

# Reinforcement learning applied to dilute combustion control for increased fuel efficiency

Bryan P Maldonado<sup>1</sup> , Brian C Kaul<sup>1</sup> , Catherine D Schuman<sup>2</sup>  
and Steven R Young<sup>3</sup>

International J of Engine Research

2024, Vol. 25(6) 1157–1173

© IMechE 2024



Article reuse guidelines:

sagepub.com/journals-permissions

DOI: 10.1177/14680874241226580

journals.sagepub.com/home/ijer



## Abstract

To reduce the modeling burden for control of spark-ignition engines, reinforcement learning (RL) has been applied to solve the dilute combustion limit problem. Q-learning was used to identify an optimal control policy to adjust the fuel injection quantity in each combustion cycle. A physics-based model was used to determine the relevant states of the system used for training the control policy in a data-efficient manner. The cost function was chosen such that high cycle-to-cycle variability (CCV) at the dilute limit was minimized while maintaining stoichiometric combustion as much as possible. Experimental results demonstrated a reduction of CCV after the training period with slightly lean combustion, contributing to a net increase in fuel conversion efficiency of 1.33%. To ensure stoichiometric combustion for three-way catalyst compatibility, a second feedback loop based on an exhaust oxygen sensor was incorporated into the fuel quantity controller using a slow proportional-integral (PI) controller. The closed-loop experiments showed that both feedback loops can cooperate effectively, maintaining stoichiometric combustion while reducing combustion CCV and increasing fuel conversion efficiency by 1.09%. Finally, a modified cost function was proposed to ensure stoichiometric combustion with a single controller. In addition, the learning period was shortened by half to evaluate the RL algorithm performance on limited training time. Experimental results showed that the modified cost function could achieve the desired CCV targets, however, the learning time was reduced by half and the fuel conversion efficiency increased only by 0.30%.

## Keywords

Internal combustion engines, reinforcement learning, fuel efficiency, real-time learning

Date received: 26 September 2023; accepted: 13 December 2023

## Introduction

Advances in computing technology, as well as cost reduction of sensors and data storage, have allowed the transportation sector to move away from simple rule-based static controllers to adaptive machine-learning-based strategies. Big data analytic tools have allowed the optimization of traffic patterns using real-time traffic feedback.<sup>1</sup> Computer vision and wireless communication have enabled a plethora of modern adaptive control strategies to optimize and guarantee a safe operation of connected and automated vehicles.<sup>2,3</sup> Hardware acceleration, sensing technology, and adaptive methods have enabled energy optimization at a variety of operating conditions for different powertrain architectures, ranging from hybrid powertrains<sup>4</sup> to advanced combustion engines.<sup>5–7</sup> This study focuses on the latter, applying data-driven adaptive optimal

control strategies to improve the efficiency of internal combustion engines.

Spark-ignition (SI) combustion engines dominate the current light-duty vehicle market, corresponding to 94% market share. In recent years, battery electric vehicles and plug-in hybrid electric vehicles have steadily

<sup>1</sup>Buildings and Transportation Science Division, Oak Ridge National Laboratory, Oak Ridge, TN, USA

<sup>2</sup>Department of Electrical Engineering and Computer Science, University of Tennessee, Knoxville, TN, USA

<sup>3</sup>Computer Science and Mathematics Division, Oak Ridge National Laboratory, Oak Ridge, TN, USA

### Corresponding author:

Bryan P Maldonado, Buildings and Transportation Science Division, Oak Ridge National Laboratory, 1 Bethel Valley Road, Oak Ridge, TN 37830, USA.

Email: maldonadopbp@ornl.gov

gained market penetration. However, the U.S. Energy Information Administration forecasts that gasoline vehicles will remain the dominant vehicle type, comprising over 70% of the market through 2050.<sup>8</sup> Medium- and heavy-duty vehicles will remain dependent on internal combustion engines for the foreseeable future due to their heavier duty cycles, with dilute SI engines offering the potential for efficiency and emissions improvement, particularly in the medium-duty sector.

Dilute combustion accomplished with exhaust gas recirculation (EGR) is a technologically proven and cost-effective way to reduce fuel consumption in a variety of engine platforms.<sup>9,10</sup> However, EGR affects combustion kinetics, reducing the combustion rate and making stable combustion more difficult to achieve. At the combustion stability limit, also called the dilute limit, the ignition becomes highly sensitive to the in-cylinder charge composition and sporadic misfires and partial burns occur, exacerbating cycle-to-cycle variability (CCV). These issues may become more pronounced for future medium-duty SI gaseous and low-carbon fuels with reduced flame speeds. Moreover, similar stability limits have been found in diesel combustion during cold starts,<sup>11,12</sup> diesel combustion in aerial powertrains,<sup>13</sup> spark-assisted compression ignition,<sup>14</sup> and gasoline compression ignition.<sup>15</sup> Thus, a robust and adaptable control strategy for minimizing combustion CCV can have wide applications within the transportation sector to improve vehicle efficiencies.

Current production engines avoid regions of unstable combustion through extensive manual calibration efforts, utilizing tables to define acceptable operating conditions, leaving efficiency opportunities on the table. Active controls to stabilize combustion at the limits would enable further efficiency gains, but robust methods for achieving this have thus far proven elusive. Previous studies have determined that CCV at the dilute limit presents deterministic patterns,<sup>16</sup> with cycle-to-cycle communication based on the composition of residual gases carried over from the previous cycle.<sup>17,18</sup> The feasibility of using the dynamics of dilute CCV for combustion stability control was demonstrated using a proportional feedback control to adjust the fuel injection quantity,<sup>19</sup> providing a proof of concept that changing the quantity of fuel injected on a cycle-resolved basis can impact combustion stability. Model-based predictive control has also been attempted,<sup>20</sup> but was limited by the difficulty of achieving a control-oriented model that is both predictive and computationally inexpensive.<sup>21</sup> Machine learning approaches, including artificial neural networks (ANNs), offer the potential to address these difficulties. Early attempts at using ANNs for combustion stability control were demonstrated for lean combustion<sup>22</sup> and for EGR dilution,<sup>23–25</sup> using an ANN-based observer for state estimation and an ANN-based controller to determine next-cycle fueling. This approach was later improved by a reinforcement learning (RL)-based controller that used an ANN-based adaptive-critic structure.<sup>26,27</sup> All

these applications, however, were trained offline using the model of Daw et al.<sup>28</sup> While they showed improvements in CCV, they also exhibited fuel enrichment, and it was not possible to determine how much of the improvement was due to next-cycle control actions. Even though the model has been improved<sup>29</sup> and recent model-based controllers have been designed,<sup>30,31</sup> offline training suffers from inaccuracies and uncertainties of the model. Additionally, the resulting control policy is restricted to the domain from where the data used to calibrate the model were collected, rendering a suboptimal solution.

The implementation of model-free control with online learning for internal combustion engines remains a challenging problem,<sup>32</sup> with the bigger advantage of providing globally optimal solutions for the problem of combustion CCV control. Modern engine control units (ECUs) are now capable of online calculation of advance control strategies for cycle-to-cycle<sup>33,34</sup> and even in-cycle<sup>35</sup> combustion control. Online learning has proven useful to avoid the dilute limit<sup>36</sup> and to exploit the stochastic properties of CCV for optimized combustion.<sup>37</sup> Even though online learning has been part of the control design, the control command itself was calculated thanks to a simplified model of the SI engine. Recently, a model-free RL controller has been designed by Henry de Frahan et al. to adaptively adjust the fuel injection timing in diesel engines.<sup>38</sup> However, the controller was trained offline on a computer model rather than using an experimental engine in real-time.

This study introduces a novel reinforcement learning strategy for optimal fuel quantity control, leveraging online model-free learning of combustion dynamics during real-time engine operation. Reinforcement learning was chosen as the preferred data-driven adaptive optimal control method based on its strong theoretical foundation and convergence guarantees.<sup>39</sup> The approach involved operating the engine under increased dilution levels beyond the dilute limit and utilizing a control-oriented model for data-efficient reinforcement learning. The proposed algorithm, tested under three scenarios, demonstrated significant improvements in fuel conversion efficiency and combustion stability, effectively extending the dilute limit. The results highlight the effectiveness of model-free learning in comparison to traditional model-based approaches, paving the way for future research to explore the method's applicability across diverse engine operating conditions.

The remaining sections are organized as follows. The experimental setup is first introduced. The combustion conditions at the dilute limit and the definition of the objectives of the controller are later discussed. The control-oriented model used for selecting the system states and the cost function is then presented. The RL algorithm used for model-free learning of the optimal control policy follows. The experimental results for three different control designs and the comparison between open-loop and closed-loop performance are later presented. Finally, the paper presents conclusions.

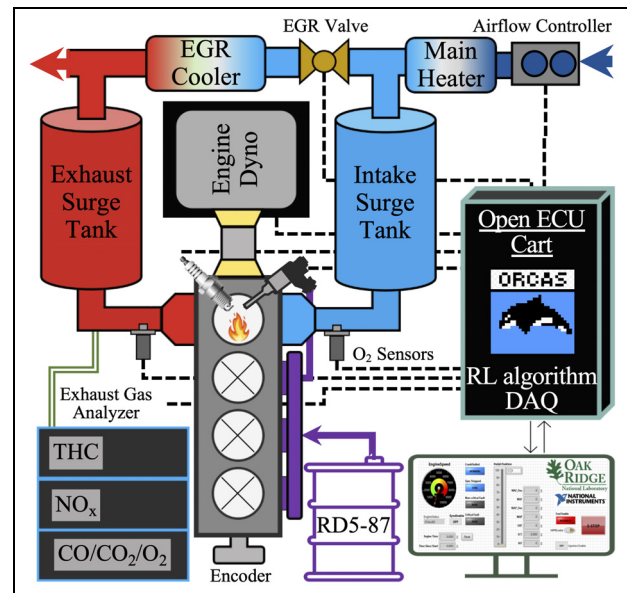
**Table 1.** Engine geometry and settings.

Displaced Volume	565.25 cm <sup>3</sup>
Bore × Stroke	87.5 mm × 94 mm
Connecting Rod	149 mm
Compression Ratio	10 : 1
Intake Valve Opening (IVO)	−359°aTDC
Intake Valve Closing (IVC)	−100°aTDC
Exhaust Valve Opening (EVO)	120°aTDC
Exhaust Valve Closing (EVC)	359°aTDC
Start of Injection	−280°a TDC
Engine Speed	2000rpm
Fresh Airflow	276.5g/s
Coolant & Oil Temperature	90°C
EGR Cooler Temperature	50°C
Intake Manifold Temperature	40°C

## Experimental setup

A single-cylinder version of a 2.3L Ford Ecoboost engine was used for experimental demonstration. The SI direct-injection experimental engine was equipped with an external, cooled, low-pressure EGR system. The airflow was kept constant using an Alicat mass flow controller. The cam timing was chosen to avoid positive or negative valve overlap. The engine speed was kept constant using an engine dynamometer. The spark advance was controlled by a slow proportional-integral (PI) controller targeting an optimal combustion phasing. A wide-band exhaust oxygen sensor was used to monitor the air-fuel equivalence ratio  $\lambda$ . The EGR fraction in the intake manifold was calculated using an intake oxygen sensor. Research-grade E10 gasoline known as RD5-87, which was designed to be representative of regular-grade market gasoline, was utilized. The start of injection was set to  $-280^\circ\text{aTDC}$  to contribute to a homogeneous air-fuel mixture and avoid possible fuel stratification that could affect the CCV. In-cylinder pressure was measured using a Kistler 6125C flush-mounted piezoelectric pressure transducer. Exhaust gas analyzers were used to estimate the average combustion efficiency. Table 1. summarizes the combustion chamber geometry and the main settings used during all experiments.

