11-29-2022

Machine Learning-based Classification of Combustion Events in an RCCI Engine Using Heat Release Rate Shapes

Radhika Sitaraman
Cummins Technical Center

Sadaf Batool
Michigan Technological University, batool@mtu.edu

Hoseinali Borhan
Cummins Technical Center

Javad Mohammadpour Velni
University of Georgia

Jeffrey Naber
Michigan Technological University, jnaber@mtu.edu

See next page for additional authors

Recommended Citation

Sitaraman, R., Batool, S., Borhan, H., Velni, J. M., Naber, J., & Shahbakhti, M. (2022). Machine Learning-based Classification of Combustion Events in an RCCI Engine Using Heat Release Rate Shapes. *IFAC PapersOnLine, 55*(37), 601-607. http://doi.org/10.1016/j.ifacol.2022.11.248
Retrieved from: https://digitalcommons.mtu.edu/michigantech-p/17301

Follow this and additional works at: https://digitalcommons.mtu.edu/michigantech-p
Part of the Mechanical Engineering Commons

Authors

Radhika Sitaraman, Sadaf Batool, Hoseinali Borhan, Javad Mohammadpour Velni, Jeffrey Naber, and Mahdi Shahbakhti

Available online at www.sciencedirect.com

ScienceDirect

IFAC PapersOnLine 55-37 (2022) 601–607

Machine Learning-based Classification of Combustion Events in an RCCI Engine Using Heat Release Rate Shapes

Radhika Sitaraman *, **Sadaf Batool** **, **Hoseinali Borhan** ***, **Javad Mohammadpour Velni** ****, **Jeffrey D. Naber** †, **Mahdi Shahbakhti** ‡

* *Cummins Technical Center, Columbus, IN 47201 USA (e-mail: sitaraman.radhika@cummins.com)*
** *Michigan Technological University, Houghton, MI 49931 USA (e-mail: batool@mtu.edu)*
*** *Cummins Technical Center, Columbus, IN 47201 USA (e-mail: hoseinali.borhan@cummins.com)*
**** *University of Georgia, Athens, GA 30602 USA (e-mail: javadm@uga.edu)*
† *Michigan Technological University, Houghton, MI 49931 USA (e-mail: jnaber@mtu.edu)*
‡ *University of Alberta, Edmonton, AB Canada (e-mail: mahdi@ualberta.ca)*

Abstract: Reactivity controlled compression ignition (RCCI) mode offers high thermal efficiency and low nitrogen oxides (NOx) and soot emissions. However, high cyclic variability at low engine load and high pressure rise rates at high loads limit RCCI operation. Therefore, it is important to control the combustion event in an RCCI engines to prevent abnormal engine combustion. To this end, combustion in RCCI mode was studied by analyzing the heat release rates calculated from the in-cylinder pressure data at 798 different operating conditions. Five distinct heat release shapes are identified. These different heat release traces were characterized based on start of combustion, burn duration, combustion phasing, maximum pressure rise rate, maximum amount of heat release, maximum in-cylinder gas temperature and pressure. Both supervised and unsupervised machine learning approaches are used to classify different types of heat release rates. K-means clustering, an unsupervised algorithm, could not cluster the heat release traces distinctly. Convolution neural network (CNN) and decision trees, supervised classification algorithms, were designed to classify the heat release rates. The CNN algorithm showed 70% accuracy in predicting the shapes of heat release rates while decision tree resulted in 74.5% accuracy in predicting different heat release rate traces.

Keywords: Low Temperature Combustion (LTC); Reactivity Controlled Compression Ignition (RCCI); Machine Learning; Multi-class Classification

1. INTRODUCTION

Increased air pollution and stringent emission regulations have shifted the focus of automotive manufacturers and researchers towards advanced combustion technologies for better fuel economy and lower emissions. Low temperature combustion (LTC) is one of the advanced combustion technologies evolved over the last two decades. As the name suggests, combustion temperatures in LTC engine are usually low enough to prevent NOx formation (Batool et al., 2022b). Furthermore, lean air-fuel mixture and early injection of fuel result in more premixed air-fuel mixture that reduces fuel rich zones inside the combustion chamber which prevents soot formation (Agarwal et al., 2017).

Several strategies have been proposed to achieve LTC modes including intake air heating, variable compression ratio, variable valve actuation, exhaust gas recirculation (EGR), dual fuels, split fuel injections, and direct dual fuel stratification (Shaver et al., 2004; Batool et al., 2022a; Ravi et al., 2012; Wissink and Reitz, 2015). Homogeneous charge compression ignition (HCCI), premixed charge compression ignition (PCCI), partially premixed charge compression ignition (PPCI), and reactivity controlled compression ignition (RCCI) are the common LTC modes.

