

# Influence of EGR Quality and Unmixedness on the High-Load Limits of HCCI Engines

Magnus Sjöberg and John E. Dec  
Sandia National Laboratories

## ABSTRACT

This work explores how the high-load limits of HCCI are affected by fuel autoignition reactivity, EGR quality/composition, and EGR unmixedness for naturally aspirated conditions. This is done for PRF80 and PRF60. The experiments were conducted in a single-cylinder HCCI research engine (0.98 liters) with a CR = 14 piston installed. By operating at successively higher engine loads, five load-limiting factors were identified for these fuels: 1) Residual- $\text{NO}_x$ -induced run-away advancement of the combustion phasing, 2) EGR- $\text{NO}_x$ -induced run-away, 3) EGR- $\text{NO}_x$ /wall-heating induced run-away 4) EGR-induced oxygen deprivation, and 5) excessive partial-burn occurrence due to EGR unmixedness.

The actual load-limiting factor is dependent on the autoignition reactivity of the fuel, the EGR quality level (where high quality refers to the absence of trace species like NO, HC and CO, *i.e.* simulated EGR), the level of EGR unmixedness, and the selected pressure-rise rate (PRR). For a reactive fuel like PRF60, large amounts of EGR are required to control the combustion phasing. Therefore, for operation with simulated EGR, the maximum IMEP becomes limited by the available oxygen. When real EGR (with trace species) is used instead of the simulated EGR, the maximum IMEP becomes limited by EGR- $\text{NO}_x$ /wall-heating induced run-away. For the moderately reactive PRF80 operated with simulated EGR, the maximum IMEP becomes limited by residual- $\text{NO}_x$ -induced run-away. Furthermore, operation with real EGR lowers the maximum steady IMEP because of EGR- $\text{NO}_x$ -induced run-away. This is similar to PRF60.

Finally, the data show that EGR/fresh-gas unmixedness can lead to a substantial reduction of the maximum stable IMEP for operation with a low PRR. This happens because the EGR unmixedness causes occasional partial-burn cycles due to excessive combustion-phasing retard for cycles that induct substantially higher-than-average level of EGR gases.

## INTRODUCTION

The homogeneous charge compression ignition (HCCI) combustion process offers good fuel economy and very low nitrogen-oxide ( $\text{NO}_x$ ) and particulate emissions. Therefore, HCCI combustion is one option for improving the efficiency of the spark-ignited gasoline engine. HCCI is also pursued in order to reduce the requirement for  $\text{NO}_x$  aftertreatment in the latest generation of compression-ignited diesel engines. However, several technical hurdles need to be overcome before HCCI can be broadly applied in production engines. For example, controlling the combustion phasing and increasing the power output while maintaining acceptable PRR are two important research areas.

The autoignition reactivity of the fuel is a very important parameter impacting the potential high-load performance of the HCCI engine. If the fuel has high resistance to autoignition (*i.e.* high octane number), high charge temperatures are required during the compression stroke in order to ensure autoignition. For a given intake pressure ( $P_{in}$ ), higher intake temperature ( $T_{in}$ ) leads to lower charge density, which typically has a negative influence on the peak load that can be obtained [1,2]. In addition, the high temperature required for autoignition also leads to high peak-combustion temperatures, so

excessive  $\text{NO}_x$  can become the load-limiting factor, even for well-mixed “classical” HCCI [1].

On the other hand, if the fuel is very prone to autoignite (*i.e.* has a low octane number), large quantities of exhaust-gas recirculation (EGR) may be required to retard the combustion phasing sufficiently to obtain an acceptable PRR. Since EGR gases displace air, the intake mole fraction of oxygen,  $[\text{O}_2]$ , is reduced with increasing amount of EGR. Therefore, the load-limiting factor for very reactive fuels can be oxygen deprivation [1].

To illustrate these points, Fig. 1 reproduces some results from a recent study of high-load HCCI operation by the current authors [1]. It shows load sweeps at  $\text{PRR} = 5 \text{ bar}/^\circ\text{CA}$  for iso-octane, PRF80, and PRF60. The larger black symbols represent the highest steady load for each of the fuels. It can be seen that the two-stage ignition fuels PRF80 and PRF60 reach significantly higher loads than the single-stage ignition iso-octane. Therefore, PRF80 and PRF60 are the fuels of choice for the current study of EGR effects on the high-load limits. In addition, under the conditions studied here, iso-octane does not require EGR to suppress autoignition, and this is the second reason why it was not included in the current study on EGR effects.

