Preprints, 10th IFAC International Symposium on Advances in Automotive Control Columbus, Ohio, USA, August 28-30, 2022

Data-Based In-Cylinder Pressure Model including Cyclic Variations of an RCCI Engine

Maarten Vlaswinkel * **Bram de Jager** * **Frank Willems** *,**

* *Control Systems Technology, Eindhoven University of Technology, Netherlands (email: m.g.vlaswinkel@tue.nl)*
** *TNO Automotive, Helmond, The Netherlands*

Abstract For advanced pre-mixed combustion concepts, Cylinder Pressure-Based Control is a key concept for robust operation. It also opens the possibility for on-line heat release shaping. For cost and time efficient development of these controllers, fast control-oriented combustion models that predict average in-cylinder pressure traces have been proposed. However, they are not able to capture cyclic variations. In this study, a data-based modelling procedure is proposed to predict the in-cylinder pressure trace and cyclic variation during the combustion cycle. The inputs to the model are the in-cylinder conditions at intake valve closing and the fuelling settings. The proposed model is based on experimental data, Principal Component Analysis and Gaussian Process Regression. This new data-driven approach is applied to model the combustion behaviour of a Reactivity Controlled Compression Ignition engine running on Diesel and E85. The resulting model has a root-square-mean-error of average behaviour and cyclic variance of 0.8° and $0.2^{\circ 2}$ in CA50, 0.1 bar and $0.03\,\text{bar}^2$ in Gross Indicated Mean Effective Pressure, and 0.1 % and $0.001\,\%^2$ in the Gross Indicated Efficiency, respectively.

Keywords: Advanced combustion concepts; Dual-fuel combustion; Empirical model; Control-oriented model; Principle Component Analysis; Gaussian Process Regression

1. INTRODUCTION

Internal Combustion Engines (ICEs) will remain the main power source for heavy-duty applications (e.g., overseas and long-distance on-road transportation, agriculture, construction and mining). To support the transition towards sustainable solutions, Reactivity Controlled Compression Ignition (RCCI) is a promising concept. In RCCI both a low and high reactive fuel are used in the combustion process (Reitz and Duraisamy, 2015). By changing the ratio between low and high reactivity fuels it is possible to optimize combustion phasing, duration and magnitude. It is more efficient and cleaner compared to traditional diesel engines and an enabler for using renewable fuels. However, to make it commercially viable, next-cycle control is necessary to guarantee robust and safe operation (Paykani et al., 2021).

Cylinder Pressure-Based Control (CPBC), as a next-cycle control method, is a key concept for enabling RCCI (Willems, 2018). In CPBC, the in-cylinder pressure is used to generate new actuator settings. Acquiring CPBC-oriented models is a major hurdle for model-based control and control evaluation. Especially for model-based control, computation times should be below the duration of a combustion cycle. On top of that, Mean-Value Engine Models (MVEMs) will not provide enough information in the cases of large cyclic variations. Therefore, useful models should include a measure for these variations.

Basic physics-based models have been presented to model the important combustion measures (e.g., Gross Indicated Mean Effective Pressure (IMEP_g) or crank angle where $x\%$ of the total heat is released ($\text{CA}x$)) Khodadadi Sadabadi et al. (2016); Guardiola et al. (2018); Raut et al. (2018); Kakoee et al. (2020). These models do not provide the full in-cylinder pressure trace, but only provide a deterministic dynamic model for combustion measures. As a consequence, new models need to be derived when new measures need to be included and are not suited to be used for Combustion Rate Shaping. This will reduce the amount of flexibility these models provide while developing new control strategies.

More complex physics-based models, like the multi-zone model of Bekdemir et al. (2015) and computational fluid dynamics models of Klos and Kokjohn (2015), have the ability to predict the full in-cylinder pressure trace. However, these models exceed the computation time constraint. Therefore, static deterministic regression models are used for the important combustion measures to meet the required computation time. These regression models have the in-cylinder conditions at Intake Valve Closing (IVC) as input parameters.

By contrast, data-based combustion models have been developed. Xia et al. (2020) propose a data-based model using Gaussian Process Regression (GPR) to map in-cylinder conditions at IVC to important combustion measures. They have used this model to optimise engine performance. Basina et al. (2020) used a data-based state-space model with linear varying parameters to model combustion phasing and Peak Pressure Rise Rate (PPRR). Both methods only provided a model for specific combustion measures, thus including additional combustion measures requires an extension of these models.

A data-based model to model the full in-cylinder pressure trace for varying operating conditions is presented by Pan et al. (2019). Using Principal Component Analysis

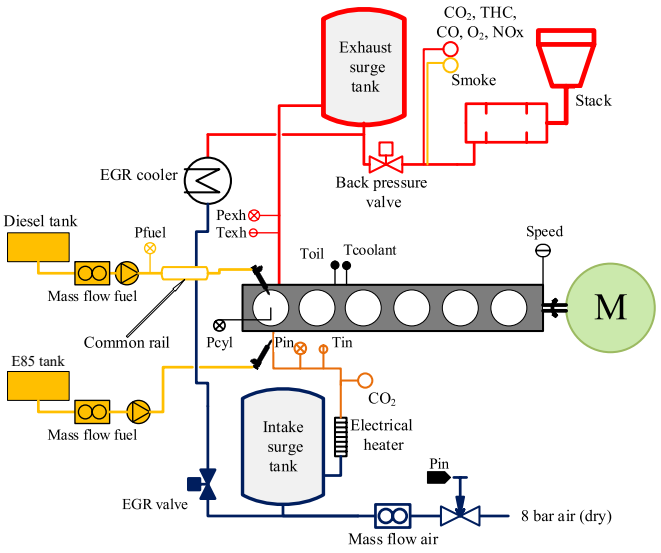


Figure 1. Experimental 6-cylinder PACCAR MX13 engine equipped with a PFI-injector (Willems et al., 2020)

(PCA), they reduce a set of measured in-cylinder pressure traces at different operating conditions to several Principal Components (PCs) and use artificial neural networks to predict the coefficient for each component. With this method, they are able to predict the in-cylinder pressure in real-time.