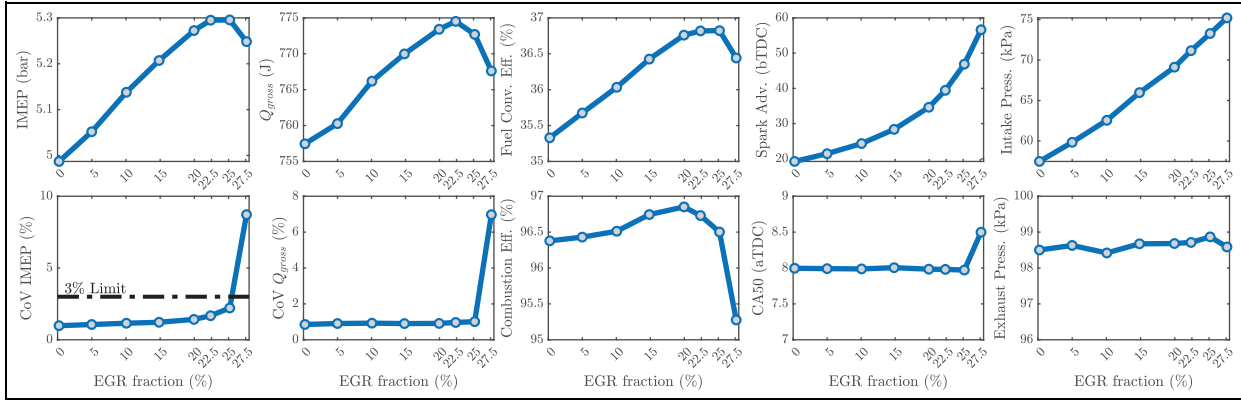
An open LabVIEW-based ECU implemented on a National Instruments Powertrain Controls (formerly Driven, Inc.) platform was used for low-level and high-level control of the engine actuators. In-cylinder pressure data processing needed for control design was done using an in-house LabVIEW-based Oak Ridge Combustion Analysis System (ORCAS) embedded in the ECU. In order to guarantee cycle-to-cycle control at an engine speed of 2000 rpm, equivalent to 60 ms per combustion cycle, the proposed learning algorithm was integrated into the ECU without communication overhead.<sup>40</sup> Finally, the control parameters of the ECU were modified through a host computer with a LabVIEW interface. Figure 1 shows the setup and instrumentation of the experimental engine test cell.

**Figure 1.** Experimental setup with rapid prototyping ECU.

## Dilute limit in spark-ignition engines

While EGR increases the overall fuel efficiency, it does so up to the stability limit. At high levels of EGR, ignition becomes highly sensitive to changes in cylinder composition. This results in a decrease of efficiency past the dilute limit due to either poor flame initiation (misfires)<sup>41,42</sup> or propagation (partial burns)<sup>43</sup> which increase the combustion CCV even if the spark advance is optimized.<sup>44</sup> Moreover, the fundamental characteristics of combustion CCV can change depending on engine speed, load, and phasing.<sup>45–47</sup> Since SI engines operate over a wide range of speeds and loads, complete characterization of the dilute limit over the entire operating regime requires a large amount of time and effort. To reduce the testing burden, the Advanced Combustion and Emission Control (ACEC) tech team of the United States Council for Automotive Research (USCAR) has chosen a relevant engine speed and load condition for different technology applications.<sup>48</sup> In particular, for a downsized boosted engine that seeks to operate at heavier loads, an engine speed of 2000 rpm and 20% load is suggested. For this engine, that load equates to 4 bar brake mean effective pressure (BMEP), or about 5 bar indicated mean effective pressure (IMEP).

Figure 2 shows different combustion metrics at 2000 rpm, 5 bar IMEP, 0% EGR fraction, and stoichiometric combustion, approximately recreating the operating condition suggested by the ACEC tech team. EGR fraction was then increased until efficiency losses due to high CCV were detected. Each marker corresponds to an average value over 2000 engine cycles. At each cycle  $k$ , the combustion features have been estimated from the in-cylinder pressure sensor as follows:



**Figure 2.** Averages over 2000 cycles of different combustion metrics as EGR increases toward the dilute limit.

- Indicated mean effective pressure:

$$\text{IMEP}_k = \frac{1}{V_d} \int_{\text{IVO}}^{\text{EVC}} P_\theta[k] \frac{dV_\theta}{d\theta} \quad (1)$$

- Gross heat release (neglecting crevices):

$$Q_{\text{gross}}[k] = \int_{\text{IVC}}^{\text{EVO}} \frac{1}{\gamma - 1} V_\theta \frac{dP_\theta[k]}{d\theta} + \frac{\gamma}{\gamma - 1} P_\theta[k] \frac{dV_\theta}{d\theta} + \frac{dQ_{\text{HT}}[k]}{d\theta} d\theta \quad (2)$$

- Combustion efficiency:

$$\eta_c[k] = \frac{Q_{\text{gross}}[k]}{M_{\text{fuel}}[k] Q_{\text{LHV}}} \quad (3)$$

- Fuel conversion efficiency:

$$\eta_i[k] = \frac{\text{IMEP}[k]}{M_{\text{fuel}}[k] Q_{\text{LHV}}} \quad (4)$$

- Combustion phasing (50% mass burned):

$$\int_{\text{IVC}}^{\text{CA50}_k} \frac{dQ_{\text{gross}}[k]}{d\theta} d\theta = 0.5 \quad (5)$$

Here,  $P_\theta$  is the in-cylinder pressure as a function of crank angle  $\theta$ ,  $V_\theta$  is the in-cylinder volume as function of crank angle,  $V_d$  is the displacement volume,  $\gamma = 1.3$  is the polytropic coefficient,  $Q_{\text{HT}}$  is the convective heat transfer to the cylinder walls,  $M_{\text{fuel}}$  is the in-cylinder fuel mass, and  $Q_{\text{LHV}} = 41.61$  MJ/kg is the lower heating value of the RD5-87 fuel. In addition, the average spark advance, intake manifold pressure, and exhaust manifold pressure were included.

As the EGR fraction increases, the combustion kinetics change and the combustion duration elongates, retarding the combustion phasing. A proportional-integral (PI) controller was used to increase the spark advance to compensate for the longer combustion duration and maintain an optimal combustion phasing of  $\text{CA50} = 8^\circ\text{aTDC}$ .<sup>49,50</sup> Since the intake airflow was

kept constant at 276.5 g/s, the increase in EGR fraction produced the increase in intake manifold pressure observed in Figure 2. The exhaust manifold pressure, however, remained fairly constant at different EGR levels, indicating that the observed increase in IMEP and fuel conversion efficiency are due to improved combustion rather than improved gas exchange. On the other hand, the combustion efficiency showed a minor increase with EGR up to the dilute limit. Combustion CCV was characterized by the coefficient of variation (CoV) of IMEP and gross heat release. The increase in IMEP and gross heat release are followed by an increase in CoV due to the increased sensitivity to in-cylinder composition. The industry CoV of IMEP limit is 3%, based on the levels of noise, vibration, and harshness experienced by the driver. For this operating condition, the dilute limit was reached at approximately 25% EGR fraction at optimal spark advance. Furthermore, the experimental data in Figure 2 showed that peak IMEP and peak fuel conversion efficiency also occur at the dilute limit of 25% EGR fraction.

When the dilution levels were increased past the limit, to 27.5% EGR fraction, a drastic increase in CCV driven by sporadic partial burns and misfires rapidly increased the CoV of IMEP and the CoV of  $Q_{\text{gross}}$ . The low-energy combustion events reduced the average values of the IMEP, gross heat release, fuel conversion efficiency, and combustion efficiency. Even though the PI controller for the spark advance was tasked to maintain an average of  $\text{CA50} = 8^\circ\text{aTDC}$ , partial and slow burns contributed to a slightly retarded combustion phasing. The same high sensitivity of ignition with respect to the in-cylinder composition that generates sporadic misfires and partial burns, however, can be used to reduce the CCV and recover the efficiency gains at high EGR. Previous studies have shown experimentally that small perturbations in fuel quantity, ranging from 0.5% to 2% additional fuel, can effectively reduce the combustion CCV.<sup>51,52</sup> Therefore, we propose an injection quantity controller to alter the in-cylinder mass composition and stabilize the charge to avoid partial burns and misfires.

## Model-based cost function

For the cost function in the RL algorithm, a control-oriented model was considered. The physics-based approach for lean combustion modeling presented by Daw et al.<sup>28</sup> was used. The coupling between combustion cycles is through the carryover of residual gas from the previous cycle to the next and can be described by the following discrete-time system:

$$\begin{bmatrix} M_{\text{fuel}} \\ M_{\text{gas}} \end{bmatrix}_{k+1} = X_{\text{res}}[k] \begin{bmatrix} 1 - \eta_c[k] & 0 \\ \eta_c[k] & 1 \end{bmatrix} \begin{bmatrix} M_{\text{fuel}} \\ M_{\text{gas}} \end{bmatrix}_k + \begin{bmatrix} m_{\text{fuel, in}}[k] \\ \frac{m_{\text{air, in}}}{1 - X_{\text{EGR}}} \end{bmatrix} \quad (6)$$

$$Q_{\text{gross}} = \eta_c[k] M_{\text{fuel}}[k] Q_{\text{LHV}} \quad (7)$$

where  $M_{\text{gas}}$  is the in-cylinder gas mass (air + burned gas),  $X_{\text{res}}$  is the residual gas fraction,  $X_{\text{EGR}}$  is the EGR fraction,  $m_{\text{air, in}}$  is the mass of fresh air per combustion cycle, and  $m_{\text{fuel, in}}$  is the fuel injection quantity to be controlled. The first term of equation (6) corresponds to the residual mass in the cylinder, while the second term corresponds to the fresh mass available for the next combustion cycle. The residual gas fraction was calculated assuming an isentropic exhaust process and zero valve overlap.<sup>53</sup>

$$X_{\text{res}}[k] = \frac{V_{\text{EVC}}}{V_{\text{EVO}}} \left( \frac{P_{\text{EVC}}}{P_{\text{EVO}}[k]} \right)^{1/\gamma} \quad (8)$$