Nonetheless, the iso-octane data set is included in Fig. 1 since it can serve as a point of reference for the discussion in this paper. As indicated by the annotation “ $T_{\text{wall}}$ ” in Fig. 1, operation beyond the highest steady load of  $\text{IMEP}_g = 481 \text{ kPa}$  leads to wall-heating-induced run-away of the combustion phasing. An example of this is shown in Fig. 2 for an  $\text{IMEP}_g$  of  $485 \text{ kPa}$ . At first, the average position of CA50 is fairly steady. However, beyond about  $140 \text{ s}$  it starts to advance at an increasing rate. Due to the advancement of CA50, the PRR

increases greatly, and at the end of the 2000-cycle sample it well exceeds the limit for acceptable PRR at  $8 \text{ bar}/^\circ\text{CA}$ . The cause of this run-away situation was examined by Olsson *et al.* [3]. They concluded that the increased wall heating associated with higher PRR for more advanced cycles feeds back to increasing charge gas temperatures, and thus further advances the combustion phasing. This coupling eventually leads to a run-away situation where the combustion phasing continues to advance in an uncontrollable manner. However, it can be noted that the run-away is not very fast. In the main part of the run-away (*i.e.*  $160 - 190 \text{ s}$ ), the average PRR increases by  $0.11 \text{ bar}/^\circ\text{CA}$  per second. Consequently, implementation of fast closed-loop control of the combustion phasing could possibly enable steady HCCI combustion at  $\text{IMEP}_g > 485 \text{ kPa}$ , as demonstrated in Ref. [3]. However, such operation is beyond the scope of the current study.

Furthermore, Fig. 1 shows that there are differences in thermal efficiency between the fuels. This is mostly related to differences in the thermodynamic efficiency due to differences of the specific heat capacity of the gases used. With higher combustion temperatures, the work-extraction efficiency during the expansion stroke becomes lower due to higher specific heat capacity. This explains why the iso-octane case has the lowest efficiency, since it requires a high compressed-gas temperature to ensure autoignition. Furthermore, higher EGR rates lead to higher specific heat capacity as well, and this explains why the reactive PRF60 has a lower efficiency than PRF80. Finally, it can be noted from the slopes of the PRF80 and PRF60 curves in Fig. 1, that the need to use extensive CA50 retard at the highest

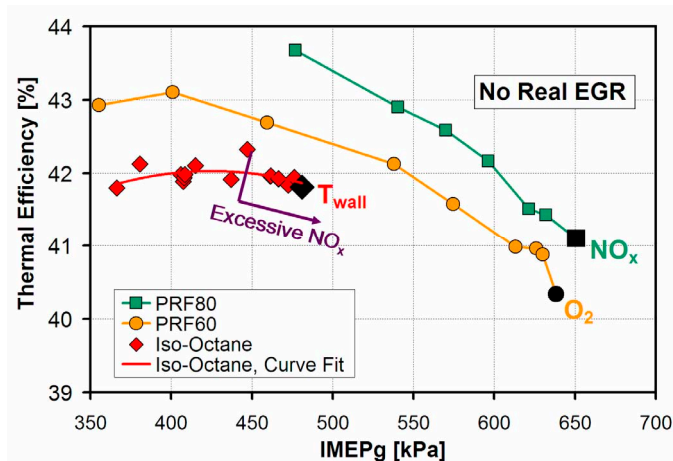


Figure 1. Reproduction of data from Ref. [1] showing thermal efficiency plotted against  $\text{IMEP}_g$  for each fuel at  $\text{PRR} = 5 \text{ bar}/^\circ\text{CA}$ .  $\text{CR} = 14$ . The load-limiting factor is indicated by the annotation located next to each data set:  $\text{NO}_x$  =  $\text{NO}_x$ -induced rapid run-away,  $\text{O}_2$  = oxygen-limited combustion,  $T_{\text{wall}}$  = wall-heating-induced run-away, and excessive level of exhaust  $\text{NO}_x$ .