Except for Xia et al. (2020), all of the above mentioned methods are deterministic, thus do not include information about cyclic variations. This is an important measure for control development, where in regions with large cyclic variations controllers can be tuned to be more conservative. However, in regions with small cyclic variations it becomes possible to tune controllers to be more aggressive. Thus fast convergence can be achieved in regions with small cyclic variations, while still guaranteeing stable operation in regions with large cyclic variations.

In this paper, we will present a data-based CPBC-oriented modelling method that is able to describe the full in-cylinder pressure trace including a measure for cyclic variations. The combustion nucleus is modelled using PCA similar to Pan et al. The artificial neural networks they use are replaced with GPR models to capture the inaccurate measurements of the in-cylinder conditions at IVC and the cyclic variations.

2. ENGINE SETUP

In this paper, the same setup and experimental data has been used as in Willems et al. (2020). Below only the most relevant information is provided.

Figure 1 shows the 6-cylinder PACCAR MX13 engine used in this study. Only cylinder 1 is active. The cylinder heads of cylinders 2-6 have been removed and no fuel is injected into these cylinders. To enable RCCI operation, the engine uses a Direct Injection (DI) system to inject diesel directly into the cylinder and a Port Fuel Injection (PFI) system to inject E85 in the intake port. An electric machine is used to generate the torque required to keep the engine running at a constant speed.

In this paper, the focus is on RCCI with a single injection of diesel to auto-ignite the well-mixed charge of E85, air and recirculated exhaust gas. Table 1 shows the nominal operating conditions. In this study, we have performed $m = 105$ training experiments and $\bar{m} = 45$ validation experiments at different in-cylinder conditions at IVC and fuel setting. Each experiment contains $M = 100$ combustion cycles.

It is not possible to exactly know the in-cylinder conditions at IVC as a result of sensor noise, when one sensor is used for multi-cylinder operation or difference in injected fuel quantity and the fuel quantity available for combustion caused by physical phenomena (e.g., wall-wetting or inadequate vaporization). To give an estimate of the in-cylinder conditions at IVC $\hat{s}_{\text{IVC}}$, we have chosen to use $n = 6$ measurable parameters that capture these conditions, namely:

- Total injected energy
$$Q_{\text{total}} = m_{\text{PFI}}\text{LHV}_{\text{PFI}} + m_{\text{DI}}\text{LHV}_{\text{DI}};$$
- Energy-based blend ratio
$$\text{BR} = \frac{m_{\text{PFI}}\text{LHV}_{\text{PFI}}}{Q_{\text{total}}};$$
- Start-of-injection of the directly injected fuel SOI_{DI};
- Pressure at the intake manifold p_{im};
- Temperature at the intake manifold T_{im}; and
- Exhaust Gas Recirculation (EGR) ratio
$$X_{\text{EGR}} = \frac{\text{CO}_{2,\text{in}}}{\text{CO}_{2,\text{out}}}$$
with $\text{CO}_{2,\text{in}}$ and $\text{CO}_{2,\text{out}}$ the concentration of CO_2 at the intake and exhaust, respectively.

Figure 2 shows the distribution of the variation in in-cylinder conditions at IVC and fuel settings for each measure pressure trace in the training data. On the diagonal the distribution of each in-cylinder conditions at IVC and fuel settings is shown, while the off-diagonal shows the joint distribution between in-cylinder conditions at IVC and fuel settings. A sweep over a range of SOI_{DI} has been applied, changes in other parameters are caused by random uncontrolled variations on the setup.

3. COMBUSTION MODEL

The combustion model described in this section is used to predict the in-cylinder pressure during the combustion cycle. In this approach, PCA is used to reduce the amount of required stored data by decomposing the measured in-cylinder pressure traces into a weighted sum of basis functions. Thereafter, Gaussian Process Regression (GPR) is used to go from a discrete set of measured operating conditions to a continuous operating space by creating a continuous mapping from in-cylinder conditions at IVC and fuel setting to weights used in the weighted sum of basis functions.