The subscripts indicate that only the values EVO and EVC are needed. However, in order to perform online calculations of the in-cylinder mass, it was assumed that  $P_{\text{EVC}} = 1 \text{ bar}$ . This allowed the ORCAS to start the calculations at EVO and enabled the host computer to send the control command to the ECU before EVC. A previous study by the authors showed that this system is stable and fully controllable.<sup>37</sup> During steady-state and stoichiometric combustion, the fuel injection quantity is determined by the amount of fresh air and the stoichiometric air-fuel ratio  $\text{AFR}_s = 14.2$  as follows:

$$m_{\text{fuel, in}}^s = \frac{m_{\text{air, in}}}{\text{AFR}_s} \quad (9)$$

To avoid engineering units and maintain the variables at similar scales, consider the following normalization:

$$F_k = M_{\text{fuel}}[k] / m_{\text{fuel, in}}^s \quad (10)$$

$$G_k = M_{\text{gas}}[k] / m_{\text{air, in}}^s \quad (11)$$

$$u_k = m_{\text{fuel, in}}[k] / m_{\text{fuel, in}}^s \quad (12)$$

$$H_k = Q_{\text{gross}}[k] / (m_{\text{fuel, in}}^s Q_{\text{LHV}}) \quad (13)$$

Finally, one can show that the normalized system obeys the following set of equations:

$$x_{k+1} = X_{\text{res}}[k] \begin{bmatrix} 1 - \eta_c[k] & 0 \\ \eta_c[k] & 1 \end{bmatrix} x_k + \begin{bmatrix} u_k \\ \frac{1}{1 - X_{\text{EGR}}} \end{bmatrix} \quad (14)$$

$$H_k = [\eta_c[k] \quad 0] x_k \quad (15)$$

where  $x_k = [F_k \quad G_k]^T$  denotes the normalized state. Figure 3 illustrates the block diagram for online calculation of cycle-to-cycle variables. The figure emphasizes that the model-based combustion cycle (gray shaded area) is not synchronized with the ECU cycle. For this reason, the control command  $u_k$  needs to be calculated before the end of the ECU cycle  $k$ .

Equation (14) can be written as  $x_{k+1} = f(x_k, u_k, w_k)$ , where  $w_k$  is a stochastic variable representing the combustion CCV. Note that the stochastic variable  $w_k$  is not explicitly written in equation (14). Rather, this variable causes the CCV observed in the experimentally calculated values of  $\eta_c[k]$  and  $X_{\text{res}}[k]$ . Also note that  $u_k = 1$  corresponds to stoichiometric combustion, while  $u_k > 1$  causes a rich mixture and  $u_k < 1$  causes a lean mixture. At the operating condition selected for this study, the steady-state values of the residual gas fraction and combustion efficiency during nominal combustion are  $X_{\text{res}}^{\text{ss}} \approx 5\%$  and  $\eta_c^{\text{ss}} \approx 96\%$ . Therefore, the steady-state values at stoichiometric conditions of the normalized states can be approximated by solving the following algebraic equation:

$$\begin{bmatrix} F^{\text{ss}} \\ G^{\text{ss}} \end{bmatrix} \approx \begin{bmatrix} 0 & 0 \\ 0 & X_{\text{res}}^{\text{ss}} \end{bmatrix} \begin{bmatrix} F^{\text{ss}} \\ G^{\text{ss}} \end{bmatrix} + \begin{bmatrix} 1 \\ 1/(1 - X_{\text{EGR}}) \end{bmatrix} \quad (16)$$

Then,  $F^{\text{ss}} \approx 1$  and  $G^{\text{ss}} \approx 1.45$  at 27.5% EGR fraction. The normalized heat release then becomes  $H^{\text{ss}} \approx \eta_c^{\text{ss}}$ .

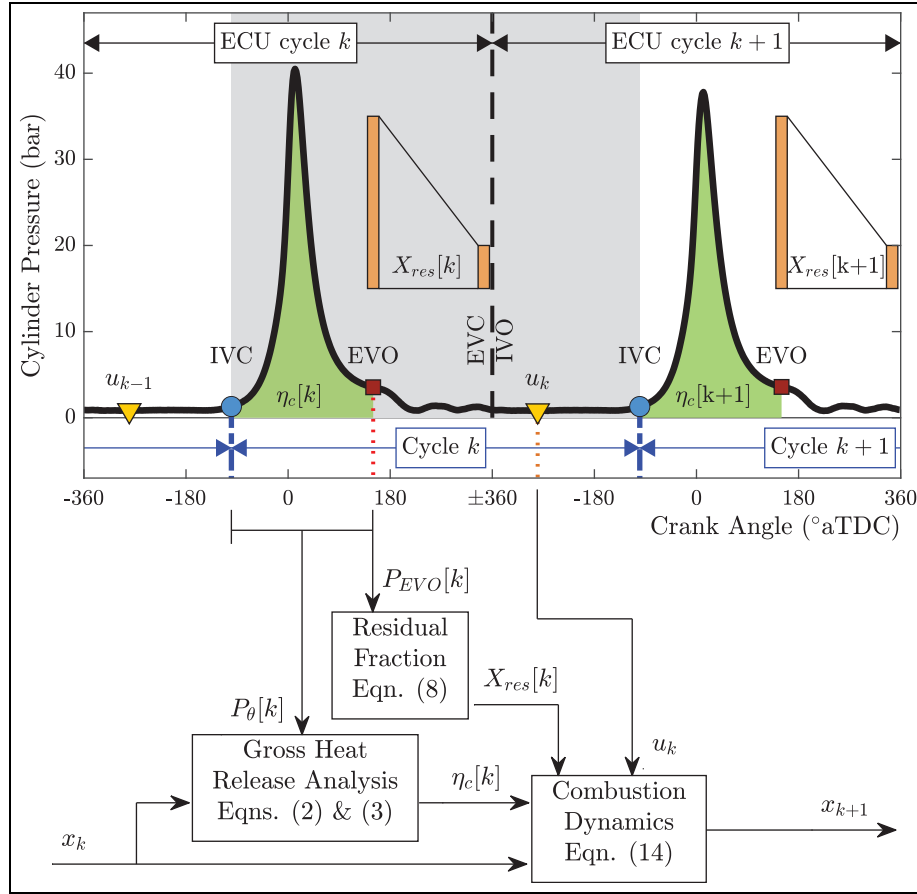
The steady-state analysis only applies to nominal combustion cycles. During misfires, the combustion efficiency drops to zero (no combustion) increasing the combustion CCV. Therefore,  $H_k$  can be used as a proxy variable to quantify and control CCV. Consider the following model-based cost function to be minimized by the RL controller at each combustion cycle:

$$g(x_k, u_k, w_k) = (u_k - 1)^2 + \rho (H_k - \eta_c^{\text{ss}})^2 \quad (17)$$

The first term penalizes the deviations from stoichiometric combustion while the second term penalizes the deviations from nominal combustion without misfires or partial burns. The parameter  $\rho$  was introduced to balance the trade-off between the two terms of the cost function. A previous study by the authors determined that a value of  $\rho = 0.2$  can strike a good balance between the two components.<sup>37</sup> Such cost was calculated by ORCAS and readily available for the learning algorithm at each combustion cycle. The state  $x_k$ , however, is not measurable by engine sensors. Therefore, cycle-to-cycle estimates were generated by the combustion model.

## Reinforcement learning algorithm

Consider the control problem of finding an optimal stationary policy  $u_k = \pi(x_k)$  such that the following infinite horizon total cost is minimized:



**Figure 3.** Calculations needed for cycle-to-cycle analysis.

$$J(x_0) = \mathbb{E}_w \left[ \sum_{k=0}^{\infty} \alpha^k g(x_k, \pi(x_k), w_k) \right] \quad (18)$$

where  $\alpha \in (0, 1)$  is a discounting factor included to guarantee the existence of a solution.<sup>39</sup> The exact solution to the problem is given by Bellman's equation:

$$J(x) = \min_u \mathbb{E}_w [g(x, u, w) + \alpha J(f(x, u, w))]. \quad (19)$$

This is a challenging functional equation that could not be solved analytically for the problem at hand. RL was used to approximate the solution of the infinite horizon discounted cost problem. Q-learning was the method of choice for the following reasons: (1) simplicity in implementation, given that we wrote the algorithm within the native LabVIEW programming environment, (2) convergence, given that the update rule ensures that the Q-values eventually converge to the optimal value, and (3) off-Policy Learning, given that it can learn from data generated by a different policy. The third aspect is particularly important since one proposed method merges RL and  $\lambda$  feedback to guarantee stoichiometric combustion. However, implementing Q-learning requires the discretization of the state and control spaces to a finite-state Markov decision process (MDP). Although there exist RL algorithms that can deal with continuous state and control spaces, such as the actor-critic

methods that have been successfully applied to a wide range of problems, they have significantly higher complexity and their sensitivity to hyperparameters may lead to slow convergence, instability, or poor generalization.<sup>54</sup>

Let  $x_k \in \mathcal{X} \subset \mathbb{R}^{20 \times 20}$  be the discretized state space. For each state, the intervals chosen were  $0.96 \leq F_k \leq 1.18$  and  $1.43 \leq G_k \leq 1.61$ , each of which was divided into 20 equidistant points. Note that the steady-state values  $F^{ss} \approx 1$  and  $G^{ss} \approx 1.45$  previously calculated do not lie at the center of the intervals. This is necessary because  $F_k$  and  $G_k$  increase significantly after a misfire occurs due to the unburned air-fuel residual mixture. Let  $u_k \in \mathcal{U} \subset \mathbb{R}^{15}$  be the discretized control space. The interval chosen for the control space was  $0.975 \leq u_k \leq 1.025$ , centered at the stoichiometric value and allowing a decrease/increase of maximum 2.5% fuel. Let  $p_{k+1|k}(u_k) = \Pr(x_{k+1}|x_k, u_k)$  be the transition probability of the MDP, then the expected value in equation (19) can be simplified to the Q-factor:

$$Q(x_k, u_k) = \sum_{x_{k+1} \in \mathcal{X}} p_{k+1|k}(u_k) (g(x_k, u_k) + \alpha J(x_{k+1})) \quad (20)$$

Since the optimal cost satisfies  $J(x) = \min_u Q(x, u)$ , then the optimal Q-factor solves the equation<sup>39</sup>:



**Algorithm 1** Q-learning for fuel quantity control

---

```

 $n = 0, k = 0, Q_0 \sim \mathbb{U}(1, 2)$ 
while  $n \leq N$  do
  Intake phase (IVO to IVC):
  if  $k > 0$  then
    apply  $u_{k-1}$  command
  else
    apply open-loop  $u_{k-1} = 1$  command
  end if
  Combustion phase (IVC to EVO):
  calculate  $H_k$  from in-cylinder pressure
  generate random variable  $p \sim \mathbb{U}(0, 1)$ 
  if  $p < \varepsilon_n$  then
    pick randomly  $u_k \sim \mathbb{U}\{\mathcal{U}\}$ 
  else
    optimize  $u_k = \arg \min_{u \in \mathcal{U}} Q_k(x_k, u)$ 
  end if
  Exhaust phase (EVO to EVC):
  calculate cycle cost  $g(x_k, u_k)$ 
  calculate state  $x_{k+1}$  using  $u_k$  and equation (14)
  update  $Q_{k+1}(x_k, u_k)$  using equation (22)
  move to next combustion cycle:  $k = k + 1$ 
  if  $k \equiv 0 \pmod{2000}$  then
    move to next episode:  $n = n + 1$ 
  end if
end while