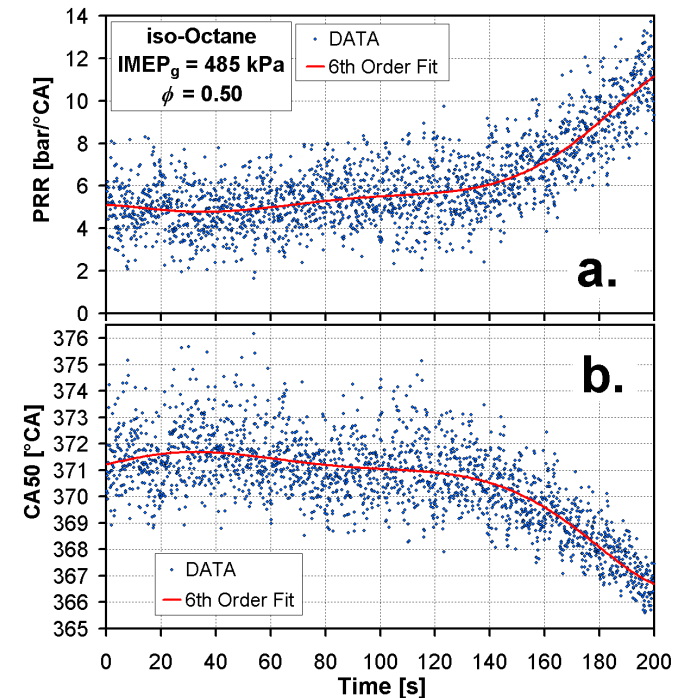


Figure 2. Example of wall-heating-induced run-away for iso-octane operation at a nominal  $\text{PRR}$  of  $5 \text{ bar}/^\circ\text{CA}$  and  $\text{IMEP}_g = 485 \text{ kPa}$ .  $\text{CR} = 14$ . From Ref. [1].

loads in order to control PRR leads to a small but significant efficiency penalty.

## SCOPE AND OBJECTIVES

The main objective of this study is to experimentally investigate how the use of EGR influences the maximum obtainable IMEP<sub>g</sub> while maintaining steady operation with acceptable PRR. Furthermore, and equally important, the load-limiting factor(s) (and the mechanisms behind them) will be identified for each operating condition. The fuels considered are PRF80 and PRF60, both of which exhibit two-stage ignition under the conditions of this study.

First, the experimental setup, data acquisition and analysis are described. This is followed by a discussion of some basic EGR considerations. Then the engine-operating strategy and the experimental approach to the determination of the load limits are explained. The results are divided into three main parts:

1. With adjustments of CA<sub>50</sub> to maintain PRR at 3, 5, or 7 bar/°CA<sup>1</sup>, the IMEP<sub>g</sub> is pushed to the limit for PRF80 when operated with simulated and well-mixed real EGR. The various responses of the engine to increasing the load above the steady-state high-load limit are observed. This is followed by an examination of the contribution from various mechanisms to the observed run-away for PRF80.
2. The load-limits are determined for PRF60 when operated with both simulated EGR and well-mixed real EGR. The causes of the observed run-away for PRF60 with real EGR are examined following the stepwise analysis of PRF80.
3. The implications of a slight EGR/fresh gas unmixedness on the high-load limits are determined for PRF60. This includes quantification of the level of unmixedness and identification of the mechanisms responsible for the observed deterioration of combustion stability.

For a given displacement, the power output is proportional to both the brake mean effective pressure (BMEP) and the engine speed. However, to limit the parameter space, this study compares the performance of the different fuels and EGR systems at a fixed engine speed of 1200 rpm and for  $P_{in} = 100$  kPa, simulating naturally aspirated operation. Both the engine speed and the intake-boost pressure influence the observed autoignition reactivity and the amount of low-temperature heat release (LTHR) [4,5]. Because of this, the differences between the fuels may grow or diminish for operation with boost or at other engine speeds. A CR = 14 piston was installed for all tests. However, the choice of compression ratio (CR) influences both  $T_{in}$  and EGR requirements, thus affecting the performance of a given

fuel [1,6,7]. These aspects should be considered when comparing the data presented here with data acquired at other operating conditions in other engines.

