Machine learning–based analysis of in-cylinder flow fields to predict combustion engine performance

Alexander Hanuschkin[1], Steffen Schober[2], Johannes Bode[3], Jürgen Schorr[4], Benjamin Böhm[3], Christian Krüger[4] and Steven Peters[1] (iD)

Abstract

Cycle-to-cycle variations in an optically accessible four-stroke direct injection spark-ignition gasoline engine are investigated using high-speed scanning particle image velocimetry and in-cylinder pressure measurements. Particle image velocimetry allows to measure in-cylinder flow fields at high spatial and temporal resolution. Binary classifiers are used to predict combustion cycles of high indicated mean effective pressure based on in-cylinder flow features and engineered tumble features obtained during the intake and the compression stroke. Basic in-cylinder flow features of the mid-cylinder plane are sufficient to predict combustion cycles of high indicated mean effective pressure as early as 180 degree crank angle before the top dead center at 0 degree crank angle. Engineered characteristic tumble features derived from the flow field are not superior to the basic flow features. The results are independent of the tested machine learning method (multilayer perceptron and boosted decision trees) and robust to hyper-parameter selection.

Introduction

Optimizing combustion engines is an important field of ongoing research to increase the degree of efficiency and to reduce emissions. Cycle-to-cycle variations (CCVs) of in-cylinder flows result in significant variations of mixture distribution and lead to overall reduction in combustion performance[1–3] and increased emissions.

Experimentally, CCVs can be investigated on the basis of time resolved data from a few measurement locations. A single pressure sensor is commonly used, mapping the complex spatio-temporal system to a single temporal thermodynamic variable. Time-series data of CCVs have been analyzed statistically[4–6] or with wavelets.[7] Next-cycle predictions in homogeneous charge compression ignition (HCCI) engines are possible based on online adaptive extreme learning machines.[8–10]

Optically accessible engines allow investigations of in-cylinder processes based on spatial and time resolved data.[11–14] High-speed particle image velocimetry (PIV) extracts flow dynamics in cross-sections of the

combustion chamber.[11] Different cross-section planes can be investigated sequentially to extract the flow in the whole volume.[15,16] These spatio-temporal data have been investigated using engineered features,[17–19] conditional statistics,[19,20] or proper orthogonal decomposition.[21–23] Graftieaux et al.[17] engineered a tumble feature from the flow field. Tracking the tumble trajectory over time has been successfully applied to visualize, understand, and optimize the flow in combustion engines late in the compression stroke.[20] Standard statistical tools turned out to be not suitable for

[1]Group Research, Future Technologies, Daimler AG, Stuttgart, Germany
[2]University of Applied Sciences Esslingen, Esslingen, Germany
[3]Reaktive Strömungen und Messtechnik, Technische Universität Darmstadt, Darmstadt, Germany
[4]Group Research, Gasoline & Hybrid Powertrains, Daimler AG, Stuttgart, Germany

Corresponding author:
Steven Peters, Group Research, Future Technologies, Daimler AG, 71065 Sindelfingen, Germany.
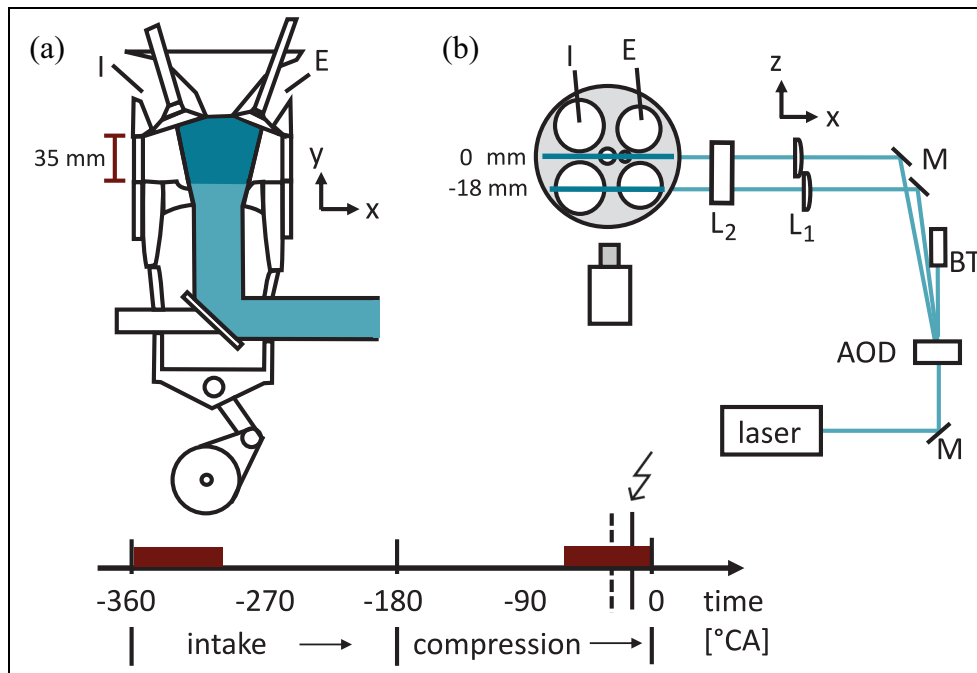Email: steven.peters@daimler.com

Figure 1. Optical engine for high-speed scanning PIV recordings. (a) The optical engine: Intake (I) and exhaust (E) valves are illustrated. The optical window has a height of 35 mm and is located below the cylinder head. Laser beams are directed from the bottom through an optical transparent piston. The imaged area is indicated by dark petrol blue color. (b) Top view of the burning chamber and orientation of measurement planes. The camera records both planes. Shown is the setup for the high-speed scanning PIV with laser, mirrors (M), beam trap (BT), lenses (L1 and L2), and the acousto-optical deflector (AOD). Bottom panel: Time in units of degree crank angle (°CA), relative to the top dead center (TDC) at 0 °CA (first injection (dotted line); ignition (solid line and arrow)). The combustion volume is optically accessible without vertical restriction during 70 °CA around TDC (deep red area).

investigating rare CCV events[20] but require new analysis methods.

Few studies have explored the potential of machine learning (ML) methods to the spatio-temporal combustion process. Flame images were investigated to detect combustion instabilities in gas turbine engine.[24,25] While Sarkar et al.[25] applied deep-belief networks in combination with symbolic time-series analysis, Akintayo et al.[24] used convolutional autoencoders. Jha et al.[26] investigated combustion flame dynamics in a swirl-stabilized combustor using a combination of convolutional neural network and Gaussian processes.

Published ML approaches to understand and predict CCVs in internal combustion (IC) engines are rare. Multivariate statistical analysis and in particular linear regression have been employed to investigate relevant features and predict the point in time of 10% mass fraction burned and the indicated mean effective pressure (IMEP) value.[27] Truffin et al.[28] used correlation analysis and linear multivariate regression to investigate the impact of flow-field and fuel distribution of a spark-ignition (SI) engine on CCVs. Recently, Kodavasal et al.[29] studied CCVs in a gasoline SI engine using random forest. They found that features from the flow fields before ignition and flame shapes can be used for peak cylinder pressure (PCP) classification. Both latter studies are based on large eddy simulation (LES) data.

In this study, we use spatio-temporal resolved PIV data from an optically accessible single-cylinder production-type direct injection spark-ignition (DISI) gasoline engine (Mercedes-Benz M254). Three different ML methods are applied to predict high IMEP cycles from the in-cylinder flow fields.

Methods

Flow-field measurements using time resolved multi-plane high-speed scanning PIV

The data for this study are chosen from two slightly different variants of a multi-plane (either $n = 3$ or 5 planes) PIV experiment of an optically accessible engine operated under identical conditions. The experimental setup of the scanning PIV technique is described in more detail in Bode et al.[30] An overview of the experimental setup is depicted in Figure 1.

A summary highlighting experimental setup, optical engine and its operating point as well as the basic principles and parameters necessary to understand the ML task, preprocessing, and training is given in the following.

Optically accessible engine

All measurements are carried out at an optically accessible single-cylinder DISI engine. Its four-valve pent-

Table 1. Optical engine specifications and operational parameters.

Engine parameter	Value
Bore	83 mm
Stroke	92 mm
Compression ratio	10.5
Engine speed	2000 rpm
IMEP (mean $\pm$ SD)	3.05 $\pm$ 0.13 bar
Intake manifold pressure	985 mbar
Intake valve opening (3 mm valve lift)	−322 °CA
Intake valve closing (3 mm valve lift)	−140 °CA
Triple injection	
$\quad$ SOI$_1$, t_1	−36.25 °CA, 213 μs
$\quad$ SOI$_2$, t_2	−29.50 °CA, 130 μs
$\quad$ SOI$_3$, t_3	−24.25 °CA, 114 μs
Time of ignition	−26 °CA

IMEP: indicated mean effective pressure; SD: standard deviation; °CA: degree crank angle; SOI: start of injection.

roof cylinder head is equipped with a piezo-actuated pintle-type injector (Bosch HDEV 4.1), which is centrally mounted, and a spark plug located between the exhaust valves. In the investigated stratified-lean operating point with a resulting global air/fuel equivalence ratio of 2.9, the fuel (Gasoline E10) is injected with 200 bar late in the compression stroke into the burning chamber. Oil and coolant water temperature are kept constantly at 50 °C and 80 °C, respectively. The engine geometry and the parameters of the stratified-lean operating point are summarized in Table 1.

Optical access to the combustion chamber of the engine is given by two pent-roof windows in the cylinder head, a quartz glass cylinder with a height of 35 mm below the cylinder head, and a piston window within a Bowditch piston arrangement (Figure 1(a)). Note that the optical access is restricted to these windows and hence does not cover the whole combustion chamber. The vertical restriction is 35 mm in its total height of 92 mm. The whole vertical combustion volume is only optically accessible during 70 degree crank angle (°CA) around top dead center (TDC), that is, the early intake and the late compression stroke (time period marked in Figure 1, bottom). Horizontally, the optical access is restricted to the area illuminated by the laser light sheet. Its outer border is given by the crank angle (CA)–dependent refraction from the piston window. The observed horizontal plane size is decreasing with increasing distance from the mid-cylinder plane. The optical access is illustrated in Figure 1(a).

