKK\FACE Fuels\\","C:\Users\Shane\Documents\IRKK\FACE Fuels\\"]
comment_line= "FB30"
absorbance_filename1=[comment_line+"015_.SPC",comment_line+"000_.SPC"]
absorbance_filename2=comment_line+"5_.SPC"
absorbance_filename3=[comment_line+"015c.SPC",comment_line+"000c.SPC"]
absorbance_filename4=[comment_line+"015n.SPC",comment_line+"000n.SPC"]
baseline_wavenumber=['450', '942', '990', '1184', '1224', '1587', '1736', '2050', '2310', '2440', '2558', '2646', '3676', '3854', '4000']
baseline_wavenumber = baseline_wavenumber[::-1]
num_anchor=str(len(baseline_wavenumber))
pathlength1 =['14.7323','.6039'] #in micrometers
pathlength2 ='513.557' #in micrometers
n_infinity= '1.4' #index of refraction at n-infinity
kk_iterations='9' #maximum iterations for kramers kronig convergence
kk_method='2'
slide='6' #slide index (6 for Kbr)

#comment_line= "MCH Fuel"
#absorbance_filename1=["MCH000_.SPC"]
#absorbance_filename2="MCH5_.SPC"
#absorbance_filename3=["MCH000c.SPC"]
#absorbance_filename4=["MCH000n.SPC"]
#baseline_wavenumber=['2733','3060']
#num_anchor=str(len(baseline_wavenumber))
#pathlength1 =['9.4847'] #in micrometers
#pathlength2 ='514.72' #in micrometers
#n_infinity= '1.4235' #index of refraction at n-infinity
#kk_iterations='9' #maximum iterations for kramers kronig convergence
#kk_method='2'
#slide='6' #slide index (6 for Kbr)


for i in range(len(wdir)):
    Anchor_Baseline = []
    correction_collected = []
    for j in range(len(baseline_wavenumber)):
        try:
            Anchor_Baseline = Anchor_Baseline + str(baseline_wavenumber[j]) + " "
        except TypeError:
            Anchor_Baseline = str(baseline_wavenumber[j]) + " "

    #-------------STEP 1-------------------
    #Defining BASELINE.ASC (It never changes)
    file1 = open(wdir[i] + "anchorpt.ASC", "w")
    file1.write("1\nData for Optically Thick "+comment_line+"\n" + absorbance_filename4[i] + "\n" + str(len(baseline_wavenumber)) + " " + "1" "\n"+Anchor_Baseline +"\n"+absorbance_filename2 + " " + pathlength2 +" 6")
    file1.close()

    #Running expabs2k with no baseline correction for initial n-spectrum estimate
    file2 = open(wdir[i] + "expabs2k.ASC", "w")
```

```
    file2.write("1\n----
data run 1\n"+ absorbance_filename1[i] +"\n"+comment_line+" 15 micron KBr Cell\n1/29/15
\n2 "+kk_iterations+" 6 0 0 0 "+ pathlength1[i] + " 0.0\n" + n_infinity + "\n0")
    file2.close()
    Expa_1 = Popen(wdir[i] + 'expabs2k.exe',cwd=wdir[i] )
    sleep(1)
    Expa_1.kill()
    #Baseline correction with initial n-spectrum estimate
    Anch = Popen(wdir[i] + 'anchorpt.exe',cwd=wdir[i] )
    sleep(1)
    Anch.kill()


    residual=[]
    tolerance = 0.0001
    loop_count=0


    #-------------STEP 2-------------------
    #Creating new expabs2k ASC file to read baseline correction data
    file3=open(wdir[i] + "anchorpt.log",'r')
    anchorlog=file3.readlines()
    file3.close()

    correction = []
    Expa_Baseline = []
    ind = 21

    for j in range(len(baseline_wavenumber)):