## EXPERIMENTAL SETUP

The engine used for this study is based on a Cummins B-series diesel engine, which is a typical medium-duty diesel engine with a displacement of 0.98 liters/cylinder. A schematic of the setup is shown in Fig. 3. The configuration of the engine and facility is nearly the same as for previous studies (e.g. Refs. [2,4]). However, one important addition is the new EGR loop, which features an EGR cooler. This is the EGR system used for all “well-mixed real EGR” data. The old EGR loop does not provide perfect mixing between the EGR gases and the air, so this old EGR loop is used to study the effects of “heterogeneous EGR” in the last section of the paper.

Furthermore, the schematic in Fig. 3 shows that the installed piston features a hemispherical bowl in the center. This custom-design piston provides a small top-land-ring crevice, amounting to only 2.1% of the top-dead-center (TDC) volume, including the volume behind the top piston ring.

For all operating conditions, except for the NO<sub>x</sub> sensitivity measurements presented in the Appendix, the engine was fueled using a fully premixed fueling system featuring an electrically heated fuel vaporizer. The intake gases were metered with separate sonic nozzles for air, CO<sub>2</sub>, and N<sub>2</sub>. To supply H<sub>2</sub>O to the intake charge, a metering pump was used to draw water from a bottle placed on a digital scale and feed it into the fuel vaporizer. A computer continually monitored the scale to verify the water-supply rate. A main air heater (Fig. 3) and tank heaters (not shown) were used to preheat the intake gases. An auxiliary air heater mounted close to the engine was used to precisely control the intake temperature. (The electric energy used by the intake heaters has not been included when computing the thermal efficiency. In a practical application, intake air heating can be accomplished by exploiting the available exhaust enthalpy.) Engine specifications are listed in Table 1. A detailed description of the engine modifications for HCCI operation can be found in Ref. [8].

Cylinder pressure measurements were made with a transducer (AVL QC33C) mounted in the cylinder head approximately 42 mm off center. The pressure transducer signals were digitized and recorded at ¼° crank angle (CA) increments for at least 100 consecutive cycles (except for the NO<sub>x</sub> sensitivity data in the Appendix, for which 40 alternate-fire cycles were recorded.) To compute the PRR, each individual cycle is first low-pass filtered (< 2.5 kHz) and then analyzed with a linear fit over a moving ±0.5°CA window to extract the maximum PRR. The reported PRRs are computed by averaging at least 100 individual cycles (except for graphs that show how the PRR varies from one cycle to the next, e.g. Fig. 2a). Similarly, the 50% burn point

<sup>1</sup> At the conditions studied, the engine knock/ringing is considered unacceptable when PRR > 8 bar/°CA.

(CA50) was first computed from the low-pass filtered pressure trace for each individual cycle (without heat-transfer correction), and then averaged. However, it should be noted that for operation with the presence of LTHR, the LTHR is excluded when the cumulative heat release is computed. Effectively, the reported combustion phasings refer to CA50 for the main combustion event, starting at the crank angle of minimum heat-release rate between the low- and high-temperature combustion events. Presenting the data referring to the main combustion event alone is considered more relevant from the standpoint of quantifying the timing of the main combustion event.

Mass-averaged charge temperatures during the closed part of the cycle (*i.e.* compression/expansion) are computed using the ideal-gas law in combination with the low-pass filtered measured pressure, the known cylinder volume, and the trapped mass. The average molecular weight used for the calculation during the compression stroke corresponds to that of the trapped gases - including fresh gases, EGR and retained residuals, and the fuel. During the combustion event, and in proportion to the mass fraction burned, the molecular weight gradually changes to that of the measured exhaust composition. This exhaust molecular weight is then used for the remainder of the expansion