Experimental setup

A high-speed scanning PIV technique is established to measure the in-cylinder flow field quasi-simultaneously in multiple parallel planes (Figure 1(b)). Therefore, a dual-cavity frequency-doubled Nd:YVO$_4$ laser (Edgewave IS100, 532 nm, 1.1 mJ/pulse) is operated at 9 kHz. The laser beam is guided to an acousto-optical deflector (AOD) scanning system (AA Opto-Electronix). The AOD scanning system is adapted to engine measurements from Li et al.[31] It consisted mainly of a 1-axis deflector (AA Opto-Electronix, DTSX-400, TeO$_2$ crystal), in which an acoustic wave propagates. The angle of the deflected beam depends on the oscillation frequency in the AOD crystal. The number of frequency steps defines the number of image planes. Maximum separation distance of planes used in this experiment is from one mid-intake valve plane to the other.

The undeflected light (zero-order deflection angle) is guided into a beam dump. After a certain distance, the deflected beams reach the required distance and are parallelized with separate mirrors. The light sheet is formed with the convex cylindrical lens L1 ($f = 1000$ mm) and the concave cylindrical lens L2 ($f = -50$ mm). The beams are then mirrored through the glass piston into the cylinder with a resulting thickness at full width half maximum of 0.5 mm.

To visualize the in-cylinder flow, silicone oil droplets with a mean diameter of 0.5 μm are seeded into the intake plenum upstream of the intake port providing a homogeneous seeding density. The droplet-scattered light is recorded with a 12-bit CMOS camera (Vision Research Phantom v1610, 28 μm pixel size, 768 $\times$ 624 active pixel). The camera operates at 9 kHz in double frame mode. The camera, the laser, and the AOD scanning system are synchronized with the engine at 2000 rpm via a high-speed controller and the image recording software DaVis 8.3.1 (LaVision).

PIV data for this study are taken from two different experiments recorded in chunks of 90 or 91 consecutive cycles. Operational parameters of the engine (Table 1) and operation conditions are consistent. The experiments had the purpose to capture the three-dimensional characteristics of the in-cylinder flow. PIV measurements are done sequentially in $n = 3$ or 5 planes. The central plane ($z = 0$ mm, below the spark and injector) and the mid-intake valve plane ($z = -18$ mm) are illuminated and imaged in every experiment and are combined to get a consistent data set with a sufficient high number of flow fields. With the engine speed set to 2000 rpm every 1.33 °CA, a double frame image is recorded. Since the n planes are sequentially recorded, this results in a flow-field measurement every 4 °CA (three planes) and 6.66 °CA (five planes) for images of the same plane (intra-plane time difference, Δt_{ip}).

In addition, for each combustion cycle, the in-cylinder pressure trace is recorded with a commercial indicating system (AVL IndiCom), which provides cyclic resolved characteristic thermodynamic values, for example, IMEP.

Flow-field calculation

The PIV processing is conducted with the commercial software DaVis 8.3.1 (LaVision), and the same procedure is used for all planes due to comparable velocity
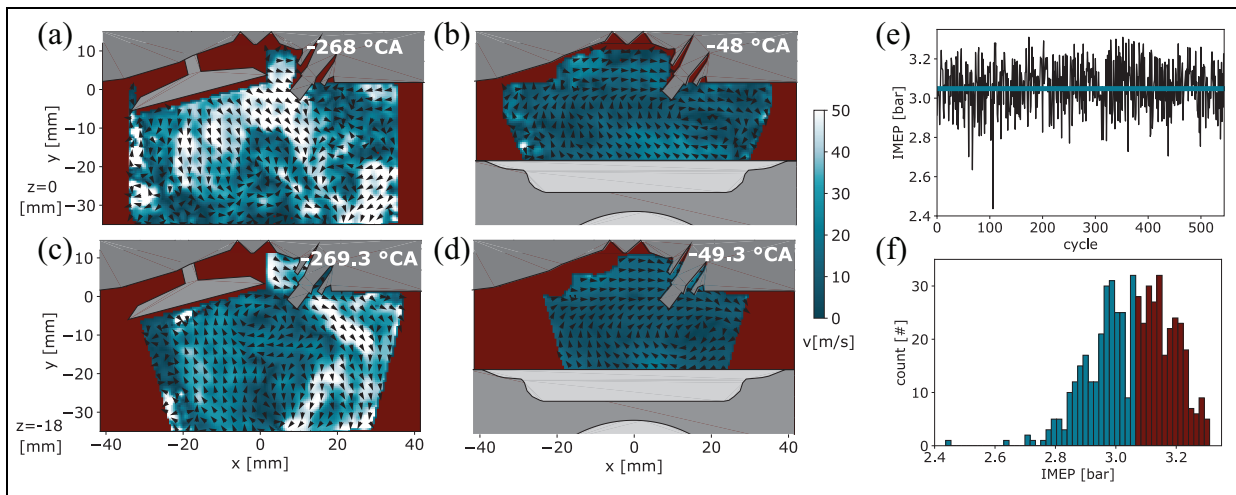
Figure 2. Exemplary flow fields of a single combustion cycle from PIV measurements; (a) mid-cylinder plane ($z = 0$ mm) at $-268\,°$CA (intake stroke), (b) mid-cylinder plane ($z = 0$ mm) at $-48\,°$CA (compression stroke), (c) mid-valve plane ($z = -18$ mm) at $-269.33\,°$CA (intake stroke), (d) mid-valve plane ($z = -18$ mm) at $-49.33\,°$CA (compression stroke). Every fifth vector in each direction is shown. Velocity magnitude is color-coded (deep red = no data). Contours of cylinder head, spark plug, intake valve (a, c), and piston are given (b, d) for orientation. (e) IMEP values of individual cycles. Median value indicated by horizontal line (petrol blue). (f) Histogram of IMEP values. Median split into low (petrol blue) and high (deep red) IMEP values.

magnitudes. Image preprocessing is carried out to improve raw image quality. First, to eliminate reflections, a time filter is used for each CA degree, and second, to improve the contrast, an 8×8 pixel sliding Gaussian background subtraction and a local intensity normalization (5×5 pixel) are performed. For vector calculation, cross-correlation with an adaptive decreasing interrogation-window size and multi-pass iteration from 64×64 to 24×24 pixel with an overlap of 75%, are used. This results in an interrogation-area size of 2.5 mm for the mid-valve plane at $z = -18$ mm and 2.6 mm for the central plane ($z = 0$ mm). Post-processing of the evaluated vector fields is done to remove inaccurate vectors. They are identified by neighbor comparison of the median and replaced by vector choices of higher order. Finally, local noise in the order of the spatial resolution is reduced with a 3×3 top-hat filter. The extracted vector fields are flow fields (i.e. velocity fields).

The flow fields and the thermodynamic data are then imported and combined to a common database with MATLAB (MathWorks). For standardization, all flow fields are transferred to the same grid with a vector spacing of 0.5 mm via linear interpolation. The resulting flow fields have dimensions 167×95 ($x : [-41.5, 41.5]$ mm; $y : [-35, 12]$ mm) and are the basis for the investigation described in the following. Exemplary flow fields from a single combustion cycle are presented in Figure 2. Flow fields from the central plane ($z = 0$ mm; Figure 2, top) and mid-valve plane ($z = -18$ mm; Figure 2, bottom) are shown. During the intake stroke (Figure 2, left), the tumble is not fully developed yet. Later in the compression stroke (Figure 2, right), the tumble flow is clearly visible in both planes. In this exemplary cycle, an out-of-center tumble

position leads to an upward directed flow under the spark plug. This flow configuration influences the fuel spray, mixture formation, and hence combustion efficiency in a different way than a centered tumble position. The occurrence of different tumble center positions at the same operating point may result in high CCVs of IMEP values. Detailed flow-field analysis of this operation point can be found in Bode et al.[30]

ML approach

Supervised ML methods train a model with a set of input (feature) and a corresponding set of output (label) values. The set of input and output values provided during training is called the training set. Appropriately, trained models generalize if they predict the correct output for given input values not used during training. A set of input and output values not used during training but to test the performance of the model is called the test set. ML methods can be used for classification, where labels are categorical. In case of only two classes (categories), the problem is called binary classification task.

A binary classification task is formulated to predict good/high IMEP combustion cycles and identify important feature pattern (Figure 3(a)). Therefore, the set of all IMEP values is median split into two classes. One class represents the high/good and the second class contains the low/poor IMEP combustion cycles. In each PIV experiment, at least 46 ($\Delta t_{ip} = 6.66\,°$CA) time points between $-312\,°$CA and $-8\,°$CA are recorded. A reference time sequence with $\Delta t_{ip} = 6.66\,°$CA for all experiments is set. Experiments with $\Delta t_{ip} = 4\,°$CA are down-sampled and experiments with $\Delta t_{ip} = 6.66\,°$CA are mapped; the experimental point in time closest
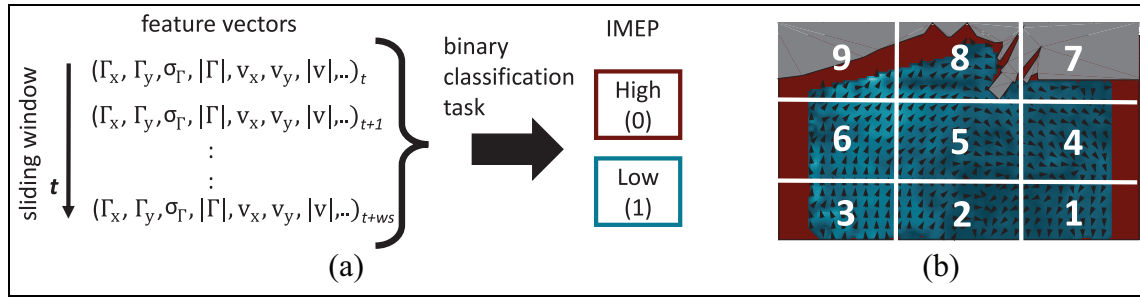
Figure 3. Feature extraction and model overview. (a) Whether a combustion cycle results in a high/good (0) or low/poor (1) IMEP value is predicted by a binary classifier. For each time point t, a feature vector is generated. Time-range models: Feature vectors of consecutive time points are combined (sliding window, window size = ws). (b) Velocity features are calculated in nine spatial sections (numbers, checkerboard regions). Exemplary flow field at $t = -136\,°CA$ and plane $z = 0\,mm$.

Table 2. Data set parameters.