        correction_indice=anchorlog[ind].find('\n')
        correction_indice2=anchorlog[ind][(correction_indice-
12):correction_indice].find(' ')
        correction.append(anchorlog[ind][(correction_indice-
12+correction_indice2):correction_indice]) #Linear absorption coefficient first value
        ind = ind + 6
        try:
            Expa_Baseline = Expa_Baseline + str(baseline_wavenumber[j]) + " " + str(cor
rection[j]) +"\n"
        except TypeError:
            Expa_Baseline = str(baseline_wavenumber[j]) + " " + str(correction[j]) +"\n
"

    correction_collected.append(correction)
    #With new linear absorption coefficient values, we can now specify these values in
expabs2k.ASC

    file4 = open(wdir[i] + "expabs2k.ASC", "w")
    file4.write("1\n----
data run 1\n"+ absorbance_filename1[i] +"\n"+comment_line+" in 15 micron KBr Cell\n1/29
/15\n2 "+kk_iterations+" 6 0 0 0 "+pathlength1[i]+ " 0.0\n"+n_infinity+"\n" + str(len(b
aseline_wavenumber)) + "\n"+Expa_Baseline)
    file4.close()
    Expa = Popen(wdir[i] + 'expabs2k.exe',cwd=wdir[i])
    sleep(1)
            Expa.kill()


            # Up to here, we have successfully estimated an initial n-
spectrum, and used that information to do our first baseline correction on uncorrected
data. Now, we must correct the "corrected" data until convergence
```

```
104.
105.           #------------------STEP 3 ITERATIVE PROCESS -------------------
106.           #EXPAS2K now needs to correct its "corrected" absorbance spectrum until conv
    ergence of linear absorption coefficient is found
107.           x=1
108.           while x==1:  #While TRUE
109.
110.               #First, we run baseline with the new n-
    spectrum after first correction
111.               anch_i = Popen(wdir[i] + 'anchorpt.exe',cwd=wdir[i] )
112.               sleep(1)
113.               anch_i.kill()
114.               #Next, update the expas2k with new linear absorption coefficient values
    from anchorpt.log, and now use the previously corrected absorbance spectrum
115.
116.               file5=open(wdir[i] + "anchorpt.log",'r')
117.               anchorlogi=file5.readlines()
118.               file5.close()
119.
120.               correction = []
121.               correction_indice = []
122.               correction_indice2 = []
123.               Expa_Baseline = []
124.               ind = 21
125.
126.               for j in range(len(baseline_wavenumber)):
127.                   correction_indice=anchorlogi[ind].find('\n')
128.                   correction_indice2=anchorlogi[ind][(correction_indice-
    12):correction_indice].find(' ')
129.                   correction.append(anchorlogi[ind][(correction_indice-
    12+correction_indice2):correction_indice]) #Linear absorption coefficient first value
130.                   ind = ind + 6
131.                   try:
132.                       Expa_Baseline = Expa_Baseline + str(baseline_wavenumber[j]) + "
    " + str(correction[j]) +"\n"
133.                   except TypeError:
134.                       Expa_Baseline = str(baseline_wavenumber[j]) + " " + str(correcti
    on[j]) +"\n"
135.
136.               correction_collected.append(correction)
137.               filei = open(wdir[i] + "expabs2k.ASC", "w")
138.               filei.write("1\n----
    data run 1\n"+ absorbance_filename1[i] +"\n"+comment_line+" in 15 micron KBr Cell\n1/29
    /15\n2 "+kk_iterations+" 6 0 0 0 "+pathlength1[i]+ " 0.0\n"+n_infinity+"\n" + str(len(b
    aseline_wavenumber)) + "\n"+Expa_Baseline)
139.               filei.close()
140.               Expa_i = Popen(wdir[i] + 'expabs2k.exe',cwd=wdir[i])
141.               sleep(1)
142.               Expa_i.kill()
143.
144.               for j in range(len(baseline_wavenumber)):
145.                   residual.append(abs(py.array(correction_collected[loop_count+1][j],d
    type=py.float32)-py.array(correction_collected[loop_count][j],dtype=py.float32)))
146.
147.               loop_count=loop_count+1
148.               if all(j <= tolerance for j in residual): #If all residuals have converg
    ed, while loop breaks
149.                   x=0
150.               residual = []
```