Table 1. Specifications and nominal operating conditions of the engine setup

Parameters	Value
PFI fuel	E85
DI fuel	Diesel (EN590)
Compression ratio	15.85
Intake valve closure	−153°CA aTDC
Exhaust valve opening	128°CA aTDC
Intake temperature	40 °C
Gross IMEP	(8.5 ± 1.0) bar
Fuel energy input	(3850 ± 100) J/cycle
Engine speed	1200 rpm
In-cylinder pressure sensor	Kistler 6125C
Crank angle resolution	0.2°

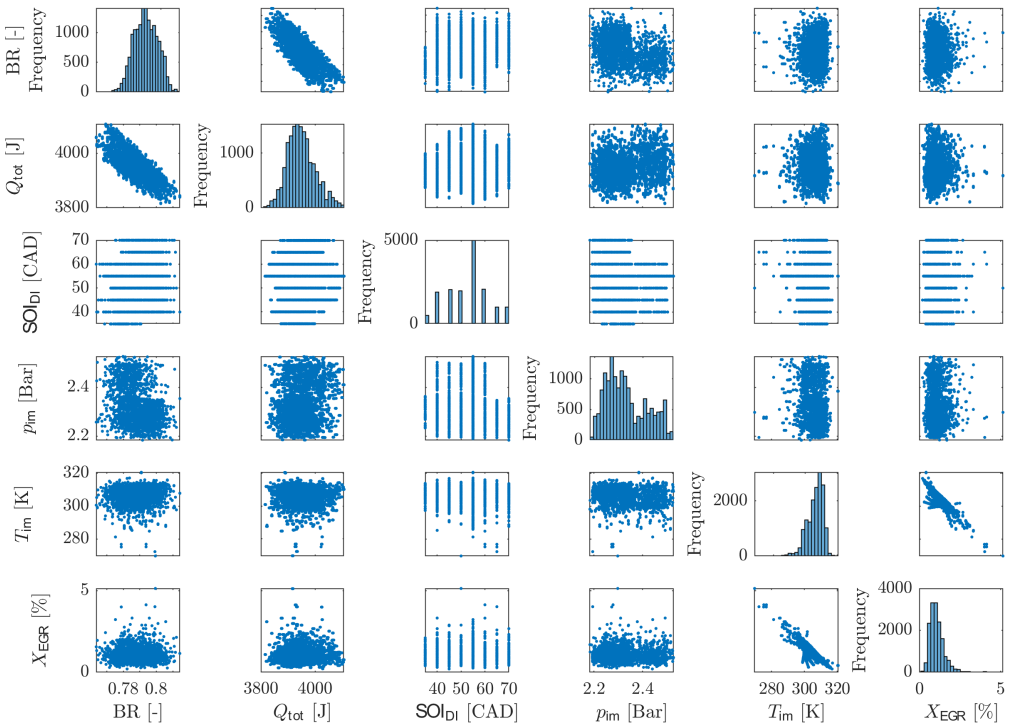


Figure 2. Distribution of in-cylinder conditions at Intake Valve Closing and fuel settings over the mM combustion cycles used during model training and validation

Having mM measured combustion cycles and using PCA similar to Pan et al. (2019), the in-cylinder pressure can be decomposed into $mM-1$ PCs such that the measured pressure trace can be represented by

$$p(\theta,\, \boldsymbol{s}_{\text{IVC}}) = \boldsymbol{w}^{\mathsf{T}}(\hat{\boldsymbol{s}}_{\text{IVC}})\boldsymbol{f}(\theta) + f_{\mu}(\theta) \tag{1}$$

with crank angle $\theta \in (-180,\, 180)$, estimated in-cylinder conditions at IVC and fuel settings $\hat{\boldsymbol{s}}_{\text{IVC}}$ as described in Section 2. The mean pressure over all cycles $f_{\mu}(\theta)$ is determined as

$$f_{\mu}(\theta) := \frac{1}{mM}\sum_{\boldsymbol{s}_{\text{IVC}}} p(\theta,\, \boldsymbol{s}_{\text{IVC}}). \tag{2}$$

The PCs $\boldsymbol{f}(\theta) := [f_1(\theta)\ f_2(\theta)\ \cdots\ f_{mM-1}(\theta)]^{\mathsf{T}}$ being the eigenvectors of $P^{\mathsf{T}}P$, with P a matrix consisting of the elements $p(\theta,\, \boldsymbol{s}_{\text{IVC}}) - f_{\mu}(\theta)$ where the columns span all measured $\boldsymbol{s}_{\text{IVC}}$ and the rows span all measured $\theta \in (-180,\, 180)$. Using $f_{\mu}(\theta)$ and $\boldsymbol{f}(\theta)$, the weights $\boldsymbol{w}(\hat{\boldsymbol{s}}_{\text{IVC}}) := [w_1(\hat{\boldsymbol{s}}_{\text{IVC}})\ w_2(\hat{\boldsymbol{s}}_{\text{IVC}})\ \cdots\ w_{mM-1}(\hat{\boldsymbol{s}}_{\text{IVC}})]^{\mathsf{T}}$ for each measured pressure trace can be found by solving (1).

The PCs are ordered from most relevant to least relevant by ordering the eigenvalues of $P^{\mathsf{T}}P$ in descending order. For illustration, Figure 3 shows $f_{\mu}(\theta)$ and the first four PCs. It shows that the higher-order PCs add higher frequency components to the decomposition. This makes it possible to reduce the required amount of information by selecting the $\tilde{m}$ most relevant PCs based on the accuracy of performance and combustion measures. This gives an estimation of the in-cylinder pressure as

$$p(\theta,\, \hat{\boldsymbol{s}}_{\text{IVC}}) \approx \tilde{p}(\theta,\, \hat{\boldsymbol{s}}_{\text{IVC}}) = \tilde{\boldsymbol{w}}^{\mathsf{T}}(\boldsymbol{s}_{\text{IVC}})\tilde{\boldsymbol{f}}(\theta) + f_{\mu}(\theta) \tag{3}$$

with the reduced set of PCs $\tilde{\boldsymbol{f}}(\theta) := \left[f_1(\theta)\ f_2(\theta)\ \cdots\ f_{\tilde{m}}(\theta)\right]^{\mathsf{T}}$ and weights $\tilde{\boldsymbol{w}}(\boldsymbol{s}_{\text{IVC}}) := \left[w_1(\boldsymbol{s}_{\text{IVC}})\ w_2(\boldsymbol{s}_{\text{IVC}})\ \cdots\ w_{\tilde{m}}(\boldsymbol{s}_{\text{IVC}})\right]^{\mathsf{T}}$.