```

---

$$Q(x_k, u_k) = \sum_{x_{k+1} \in \mathcal{X}} p_{k+1|k}(u_k)(g(x_k, u_k) + \alpha \min_{u \in \mathcal{U}} Q(x_{k+1}, u)). \quad (21)$$

Let  $Q \in \mathbb{R}^{20 \times 20 \times 15}$  be the Q-table with Q-factors  $Q(x, u)$ . If the transition probability is known, then equation (21) can be solved for  $Q$  using algebraic methods. Even though previous studies have focused on modeling and simulating the system described by equation (14), the transition probability is not explicitly known but rather the result of simulations. More importantly, the simple physics-based model does not capture all the complexity of the combustion process typically recreated by computational fluid dynamics simulations.<sup>55</sup>

The Q-learning algorithm introduced by Watkins<sup>56</sup> was used to adaptively learn the Q-table without knowledge of the transition probability (model-free learning). The Q-factor update law used a constant learning rate  $\kappa$  as follows:

$$Q_{k+1}(x_k, u_k) = (1 - \kappa)Q_k(x_k, u_k) + \kappa(g(x_k, u_k) + \alpha \min_{u \in \mathcal{U}} Q_k(x_{k+1}, u)) \quad (22)$$

A  $\varepsilon$ -greedy approach was used to balance the exploration-exploitation trade-off and avoid sub-optimal solutions. Let  $n = \{1, 2, \dots, N\}$  be the episode number where  $N$  is the total number of episodes used in the Q-learning algorithm. Due to the low probability of misfire events at the dilute limit, each episode consisted of 2000 engine cycles in order to guarantee that the total cost includes misfire events at each episode. Let  $Q_0$  be the initial condition for the Q-table where each Q-factor is

chosen randomly from the interval  $(1, 2)$  according to a uniform distribution, in other words:

$$Q_0 = [Q(x, u) \sim \mathbb{U}(1, 2)], \forall x \in \mathcal{X}, u \in \mathcal{U}. \quad (23)$$

Such an interval was chosen to guarantee a large enough initial condition. For each episode, let  $\varepsilon_n = 1 - n/N$  be the probability of choosing a random action over the optimal. Finally, algorithm 1 shows the Q-learning approach used for fuel quantity control in this study.

## Experimental results

The Q-learning hyperparameters  $N$ ,  $\alpha$ , and  $\kappa$  were calibrated using the computer model previously described. Simulations were done in Python, and the code can be found in Code Ocean <https://codeocean.com/capsule/8130780/tree/v1>. The parameter  $N$  was constrained by the length of the experiment, while the parameters  $\alpha = 0.1$  and  $\kappa = 0.3$  showed acceptable convergence behavior in the simulations. Note that simulations were used only to determine the best set of hyperparameters. During experimental testing, the RL algorithm started from the random initial condition  $Q_0$ , as defined by equation (23), and algorithm 1 was fully implemented and executed in the ECU.

Two training scenarios were considered, pictured in Figure 4. In the first, the cylinder-pressure-based controller was given full authority over the control action, resulting in  $m_{\text{fuel}, \text{in}}[k] = m_{\text{fuel}, \text{in}}^s u_k$ . In the second scenario, the RL controller coordinated the fuel command with a PI controller based the exhaust oxygen sensor. A slow PI controller was tuned to maintain, on average, stoichiometric combustion based on  $\lambda$  feedback. Then, the amount of fuel injected per cycle was calculated as  $m_{\text{fuel}, \text{in}}[k] = m_{\text{fuel}, \text{in}}^s u_k \lambda_{\text{FB}}[k]$ , where  $\lambda_{\text{FB}}[k]$  is the output of the PI controller driven by the  $\lambda_k$  error.

### RL controller without $\lambda$ feedback

The first model-free online learning experiment was conducted using the cost function in equation (17) without  $\lambda$  feedback over  $N = 195$  episodes. Given that each episode (2000 engine cycles) took a total of 2 min to finish, the total training time in the experimental engine was 6h30min. Figure 5 shows the total cost per episode and the average values of several variables of interest during training. The total cost shows a steady decrease during training. The sudden drop in total cost observed before episode 50 was due to an adjustment of the EGR valve. Note that  $\mathbb{E}[\lambda] > 1$  in the early stages of learning. This generates an overly lean mixture which exacerbates variability. This can also be observed in the initial increase of the average EGR fraction. After the EGR valve was adjusted, the average EGR fraction remained fairly constant during the rest of the experiment. The relatively constant amount of burned gas in the cylinder can also be observed by the

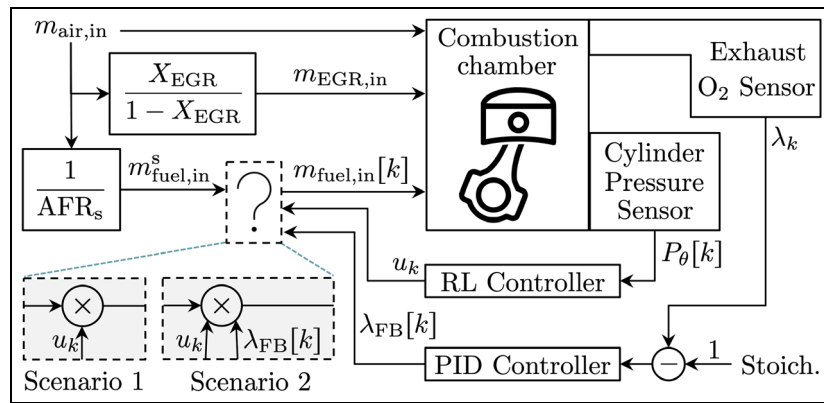


Figure 4. Feedback controllers considered during experiments.

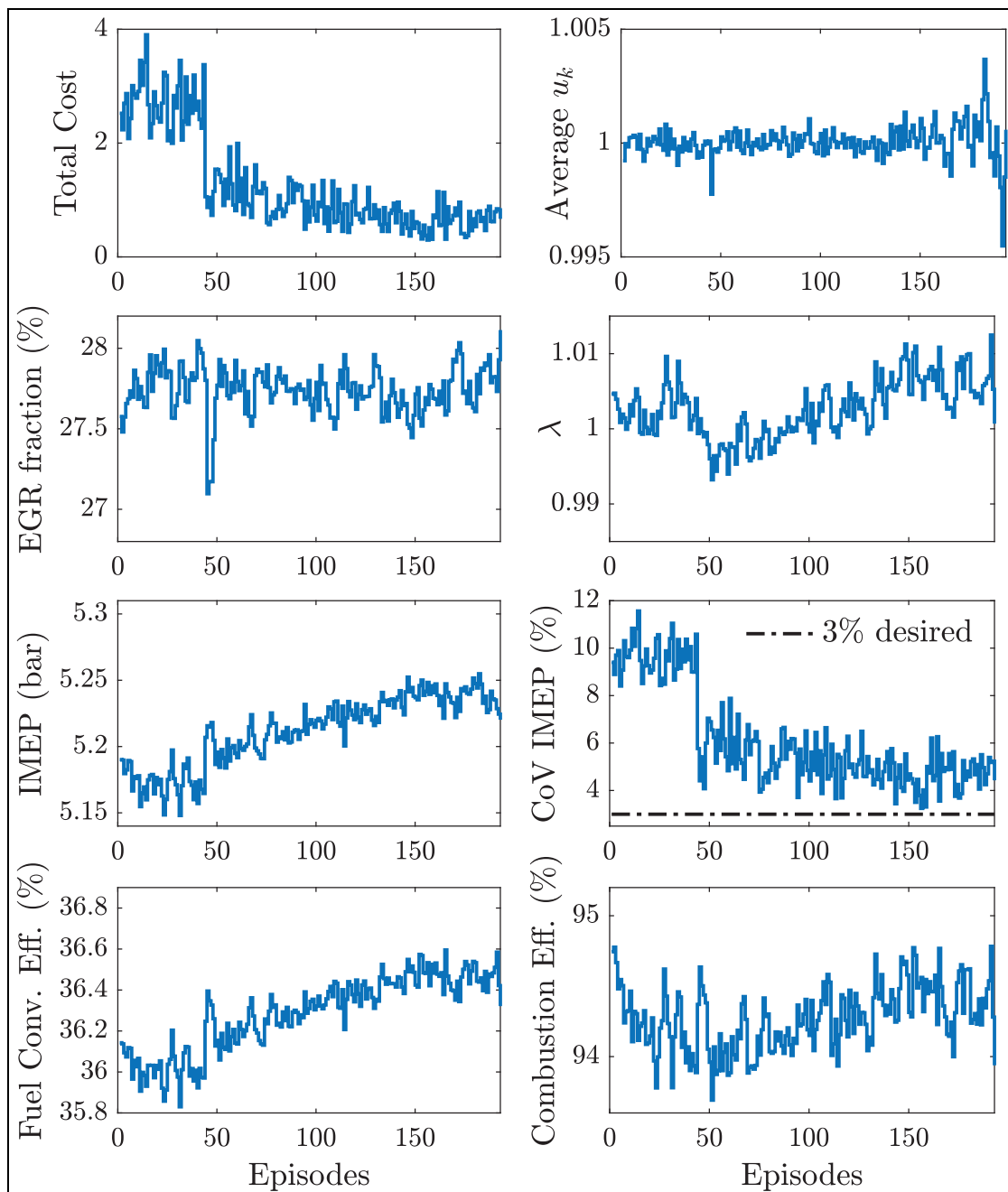
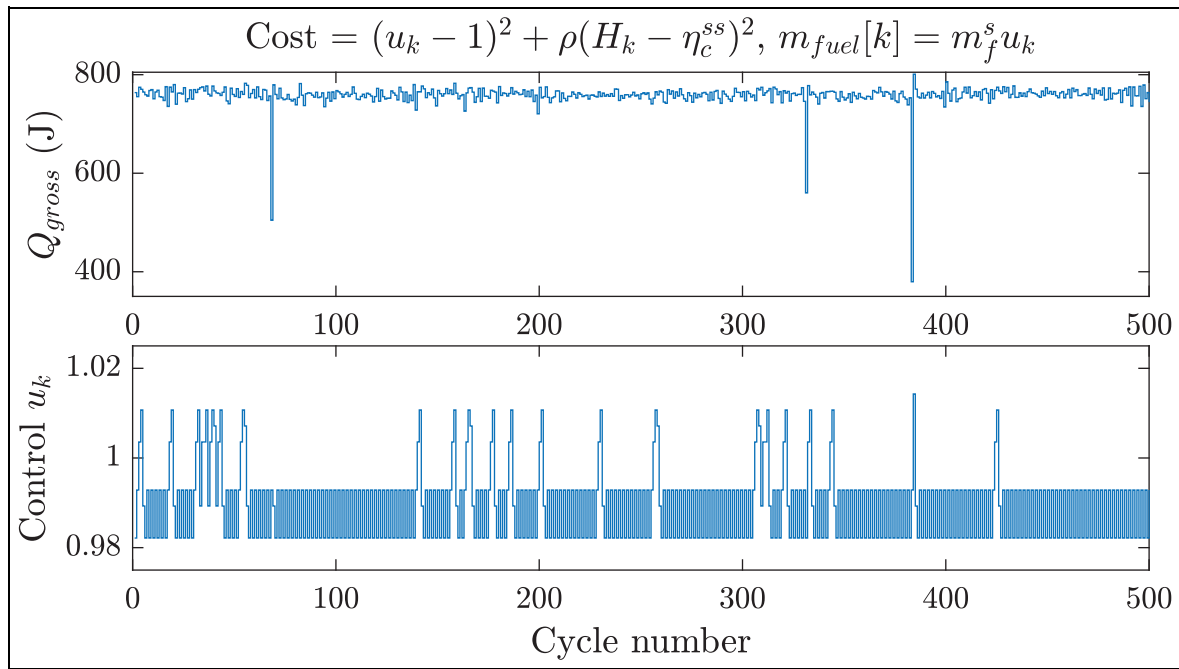


Figure 5. Training of the RL controller without  $\lambda$  feedback.