TABLE 1. Engine Specifications

|                                                  |              |
|--------------------------------------------------|--------------|
| Displacement (single-cylinder) .....             | 0.981 liters |
| Bore .....                                       | 102 mm       |
| Stroke.....                                      | 120 mm       |
| Connecting Rod Length .....                      | 192 mm       |
| Nominal Geometric Compression Ratio .....        | 14:1         |
| No. of Valves.....                               | 4            |
| IVO.....                                         | 0°CA*        |
| IVC.....                                         | 202°CA*      |
| EVO .....                                        | 482°CA*      |
| EVC.....                                         | 8°CA*        |
| Intake Air Swirl Ratio, Both Ports Combined..... | 0.9          |

\* 0°CA is taken to be TDC intake, so TDC combustion is 360°CA. Valve closing and opening timings refer to 0.1 mm lift.

stroke. The peak temperatures reported in this paper are extracted from these mass-averaged temperature traces.

The fuels used were blends of the primary reference fuels (PRF), iso-octane and n-heptane. The PRF blends allow selection of autoignition reactivity by varying the blending proportions. They are also fairly volatile, and this facilitates the charge preparation. Table 2 shows the fuel compositions and other relevant parameters.

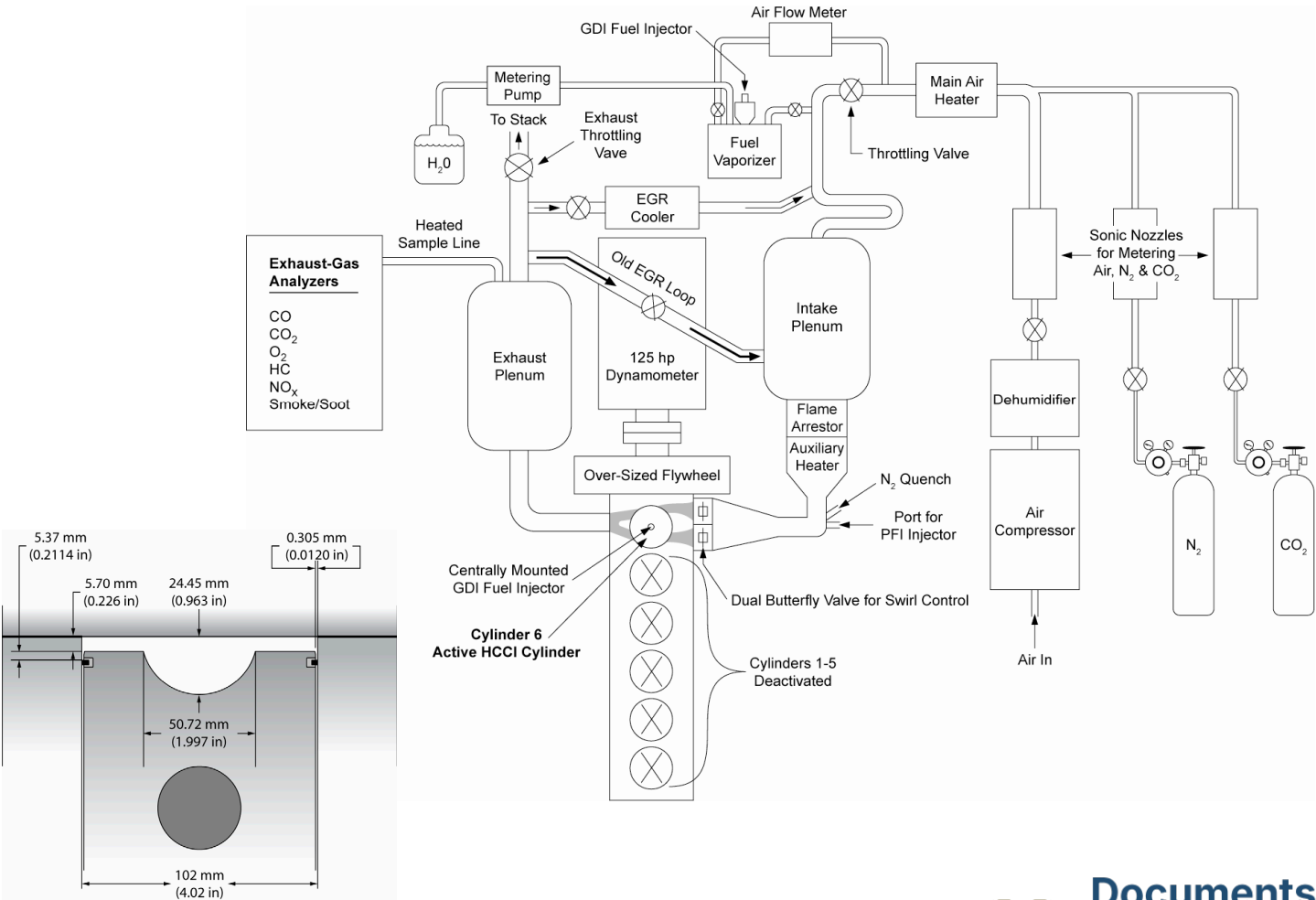


Figure 3. Schematic of the HCCI engine facility, and combustion-chamber dimensions at TDC with the CR = 14 piston

TABLE 2. Fuel Properties

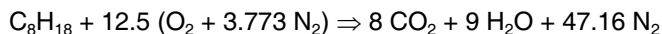
| Fuel                                                                   | Iso-Octane | PRF80        | PRF60        |
|------------------------------------------------------------------------|------------|--------------|--------------|
| Iso-Octane %vol.                                                       | 100        | 80           | 60           |
| n-Heptane %vol.                                                        | 0          | 20           | 40           |
| C-atoms                                                                | 8          | 7.78         | 7.57         |
| H-atoms                                                                | 18         | 17.56        | 17.14        |
| Molecular weight                                                       | 114.23     | 111.15       | 108.21       |
| A/F Stoichiometric                                                     | 15.13      | 15.14        | 15.15        |
| Lower heating value for gas-phase fuel [MJ/kg fuel]                    | 44.65      | 44.71        | 44.76        |