Name	D1	D2	D3
Planes (mm)	$z = 0$	$z = 0, -18$	$z = -18$
Features per time frame	134	268	134
Unique frames	25,024	50,048	25,024
Combustion cycles (sample size)	544	544	544
IMEP values (labels)	544	544	544
Time resolution	6.66 °CA	6.66 °CA	6.66 °CA

IMEP: indicated mean effective pressure; °CA: degree crank angle.

Table 3. Machine learning methods and parameters.

Model	Method	Parameter	Preprocessing
M1	MLP	$hls = \{50, 20\}$, $\alpha = 1e-2$, $bs = 50$, $n_{max} = 20$, $bos = 100$	$n_{PCA} = 10$, $vt = 0.05$
M2	AdaBoost	$ne = 10$, $lr = 0.6$, $bos = 100$	nfr
M3	GBoost	$ne = 4$, $lr = 0.5$, $bos = 100$	$vt = 0.05$
M4	MLP	$hls = \{100, 50\}$, $\alpha = 1e-5$, $bs = 200$, $n_{max} = 100$, $bos = 100$	nfr
M5	AdaBoost	$ne = 10$, $lr = 0.2$, $bos = 100$	nfr
M_{Γ_x, Γ_y}	AdaBoost	Same as M5	Two input features
$M_{\bar{v}_{global, x, y}}$	AdaBoost	Same as M5	Two input features
M_v	AdaBoost	Same as M5	Features subset
M_Γ	AdaBoost	Same as M5	Features subset
M6	MLP	$hls = \{50, 20\}$, $\alpha = 3$, $bs = 50$, $n_{max} = 20$, $bos = 100$	$n_{PCA} = 10$, $vt = 0.05$

MLP: multilayer perceptron; *hls*: hidden layer sizes; α: L2 penalty (regularization); *bs*: batch size; n_{max}: maximum number of optimization steps; *bos*: times bootstrapping; n_{PCA}: number of PCA components; *vt*: variance threshold; *ne*: number of estimators; *lr*: learning rate; nfr: no feature reduction; PCA: principal components analysis.

to the reference time point is always selected without flow-field interpolation. Three different data sets D1, D2, and D3 are generated (Table 2). Data set D1 includes plane $z = 0\,mm$ only, D2 planes $z = 0\,mm$ and $z = -18\,mm$, and D3 plane $z = -18\,mm$ only. All data sets consist of $\geqslant 25,024$ unique frames and 544 combustion cycles (sample size) with the corresponding 544 IMEP values (labels). For each time window, the selected data set is split into training (90%) and test (10%) sets.

Multilayer perceptrons (MLPs) and boosted decision trees (AdaBoost[32] and Gradient Boost) are trained on the training set and tested on the test set. The ML methods and parameter sets used are summarized in Table 3. A brief introduction to the ML methods, model details, and a rigorous performance analysis is given in Appendix.

Decision-tree methods allow direct feature importance calculation, for example, using the Gini importance. Heuristic methods[33,34] or, for example, methods based on relevance propagation[35] to extract feature importance from neural networks exist but are not utilized in this study.

Python is used throughout the study with MLP, AdaBoost, and Gradient Boost implementations in Scikit-learn[36] v. 0.19. Calculations are performed on i7-

Table 4. List of features.

ID(s)	Name	Region
0	Tumble center x: (Γ_x)	Global
1	Tumble center y: (Γ_y)	Global
2–41[a]	Tumble distribution (σ_Γ), histogram (40 bins)	Global
42	Tumble histogram max value	Global
43	Tumble center value (Γ_{max})	Global
44	v_x max	Global
45	v_x min	Global
46	v_x mean	Global
47	v_y max	Global
48	v_y min	Global
49[b]	v_y mean	Global
50	$|v|$ max	Global
51	$|v|$ min	Global
52	$|v|$ mean	Global
53–55[c]	v_x (max/min/mean)	Section 1
56–58	v_y (max/min/mean)	Section 1
59–61	$|v|$ (max/min/mean)	Section 1
62–70	v_x, v_y, $|v|$: (max/min/mean)	Section 2
71–79	v_x, v_y, $|v|$: (max/min/mean)	Section 3
80–88	v_x, v_y, $|v|$: (max/min/mean)	Section 4
89–97	v_x, v_y, $|v|$: (max/min/mean)	Section 5
98–106	v_x, v_y, $|v|$: (max/min/mean)	Section 6
107–115	v_x, v_y, $|v|$: (max/min/mean)	Section 7
116–124	v_x, v_y, $|v|$: (max/min/mean)	Section 8
125–133	v_x, v_y, $|v|$: (max/min/mean)	Section 9

[a]σ_Γ also written as σ_i, with $i \in \{1, \ldots, 40\}$.
[b]Global mean v_y also written as $\bar{v}_{global,y}$.
[c]v_x (max/min/mean) to abbreviate v_x max, v_x min, v_x mean.

6800K or i76820HQ (Intel) multi-core processors. Data plots are generated with Matplotlib[37] v. 2.0.0 and post-processed with Inkscape v. 0.92.3.

Feature extraction

As described in the previous section, supervised ML methods generate an ML model given a set of input feature values and corresponding output labels. Features have to be chosen or engineered for the given problem. Which input features to select is of crucial importance because it influences model performance to a large degree. Supervised ML methods (e.g. deep neural networks (DNNs)) which do not need feature engineering exist and are discussed in the "Discussion" section.

For each image plane, 134 distinct features are extracted for each time point (Table 4). Features of different planes are concatenated for each reference time point in plane order $z = 0, -18$ mm. For example, for data set D2 (compare Table 2), planes $z = 0$ mm and $z = -18$ mm are input to the classifier; the first 134 features are from plane $z = 0$ mm and features 135 to 268 are from plane $z = -18$ mm.

The first 44 features (tumble features) are derived from the engineered in-plane tumble flow.[17] This characterization is called tumble criteria in the following. For each position $\mathbf{x}_i$ in the flow field, its gamma value is given by

$$\Gamma(i) = \frac{1}{N}\sum_{j=1}^{N} \sin(\gamma_{ij}) \tag{1}$$

where γ_{ij} is the angle between the vector $\mathbf{r}_{ij} = \mathbf{x}_j - \mathbf{x}_i$ and the flow vector $\mathbf{u}_j$ at $\mathbf{x}_j$ for all elements j (i.e. positions) in the flow field. The highest absolute gamma value is defined as tumble center. For each image plane, the following features are extracted. The x position and y position of the tumble center denote by Γ_x and Γ_y (features 0 and 1), the tumble value distribution σ_Γ of this frame binned into 40 bins (σ_i, features 2–41), and the values of the highest absolute tumble Γ_{max} derived from the histogram or the formula value (feature 42 or 43, respectively). The tumble center trajectory during the compression stroke after $-150\,°$CA has been successfully applied to engine performance analysis.[20] During this time period, tumble flow with a single clear tumble center is observed in the central plane (Figure 2(b)) and the mid-valve plane (Figure 2(d)). During the intake stroke, the flow is less structured and multiple vortex centers can exist as the tumble is not fully developed yet. Although the tumble center position (features 0 and 1) is only meaningful with a developed and observable tumble, it is computed also in the intake stroke to get a consistent feature map that allows feature importance to be compared over time.

The remaining features (IDs 44–133, basic flow features) are directly derived from the flow field. Velocity magnitudes during the compression stroke after –150 °CA have been successfully correlated to IMEP values before.[15,20] The flow field is divided into nine sections (compare Figure 3(b)). For the total area and in each section, the minimum, maximum, and mean value of the velocity in x-direction, y-direction, and Euclidean vector norm are evaluated.

More than one reference time point must be evaluated to investigate derivatives with respect to time (e.g. gas acceleration). To cover this, a sliding window approach is used. All features at all reference time points within the sliding window of window size ws are combined (Figure 3(a)). Models with ws = 1 are called (single) time-point models while models with ws > 1 are called time-range models in the following. Time-range models have a second parameter (non-overlap, set to 1 throughout this study) defining how the sliding windows overlap. All feature values are mean subtracted and normalized to one standard deviation (SD). To reduce the number of features in a subset of the MLP, AdaBoost and Gradient Boost model features of low variance (threshold = 0.05) are neglected and/or principal components analysis[38,39] (PCA) is used for dimensionality reduction (10, 25 components). For PCA, the data are orthogonally transformed, and the axes of the new basis set are called principal components. The projection is chosen such that the principal components are orthogonal and explain the variance of the data in decreasing order. Technically, the eigenvalue decomposition of the data covariance matrix is carried out, and the eigenvectors of the highest n eigenvalues are used as principal components. A dimension reduction is achieved by choosing n smaller than the data dimension. Preprocessing information for each model is given in Table 3.

Results

The flow fields A to D in Figure 2 are instantaneous representations of the in-cylinder flow (A and B, mid-cylinder plane; C and D, mid-valve plane). The flow fields in both planes are characterized by strong spatial fluctuations and high velocity magnitudes during the intake stroke (A and C). In the mid-valve plane (C), the intake flow on the right-hand side of the intake valve should be noted since this feature mainly forms the large-scale tumble vortex. The tumble is clearly apparent during compression (B and D). The spatial fluctuations are considerably reduced, while the position of the tumble center is shifted toward the right side of the cylinder. This yields an upward directed flow feature in the middle of the cylinder above the piston. But, regarding the complete entity of cycles, the flow and especially the tumble center is distributed from a centered position to the shown one. These different realizations of individual flow field lead to different flow–spray

interactions, fuel distributions, and combustion efficiencies.[30] Figure 2(e) shows the IMEP over all measured cycles. The high fluctuations of the IMEP highlight the CCV of the engine combustion performance. The fluctuations are summarized in a histogram (Figure 2(f)). With a median split, the cycles are separated to good cycles (50% highest IMEP, red) and poor cycles (50% lowest IMEP, blue).

Binary classifiers (MLP and boosted decision trees) are trained to distinguish good/high from poor/low IMEP combustion cycles (Figure 3(a)). A single time-point classifier is trained for each time point ($\Delta t = 6.66$ °CA) during the intake and compression stroke of the combustion cycle.