**Figure 6.** Closed-loop response of RL without  $\lambda$  feedback.

**Table 2.** Performance of controller trained without  $\lambda$  feedback.

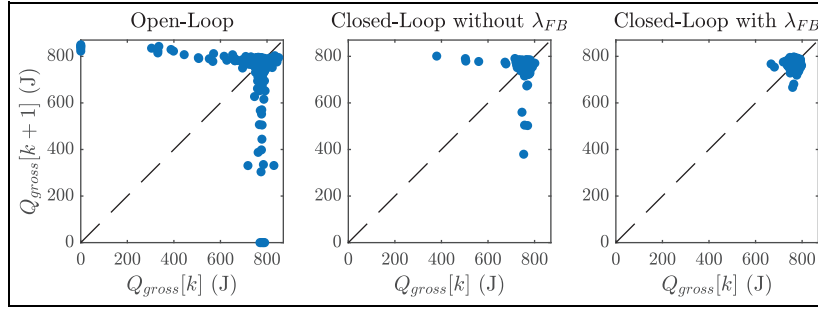
	OL	CL	Change (%)	CL + $\lambda_{FB}$	Change (%)
Total Cost	1.67	0.7	−58.0	0.51	−69.2
$\mathbb{E}[u]$	1	0.99	−0.86	0.99	−0.52
$\mathbb{E}[\lambda]$	1	1.01	0.93	1.01	0.33
$\mathbb{E}[\text{IMEP}]$	5.18	5.2	0.35	5.26	1.60
CoV IMEP	8.55	4.23	−50.5	3.01	−64.8
$\mathbb{E}[\eta_i]$ (%)	36.05	36.49	1.22	36.53	1.33
$\mathbb{E}[\eta_c]$ (%)	94.54	94.55	0.01	94.41	−0.13

stationary combustion efficiency. If the EGR levels are too high, generating an excessive number of misfires and partial burns, the combustion efficiency is expected to drop significantly. As the learning phase evolves and the  $\epsilon$ -greedy strategy favors exploitation over exploration, the average IMEP increase, increasing the fuel conversion efficiency. Toward the end of the Q-learning, where most of the optimization was done, the average command  $u_k$  slightly incurs into rich conditions ( $\mathbb{E}[u] > 1$ ) but ultimately settles for a lean condition ( $\mathbb{E}[u] < 1$ ). The advantages of lean combustion for increasing efficiency have been documented,<sup>57</sup> and the RL strategy was able to learn it. Therefore, the controller favors lean combustion with an increased fuel conversion efficiency over a stoichiometric condition. Nonetheless, the CoV of IMEP was significantly reduced compared to the open-loop conditions.

Figure 6 shows the closed-loop response of the controller after the learning phase. Several partial burns were identified, but no misfire occurred during closed-

loop. The control  $u_k$  remained most of the time oscillating between lean values. Occasionally, the controller issued rich commands ( $u_k > 1$ ) to the injector. As previously mentioned, rich conditions help stabilize the combustion process to avoid misfires. Therefore, the controller operated at lean conditions to achieve high  $Q_k$ , generating high fuel conversion efficiency, but occasionally issued rich commands to avoid misfires.

The average lean condition that the controller generated for combustion, although effective for increasing efficiency, is not compatible with a current three-way catalyst (TWC) that requires stoichiometric combustion to work effectively. To maintain TWC compatibility, consider using the exhaust oxygen sensor feedback controller  $\lambda_{FB}[k]$  after the training phase. In such a case,  $\lambda_{FB}[k]$  will be commanding a fuel increase when  $\mathbb{E}(u) < 1$  and vice-versa. Table 2 shows the comparison between the open-loop (OL) response before learning and closed-loop (CL) response after learning. The second column shows the closed-loop average values without  $\lambda$  feedback, while the fourth column (CL +  $\lambda_{FB}$ ) shows the average response when the RL controller and  $\lambda$  feedback were both enabled. Table 2 also shows the average change of the combustion variables when operated in closed-loop with and without  $\lambda$  feedback. Although the original RL controller showed a 58% reduction in the total cost due to an increase of 1.22% in fuel conversion efficiency and a 50% reduction in CoV of IMEP, the CL +  $\lambda_{FB}$  strategy showed a superior performance. This was due to the near-stoichiometric operation achieved with the additional  $\lambda$  feedback, which is less prone to partial burns than lean combustion. The reduction of partial burns further reduced the CoV of IMEP to the desired 3% target



**Figure 7.** Return maps of gross heat release at OL (left), CL without  $\lambda$  feedback (center), and CL with  $\lambda$  feedback (right).

and increased the average fuel conversion efficiency by 1.33% compared to open-loop.

The previous study by the authors solved the one-step-ahead control problem with the same cost function proposed here.<sup>37</sup> In such a study, online learning was done to model the stochastic properties of CCV but the solution was calculated using the model described by equation (14). An increase of 0.5% in fuel conversion efficiency was reported under rich conditions. The current result of a 1.33% increase in fuel conversion efficiency at near-stoichiometric conditions highlights the benefits of the model-free RL strategy. A graphical comparison of the improvements made by the RL controller is depicted in Figure 7, which shows the return maps of  $Q_{gross}$  for the three different tests summarized in Table 2. The return map is centered at the nominal value with “arms” extending toward  $Q_{gross} = 0$  representing misfires. The density of the “arms” in the return map indicates the number of partial burns and misfires. Finally, the asymmetric shape is due to the combustion cycle-to-cycle dynamics. The return map during closed-loop without  $\lambda$  feedback shows no misfires and a more symmetric shape. However, the last return map where the  $\lambda$  feedback was enabled shows a truly symmetric and compact distribution which correlates with the reduction of combustion CCV.

### RL controller with $\lambda$ feedback

Given the benefits of using the  $\lambda$  feedback controller in addition to the RL controller, online learning was conducted with the feedback command  $\lambda_{FB}[k]$  enabled at all times. In addition, a third term in the cost function defined in equation (17) was introduced. Previous research trained spiking neural networks to control fuel quantity commands, extending the dilute limit by penalizing heavily the occurrence of misfires and partial burns.<sup>58,59</sup> Therefore, consider the following cost function:

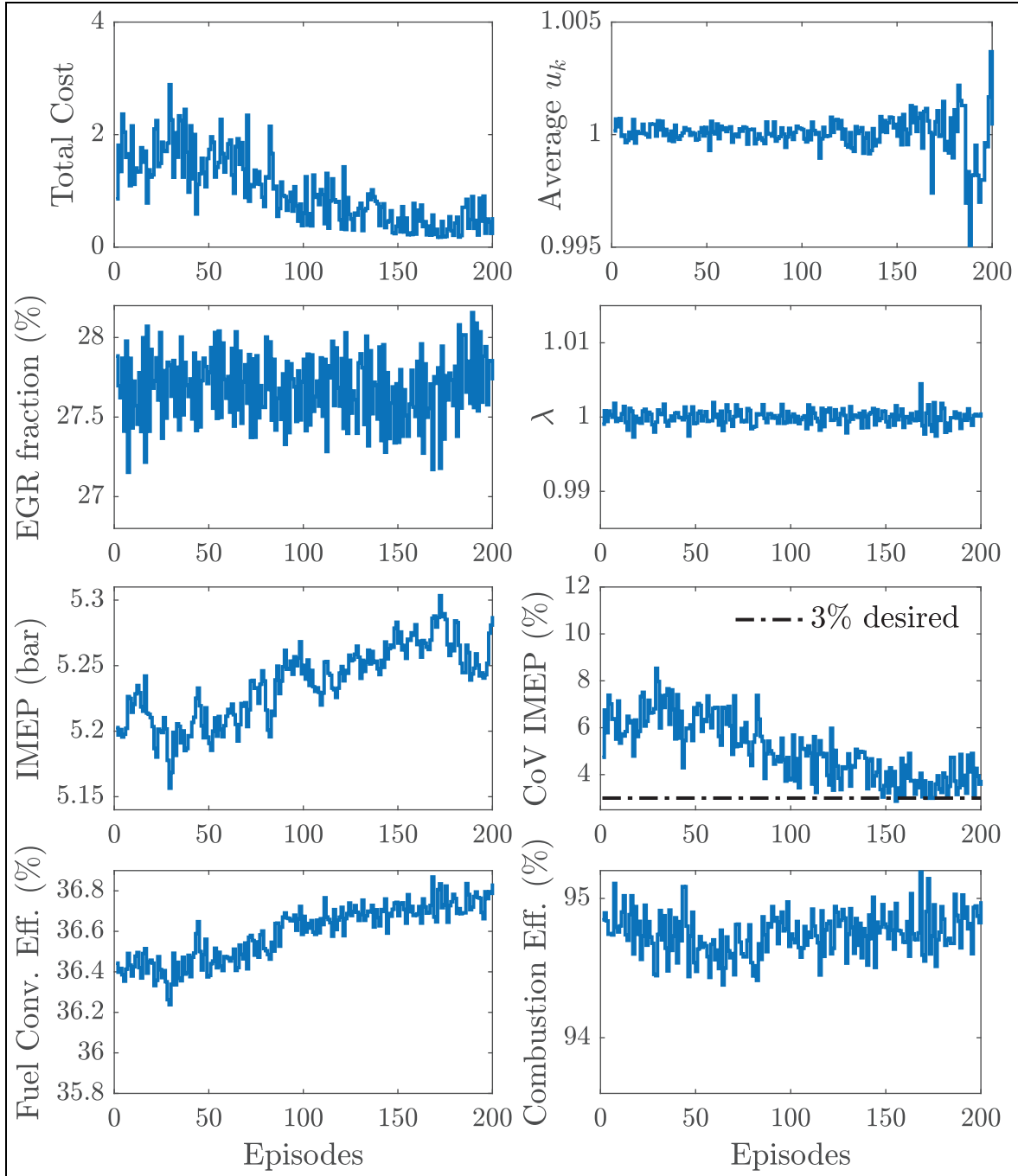
$$g(x_k, u_k, w_k) = (u_k - 1)^2 + \rho(H_k - \eta_c^{ss})^2 + \alpha \mathbb{I}(H_k < 0.6) \quad (24)$$

where  $\mathbb{I}(\cdot)$  is the indicator function and  $\alpha = 0.05$  penalizes the occurrence of misfires and partial burns. The Q-learning strategy described by algorithm 1 was used

to find the optimal  $u_k$ . However, the actual fuel command at any point in the experimental engine was calculated as  $m_{fuel,in}[k] = m_{fuel,in}^s u_k \lambda_{FB}[k]$ . In this case, a total of  $N = 200$  episodes were used during the learning phase, taking a total of 6h40min. Figure 8 shows the average training results per episode. Note how  $\mathbb{E}[\lambda] = 1$  during the learning phase, implying that  $\mathbb{E}[u_k \lambda_{FB}[k]] = 1$ . Similarly, the EGR fraction and the combustion efficiency remained fairly constant, signaling the stationarity of the exhaust residual gases. The reduction of the total cost is driven by the increase in fuel conversion efficiency and the reduction of the CoV of IMEP. Toward the end of the learning phase, we observe that the RL controller alternates between rich and lean, ultimately settling for an average rich fuel command. However, this was counteracted by the  $\lambda$  feedback, maintaining an stoichiometric condition.