Small unmeasurable deviations in in-cylinder conditions at IVC and fuel settings make it impossible to have exact knowledge of $\tilde{\boldsymbol{w}}(\hat{\boldsymbol{s}}_{\text{IVC}})$. Therefore, we will describe $\tilde{\boldsymbol{w}}(\hat{\boldsymbol{s}}_{\text{IVC}})$ as a stochastic process, such that

$$\tilde{\boldsymbol{w}}(\hat{\boldsymbol{s}}_{\text{IVC}}) := \mathcal{N}(\hat{\boldsymbol{w}}(\hat{\boldsymbol{s}}_{\text{IVC}}),\, W(\hat{\boldsymbol{s}}_{\text{IVC}})) \tag{4}$$

with $\hat{\boldsymbol{w}}(\hat{\boldsymbol{s}}_{\text{IVC}}) := \mathbb{E}[\boldsymbol{w}(\hat{\boldsymbol{s}}_{\text{IVC}})]$ and $W(\hat{\boldsymbol{s}}_{\text{IVC}}) := \mathbb{E}[(\tilde{\boldsymbol{w}}(\hat{\boldsymbol{s}}_{\text{IVC}}) - \hat{\boldsymbol{w}}(\hat{\boldsymbol{s}}_{\text{IVC}}))(\tilde{\boldsymbol{w}}(\hat{\boldsymbol{s}}_{\text{IVC}}) - \hat{\boldsymbol{w}}(\hat{\boldsymbol{s}}_{\text{IVC}}))^{\mathsf{T}}]$. This makes it possible to describe the in-cylinder pressure as a stochastic process with mean and covariance as

$$\mathbb{E}\left[\tilde{p}(\theta,\, \hat{\boldsymbol{s}}_{\text{IVC}})\right] = \hat{\boldsymbol{w}}^{\mathsf{T}}(\hat{\boldsymbol{s}}_{\text{IVC}})\tilde{\boldsymbol{f}}(\theta) + f_{\mu}(\theta) \tag{5}$$

and

$$\mathbb{E}\left[\left(\tilde{p}(\theta, \hat{\boldsymbol{s}}_{\text{IVC}}) - \mathbb{E}[\tilde{p}(\theta, \hat{\boldsymbol{s}}_{\text{IVC}})]\right)^2\right] = \tilde{\boldsymbol{f}}^{\mathsf{T}}(\theta)W(\hat{\boldsymbol{s}}_{\text{IVC}})\tilde{\boldsymbol{f}}(\theta). \tag{6}$$

Notice that in (6) $w_i(\hat{\boldsymbol{s}}_{\text{IVC}})$ and $w_j(\hat{\boldsymbol{s}}_{\text{IVC}})$ $\forall i,\, j \in \{1,\, 2,\, \ldots,\, \tilde{m}\}$ are not assumed to be independent.

Figure 4 shows the (joint) distribution of $\tilde{\boldsymbol{w}}(\hat{\boldsymbol{s}}_{\text{IVC}})$ over M cycles at the same operating point $\hat{\boldsymbol{s}}_{\text{IVC}}$ for the case $\tilde{m} = 5$. Comparable results are obtained for different operating points. On the diagonal, it shows that $w_i^{\mathsf{T}}(\hat{\boldsymbol{s}}_{\text{IVC}})$ $\forall i \in \{1,\, 2,\, \ldots,\, \tilde{m}\}$ can be approximated by a normal distribution. We will use the Pearson correlation coefficient matrix R to investigate the degree of linear dependence of $w_i(\hat{\boldsymbol{s}}_{\text{IVC}})$ and $w_j(\hat{\boldsymbol{s}}_{\text{IVC}})$ $\forall i,\, j \in \{1,\, 2,\, \ldots,\, \tilde{m}\}$. The Pearson correlation coefficient is given by