| Lower heating value for stoichiometric charge [MJ/kg air-fuel mixture] | 2.769      | 2.770        | 2.772        |
| Boiling point / range at 1 atm [°C]                                    | 99.2       | 98.4* - 99.2 | 98.4* - 99.2 |

\* The boiling point for pure n-heptane is 98.4°C.

Because HCCI combustion is very sensitive to temperature, the engine coolant and oil were both fully preheated to 100°C before starting the experiments. In addition, the intake tanks and plumbing were preheated to 60°C to avoid any condensation of the water present in the EGR gases. For all data presented, 0°CA is defined as TDC intake (so TDC compression is at 360°). This eliminates the need to use negative crank angles or combined bTDC, aTDC notation.

## BASIC EGR CONSIDERATIONS

In an engine operating lean with complete combustion, EGR consists of air and complete stoichiometric products. Typically, there are combustion inefficiencies, and these generate trace species like carbon monoxide (CO), hydrocarbons (HC) and partially-oxidized hydrocarbons (OHC). In addition, NO<sub>x</sub> can be formed if the combustion temperatures are sufficiently high. All these species can potentially influence the autoignition. To investigate the influence of these combustion by-products, this study compares the use of real EGR and simulated EGR. The simulated EGR consists of a mixture of air with the three components of complete stoichiometric products (CSP), namely CO<sub>2</sub>, H<sub>2</sub>O and N<sub>2</sub>. The relative proportions of these components vary slightly with fuel-type and can be computed based on the average number of C and H atoms listed in Table 2. For iso-octane, complete stoichiometric combustion can be written as:



where atmospheric argon and CO<sub>2</sub> have been lumped into the atmospheric N<sub>2</sub>, following Ref. [9]. For complete combustion with  $\phi = 1$ , the gas composition (excluding fuel) changes from: 20.95% O<sub>2</sub> and 79.05% N<sub>2</sub> for intake air, to products with 12.47% CO<sub>2</sub>, 14.03% H<sub>2</sub>O and 73.50% N<sub>2</sub> for wet exhaust. This exhaust composition

TABLE 3. Composition of CSP for all three fuels (% mole fraction). Average molecular weight (M) is also given.

| Fuel:            | Iso-Octane | PRF80 | PRF60 |
|------------------|------------|-------|-------|
| CO <sub>2</sub>  | 12.47      | 12.45 | 12.44 |
| H <sub>2</sub> O | 14.03      | 14.05 | 14.08 |
| N <sub>2</sub>   | 73.50      | 73.49 | 73.48 |
| M                | 28.71      | 28.71 | 28.70 |

will be referred to as CSP. As Table 3 shows, the CSP composition varies only marginally with changes in PRF number.