Figure 9 shows the closed-loop response of the RL controller within the proposed coordinated control strategy after the learning phase. Once more, no misfires were observed during the experiment. This highlights the importance of the additional penalty term in the cost function. Therefore, the CoV of IMEP was reduced fairly close to the desired 3% limit. For this control strategy, the injection quantity command does not follow a simple pattern. For some cycles, the controller switches between stoichiometric and lean conditions. After some time, the average command switches to the rich region ( $u_k > 1$ ). On average, however, the combined controller maintains the stoichiometric command.

Table 3 shows the comparison between the open- and closed-loop systems after being trained with  $\lambda$  feedback enabled. Here, a 60% reduction in the total cost was achieved due to a 51% reduction of the CoV of IMEP and a 1.09% increase in fuel conversion efficiency. Note how the desired  $\mathbb{E}[\lambda] = \mathbb{E}[u] = 1$  was achieved with a small difference between open- and closed-loop. The increase in fuel conversion efficiency was not as high as the previous controller because the previous strategy produced slightly leaner combustion, contributing to higher efficiency. The complexity of engine experiments and the stochasticity of the system makes it difficult to guarantee the same open-loop condition for both experiments. Thus, the use of



**Figure 8.** Training of the RL controller with  $\lambda$  feedback.

percentage change relative to the corresponding open-loop was deemed as an acceptable metric to compare the different control strategies presented in this study.

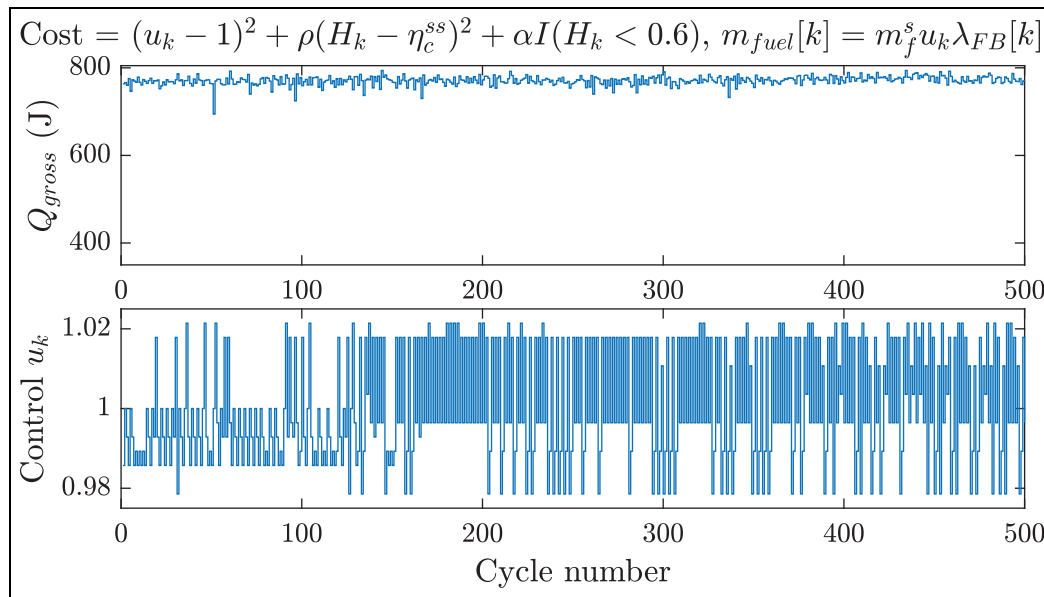
#### *RL controller with modified cost function*

Recall that the cost functions defined by equations (17) and (24) are based on the physical understanding of the cycle-to-cycle dynamics. Although they were designed to achieve the goals of stoichiometric combustion, higher fuel efficiency, and reduction of combustion CCV, they rely on the explicit variables  $u_k$  and  $H_k$  from

the control-oriented model. The model-free approach of RL, however, allows for the inclusion of other, more suitable variables, that implicitly depend on  $u_k$  and  $H_k$ , but for which the explicit relations are not necessarily known. Consider the following modified cost function to achieve the control objectives previously mentioned:

$$g(x_k, u_k) = (\lambda_k - 1)^2 + \rho^*(\eta_i[k] - \eta_i^*)^2 + \alpha \mathbb{I}(H_k < 0.6) \quad (25)$$

The first term uses the exhaust oxygen sensor to penalize deviations from stoichiometric conditions and



**Figure 9.** Closed-loop response of RL and  $\lambda$  feedback.

**Table 3.** Performance of controller trained with  $\lambda$  feedback.

	OL	CL + $\lambda_{FB}$	Change (%)
Total Cost (-)	1.27	0.51	-60.0
$\mathbb{E}[u]$ (-)	1	1	0.25
$\mathbb{E}[\lambda]$ (-)	1	1	-0.21
$\mathbb{E}[\text{IMEP}]$ (bar)	5.2	5.29	1.68
CoV IMEP (%)	6.47	3.17	-50.9
$\mathbb{E}[\eta_i]$ (%)	36.35	36.75	1.09
$\mathbb{E}[\eta_c]$ (%)	94.67	94.73	0.06

serves as the replacement of the PI controller previously used. The second term attempts to directly increase the fuel conversion efficiency to a maximum value of  $\eta_i^* = 0.37$ , where the weighting factor used was  $\rho^* = 0.05$ . Finally, the third term is similar to the previous cost function, where  $\alpha = 0.05$  and the indicator function  $\mathbb{I}(H_k < 0.6)$  heavily penalizes misfires and partial burns. Given that the  $\lambda$  feedback is now embedded in the cost function, the fuel quantity issued was calculated only using the RL controller as  $m_{\text{fuel, in}}[k] = m_{\text{fuel, in}}^s u_k$ . For this controller, the Q-learning algorithm was trained using  $N = 100$  episodes during a period of 3h20min in the experimental engine. The goal of the limited number of episodes was to test the effectiveness of the proposed engine control strategy in a shorter training period.

Figure 10 shows the experimental results over the training period using the modified cost function. The results display a reduction of the total cost due to (1) average stoichiometric conditions, (2) a moderate increase in fuel conversion efficiency, and (3) elimination of misfires and partial burns. Note that, toward the end of the learning phase, the controller consistently maintained rich conditions ( $\mathbb{E}[u] > 1$ ). However, the

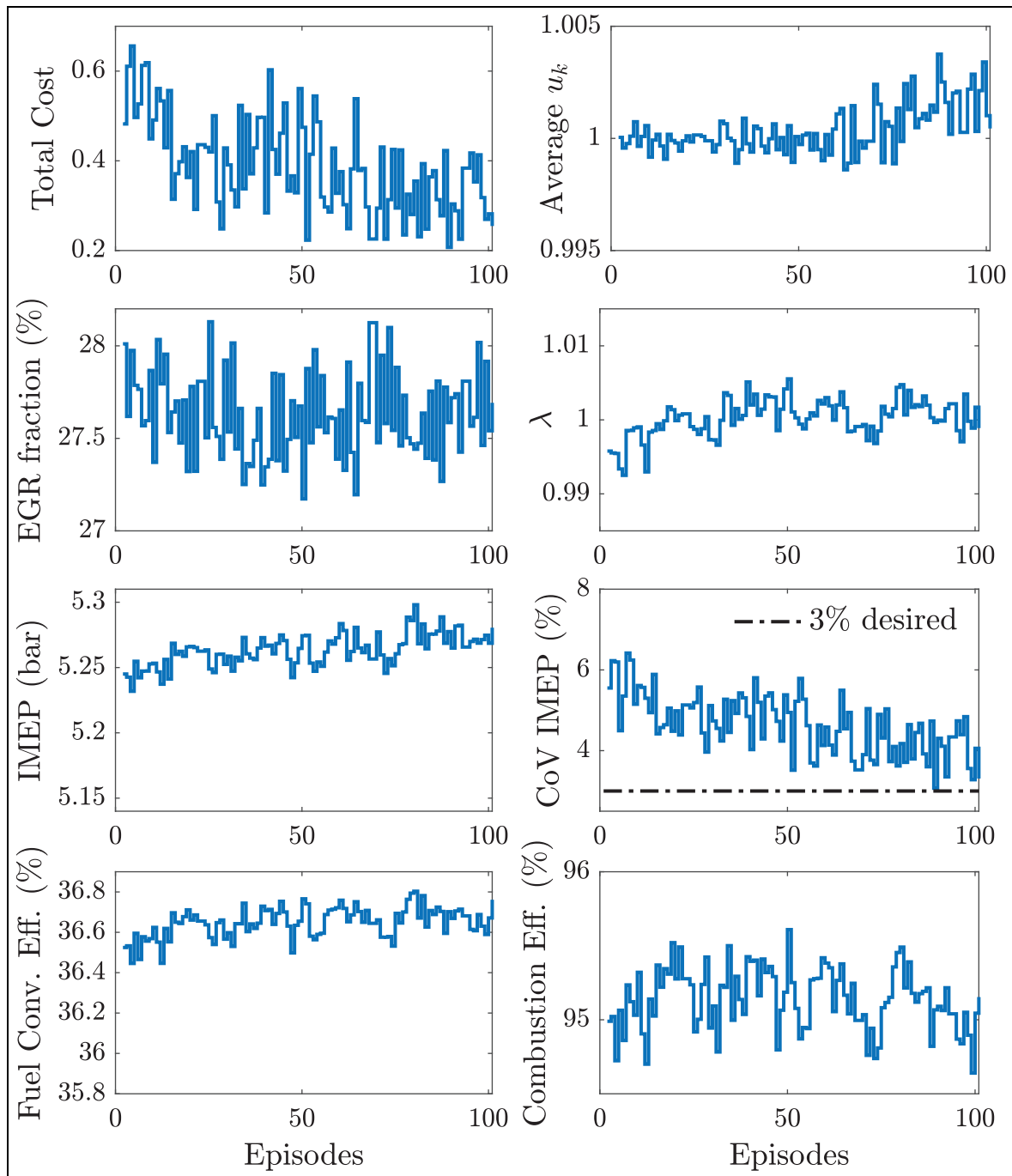
exhaust oxygen sensor indicated near-stoichiometric conditions at the end of the learning. This issue was probably caused by a miss-calibration of the open-loop command  $m_{\text{fuel, in}}^s$  at the beginning of the learning phase. However, this small mistake shows the robustness of the controller under perturbations on initial conditions and highlights the advantages of online learning and adaptation. Ultimately, the proposed controller with the modified cost function was able to approach the 3% CoV of IMEP desired with a net fuel efficiency gain.