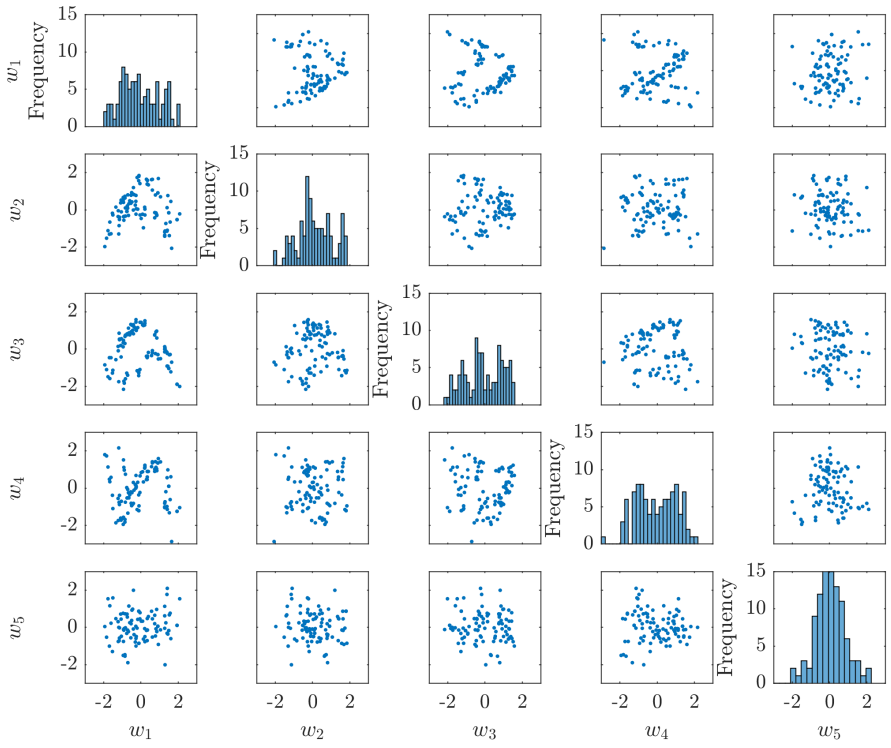


Figure 4. Distribution of scaled $w^{\mathsf{T}}(\hat{\boldsymbol{s}}_{\mathrm{IVC}})$ for 100 cycles at the same operating point $\hat{\boldsymbol{s}}_{\mathrm{IVC}}$ for the first five principle components. The diagonal elements show the histogram of $w_i(\hat{\boldsymbol{s}}_{\mathrm{IVC}})$ and the off diagonal elements show the distribution of $w_i(\hat{\boldsymbol{s}}_{\mathrm{IVC}})$ with $w_j(\hat{\boldsymbol{s}}_{\mathrm{IVC}})$.

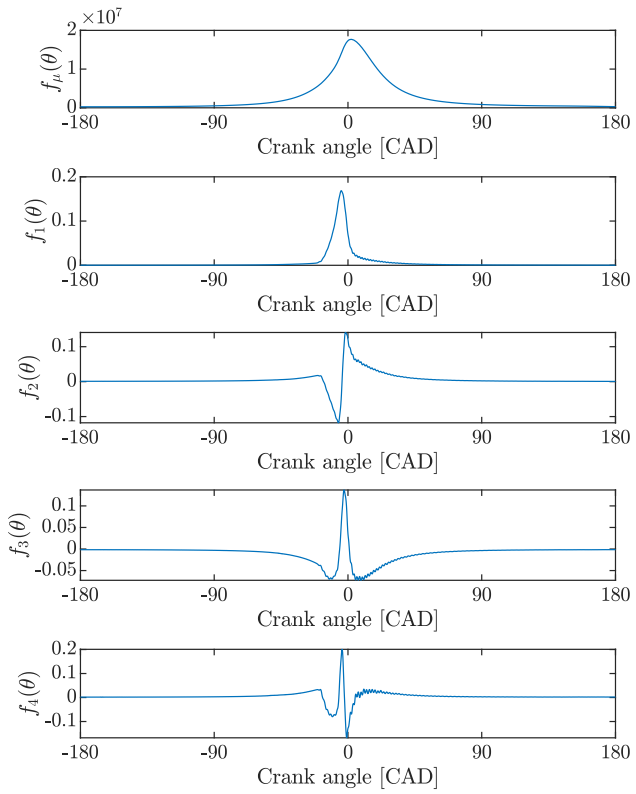


Figure 3. $f_\mu(\theta)$ and the first four PCs

$$r_{ij} = \frac{1}{M-1}\sum_{k=1}^{M} S\left\{\frac{w_{i,k}(\hat{\boldsymbol{s}}_{\mathrm{IVC}}) - \mu(w_i(\hat{\boldsymbol{s}}_{\mathrm{IVC}}))}{\sigma(w_i(\hat{\boldsymbol{s}}_{\mathrm{IVC}}))} \times \frac{w_{j,k}(\hat{\boldsymbol{s}}_{\mathrm{IVC}}) - \mu(w_j(\hat{\boldsymbol{s}}_{\mathrm{IVC}}))}{\sigma(w_j(\hat{\boldsymbol{s}}_{\mathrm{IVC}}))}\right\},$$

where $\mu(w_i(\hat{\boldsymbol{s}}_{\mathrm{IVC}}))$ and $\sigma(w_i(\hat{\boldsymbol{s}}_{\mathrm{IVC}}))$ are the mean and standard deviation of the weights for each cycle $w_{i,k}(\hat{\boldsymbol{s}}_{\mathrm{IVC}})$, respectively. For the shown operating point, the symmetric Pearson correlation coefficient matrix is given by

$$R = \begin{bmatrix} 1 & & & & \\ -0.10 & 1 & & & \\ 0 & 0.17 & 1 & & \\ 0.14 & 0.09 & -0.02 & 1 & \\ -0.06 & 0.10 & 0 & 0.14 & 1 \end{bmatrix}$$

and shows small linear dependence between $w_i(\hat{\boldsymbol{s}}_{\mathrm{IVC}})$ and $w_j(\hat{\boldsymbol{s}}_{\mathrm{IVC}})$ $\forall i, j \in \{1, 2, \ldots, \tilde{m}\}$. Now we can assume that $w_i(\hat{\boldsymbol{s}}_{\mathrm{IVC}})$ and $w_j(\hat{\boldsymbol{s}}_{\mathrm{IVC}})$ are linear independent ($W_{ij}(\hat{\boldsymbol{s}}_{\mathrm{IVC}}) = 0$ when $i \neq j$) and (6) can be reduced to

$$\mathbb{E}\left[(\tilde{p}(\theta, \hat{\boldsymbol{s}}_{\mathrm{IVC}}) - \mathbb{E}[\tilde{p}(\theta, \hat{\boldsymbol{s}}_{\mathrm{IVC}})])^2\right] = \textstyle\sum_{i=1}^{\tilde{m}} W_{ii}(\hat{\boldsymbol{s}}_{\mathrm{IVC}}) f_i(\theta)^2. \quad (7)$$