To simulate various amounts of EGR, air and CSP were metered into the intake system to form a homogeneous mixture. Figure 4 demonstrates how the intake gas composition changes with the CSP mole fraction. The mole fractions of CO<sub>2</sub> and H<sub>2</sub>O both increase linearly with increasing fraction of CSP, while [O<sub>2</sub>] decreases linearly. When evaluating data with simulated EGR, it can be beneficial to consider the intake [O<sub>2</sub>]. For such cases, Fig. 4 can be used to convert the given %CSP to [O<sub>2</sub>].

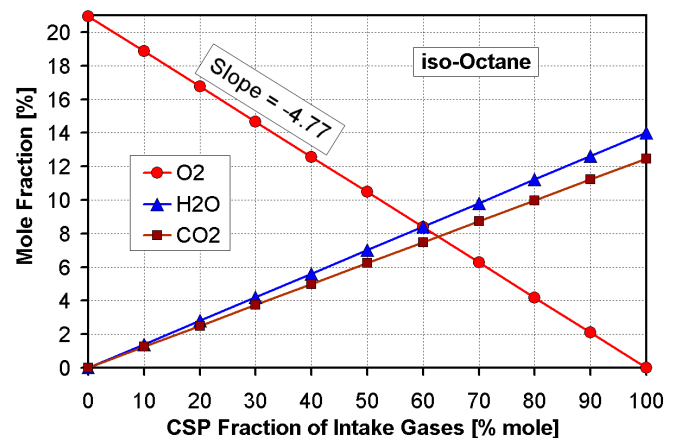


Figure 4. Intake mole fraction of O<sub>2</sub>, CO<sub>2</sub>, and H<sub>2</sub>O as a function of CSP mole fraction. Excluding fuel.

## APPROACH

As discussed in the introduction, the intake-temperature requirement can influence the load capabilities of the HCCI engine. This happens because the charge density changes with  $T_{in}$ . Both PRF80 and PRF60 are sufficiently reactive that both require low  $T_{in}$  for high-load operation. Because it is impractical to work with  $T_{in}$  lower than normal room temperature, and to avoid condensation of water in the intake system, for these tests  $T_{in}$  was held constant at 60°C. The combustion phasing (CA50) was then controlled by adjusting the amount of real or simulated EGR as the fueling was changed.

For HCCI to be practical for automotive use, many wide load ranges must be covered. Within the load range, the

combustion phasing must be easily controlled in a robust and repeatable manner. As will be clear from the results presented below, the combustion phasing generally becomes more sensitive to small disturbances in EGR rate as the load is increased towards the upper load limit. This work does not consider operation with closed-loop CA50 control, which can allow continuous operation at combinations of CA50 and fueling rate that would otherwise spontaneously lead to misfire or combustion-phasing run-away [3]. Instead, the upper load limit in this study is determined by the combination of fueling and CA50 which can be operated continuously and steadily with low cycle-to-cycle variations of the IMEP<sub>g</sub> by only applying slow (~minute timescale) adjustments of the EGR rate to correct for minor disturbances. As will be shown, it becomes impossible to obtain steady operation above this IMEP<sub>g</sub> limit. However, the factor that makes steady operation impossible often changes with both fuel type and CA50, as will be shown.

In the process of determining these steady-state high-load limits, the fueling rate was increased stepwise from moderate fueling rates to high fueling rates while data were recorded for each steady operating point. Generally, the cylinder pressure was recorded also during the transition from steady operation to unsteady run-away as the IMEP<sub>g</sub> was increased above the steady load limit.