Figure 11 shows the response of the controller after the learning phase. Similar to the previous cases, no misfire was observed during 2000 cycles, contributing to the reduction of CoV of IMEP. The control command showed some distinct patterns during sustained periods of time, oscillating between four particular values. There were, however, moments where the controller was less predictable, probably leveraging the cycle-to-cycle dynamics to avoid partial burns and misfires.

Table 4 presents the comparison between open- and closed-loop results of the average combustion indicators. The design of the new cost function guarantees average stoichiometric combustion by maintaining the control command centered at one. A net increase in fuel conversion efficiency of 0.3% was observed. The modest increase in efficiency was probably the consequence of the shortened learning phase. Nonetheless, the cost function was reduced by 53.4% with a reduction of 34.4% in the CoV of IMEP near the desired CCV target.

## Conclusion and outlook

A reinforcement learning strategy was proposed to design an optimal fuel quantity control based on online

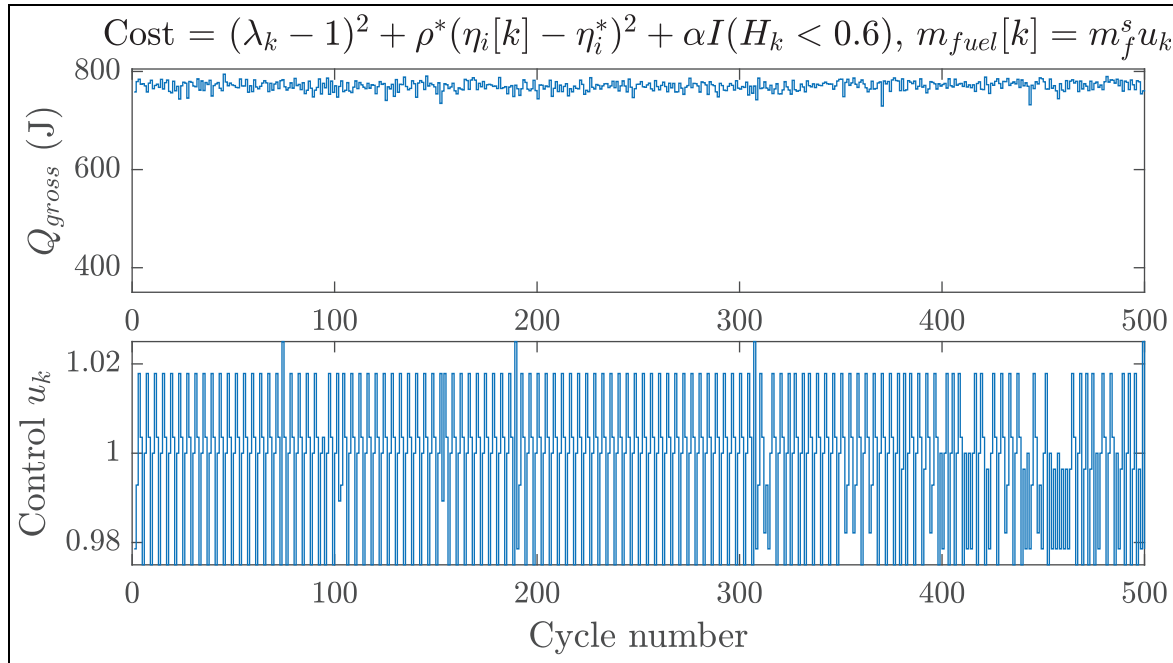


**Figure 10.** Training of RL controller with modified cost.

model-free learning of the combustion dynamics by running the engine in real-time. Full authority over the fuel injection control strategy was achieved using a LabVIEW-based open ECU on an experimental SI engine. The engine was operated at a constant speed and intake airflow according to the ACEC tech team guidelines. The dilution levels were increased to 27.5% EGR fraction, corresponding to a condition past the dilute limit. At each cycle, several combustion parameters were calculated using a simple cycle-to-cycle combustion dynamics model. The control-oriented model allowed for the identification of the system states

needed for a data-efficient RL algorithm. The cycle cost was designed to extend the dilute combustion limit by (1) reducing combustion CCV, (2) improving fuel conversion efficiency, and (3) operating at stoichiometric conditions. An infinite horizon discounting cost problem was formulated and the solution was approximated by discretizing the state and control spaces to use Q-learning. A  $\epsilon$ -greedy decision process was used to encourage exploration and interactions with the engine. The algorithm was tested under three scenarios:

1. Full authority, training for 6h30min



**Figure 11.** Closed-loop response of RL with modified cost.

**Table 4.** Performance of controller trained with modified cost.

	OL	CL	Change (%)
Total Cost (-)	0.43	0.2	-53.40
$\mathbb{E}[u]$ (-)	1	1	-0.09
$\mathbb{E}[\lambda]$ (-)	1	1	-0.03
$\mathbb{E}[\text{IMEP}]$ (bar)	5.24	5.25	0.20
CoV IMEP (%)	5.35	3.51	-34.39
$\mathbb{E}[\eta_i]$ (%)	36.53	36.64	0.30
$\mathbb{E}[\eta_c]$ (%)	95.17	94.96	-0.22

2. Coordination +  $\lambda$  feedback, training for 6h40min
3. Full authority + modified cost, training for 3h20min

In the first scenario, the controller was trained with full authority over the control space. During closed-loop, the improvement of fuel conversion efficiency was the highest observed, at 1.33% over the open-loop condition. However, this was due to a slightly leaner mixture, rendering it incompatible with standard TWCs. In the second scenario, the controller was trained in coordination with a  $\lambda$  feedback controller to maintain a stoichiometric condition. This closed-loop controller was able to extend the dilute limit and improve the fuel conversion efficiency by 1.09% compared to the open-loop. Finally, the third scenario used a modified cost function to guarantee stoichiometric conditions without the need for a second  $\lambda$  feedback controller. The experimental results showed that the controller achieved the desired goals. However, the fuel conversion efficiency improvement was limited to 0.3%, probably due to the reduced

training time. Nonetheless, all the aforementioned controllers reduced significantly the CoV of IMEP, bringing it near to the limit imposed by industry standards. The results presented here were compared against previous studies where a model-based approach was used, highlighting the advantages of model-free learning for engine control. A significant challenge faced by RL in experimental testing on real-world systems is the issue of convergence time. In our case, we limited to operating the engine for up to 7 h. This time constraint implies that if RL requires an extended duration to converge to the global minimum, the resulting control strategy may fall short of optimality. Other than observing the convergence behavior of the episodic cost, we did not have a way to validate that the policy at the end of the training was optimal. Future research will focus on the application of the proposed model-free method under different engine operating conditions and longer training times to quantify the net gain over the entire engine map.

#### Authors' note

This manuscript has been authored in part by UT-Battelle, LLC, under contract DEAC05-00OR22725 with the US Department of Energy (DOE). The US government retains and the publisher, by accepting the article for publication, acknowledges that the US government retains a nonexclusive, paid-up, irrevocable, worldwide license to publish or reproduce the published form of this manuscript, or allow others to do so, for US government purposes. DOE will provide public access to these results of federally sponsored research in accordance with the DOE Public Access Plan (<http://energy.gov/downloads/doe-publicaccess-plan>).




## Declaration of conflicting interests


The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

## Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This research was supported by the Laboratory Directed Research and Development Program of Oak Ridge National Laboratory and used resources at the National Transportation Research Center, a DOE User Facility.