Classification of high IMEP cycles is possible during late intake and compression stroke

To evaluate the classifier performance, the accuracy of the MLP model (M1, Table 3) is estimated on training and test sets (D2, Table 2). The result is shown in Figure 4(a). For time points in the compression stroke (–180 °CA to –8 °CA), the classifier can clearly identify good/high IMEP combustion cycles above chance level. The mean accuracy value is always more than 1 SD above chance level. For time points in the late intake stroke (–200 °CA to –180 °CA), the mean model accuracy is above chance level but within the 1 SD range. A strong decrease in model accuracy after –36 °CA, the time of first injection, is observed. After the time of first injection and the following flame kernel growth, the quality of the PIV measurement decreases due to tracer particle loss. For earlier time points in the intake stroke (–301 °CA to –200 °CA), the mean accuracy is at chance level. Also the training accuracy is lower for this phase. The obtained accuracy at each time point is independent of the classifier method. MLP (M1), AdaBoost (M2), and Gradient Boost classifier (M3) results are indistinguishable (Figure 4(b)). Receiver operating characteristic (ROC) analysis and F_1 scores are given in Appendix.

The central plane is important for classification accuracy

To investigate the contributions of different combustion chamber (image) planes to model accuracies (Figure 5(a)), three different data sets (Table 2) are investigated with the same MLP classifier (M4). The first data set (D1) contains the central plane ($z = 0$ mm), the second (D2) the central plane ($z = 0$ mm) and the mid-valve plane ($z = –18$ mm), and the third (D3) only the mid-valve plane ($z = –18$ mm). The results of data sets containing the central plane ($z = 0$ mm) do not differ much. For the data set containing only the mid-valve plane, the accuracy is lower in the compression stroke. For values before –180 °CA, all three models produce indistinguishable results. For
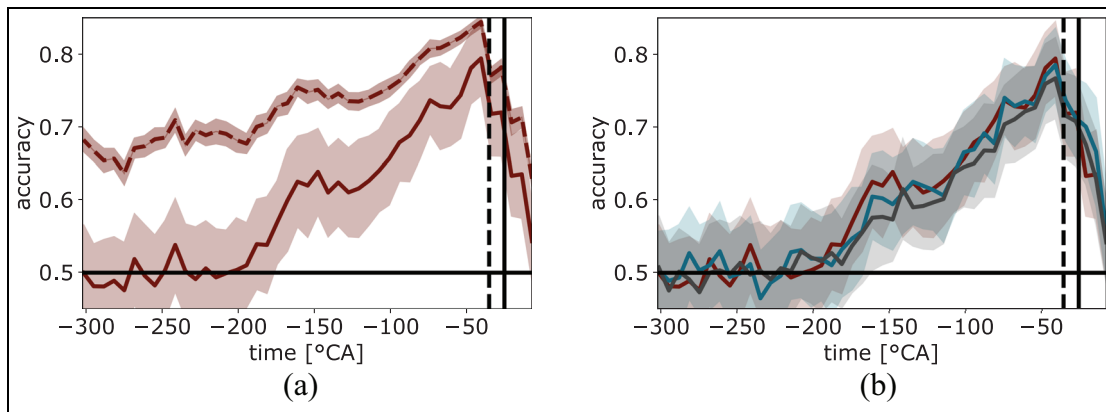
Figure 4. Model accuracy. (a) Accuracy (mean $\pm$ SD) on the training set (dotted line, deep red color) and the test set (solid line, deep red color) of the MLP classifier (M1). Chance level indicated by black line at accuracy of 0.5. First injection vertical dotted line; ignition vertical solid line. (b) Comparison of different classifier. MLP (M1, deep red color), AdaBoost (M2, petrol blue color), and GBoost (M3, gray color). Accuracy (mean) on the test set (solid line). Chance level, first injection, and ignition time are illustrated as in panel (a). All models trained on planes z = 0 mm and z = −18 mm (D2).

the compression stroke, features of the central plane are important for the classifier and just little additional information is contained in the mid-valve plane features.

This is in good agreement with correlation map analyses of Stiehl et al.[19] and Bode et al.[15] and a conditional statistical analysis of Krüger et al.[20] where the flow field in the central tumble plane prior to ignition had the highest relevance for the IMEP CCV. These authors showed in detail that the CCVs were due to flow-spray interactions during stratified operation where the flow in the central plane is directly pointing toward the spray. However, Bode et al.[15] and Krüger et al.[20] showed that CCVs of the tumble flow in the mid-cylinder plane are initially caused by cross flows from the mid-valve plane into the mid-cylinder plane. Thus, the flow in the mid-valve plane is still highly relevant for the entire chain of cause and effect.

Only little gain of time-range models

The features of consecutive time steps can be combined using a sliding window of window size $ws > 1$. These time-range models can potentially use derivatives of the flow field (accelerations). We use the same MLP classifier as above (M4) but vary the window size (Figure 5(b)). Only for large window sizes ($ws = 5$, that is, time window = 33.3 °CA), a small increase in accuracy close to the ignition time is observed. Similar results are observed for a data set with smaller intra-plane time difference ($\Delta t_{ip} = 4°CA$) but smaller (272 samples) sampling size (data not shown).

Time-range models only show minor improvements compared to single time-point models. First, the number of features increases in time-range models because feature sets for each time step are combined. It becomes harder to train a classifier with the same number of samples but with more features, and the model becomes

prone to overfitting.[40] Engineering a small (sub-)set of features or native feature reduction like PCA might be used. Second, the intra-plane time difference has to be sufficiently small. CCV has been described as nonlinear[41] and stochastic.[42] Chaotic behavior has also been reported.[43–48] Nonlinear systems can be linearly approximated in the limit of infinite small time differences. Chaotic behavior is characterized by its sensitivity to initial conditions. Two neighboring trajectories in phase space separate exponentially in time. A sufficiently small time difference in the order of one over the largest Lyapunov exponent has to be chosen for approximation.[49]

Multiple discrete planes are imaged sequentially with PIV in this study. This results in a larger intra-plane time difference. Our results indicate that in future, experimental PIV recordings to study combustion energy variations should be recorded with smaller intra-plane time difference. This data set could then be used for investigating time-range models.

Basic flow features more important than tumble features

Feature importance can be extracted from decision-tree classifier models. In the following, an AdaBoost classifier (M5) is trained on the data set D2 containing planes $z = 0$ mm and $z = −18$ mm (Figure 6(a)). The feature importance of each feature and time point is shown in Figure 6(b). Nearly all features of high importance are from the mid-cylinder ($z = 0$ mm) plane, reflecting the results shown in Figure 5(a). The topmost features (0–43) are derived from the tumble criteria in plane $z = 0$ mm. They do not show high importance (Figure 6(b) and (c)). Note that neither Γ_x nor Γ_y are selected by the classifier. In contrast, the basic flow-derived velocity features in $z = 0$ mm are of high importance and show a pattern during the compression
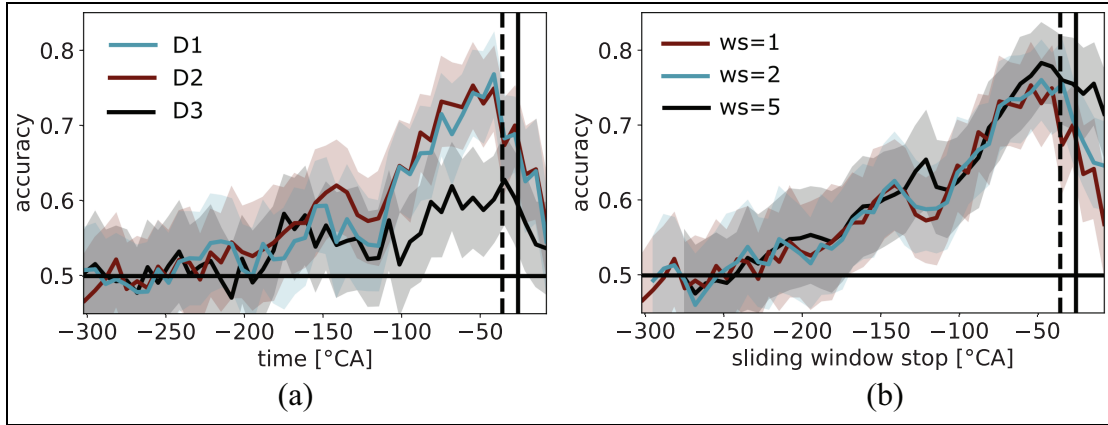
Figure 5. Plane dependence and time-range models. (a) Comparison of models using different plane information (data sets D1, D2, and D3). Accuracy (mean ± SD; test set) of MLP classification model (M4) trained on plane $z = 0$ mm (D1), planes $z = 0$ mm and $z = -18$ mm (D2), or on plane $z = -18$ mm (D3). Chance level indicated by black horizontal line. First injection vertical dotted line; ignition vertical solid line. (b) Comparing different time-range models. Accuracy (mean ± SD; test set) of MLP classifier (M4) with a sliding window size (*ws*) of 1 (identical to panel (a), time window = 6.66 °CA), *ws* = 2 (time window = 13.3 °CA), and *ws* = 5 (time window = 33.3 °CA). All models trained on planes $z = 0$ mm and $z = -18$ mm (D2).

stroke (Figure 6(b), (d), and (e)). Especially, $\bar{v}_{global,x}$ (feature 46, mean velocity in x-direction of total plane) and $\bar{v}_{global,y}$ (feature 49) are important (Figure 6(d)). To evaluate the importance pattern, the mean importance of the features belonging to the same spatial area (section) is averaged (Figure 6(e)) and a clear pattern emerges. At the beginning of the compression stroke (roughly –170 °CA), features in the first section are important. Nearly at the same time but during a longer time period (roughly –160 °CA until –68 °CA), second section features are important. Next (roughly –140 °CA until –120 °CA and –74 °CA until –24 °CA) features in the fifth and finally (after –30 °CA) in seventh and eighth sections are important. The evolution of spatio–temporal flow importance is visualized by arrows in the upper panel of Figure 6(e).

It has been reported that flow near the spark plug is most important for the CCV prediction.[15,20] This is consistent with our result where the spark plug region (region 5) in the central plane after around –44 °CA has been identified to be important for IMEP prediction. We further show that in the compression stroke, the averaged velocities are better suited for classifying high IMEP combustion cycles than features derived from the tumble criteria. This raises the questions of whether the tumble trajectory can be approximated with the averaged velocities and how these features correlate. Conditioned statistics on averaged velocity trajectories for best and worst IMEP subsets have been investigated before[20] but could be refined given the results presented in this study. These investigations could reveal why the averaged velocities are good predictive features. In combination with computational fluid dynamics (CFD) simulations, this insight could be used to optimize the overall engine performance.