Combustion in RCCI mode is achieved by using two fuels of different reactivity. The premixed ratio of two fuels and the injection timing of high reactivity fuel provide better control means to adjust the combustion process

* This work is supported by the United States National Science Foundation (awards #1762520 and #1762595) and the U.S Department of State, Bureau of Educational and Cultural Affairs, Fulbright Program.

10.1016/j.ifacol.2022.11.248

in RCCI (Dempsey, 2013). However, RCCI operation is limited due to high MPRR which causes engine knocking at high load and high cyclic variability at low load (Aliriamezani et al., 2021). Therefore, it is important to understand the combustion in an RCCI engine by analyzing the in-cylinder pressure data. Heat release rate traces provide most of the information about the combustion performance parameters. These combustion performance parameters include start of combustion (SOC), crank angle for 50% heat release (CA50), burn duration (BD), indicated thermal efficiency ($\eta_{ind,th}$). In addition, controlling heat release helps in reducing MPRR and RCCI cyclic variability. Heat release rate is predominantly affected by the variations in the operating conditions. Injection timing of fuel affects the rate of heat release and the HC and CO emissions (Kanda et al., 2005). Effects of injection timing in a dual fuel LTC operation on heat release rate were investigated by Kokjohn et al. (2012). The study suggested that injection timings around -145 CAD aTDC led to rapid combustion. Retarding the injection timing to -50 CAD aTDC resulted in reduced heat release rate. However, reducing the injection timing further (around -15 CAD aTDC) resulted in an abrupt early stage heat release with a tail towards the end of heat release (Kokjohn et al., 2012). Premixed ratio of dual fuels also affects the combustion. A study investigated the effects of variation in the premixed ratio of gasoline and diesel fuels (Lee et al., 2019). They found that prolonged ignition delay was observed for the dual fuel with the highest quantity of gasoline. This resulted in increased indicated efficiency when compared with diesel combustion while reduction in NOx emissions was observed (Lee et al., 2019).

In order to optimize the combustion in the RCCI engine for maximum indicated thermal efficiency, lower engine-out emissions, reduced MPRR and cyclic variability, it is important to control rate of heat release. This can be done by identifying the shapes of heat release rate as a function of engine operating conditions. The focus of this work is to employ machine learning (ML) algorithms for the identification of the type of heat release rate. Various combustion metrics listed in Fig. 1 can be analyzed with ML techniques. ML techniques have been utilized for engine modeling and control (Bidarvatan and Shahbakhti, 2014). Feedforward neural network (FFNN) and radial basis function neural network (RBFNN) were used for the dual-fuel HCCI engine modeling to predict the indicated mean effective pressure (IMEP), thermal efficiency, in-cylinder pressure, net total heat released and engine-out emissions (Rezaei et al., 2015). In (Batool et al., 2021), classification algorithms were developed to model coefficient of variation of indicated mean effective pressure (COV_{IMEP}) for HCCI and RCCI modes. Nonlinear model predictive controllers were developed to control CA50 and IMEP while limiting COV_{IMEP} below 3% (Batool et al., 2021).

To the best of the authors' knowledge, this is the first study undertaken to develop a learning based classification method to identify different types of heat release in RCCI engine operation. This classification is based on an extensive engine study including 798 different engine operating conditions. The contributions of this work are as follows:

(1) An unsupervised learning method, namely k-means clustering, is used for the identification of different heat release rate shapes using normalized heat release rate data;
(2) A supervised learning algorithm, namely convolution neural networks (CNN), is used to classify different heat release shapes using normalized heat release rate data;
(3) Another supervised learning algorithm, namely decision tree, is used to develop an algorithm for the classification of heat release rates based on the engine operating conditions.

Table 1. Range of operating conditions

Parameters	**Range**
Engine speed, N (RPM)	800-2300
Intake manifold pressure, P_{man} (kPa)	96
Intake manifold temperature, T_{man} (°C)	40-100
Start of injection, SOI (CAD bTDC)	15-100
Premixed ratio, PR (-)	20-60
Fuel quantity, FQ (mg/cycle)	9-40

2. RULE-BASED CLASSIFICATION OF HEAT RELEASE RATE IN RCCI

A 2.0L, 4-cylinder gasoline direct injection GM engine was used in this study. RCCI data was collected by running the engine under wide open throttle condition using iso-octane and n-heptane as dual fuels. The premixed ratio of the two fuels is determined by:

$$PR = \frac{m_{iso}LHV_{iso}}{m_{iso}LHV_{iso} + m_{nhep}LHV_{nhep}} \tag{1}$$

where m_{iso} and m_{nhep} are the masses of iso-octane and n-heptane, respectively, LHV_{iso} and LHV_{nhep} represent the lower heating values of iso-octane and n-heptane, respectively. Table 1 shows the range of operating conditions of the RCCI engine data used in this study.