To allow efficient data acquisition and subsequent comparison of the fuels, as many parameters as possible should be held constant as the fueling rate is increased. However, when spanning very wide ranges of fueling rates up to the load limit, it becomes impossible to work with a predetermined CA50 while maintaining acceptable engine performance. The reason is that the PRR increases strongly with increasing fueling when the CA50 is held fixed, thus quickly leading to knocking combustion. Knocking combustion can be avoided if CA50 is retarded with increasing fueling rate such that a reasonable PRR is maintained across the IMEP<sub>g</sub> range of interest. Retarding the combustion phasing amplifies the benefit of the existing thermal stratification, leading to a longer combustion duration. In addition, having the combustion take place away from TDC reduces the total pressure rise associated with the combustion since the in-cylinder volume is larger than at TDC. These beneficial effects of combustion-phasing retard are explained in detail in Ref. [1]. For the current engine speed of 1200 rpm, the maximum acceptable PRR  $\approx 8$  bar/°CA, which corresponds to a ringing intensity of 5 MW/m<sup>2</sup> [10]. Based on these considerations, fueling-rate sweeps were conducted while maintaining the PRR constant at 3, 5, or 7 bar/°CA, following the approach of previous work [1].

It is important to maintain a high thermal efficiency of the HCCI engine. Retarding CA50 reduces the effective expansion ratio and can potentially reduce the thermal efficiency. However, because the piston motion is small around TDC, it is not until CA50 >  $\approx 372^\circ\text{CA}$  that a substantial impact on the thermal efficiency is noted.

This effect is the main cause for the reduction of thermal efficiency with increasing load for PRF80 and PRF60 in Fig. 1. (Increasing specific heat capacity of the gases due to higher amounts of EGR and fuel also contribute to a reduction of the thermal efficiency for higher IMEP<sub>g</sub>.) Even so, it is not a viable option to limit the CA50 retard to  $\approx 372^\circ\text{CA}$  since operation with excessive PRR and knocking leads to increased wall heat transfer and even larger reductions of the thermal efficiency than does CA50 retard.

(CA50 =  $372^\circ\text{CA}$  is not a well-defined point beyond which the thermal efficiency drops distinctly. Nonetheless, this approximate number is supported by visual examination of graphs that plot thermal efficiency against CA50, both for the experimental data sets presented here and elsewhere [2], and also for a single-zone CHEMKIN model, not included here.)

## RESULTS

### PRF80 WITH REAL AND SIMULATED EGR

Figure 5a illustrates how CA50 has to be adjusted with fueling rate for three PRRs when PRF80 fuel is used. As expected, for a given IMEP<sub>g</sub>, more combustion-phasing retard is required to achieve a lower PRR. For a given PRR, increasing the fueling rate to reach a higher IMEP<sub>g</sub> requires more CA50 retard. For each PRR, IMEP<sub>g</sub> sweeps with both simulated EGR and well-mixed real EGR are included. In total, six data sets are plotted. For reference, the corresponding charge<sup>2</sup>/fuel-mass ratio (C/F) is plotted in Fig. 5b. Looking at Fig. 5a, it can be observed that for a given PRR, the CA50/IMEP<sub>g</sub> trade-off is identical for both CSP and well-mixed real EGR. This suggests that the combustion processes are nearly identical regardless of the differences in EGR quality.

Also shown in Fig. 5a are the load limits for both EGR types. With CSP, the highest IMEP<sub>g</sub> of 651 kPa is obtained at a PRR of 5 bar/°CA. Operation with a higher or lower PRR leads to a reduction of the obtainable IMEP<sub>g</sub>. If it is attempted to operate the engine beyond the CSP high-load limit, a rapid run-away event occurs for all PRRs. The run-away is particularly rapid for the PRR = 3 bar/°CA data set, as illustrated with Fig. 6. In this particular example, CA50 advances by  $10^\circ\text{CA}$  in 2.3 s, while the PRR increases from 5 to 25 bar/°CA. Thus, the PRR increases on average by 8.7 bar/°CA per second. This rate of PRR increase is roughly 80 times faster than that of iso-octane in Fig. 2. This rules out wall-heating induced run-away as a possible factor. Instead, previous work concluded that the combustion-phasing run-away occurs because of the autoignition-enhancing effect of NO recirculated from one cycle to the next with the residuals [1]. This is explained in the following.

<sup>2</sup> The charge mass is the sum of the supplied air, EGR and CSP constituents, excluding the fuel.