The influence of two important features on the model accuracy is investigated (Figure 7(a)). The two input features Γ_x and Γ_y are used to train an AdaBoost classifier (M_{Γ_x, Γ_y}) on data set D1. M_{Γ_x, Γ_y} only reaches chance level before roughly –150 °CA. For later time points, this model is able to fairly classify and is nearly 1 SD below the reference model trained on all 134 features. A second model $M_{\bar{v}_{global,x,y}}$ is trained with the two features $\bar{v}_{global,x}$ and $\bar{v}_{global,y}$ (features 46 and 49) only. $M_{\bar{v}_{global,x,y}}$ is not able to classify above chance level for time points in the intake stroke. In the early time period of the compression stroke (–180 °CA until –100 °CA), classification becomes feasible with lower accuracy as the reference model. However, for the late compression stroke > –100 °CA, the two feature classifier $M_{\bar{v}_{global,x,y}}$ is nearly as accurate as the reference classifier trained on all features.

To disentangle the influence of either tumble derived features or basic flow features, two AdaBoost classifiers on either all tumble features only (M_Γ) or all velocity features only (M_v) are generated (Figure 7(b)). Both are not able to classify in the intake stroke above chance level. In the compression stroke, the accuracy of M_v is higher than the accuracy of M_Γ. This observation is consistent with the result reported above that the velocity features are more important than engineered tumble features. The feature importance of M_Γ for each time point is shown in Figure 7(c). The tumble position (Γ_x, Γ_y) is important in the entire compression stroke. The σ_1 feature is only important at the end of the intake stroke (–200 °CA to –180 °CA). The low tumble distribution features ($\sigma_2, \ldots, \sigma_{18}$) are not used. These features only show small variations (below 0.05 SD) and are usually neglected if a feature preselection would have been done. Higher tumble distribution features ($\sigma_{19}, \ldots, \sigma_{40}$) are used with an unclear temporal pattern. Even though σ_i are calculated globally, they summarize localized velocity structures. Note that Γ_{max} (the maximum value of tumble, features 42 and 43) is not
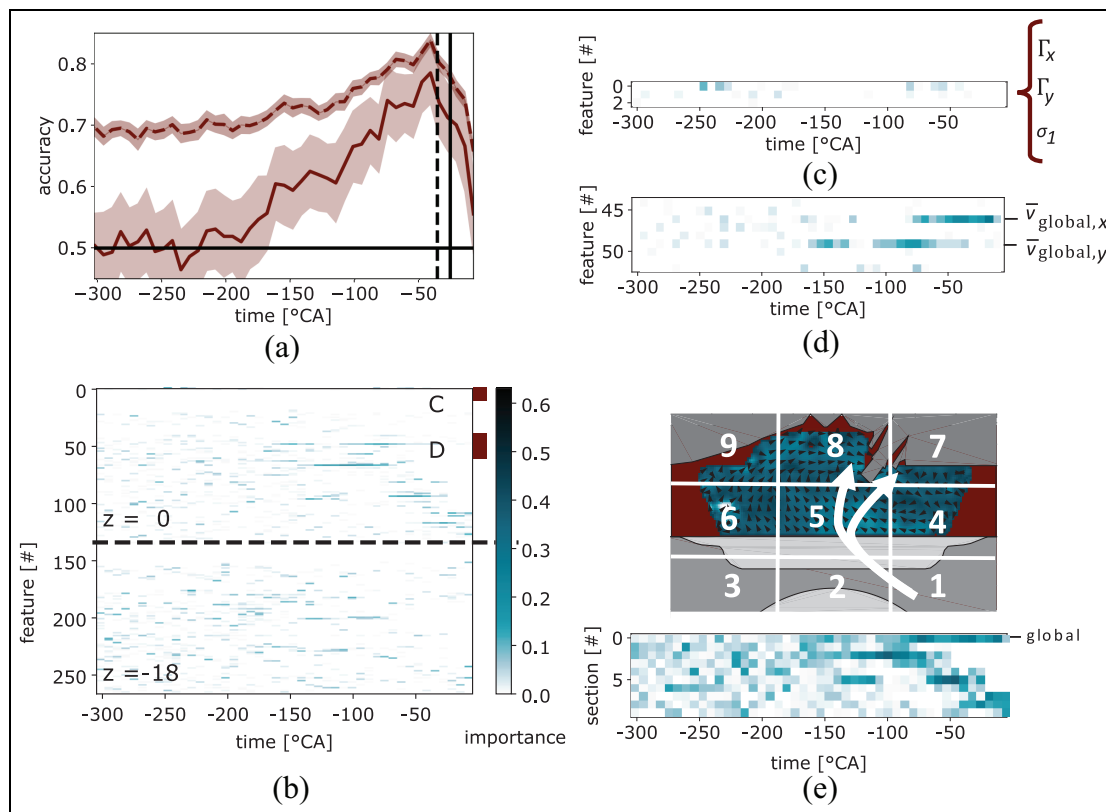
Figure 6. Feature importance for each time point. (a) Accuracy (mean $\pm$ SD) of the AdaBoost classifier (M5) on the test set (solid line, deep red color) and the training set (dotted line, black color). Chance level indicated by black horizontal line. Time of first injection (black vertical dotted line) and ignition (black vertical solid line). (b) Feature importance (color) for each feature and time point. Features 0 to 133 belong to plane $z = 0$ mm and features 134 to 268 to plane $z = -18$ mm. Features 0 to 2 and 44 to 52 are highlighted by deep red boxes (zoom is additionally presented in panels (c) and (d), respectively). (c) Tumble features ($0 = \Gamma_x$, $1 = \Gamma_y$, $2 = \sigma_1$) of the central plane ($z = 0$). (d) Global velocity features ($46 = \bar{v}_{global,x}$, $49 = \bar{v}_{global,y}$) of the central plane ($z = 0$). (e) Section averaged velocity feature importance (lower panel) of the central plane ($z = 0$). Velocity features are calculated in nine spatial sections (upper panel: numbers; $0 = $ global). Arrows (white) illustrate the change of region importance with time. Exemplary flow field at $-44\,°$CA. Panels (b), (c), and (d) share the same color bar (of panel (b)). Colors of panel (e) are linearly scaled.

valued as important by the classifier. The feature importance of M_v (Figure 7(d)) is similar to the feature importance of the reference model on all features (Figure 6(b)). The features $\bar{v}_{global,x}$ and $\bar{v}_{global,y}$ have high importance, and a clear spatio-temporal pattern is observable in the compression stroke.

The complete combustion chamber volume is not optically accessible at all times. Only the area illuminated by the laser light sheet can be imaged (horizontal restriction) and does not always reach the cylinder wall depending on the °CA (Figure 2). The light sheet has maximum width in the center plane and is narrower in the mid-valve planes due to reflection caused by the piston bowl shape. In addition, only the topmost 35 mm of the burning chamber are observed, while the cylinder is 92 mm in height (vertical restriction). In the time period after –290 °CA and before –70 °CA, parts of the combustion volume are not visible and CFD simulations reveal that the tumble center can be below the optical window.[50] This might explain why the tumble criteria derived features are not as powerful as the

raw flow-field data during this time period. However, the tumble assumable generates correlated pattern in the upper part of the cylinder which are sufficient for the ML model. Such patterns have been described before[30] but are difficult to analyze manually.

Discussion

To our knowledge, this study is the first to have successfully applied ML approaches to the in-cylinder processes during the intake and compression stroke. We show that an ML approach using a binary classifier is able to identify high IMEP combustion cycles in each time point of the compression stroke above chance level. The spatial flow-field information or features derived from the tumble criteria[17] are sufficient.

To characterize the tumble flow, Graftieaux et al.[17] engineered a tumble feature (Γ). We verified in this study that a binary classifier with input features (Γ_x, Γ_y) is capable to classify high IMEP combustion cycles in the compression stroke. It is, however, remarkable that
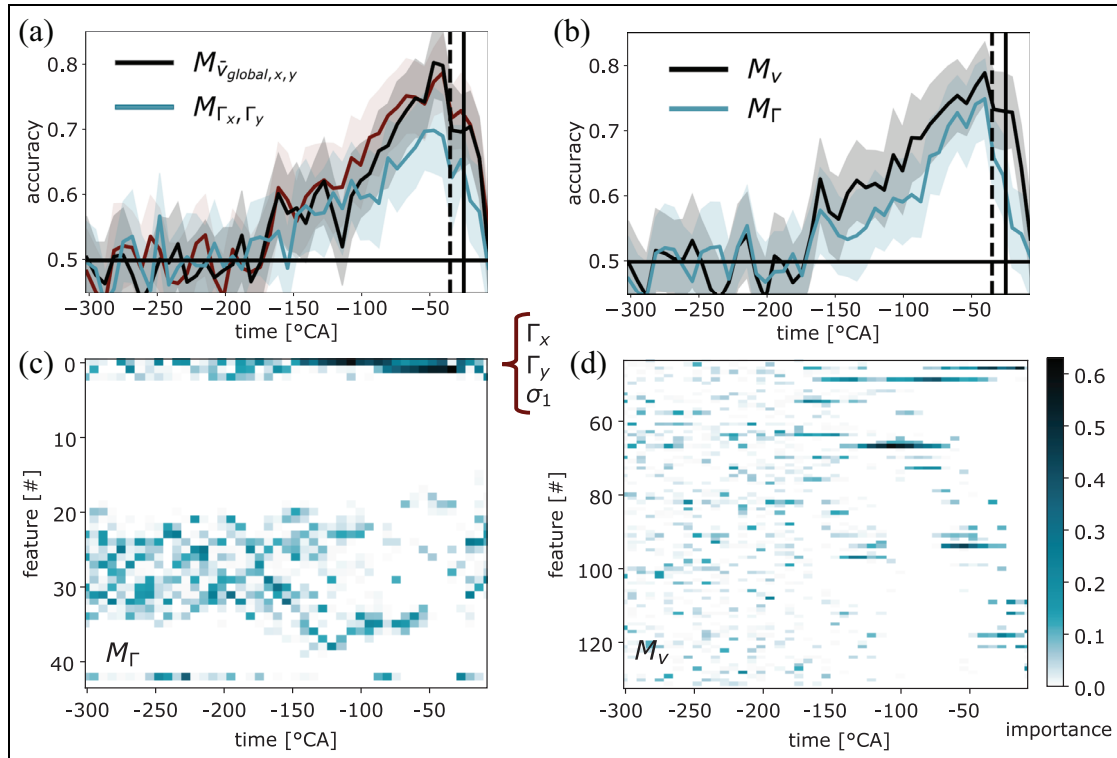
Figure 7. Models with different feature subsets. (a) Without spatial information: Accuracy (mean ± SD; test set) of AdaBoost classifiers with only two input features. $M_{\bar{v}_{global,x,y}}$ (features 46 and 49, black color) and M_{Γ_x,Γ_y} (features 0 and 1, petrol blue color). Reference AdaBoost classifier (M5) with all features (deep red color). Time-point models on data set D1. Chance level (black horizontal line), time of first injection (black vertical dotted line), and ignition (black vertical solid line). (b) With spatial information: Accuracy (mean ± SD; test set) of the AdaBoost classifier (parameter as in (a)) with either all velocity dependent features (M_v, features 44–133, black color) or all tumble derived features (M_Γ, features 0–43, petrol blue color). (c) Feature importance (color values) for each tumble derived feature and time point for M_Γ. (d) Feature importance (color values) for each velocity derived feature and time point for model M_v. Panels (c) and (d) share the same color bar. Panels (b) to (d) group together as indicated by legends.