Based on the shape, five different heat release rate patterns were observed in RCCI engine operation. These shapes can be caused by variations in the fuel stratification, fuel reactivity gradient and equivalence ratio gradient. The identified heat release rate patterns can be classified based on the fractions of early and late heat release rate. This classification serves as the basis for developing supervised and unsupervised learning algorithms for identification of heat release types. For rule-based classification, different heat release rate shapes were sorted by identifying the crank angles of the start and end of the main heat release.

Fraction of early heat release is calculated as a percentage of the total cumulative heat release from the start of injection (SOI) to the start of main heat release (SOM) with respect to the total amount of energy of the fuel injected. Fraction of early heat release is determined using (2). Similarly, cumulative late heat release is calculated from the crank angle corresponding to the end of main stage heat release (EOM) to the crank angle corresponding to 90% of heat release (CA90). Then, the fraction of late heat release is computed w.r.t. the total amount of energy of the injected fuel using (3).

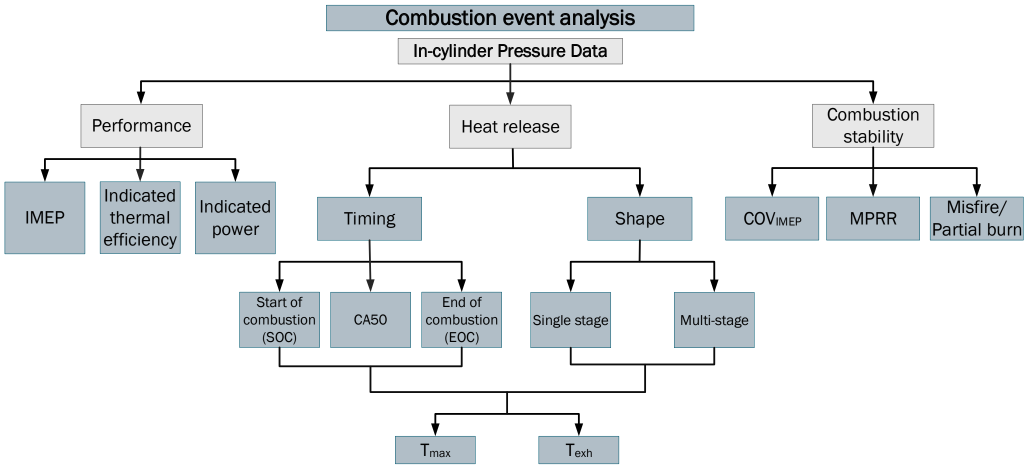

Fig. 1. Combustion metrics

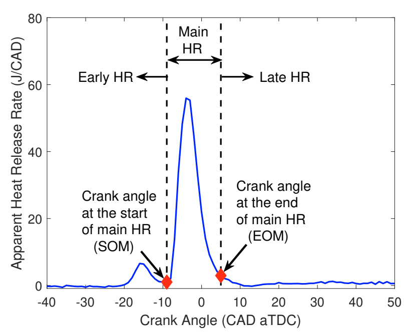

Fig. 2. Crank angles showing early, main and late heat release (HR)

$$\alpha_{HR_{early}} = \frac{\sum_{SOI}^{SOM} HR}{\text{Energy in the injected fuel}} \tag{2}$$

$$\alpha_{HR_{late}} = \frac{\sum_{EOM}^{CA90} HR}{\text{Energy in the injected fuel}} \tag{3}$$

The threshold values for the fractions of early and late heat release were determined based on the experimental data. The rule-based classification logic is shown in Fig. 3 using fractions of early and late heat release rates. In this approach, if the fraction of early heat release is ≤ 5% while the fraction of late heat release is < 17%, then the HRR belongs to type-1. If the fraction of early heat release is > 7%, then the HRR is classified as type-2. If the fraction of early heat release is ≤ 5%, however, the fraction of late heat release is > 23%, then the HRR is identified as type-3. If the fraction of early heat release is ≤ 5% while the fraction of late heat release is ≥ 17% and < 23%, then the HRR is classified as type-4. If the fraction of early heat release is > 5% and ≤ 7%, then the HRR is identified as type-5. These fractions of HR can be used to control RCCI engine operation (Sitaraman et al., 2022).

Five distinct types of heat release (HR) traces are shown in Fig. 4. Type-4 shows the transition phase between type-1 and type-3 while type-5 shows the transition between type-1 and type-2.