## ORCID iDs

Bryan P Maldonado  <https://orcid.org/0000-0003-3880-0065>

Brian C Kaul  <https://orcid.org/0000-0001-8481-3620>

## References

1. Alsrehin NO, Klaib AF and Magableh A. Intelligent Transportation and control systems using data mining and machine learning techniques: a comprehensive study. *IEEE Access* 2019; 7: 49830–49857.
2. Guanetti J, Kim Y and Borrelli F. Control of connected and automated vehicles: State of the art and future challenges. *Annu Rev Control* 2018; 45: 18–40.
3. Liu L, Lu S, Zhong R, et al. Computing Systems for autonomous driving: State of the art and challenges. *IEEE Internet Things J* 2021; 8(8): 6469–6486.
4. Zhu Q and Prucka R. Transient hybrid electric vehicle powertrain control based on iterative dynamic programming. *J Dyn Syst Meas Control* 2022; 144(2): 021003-1-021003-11. DOI: 10.1115/1.4052230
5. Kosowatz J. Rekindling the Spark. *Mech Eng* 2017; 139(11): 28–33.
6. Maldonado B, Kaul B and Szybist J. Artificial Neural Networks for In-Cycle Prediction of Knock Events. SAE Technical Paper 2022-01-0478, 2022.
7. Kaul B, Maldonado B, Michlberger A, et al. Analysis of Real-World Preignition Data Using Neural Networks. SAE Technical Paper 2023-01-1614, 2023.
8. US Energy Information Administration. Annual Energy Outlook 2021. Technical report, U.S. Department of Energy, 2021.
9. Abd-Alla GH. Using exhaust gas recirculation in internal combustion engines: a review. *Energy Convers Manag* 2002; 43(8): 1027–1042.
10. Alger T, Gingrich J, Roberts C and Mangold B. Cooled exhaust-gas recirculation for fuel economy and emissions improvement in gasoline engines. *Int J Engine Res* 2011; 12(3): 252–264.
11. Bieniek M, Stefanopoulou A, Hoard J, et al. Retard to the limit: closed-loop COVIMEP control for aggressive exhaust heating. *IFAC-PapersOnLine* 2019; 52(5): 624–629.
12. Maldonado BP, Bieniek M, Hoard J, et al. Modelling and estimation of combustion variability for fast light-off of diesel aftertreatment. *Int J Powertrains* 2020; 9(1-2): 98–121.
13. Ahmed OY, Middleton RJ, Stefanopoulou AG, Kim KS and Kweon CBM. Closed-loop diesel combustion control leveraging ignition assist. *IEEE Control Syst Lett* 2022; 6: 1628–1633.
14. Triantopoulos V, Bohac SV, Sterniak J, et al. Cycle-to-cycle variability in spark-assisted compression ignition engines near optimal mean combustion phasing. *Int J Engine Res* 2023; 24(2): 420–436.
15. Krishnamoorthi M, Agarwal A, et al. Combustion instabilities and control in compression ignition, low-temperature combustion, and gasoline compression ignition engines. In: G Kalghatgi, AK Agarwal and H Goyal (eds) *Gasoline compression ignition technology: future prospects*. Springer Singapore, 2022, pp.183–216, 7.
16. Kaul B, Wagner R and Green J. Analysis of cyclic variability of heat release for High-EGR GDI Engine Operation with Observations on implications for Effective Control. *SAE Int J Engines* 2013; 6(1): 132–141.
17. Daw C, Kennel M, Finney C and Connolly F. Observing and modeling nonlinear dynamics in an internal combustion engine. *Phys Rev E* 1998; 57: 2811–2819.
18. Finney CE, Kaul BC, Daw CS, et al. Invited review: a review of deterministic effects in cyclic variability of internal combustion engines. *Int J Engine Res* 2015; 16(3): 366–378.
19. Daw CS, Green JB and Wagner RM. Controlling cyclic combustion variations in lean-fueled spark-ignition engines. *AIP Conf Proc* 2002; 622(1): 265–277.
20. Green J, Wagner R and Daw C. Model based control of cyclic dispersion in lean spark ignition combustion. In: *Proceedings of the 2002 Technical Meeting of the Central States Section of the Combustion Institute*. pp.1–6.
21. Di Cairano S and Kolmanovsky I. Automotive applications of model predictive control. In: SV Raković and WS Levine (eds) *Handbook of model predictive control*. Cham: Springer International, 2019, pp.493–527.
22. Vance JB, Kaul BC, Jagannathan S and Drallmeier JA. Output Feedback controller for operation of spark ignition engines at lean conditions using neural networks. *IEEE Trans Control Syst Technol* 2008; 16(2): 214–228.
23. Singh A, Vance J, Kaul B, et al. Neural Network Control of Spark Ignition Engines with High EGR Levels. In: *The 2006 IEEE International Joint Conference on Neural Network Proceedings*. pp.4978–4985.
24. Vance JB, Singh A, Kaul BC, Jagannathan S and Drallmeier JA. Neural network controller development and implementation for spark ignition engines with high EGR levels. *IEEE Trans Neural Netw* 2007; 18(4): 1083–1100.
25. Vance J, Kaul B, Jagannathan S and Drallmeier J. Neuro emission controller for minimising cyclic dispersion in spark ignition engines with EGR levels. *Int J Gen Syst* 2009; 38(1): 45–72.
26. Shih P, Kaul BC, Jagannathan S and Drallmeier JA. Reinforcement learning-based dual-control methodology for complex nonlinear discrete-time systems with application to spark engine EGR operation. *IEEE Trans Neural Netw* 2008; 19(8): 1369–1388.
27. Shih P, Kaul BC, Jagannathan S and Drallmeier JA. Reinforcement-learning-based output-feedback control of nonstrict nonlinear discrete-time systems with application to engine emission control. *IEEE Trans Syst Man Cybern B* 2009; 39(5): 1162–1179.
28. Daw C, Finney C, Green J, et al. A Simple Model for Cyclic Variations in a Spark-Ignition Engine. In SAE Technical Paper 962086. pp. 2297–2306.

29. Maldonado B and Kaul B. Control-Oriented Modeling of Cycle-to-Cycle Combustion Variability at the Misfire Limit in SI Engines. In: *Proceedings of the ASME 2020 Dynamic Systems and Control Conference*. p. V002T26A001.
30. Schuman C, Young S, Mitchell J, et al. Low Size, Weight, and Power Neuromorphic Computing to Improve Combustion Engine Efficiency. In: *2020 11th International Green and Sustainable Computing Workshops (IGSC)*. pp.1–8.
31. Maldonado BP, Kaul BC, Schuman CD, Young SR and Mitchell JP. Next-cycle optimal fuel control for cycle-to-cycle variability reduction in EGR-diluted combustion. *IEEE Control Syst Lett* 2021; 5(6): 2204–2209.
32. Maldonado BP, Stefanopoulou AG, Kaul BC, et al. Chapter 8 -artificial-intelligence-based prediction and control of combustion instabilities in spark-ignition engines. In: J Badra, P Pal and Y Pei (eds) *Artificial intelligence and data driven optimization of internal combustion engines*. Cambridge, MA: Elsevier, 2022, pp.185–212.
33. Maldonado BP, Zaseck K, Kitagawa E and Stefanopoulou AG. Closed-loop control of combustion initiation and combustion duration. *IEEE Trans Control Syst Technol* 2020; 28(3): 936–950.
34. Maldonado BP and Stefanopoulou AG. Cycle-to-cycle feedback for combustion control of spark advance at the misfire limit. *J Eng Gas Turbine Power* 2018; 140(10): 102812.
35. Jorques Moreno C, Stenlås O and Tunestål P. Indicated efficiency optimization by in-cycle closed-loop combustion control of diesel engines. *Control Eng Pract* 2022; 122: 105097-1-105097-16.
36. Maldonado BP, Li N, Kolmanovsky I and Stefanopoulou AG. Learning reference governor for cycle-to-cycle combustion control with misfire avoidance in spark-ignition engines at high exhaust gas recirculation–diluted conditions. *Int J Engine Res* 2020; 21(10): 1819–1834.
37. Maldonado BP, Kaul BC, Schuman CD, Young SR and Mitchell JP. Next-cycle optimal dilute combustion control via online learning of cycle-to-cycle variability using kernel density estimators. *IEEE Trans Control Syst Technol* 2022; 30(6): 2433–2449.
38. Henry de Frahan MT, Wimer NT, Yellapantula S and Grout RW. Deep reinforcement learning for dynamic control of fuel injection timing in multi-pulse compression ignition engines. *Int J Engine Res* 2022; 23(9): 1503–1521.
39. Bertsekas D. *Reinforcement learning and optimal control. Athena scientific optimization | computation series*. Nashua, NH: Athena Scientific, 2019.
40. Luo Y, Maldonado B, Liu S, et al. Portable In-cylinder pressure measurement and signal processing system for real-time combustion analysis and engine control. *SAE Int J Adv Curr Pr Mobil* 2020; 2(6): 3432–3441.
41. Aleiferis P, Taylor A, Whitelaw J, et al. Cyclic Variations of Initial Flame Kernel Growth in a Honda VTEC-E Lean-Burn Spark-Ignition Engine. SAE Technical Paper 2000-01-1207. SAE International, pp.1340–1380.
42. Lian H, Martz JB, Maldonado BP, et al. Prediction of flame burning velocity at early flame development time with high exhaust gas recirculation and spark advance. *J Eng Gas Turbine Power* 2017; 139(8): 082801–082801.
43. Cha J, Kwon J, Cho Y and Park S. The effect of exhaust gas recirculation (EGR) on combustion stability, engine performance and exhaust emissions in a gasoline engine. *KSME Int J* 2001; 15(10): 1442–1450.
44. Maldonado B, Stefanopoulou A, Scarcelli R, et al. Characteristics of Cycle-to-Cycle Combustion Variability at Partial-Burn Limited and Misfire Limited Spark Timing Under Highly Diluted Conditions. In: *Proceedings of the ASME 2019 Internal Combustion Engine Division Fall Technical Conference*. p. V001T03A018.
45. Maldonado B and Stefanopoulou A. Non-Equiprobable Statistical Analysis of Misfires and Partial Burns for Cycle-to-Cycle Control of Combustion Variability. In: *Proceedings of the ASME 2018 Internal Combustion Engine Division Fall Technical Conference*. ASME, p. V002T05A003.
46. Di H, Zhang Y and Shen T. Chaos theory-based time series analysis of in-cylinder pressure and its application in combustion control of SI engines. *J Therm Sci Technol* 2020; 15(1): JTST0001–JTST0001.
47. Stiffler R, Kaul B and Drallmeier J. Cyclic dynamics of misfires and partial burns in a dilute spark-ignition engine. *Proc IMechE, Part D: J Automobile Engineering* 2021; 235(2-3): 333–345.
48. The Advanced Combustion and Emission Control (ACEC) Tech Team. Advanced Combustion and Emission Control Roadmap. Technical report, U.S. DRIVE Partnership, 2018.
49. Pipitone E. A comparison between combustion phase indicators for optimal spark timing. *J Eng Gas Turbine Power* 2008; 130(5): 052808-1-052808-11. DOI: 10.1115/1.2939012
50. Caton JA. Combustion phasing for maximum efficiency for conventional and high efficiency engines. *Energy Convers Manag* 2014; 77: 564–576.
51. Wallner T, Sevik J, Scarcelli R, et al. Effects of Ignition and Injection Perturbation under Lean and Dilute GDI Engine Operation. In: *JSAE/SAE 2015 International Powertrains, Fuels & Lubricants Meeting*. SAE International, pp. 1–10.
52. Jatana GS and Kaul BC. Determination of SI combustion sensitivity to fuel perturbations as a cyclic control input for highly dilute operation. *SAE Int J Engines* 2017; 10(3): 1011–1018.
53. Maldonado BP and Kaul BC. Evaluation of residual gas fraction estimation methods for cycle-to-cycle combustion variability analysis and modeling. *Int J Engine Res* 2022; 23(2): 198–213.
54. Konda V and Tsitsiklis J. Actor-critic algorithms. In: S Solla, T Leen and K Müller (eds) *Advances in neural information processing systems*. Cambridge, MA: MIT Press, Vol. 12, 2001, pp. 1008–1014.
55. Ameen MM, Mirzaei M, Millo F and Som S. Numerical prediction of cyclic variability in a spark ignition engine using a parallel large eddy simulation approach. *J Energy Resour Technol* 2018; 140(5): 052203.
56. Watkins C. Learning from delayed rewards. PhD Thesis, King's College, Cambridge, 1989.
57. Heywood JB (ed.). *Internal combustion engine fundamentals*. 2nd ed. New York: McGraw-Hill Education, 2018.
58. Schuman C, Young S, Maldonado B, et al. Real-Time Evolution and Deployment of Neuromorphic Computing at the Edge. In: *12th International Green and Sustainable Computing Workshop (IGSC)*, Pullman, WA, USA, pp.1–8.
59. Maldonado B, Kaul B, Schuman C, et al. Dilute Combustion Control Using Spiking Neural Networks. SAE Technical Paper 2021-01-0534. SAE International, pp.1–13.

**Appendix****Notation**

ANN	Artificial Neural Network	EGR	Exhaust Gas Recirculation
AFR	Air-to-Fuel Ratio	EVC	Exhaust Valve Closing
ACEC	Advanced Combustion and Emission Control	EVO	Exhaust Valve Opening
aTDC	after Top Dead Center	IMEP	Indicated Mean Effective Pressure
bTDC	before Top Dead Center	IVC	Intake Valve Closing
CA50	Crank Angle at 50% mass fraction burned	IVO	Intake Valve Opening
CCV	Cycle-to-Cycle Variability	MDP	Markov Decision Process
CoV	Coefficient of Variation	OL	Open-Loop
CL	Closed-Loop	ORCAS	Oak Ridge Combustion Analysis System
ECU	Engine Control Unit	PI	Proportional-Integral
		RL	Reinforcement Learning
		SI	Spark-Ignition
		TWC	Three-Way Catalyst