It is impossible to measure the in-cylinder pressure for every possible $\hat{\boldsymbol{s}}_{\mathrm{IVC}}$. Therefore, GPR is used to create a mapping from $\hat{\boldsymbol{s}}_{\mathrm{IVC}}$ to $\hat{w}_i(\hat{\boldsymbol{s}}_{\mathrm{IVC}})$ and $W_{ii}(\hat{\boldsymbol{s}}_{\mathrm{IVC}})$ using a limited set of measurements.

Using GPR, the expected value and covariance matrix can be computed as (Rasmussen and Williams, 2006):

$$\hat{w}_i(\hat{\boldsymbol{s}}_{\mathrm{IVC}}) = K(\hat{\boldsymbol{s}}_{\mathrm{IVC}}, \hat{\boldsymbol{s}}^*_{\mathrm{IVC}}, \phi) K(\hat{\boldsymbol{s}}^*_{\mathrm{IVC}}, \hat{\boldsymbol{s}}^*_{\mathrm{IVC}}, \phi)^{-1} w_i(\hat{\boldsymbol{s}}^*_{\mathrm{IVC}}) \quad (8)$$

and

$$W_{ii}(\hat{\boldsymbol{s}}_{\mathrm{IVC}}) = K(\hat{\boldsymbol{s}}_{\mathrm{IVC}}, \hat{\boldsymbol{s}}_{\mathrm{IVC}}, \phi) - K(\hat{\boldsymbol{s}}_{\mathrm{IVC}}, \hat{\boldsymbol{s}}^*_{\mathrm{IVC}}, \phi) K(\hat{\boldsymbol{s}}^*_{\mathrm{IVC}}, \hat{\boldsymbol{s}}^*_{\mathrm{IVC}}, \phi)^{-1} K^{\mathsf{T}}(\hat{\boldsymbol{s}}_{\mathrm{IVC}}, \hat{\boldsymbol{s}}^*_{\mathrm{IVC}}, \phi), \tag{9}$$

where $\boldsymbol{s}^*_{\mathrm{IVC}}$ contains all fuelling settings and in-cylinder conditions used during training.

For the kernel matrix K, we have chosen a squared exponential kernel with Automatic Relevance Determination (ARD), because $w_i(\boldsymbol{s}_{\mathrm{IVC}})$ is normally distributed. The kernel matrix has entries

$$K_{ij} = k(x_i, x_j) = \sigma_f^2 \exp\left(-\tfrac{1}{2}(x_i - x_j)^{\mathsf{T}} \Theta^{-2} (x_i - x_j)\right) + \delta_{ij}\sigma_n$$

with Θ a diagonal matrix of terms $\phi_{l,i}$ with $i \in \{1, 2, \ldots, n\}$. The positive scalars ϕ_f, $\phi_{l,i}$ and ϕ_n are hyperparameters grouped as ϕ, and δ_{ij} is the Kronecker delta. The hyperparameters are found by minimizing the negative log marginal likelihood for each $w_i(\boldsymbol{s}_{\mathrm{IVC}})$ as

$$-\log p(w_i|\hat{\boldsymbol{s}}_{\mathrm{IVC}}, \phi) = \tfrac{n}{2}\ln(2\pi) + \tfrac{1}{2}\ln\det(K(\boldsymbol{s}^*_{\mathrm{IVC}}, \boldsymbol{s}^*_{\mathrm{IVC}}, \phi)) + \tfrac{1}{2} w_i(\boldsymbol{s}^*_{\mathrm{IVC}})^{\mathsf{T}} K^{-1}(\boldsymbol{s}^*_{\mathrm{IVC}}, \boldsymbol{s}^*_{\mathrm{IVC}}, \phi) w_i(\boldsymbol{s}^*_{\mathrm{IVC}}). \tag{10}$$

In this study, we have used the default implementation in MATLAB's Statistics and Machine Learning Toolbox to determine the hyperparameters.

4. VALIDATION

In this section, the validation of the method in Section 3 is performed using the RCCI engine described in Section 2. First, a method to determining the required number of PCs and the quality of relevant combustion parameters is shown. Afterwards, a comparison is made between measured and modeled pressure traces.

It is well known that Gaussian process regression performs poorly at the boundary of the training data. Therefore, the experiments that can be used for validation are chosen such that

$$\|\hat{\boldsymbol{s}}_{\mathrm{IVC}} - \mu(\hat{\boldsymbol{s}}_{\mathrm{IVC}})\|_2 < \|\sigma(\hat{\boldsymbol{s}}_{\mathrm{IVC}})\|_2 \tag{11}$$

holds, where $\mu(\hat{\boldsymbol{s}}_{\mathrm{IVC}})$ and $\sigma(\hat{\boldsymbol{s}}_{\mathrm{IVC}})$ are the mean and standard deviation over each measurable in-cylinder conditions at IVC. A random selecting of 30% of the experiments, each containing 100 combustion cycles, that lie in this set are used as validation data. These experiments are not used during model training.