the boosted decision-tree classifiers value the flow-field information higher and gain more accuracy than the models with this engineered tumble feature even during the compression stroke. Several other engineered features are widely used to characterize the combustion process, but their rigorous testing is beyond the scope of this study. It seems reasonable to argue that non-invertible engineered functions discard information that an ML approach could exploit. More generally, human interpretable features derived from the vector field might be less powerful than using the vector field information directly in an ML approach. Indeed, advances in visual object recognition have shown that human engineered approaches are inferior to direct sensor input to neural networks.[51]

MLPs and boosted decision trees are used in this study for the binary classification task. Both methods produce qualitatively and quantitatively similar results, but their relative performance is not studied in detail. Other methods like logistic regression, Bayesian networks, support vector machines,[52] or random forest[29] could be applied. Ensemble ML methods[53] combining several ML models could also be used. Recently, Moiz et al.[54] employed an ensemble ML method for engine design optimization. A rigorous comparison of these

methods is beyond the scope of this study but could be performed based on a data set of large sample size. It should also be noted that the presented results should be verified with larger sample sizes in the future. Any data analysis relies on reproducible data, that is, data recorded or modeled with invariant parameters and conditions. ML methods for classification and clustering are notably vulnerable to inconsistent training sets because their objective is to separate the feature space into subregions. We deliberately restrict ourselves to minimal feature engineering. Spatial flow-field information alone is sufficient. However, sophisticated feature engineering could be based on vector field feature extraction[55] or segmentation techniques.[56–58]

Kodavasal et al.[29] used a random forest method to classify PCP of parallel LES. The authors used 10 features derived from the flame topology at 2 °CA and the average local velocity in x-, y-, and z-directions in the vicinity of the spark location at –10 °CA. They showed that a random forest classifier is able to predict high combustion cycles based on these features. A random forest with only two features (average local velocities in x- and z-directions at –10 °CA) proves to be able to predict high PCP similarly well. The latter result is consistent with our findings that the local average flow-field

features are sufficient to predict combustion cycles of high/low IMEP in the late compression stroke.

Instead of deploying a binary classifier, a multi-class classifier or a regression approach might reveal results that are more finely graded. A regression approach could uncover how individual features and their interaction determine the IMEP value at each point in time during the combustion cycle. DNNs allow multi-class classification and regression without feature engineering. The features are learned by the DNN.[59,60] This could help to identify high-level features which are not hand crafted. These approaches require large sample sizes with consistent engine parameters. It will be insightful to test these approaches in the future on larger data sets.

Conclusion

We show that ML approaches can be successfully applied to analyze flow-field data of an IC engine with minimal feature engineering. Classification of high IMEP cycles is possible within the compression stroke as early as –180 °CA before TDC.

Our study is based on high spatial and temporal resolved in-cylinder flow-field PIV measurements. Experimental data with sufficient high statistics are derived from an optically accessible single-cylinder version of a commercially available DISI gasoline engine (Mercedes-Benz M254). This allows a direct utilization of the results to future engine design.

The results are robust to classifier method and hyper-parameter settings. MLP and boosted decision-tree classifiers yield quantitatively and qualitatively indistinguishable results. In contrast to previous analysis focusing on correlation analysis or multivariate linear regression, our classifiers are potentially nonlinear. This might be the reason why predictions are possible already for time points in the early compression stroke.

It turns out that the central plane is important for classification accuracy in the compression stroke and only little gain is achieved by using off-center plane information or time-range models. Flow-field features provide more information than the engineered features from the tumble criteria. Hence, ML approaches require less engineering expertise and minimal preprocessing. Only two features (either gamma center (Γ_x, Γ_y) or mean global velocity vectors ($\bar{v}_{global,x}$, $\bar{v}_{global,y}$)) are sufficient to classify in the compression stroke.

Our study highlights how ML approaches can be successfully applied to investigate in-cylinder flow fields and predict combustion cycles of high IMEP in combustion engines. ML allows us to test well-proven engineered features and reveal flow-field feature importance. Important features can be further analyzed in detail. The insights gained into relevant flow-field features can be used for systematic engine design optimization.

ORCID iD

Steven Peters (iD) https://orcid.org/0000-0003-3131-1664

References

1. Hu Z. Non-linear instabilities of combustion processes and cycle-to-cycle variations in spark-ignition engines. SAE technical paper 961197, 1996.
2. Lyon D. Knock and cyclic dispersion in a spark ignition engine. In: *Proceedings of the international conference on petroleum based fuels and automotive applications*, London, 25–26 November 1986. London: Engineering Institution of Mechanical Engineers by Mechanical Publications.
3. Young MB. Cyclic dispersion in the homogeneous-charge spark-ignition engine—a literature survey. *SAE technical paper* 810020, 1981.
4. Litak G, Kamiński T, Rusinek R, Czarnigowski J and Wendeker M. Patterns in the combustion process in a spark ignition engine. *Chaos Soliton Fract* 2008; 35(3): 578–585.
5. Litak G, Kamiński T, Czarnigowski J, Sen AK and Wendeker M. Combustion process in a spark ignition engine: analysis of cyclic peak pressure and peak pressure angle oscillations. *Meccanica* 2009; 44(1): 1–11.
6. Selim MY. Effect of engine parameters and gaseous fuel type on the cyclic variability of dual fuel engines. *Fuel* 2005; 84(7–8): 961–971.
7. Sen AK, Ash SK, Huang B and Huang Z. Effect of exhaust gas recirculation on the cycle-to-cycle variations in a natural gas spark ignition engine. *Appl Therm Eng* 2011; 31(14–15): 2247–2253.
8. Janakiraman VM, Nguyen X and Assanis D. An elm based predictive control method for HCCI engines. *Eng Appl Artif Intell* 2016; 48: 106–118.
9. Vaughan A. *Adaptive machine learning for modeling and control of non-stationary, near chaotic combustion in real-time*. PhD Thesis, University of Michigan, Ann Arbor, MI, 2015.
10. Vaughan A and Bohac SV. Real-time, adaptive machine learning for non-stationary, near chaotic gasoline engine combustion time series. *Neural Netw* 2015; 70: 18–26.
11. Buschbeck M, Bittner N, Halfmann T and Arndt S. Dependence of combustion dynamics in a gasoline engine upon the in-cylinder flow field, determined by high-speed PIV. *Exp Fluids* 2012; 53(6): 1701–1712.

12. Müller S, Böhm B, Gleißner M, Grzeszik R, Arndt S and Dreizler A. Flow field measurements in an optically accessible, direct-injection spray-guided internal combustion engine using high-speed PIV. *Exp Fluids* 2010; 48(2): 281–290.

13. Peterson B, Reuss DL and Sick V. On the ignition and flame development in a spray-guided direct-injection spark-ignition engine. *Combust Flame* 2014; 161(1): 240–255.

14. Sick V, Drake MC and Fansler TD. High-speed imaging for direct-injection gasoline engine research and development. *Exp Fluids* 2010; 49(4): 937–947.

15. Bode J, Schorr J, Krüger C, Dreizler A and Böhm B. Influence of three-dimensional in-cylinder flows on cycle-to-cycle variations in a fired stratified DISI engine measured by time-resolved dual-plane PIV. *Proc Combust Inst* 2017; 36(3): 3477—3485.

16. Voisine M, Thomas L, Borée J and Rey P. Spatio-temporal structure and cycle to cycle variations of an in-cylinder tumbling flow. *Exp Fluids* 2011; 50(5): 1393–1407.

17. Graftieaux L, Michard M and Grosjean N. Combining PIV, POD and vortex identification algorithms for the study of unsteady turbulent swirling flows. *Meas Sci Technol* 2001; 12(9): 1422.

18. Stiehl R, Schorr J, Krüger C, Dreizler A and Böhm B. In-cylinder flow and fuel spray interactions in a stratified spray-guided gasoline engine investigated by high-speed laser imaging techniques. *Flow Turbul Combust* 2013; 91(3): 431–450.

19. Stiehl R, Bode J, Schorr J, Krüger C, Dreizler A and Böhm B. Influence of intake geometry variations on in-cylinder flow and flow–spray interactions in a stratified direct-injection spark-ignition engine captured by time-resolved particle image velocimetry. *Int J Engine Res* 2016; 17(9): 983–997.

20. Krüger C, Schorr J, Nicollet F, Bode J, Dreizler A and Böhm B. Cause-and-effect chain from flow and spray to heat release during lean gasoline combustion operation using conditional statistics. *Int J Engine Res* 2017; 18(1–2): 143–154.

21. Bizon K, Continillo G, Leistner K, Mancaruso E and Vaglieco B. POD-based analysis of cycle-to-cycle variations in an optically accessible diesel engine. *Proc Combust Inst* 2009; 32(2): 2809–2816.

22. Fogleman M, Lumley J, Rempfer D and Haworth D. Application of the proper orthogonal decomposition to datasets of internal combustion engine flows. *J Turbul* 2004; 5(23): 1–3.

23. Roudnitzky S, Druault P and Guibert P. Proper orthogonal decomposition of in-cylinder engine flow into mean component, coherent structures and random Gaussian fluctuations. *J Turbul* 2006; 7: N70.

24. Akintayo A, Lore KG, Sarkar S and Sarkar S. Early detection of combustion instabilities using deep convolutional selective autoencoders on hi-speed flame video. 2016, https://arxiv.org/abs/1603.07839

25. Sarkar S, Lore KG, Sarkar S, Ramanan V, Chakravarthy SR, Phoha S, et al. Early detection of combustion instability from hi-speed flame images via deep learning and symbolic time series analysis. In: *Proceedings of the annual conference of the prognostics and health management society* 2015, https://pdfs.semanticscholar.org/1bcd/f1fd955eea19ee8b094cec800af30c25f999.pdf

26. Jha DK, Srivastav A and Ray A. Temporal learning in video data using deep learning and Gaussian processes. *Int J Prognost Health Monit* 2016; 7(22): 11.

27. Schiffmann P, Reuss DL and Sick V. Empirical investigation of spark-ignited flame-initiation cycle-to-cycle variability in a homogeneous charge reciprocating engine. *Int J Engine Res* 2017; 19: 491–508.

28. Truffin K, Angelberger C, Richard S and Pera C. Using large-eddy simulation and multivariate analysis to understand the sources of combustion cyclic variability in a spark-ignition engine. *Combust Flame* 2015; 162(12): 4371–4390.

29. Kodavasal J, Moiz AA, Ameen M and Som S. Using machine learning to analyze factors determining cycle-to-cycle variation in a spark-ignited gasoline engine. *J Energ Resour Technol* 2018; 140: 102204.

30. Bode J, Schorr J, Krüger C, Dreizler A and Böhm B. Influence of the in-cylinder flow on cycle-to-cycle variations in lean combustion DISI engines measured by high-speed scanning-PIV. *Proc Combust Inst* 2019; 37: 4929–4936.

31. Li T, Pareja J, Becker L, Heddrich W, Dreizler A and Böhm B. Quasi-4D laser diagnostics using an acousto-optic deflector scanning system. *Appl Phys B* 2017; 123(3): 78.

32. Freund Y and Schapire RE. A decision-theoretic generalization of on-line learning and an application to boosting. *J Comput Syst Sci* 1997; 55(1): 119–139.

33. Egmont-Petersen M, Talmon JL, Hasman A and Ambergen AW. Assessing the importance of features for multi-layer perceptrons. *Neural Netw* 1998; 11(4): 623–635.

34. Ruck DW, Rogers SK and Kabrisky M. Feature selection using a multilayer perceptron. *J Neural Netw Comput* 1990; 2(2): 40–48.

35. Bach S, Binder A, Montavon G, Klauschen F, Müller KR and Samek W. On pixel-wise explanations for non-linear classifier decisions by layer-wise relevance propagation. *PLoS ONE* 2015; 10(7): e0130140.

36. Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, et al. Scikit-learn: machine learning in Python. *J Mach Learn Res* 2011; 12: 2825–2830.

37. Hunter JD. Matplotlib: a 2D graphics environment. *Comput Sci Eng* 2007; 9(3): 90–95.

38. Hotelling H. Analysis of a complex of statistical variables into principal components. *J Educ Psychol* 1933; 24(7): 498–520.

39. Pearson K. On lines and planes of closest fit to systems of points in space. *Philos Mag* 1901; 2(11): 559–572.

40. Bellman R. *Dynamic programming*. 1st ed. Princeton, NJ: Princeton University Press, 1957.

41. Daw C, Kennel M, Finney C and Connolly F. Observing and modeling nonlinear dynamics in an internal combustion engine. *Phys Rev E* 1998; 57(3): 2811.

42. Kamiński T, Wendeker M, Urbanowicz K and Litak G. Combustion process in a spark ignition engine: dynamics and noise level estimation. *Chaos* 2004; 14(2): 461–466.

43. Daily JW. Cycle-to-cycle variations: a chaotic process? *Combust Sci Technol* 1988; 57(4–6): 149–162.

44. Foakes A and Pollard D. Investigation of a chaotic mechanism for cycle-to-cycle variations. *Combust Sci Technol* 1993; 90(1–4): 281–287.

45. Kantor JC. A dynamical instability of spark-ignited engines. *Science* 1984; 224(4654): 1233–1235.

46. Wendeker M, Czarnigowski J, Litak G and Szabelski K. Chaotic combustion in spark ignition engines. *Chaos Soliton Fract* 2003; 18(4): 803–806.

47. Wendeker M, Litak G, Czarnigowski J and Szabelski K. Nonperiodic oscillations of pressure in a spark ignition combustion engine. *Int J Bifur Chaos* 2004; 14(5): 1801–1806.

48. Yang LP, Ding SL, Litak G, Song EZ and Ma XZ. Identification and quantification analysis of nonlinear dynamics properties of combustion instability in a diesel engine. *Chaos* 2015; 25(1): 013105.

49. Guegan D and Leroux J. Forecasting chaotic systems: the role of local Lyapunov exponents. *Chaos Soliton Fract* 2009; 41(5): 2401–2404.

50. Nicollet F, Krüger C, Schorr J, Nicoud E, Colin O, Angelberger C, et al. A PIV-guided large-eddy simulation of in-cylinder flows. *Oil Gas Sci Technol* 2017; 72(5): 28.

51. Krizhevsky A, Sutskever I and Hinton GE. ImageNet classification with deep convolutional neural networks. In: *Proceedings of the 25th international conference on neural information processing systems*, Lake Tahoe, NV, 3–6 December 2012, pp.1097–1105. New York: ACM.

52. Boser BE, Guyon IM and Vapnik VN. A training algorithm for optimal margin classifiers. In: *Proceedings of the 5th annual workshop on computational learning theory*, Pittsburgh, PA, 27–29 July 1992, pp.144–152. New York: ACM.

53. Polley EC and Van der Laan MJ. Super learner in prediction. Working paper 266, UC Berkeley Division of Biostatics, May 2010.

54. Moiz A, Pal P, Probst D, Pei Y, Zhang Y, Som S, et al. A machine learning-genetic algorithm (ML-GA) approach for rapid optimization using high-performance computing. SAE technical paper 2018-01-0190, 2018.

55. Hou J, Gao T and Wang P. Flow feature analysis and extraction for classifying axisymmetric vector field patterns. *Multim Tool Appl* 2017; 76(13): 14617–14634.

56. McKenzie A, Lombeyda SV and Desbrun M. Vector field analysis and visualization through variational clustering, Joint Eurographics and IEEE VGTC Symposium on Visualization, 2005. http://dx.doi.org/10.2312/VisSym/EuroVis05/029-035

57. Zhang L, Laramee RS, Thompson D, Sescu A and Chen G. An integral curve attribute based flow segmentation. *J Visual* 2016; 19(3): 423–436.

58. Zhang L, Nguyen D, Thompson D, Laramee R and Chen G. Enhanced vector field visualization via Lagrangian accumulation. *Comput Graph* 2018; 70: 224–234.

59. Dosovitskiy A, Fischer P, Springenberg JT, Riedmiller M and Brox T. Discriminative unsupervised feature learning with exemplar convolutional neural networks. *IEEE T Pattern Anal Mach Intell* 2016; 38(9): 1734–1747.

60. Yu D, Seltzer ML, Li J, Huang JT and Seide F. Feature learning in deep neural networks—studies on speech recognition tasks. 2013, https://arxiv.org/abs/ 1301.3605

61. Rosenblatt F. The perceptron: a probabilistic model for information storage and organization in the brain. *Psychol Rev* 1958; 65(6): 386–408.

62. Rumelhart DE, Hinton GE and Williams RJ. Learning representations by back-propagating errors. *Nature* 1986; 323: 533–536.

63. Cybenko G. Approximation by superpositions of a sigmoidal function. *Math Control Signal Syst* 1989; 2(4): 303–314.

64. Hornik K, Stinchcombe M and White H. Multilayer feedforward networks are universal approximators. *Neural Netw* 1989; 2: 359–366.

65. Kingma DP and Ba J. Adam: a method for stochastic optimization. 2014, https://arxiv.org/abs/ 1412.6980

66. Breiman L. Heuristics of instability and stabilization in model selection. *Ann Stat* 1996; 24(6): 2350–2383.

67. Fawcett T. An introduction to ROC analysis. *Pattern Recogn Lett* 2006; 27(8): 861–874.

68. Powers DMW. What the F-measure doesn't measure: features, flaws, fallacies and fixes. 2015, https://arxiv.org/ftp/arxiv/papers/1503/1503.06410.pdf

Appendix

Multilayer perceptrons (MLPs) and boosted decision trees are used for binary classification. In the following, a brief introduction to these methods is given. MLPs are feed-forward networks of artificial neurons (units) based on the seminal work of Rosenblatt.[61] They consist of an input layer, an output layer, and at least one hidden layer. All neurons of one layer are connected to all neurons of the next layer (fully connected) and nonlinear activation functions are used. MLPs are trained using the backpropagation algorithm.[62] For classification models, the neurons in the output layer represent the different classes and a softmax function is applied. MLPs are powerful tools as they have been proven (with sufficient many hidden units) to be universal approximators.[63,64] A decision-tree classifier basically asks a set of "if–else" questions (binary trees) until a leaf node with an associated class value is reached. The training (classification and regression tree (CART)) algorithm recursively splits a set into two subsets using the single feature and the threshold-value producing purest subsets (i.e. lowest number of false classified training instances). Decision-tree classifiers are easy to interpret (white box models) because they provide simple classification rules, are easy to visualize, and allow direct calculation of feature importance. Boosted decision-tree methods use an ensemble of weak decision-tree classifier to create a strong classifier. A single decision tree is trained and then a second decision tree is trained to correct for its errors. The latter can be achieved by fitting the second tree to a weighted version of the original data set. Additional decision trees are added in the same fashion until a strong classifier is obtained. AdaBoost is a boosting method which performs very well on binary classification problems by using decision trees with only one level (decision stumps). Gradient boosting also sequentially adds new weak classifiers. Each added weak classifier predicts the error of the prior model and uses the gradient descent algorithm to minimize the loss.

Model details

The MLP is optimized with the Adam optimizer[65] on batch sizes of 50 or 200 samples. To avoid bias due to model selection,[66] 100 bootstrapping variants are

generated, and the mean and standard deviation (SD) of the derived model accuracies (number correct predictions/total number predictions) are evaluated. The MLP hyper-parameters are optimized using grid search on regularization parameter (α), the hidden layer size, and the maximum number of iterations at two selected time points (–74.67 °CA (degree crank angle) and –241.33 °CA). The AdaBoost with decision-tree classifiers (weak learners; max depth = 1) is also optimized at this time points using grid search on learning rate and the number of estimators. All other parameters use the default values as specified in the Python Scikit-learn API v. 0.19.