The Root Mean Square (RMS) between measured and modelled average behaviour and cyclic variations in combustion timing and performance measures at 45 different operating points is shown in Figure 5 for an increasing number of PCs. We selected to evaluate the accuracy of CA50 as a measure for the accuracy of the underlying heat release, $\mathrm{IMEP_g}$ and Gross Indicated Efficiency (GIE) for its importance during engine calibration. The $\mathrm{IMEP_g}$ is given by

$$\mathrm{IMEP_g} = \frac{\int_{\theta=-180^\circ}^{180^\circ} p(\theta)\, dV(\theta)}{V_{\mathrm{d}}}$$

and GIE as

$$\mathrm{GIE} = \frac{\int_{\theta=-180^\circ}^{180^\circ} p(\theta)\, dV(\theta)}{m_{\mathrm{DI}}\mathrm{LHV_{DI}} + m_{\mathrm{PFI}}\mathrm{LHV_{PFI}}},$$

where $\mathrm{LHV_{DI}}$ and $\mathrm{LHV_{PFI}}$ are the lower heating values of the direct and port fuel injected fuels, respectively. From these figures, the required number of PCs can be determined as the point where the error does not decrease when more PCs are added. The best average performance for CA50 (Figure 5a) is achieved when 3 or more PCs are used and 4 or more PCs when cyclic variation is taken into account. This results in a RMS of 0.8° in average behaviour and $0.2\,^{\circ 2}$ in cyclic variation. Figures 5b and 5c show the error in $\mathrm{IMEP_g}$ and GIE. They show that the performance measures are less affected by the number of PCs and errors in both average behaviour (0.1 bar for $\mathrm{IMEP_g}$ and 0.1% for GIE) and cyclic variations ($0.03\,\mathrm{bar}^2$ for $\mathrm{IMEP_g}$ and $0.001\%^2$ for GIE) are small. This can be explained by the fact that higher-order PCs add high frequency information which does not alter the overall shape of the in-cylinder pressure in much detail.

Table 2. Computation time of pressure trace and related variance using four PCs for a different number of measured in-cylinder pressure traces

#Experiments m	#Cycles M	Computation time
100	100	0.36 s
100	20	0.18 s
50	100	0.09 s
50	20	0.06 s

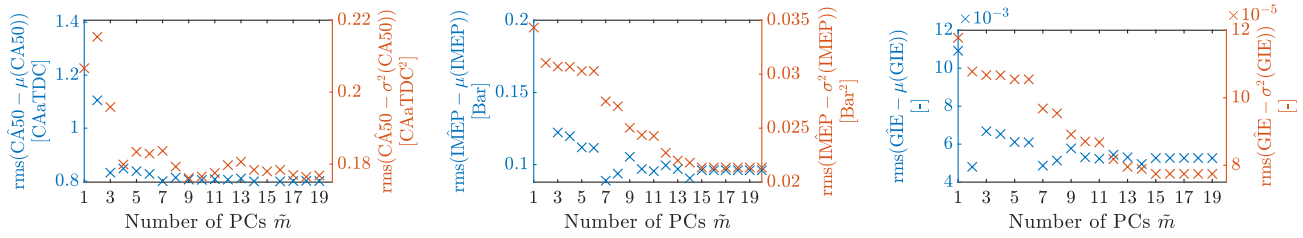
Figures 6a and 6b show the output of the GPR combustion model with 4 PCs at two different operating points. In each figure, all 100 measured cycles and the mean of these cycles at a single operating point are shown. The output of the model at this operating point includes the expected in-cylinder pressure derived using (5) and (8), and the 95%-confidence interval computed using (7) and (9). We observe two kinds of behaviours depending on the operating point: mean and cyclic variations are modelled accurately (Figure 6a) or mean behaviour fit well but the uncertainty bound is too small (Figure 6b). For the case when mean and cyclic variations are modelled accurately the operating point is far away from the border of the space described by (11). For the case when the uncertainty bound is too small the operating point lies on the edge of the space described by (11).

The computation time to compute a pressure trace and related variance increases when more measured in-cylinder pressure traces are added or more PCs are used. Table 2 shows the computation time for different amounts of experiments m and combustion cycles per experiment M. These computations times are achieved on a laptop with an Intel Core i7-9750H CPU running at 2.6 GHz, it runs Windows 10 with a MATLAB 2021b installation. It can be seen that only the models with $m = 50$ meet the computation time limitation of 0.1 s. Each added PC adds additional computation complexity, since more $\hat{w}_i(\hat{\boldsymbol{s}}_{\mathrm{IVC}})$ are required. In the original model with $m = 105$ and $M = 100$, for every PC added 0.09 s of computation time is added with the current amount of measured in-cylinder pressure traces.

5. CONCLUSION

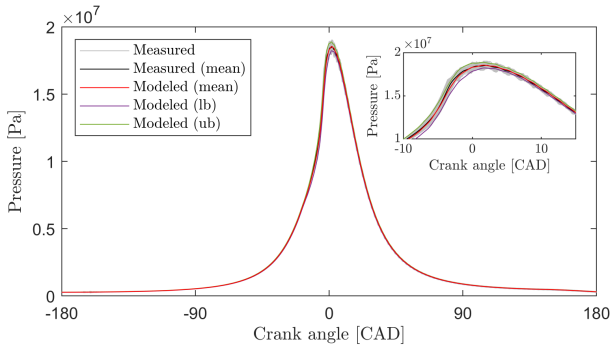
In this paper, we have presented a data-based CPBC-oriented modelling method that is able to describe the in-cylinder pressure trace during the combustion stroke including a measure for cyclic variations. Similar to Pan et al., we use a weighted sum of PCs. However, the required weights are mapped using GPR instead of neural network. This makes it possible to capture cyclic variations.