Time-point and time-range classifiers

A time-point classifier is trained on individual time points ($ws = 1$), while a time-range classifier is trained on a set of time points ($ws > 1$) using a sliding window approach (Figure 3(a)). For example, in data set D2 (Table 2), the reference time points are 6.66 °CA apart with 268 features (134 features, 2 planes: $z = 0$, –18 mm) each. Hence, for $ws = 2$ reference points (e.g. [–14.66, –8 °CA]), 536 features are used. The time point classifier trained at time point t_0 is independent from a classifier trained on another time point $t \neq t_0$, except that they share the same feature space and might share the same hyper-parameters during training (e.g. for 46 time points and 100 times bootstrapping, 4600 independent classifier models are trained and evaluated). Neighboring time-range classifiers ($ws > 1$) share parts of the feature values (non-overlap = 1). For example, neighboring time-range classifiers with $ws = 2$ share the last and first set of 134 feature values (per plane), respectively.

Receiver operating characteristic analysis

To evaluate the classifier performance, the receiver operating characteristic (ROC) is investigated using model M6 (Table 3). For each time point, one ROC curve is calculated. PCA (principal components analysis) with 10 components and a variance threshold of 0.05 are used for feature reduction. A shallow MLP classifier (M6) is trained on individual time points ($ws = 1$). The regularization parameter α is increased to 3 in comparison to model M1 ($\alpha = 0.01$). During training, the Adam optimizer with a batch size of 50 is used and the optimization is stopped after 20 epochs to prevent overfitting. To obtain an estimate of the model accuracy distribution, 100 times bootstrapping is applied.

The true positive rate (TPR) as a function of false positive rate (FPR) is plotted in Figure 8(a) using the example of time points –241 °CA, –108 °CA, and –41 °CA. Bin (20 bins) averaged TPR and FPR values are plotted. ROC curves at –108 °CA and –41 °CA are clearly above the bisection line, indicating that the model learns to classify for each of these time points. TPR and the FPR are equal on the bisection line,

indicating that the model performance is at chance level. The area under the curve (AUC) can be used as a measure for model quality.[67] The AUC is largest for –41 °CA, that is, the exemplary time point closest to the ignition. The discrimination threshold (DT) of our binary classifiers can be used to vary the TPR as a function of FPR, as indicated in Figure 8(b). Here, the (bin averaged) DT is plotted as a function of FPR (e.g. at –41 °CA for a DT of 0.5, the FPR is ~0.2 and the TPR is ~0.75, while for DT = 0.2, the FPR is ~0.6 and the TPR is ~0.92). To visualize the ROC for each time point, a heat map is generated (Figure 8(c)). As can be seen from the figure, especially for time points close to the first injection, high TPR and low FPR values are achieved. The information of the time resolved ROC is easier to analyze by plotting the difference between TPR and FPR (Figure 8(d)). The difference is equal to the distance from the TPR line to the bisecting line in each FPR point. As can be seen clearly, all ROC curves for time points after –180 °CA are above the bisecting line except for high FPR values close to 1. Note that at the FPR values of 1, the DT threshold and hence the TPR would be 1 and FPR – TPR = 0. These values are not plotted because only bin averaged values are calculated. Time points between –200 °CA and –180 °CA also exhibit large areas of (mean) positive differences. To indicate the classifier quality, the FPR for DT = 0.5 value is shown on top of the heat map (white solid line in Figure 8(d)). The accuracy for DT = 0.5 is plotted in Figure 8(e). This DT value is used throughout the study. All mean accuracy values are above chance level of 0.5 (median split). The curve of mean accuracy mirrors the shape of the FPR at DT = 0.5 curve in Figure 8(d), because the accuracy of the model depends only on TPR and FPR (prior probability = 0.5 for each class). For each time point during the compression stroke (–180 °CA to –8 °CA), the classifier can clearly identify good/high indicated mean effective pressure (IMEP) combustion cycles above chance level. The mean accuracy value is always more than 1 SD above chance level. For time points in the late intake stroke (–200 °CA to –180 °CA), the mean accuracy is above chance level but within the 1 SD range. A strong decrease in model accuracy after –36 °CA, the time of first injection, is observed. After the time of first injection and the following flame kernel growth, the quality of the particle image velocimetry (PIV) measurement decreases due to tracer particle loss. For earlier time points in the intake stroke (–301 °CA to –200 °CA), the mean accuracy is at chance level. Also the training accuracy is lower for this phase. For each time point, the hyper-parameters are constant and deliberately conservatively chosen to avoid overfitting. To further test for overfitting artifacts, the regularization parameter α in this model (M6) is increased in comparison to model M1 (compare Figures 4(a) and 8(e)). In this case, the training-set accuracy is closer to test-set accuracy during the intake stroke without changing the accuracy on the test set.
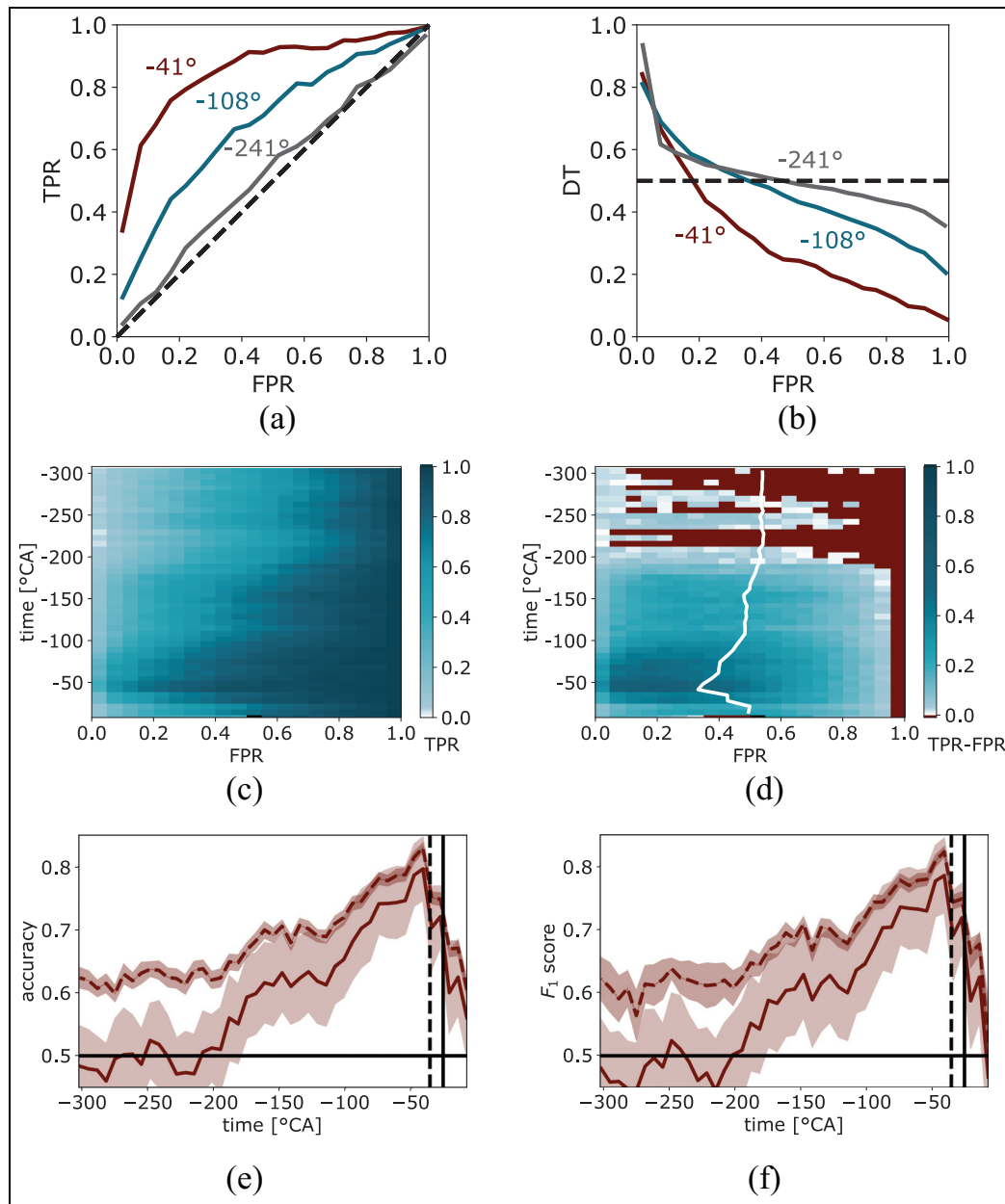
Figure 8. Receiver operating characteristic (ROC) analysis. (a) ROC curve for the MLP classifier (parameters, see text) at different time points on data set D2. False positive rate (FPR) and true positive rate (TPR). Bisecting line (dotted black) is added. (b) Discrimination threshold (DT) as a function of FPR. DT of 0.5 indicated by black dotted line. (c) ROC curve for each time point. TPR indicated by color. (d) Difference between TPR and FPR (color bar) for each time point. Negative values (deep red) are below the bisection line of the ROC curve. FPR values with DT = 0.5 are marked (white, solid line). (e) Accuracy (mean ± SD, DT = 0.5) on training (dotted line) and test sets (solid line). Chance level indicated by black line at accuracy of 0.5. First injection vertical dotted line; ignition vertical solid line. (f) F_1 score (mean ± SD). Line style, chance level, first injection, and ignition time are illustrated as in panel (e).

As shown before, the obtained accuracy at each time point is independent of the classifier method. MLP (M1), AdaBoost (M2), and Gradient Boost (M3) classifier results are indistinguishable (Figure 4(b)). The F_1 score for the MLP classifiers (M6) is shown in Figure 8(f). F_1 score and accuracy (Figure 8(e)) show the same qualitative behavior. The F_1 score is the weighted harmonic mean of recall and precision. Precision is defined

by TP/(TP + FP) and recall by TP/(TP + FN) with true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN). In this study, the classes are perfectly balanced (median split), and the F_1 score for random guessing is 0.5. It can be problematic to rely on the F_1 score alone if the classes are not balanced, because the score is biased to the majority class.[68]