The modelled operating region is relatively small, we will extend this region to cover the full operating space. Including a smart method for collecting the data to meet computation time limitation. We will also provide a more thorough validation using real engine data in future work.

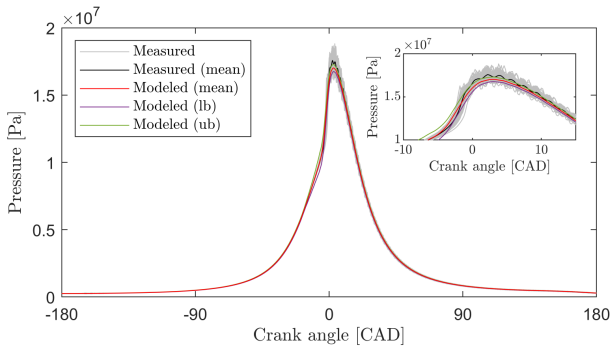


(a) CA50 (b) IMEP_g (c) GIE

Figure 5. Validation of the average behaviour (blue) and cyclic variation (red) in combustion timing (CA50) and performance measures (IMEP_g and GIE) at 45 different operating points



(a) Accurate fit



(b) Uncertainty bound too small

Figure 6. Measured and modelled in-cylinder pressure where $\tilde{m} = 4$ at two different in-cylinder conditions at IVC and fuel setting

Furthermore, to capture the full engine capabilities, an air-path model (e.g., based on a mean-value engine model) should be included.

We will also show how the presented framework can be used in the development of advanced feedback control methods. These methods can include information about the full in-cylinder pressure and therefore maximize efficiency. Furthermore, the information about cyclic variations can be used to create faster controllers while still being robust.

With our framework, we are able to speed up the development of advanced feedback controllers. Since, a new static mapping or dynamic model is not required if a new combustion measure (e.g., IMEP_g or CAx) needs to be added.

ACKNOWLEDGEMENTS

We would like to thank Robbert Willems for providing the experimental data used during this study.

REFERENCES

Basina, L.N.A., Irdmousa, B.K., Velni, J.M., Borhan, H., Naber, J.D., and Shahbakhti, M. (2020). Data-driven Modeling and Predictive Control of Maximum Pressure Rise Rate in RCCI Engines. In *2020 IEEE Conference on Control Technology and Applications (CCTA)*, 94–99. IEEE, Montreal, QC, Canada.

Bekdemir, C., Baert, R., Willems, F., and Somers, B. (2015). Towards Control-Oriented Modeling of Natural Gas-Diesel RCCI Combustion. In *SAE 2015 World Congress & Exhibition*, 2015–01–1745.

Guardiola, C., Pla, B., Bares, P., and Barbier, A. (2018). A combustion phasing control-oriented model applied to an RCCI engine. *IFAC-PapersOnLine*, 51(31), 119–124.

Kakoee, A., Bakhshan, Y., Barbier, A., Bares, P., and Guardiola, C. (2020). Modeling combustion timing in an RCCI engine by means of a control oriented model. *Control Engineering Practice*, 97, 104321.

Khodadadi Sadabadi, K., Shahbakhti, M., Bharath, A.N., and Reitz, R.D. (2016). Modeling of combustion phasing of a reactivity-controlled compression ignition engine for control applications. *International Journal of Engine Research*, 17(4), 421–435.

Klos, D. and Kokjohn, S.L. (2015). Investigation of the sources of combustion instability in low-temperature combustion engines using response surface models. *International Journal of Engine Research*, 16(3), 419–440.

Pan, W., Korkmaz, M., Beeckmann, J., and Pitsch, H. (2019). Unsupervised learning and nonlinear identification for in-cylinder pressure prediction of diesel combustion rate shaping process. *IFAC-PapersOnLine*, 52(29), 199–203.

Paykani, A., Garcia, A., Shahbakhti, M., Rahnama, P., and Reitz, R.D. (2021). Reactivity controlled compression ignition engine: Pathways towards commercial viability. *Applied Energy*, 282, 116174.

Rasmussen, C.E. and Williams, C.K.I. (2006). *Gaussian Processes for Machine Learning*. Adaptive Computation and Machine Learning. MIT Press, Cambridge, Mass.

Raut, A., Irdmousa, B.K., and Shahbakhti, M. (2018). Dynamic modeling and model predictive control of an RCCI engine. *Control Engineering Practice*, 81, 129–144.

Reitz, R.D. and Duraisamy, G. (2015). Review of high efficiency and clean reactivity controlled compression ignition (RCCI) combustion in internal combustion engines. *Progress in Energy and Combustion Science*, 46, 12–71.

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

Willems, R., Willems, F., Deen, N., and Somers, B. (2020). Heat release rate shaping for optimal gross indicated efficiency in a heavy-duty RCCI engine fueled with E85 and diesel. *Fuel*, 119656.

Xia, L., de Jager, B., Donkers, T., and Willems, F. (2020). Robust constrained optimization for RCCI engines using nested penalized particle swarm. *Control Engineering Practice*, 99, 104411.