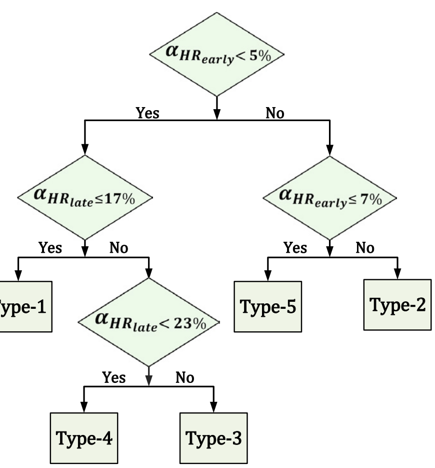

Fig. 3. Flowchart of classification algorithm for HR shapes

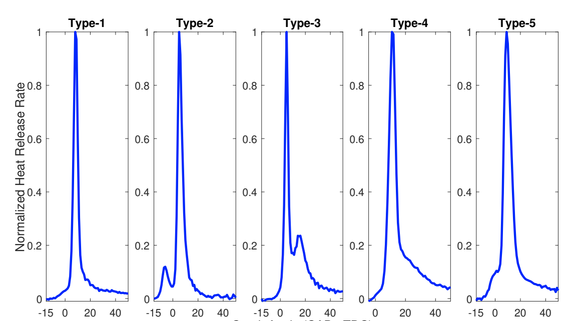

Fig. 4. Identified shapes of heat release rate traces

Summary of the count of HRR traces identified into each type is shown in Table 2.

2.1 Characteristics of HRR Types

Based on the experimental data, heat release rate traces were analyzed on the basis of combustion performance parameters. Table 3 shows the average values of the combustion parameters characterizing each heat release rate type. Type-1 is characterized by its shorter burn duration and higher MPRR compared to other types. This can be associated with the single stage combustion

Table 2. Summary of the classified HRR traces

Type of HRR traces	Count of traces
Type 1	131
Type 2	71
Type 3	373
Type 4	162
Type 5	61

event in type-1. Moreover, type-1 HRR corresponds to the SOI ranging between 35-60 CAD bTDC which may cause fuel stratification. Li et al. (2016) made a similar observation for gasoline fuel (RON 87) in which start of injection in the range of 35 to 50 CAD (bTDC) caused rapid HRR due to mixture stratification. The maximum in-cylinder gas temperature and pressure were higher in type-1 combustion events as compared to the other types. Type-2 HRR showed the most advanced start of combustion and combustion phasing followed by type-5 HRR. The SOI range associated with types 2 and 5 is mainly between 60-100 CAD bTDC. The air-fuel mixture in type-2 and type-5 are well premixed which resulted in advanced combustion. Even with advanced combustion, type-2 showed lower mean MPRR as compared to type-1. This can be explained by the fact that type-2 and type-5 showed staged combustion. That is why the magnitude of peak heat release is lower in type-2 and type-5 as compared to type-1. Types 3 and 4 HRR are characterized as main stage heat release followed by a diffusion type heat release rate. Type-3 showed relatively low P_{max}, T_{max}, HR_{max} and MPRR followed by type-4. A shorter burn duration in type-4 as compared to type-3 can be the reason of slightly higher magnitudes of P_{max}, T_{max}, HR_{max} and MPRR in type-4 HRR. The main purpose of this classification and characterization is to achieve optimal engine operation. The optimal engine operation may target for the maximum thermal efficiency and low engine-out emissions while preventing high MPRR. Among different types of HRR, type-2 showed the maximum indicated thermal efficiency with the maximum ignition delay. Ignition delay is defined as the time between start of injection and the onset of combustion. Higher ignition delay means more premixed air-fuel mixture which helps in preventing NOx formation Lee et al. (2019). However, due to more premixing, some of the operating conditions resulted in higher MPRR which is not a favorable condition. Therefore, for the required load and speed, selection of an optimal HRR type is important to provide a safe engine operation.

3. CLASSIFICATION METHODS FOR IDENTIFICATION OF HRR SHAPES

3.1 Unsupervised Learning using K-means Clustering

K-means clustering is a popular technique for clustering problem, where each centroid represents a data point in a 2-dimensional data frame. In this study, the centroid corresponds to a complete HRR trace. K-means clustering starts with random initialization of centroids, c_1, c_2,...,c_k, of heat release rate data. K is initialized to 5 based on the aim of clustering the HRR traces into five bins. To

Table 3. Characteristics of HRR types based on average combustion parameters

Parameters	HRR Types				
	1	2	3	4	5
$\overline{CA10}$ (CAD aTDC)	4.4	-2.7	3.8	4.8	0.7
$\overline{CA50}$ (CAD aTDC)	7.6	4.6	7.4	8.0	5.4
$\overline{BD}$ (CAD)	12.2	16.7	21.2	18.0	17.4
$\overline{MPRR}$ (bar/CAD)	5.7	5.2	4.0	4.4	4.2
$\overline{Tmax}$ (K)	1780	1509	1542	1592	1446
$\overline{Pmax}$ (kPa)	4204	3998	3561	3697	3723
$\overline{HR_{max}}$ (J/CAD)	115.6	106.7	79.2	87.8	84.2

achieve the centroid convergence, the following two steps are iterated:

(1) In the first step, each data point is designated based on the minimum Euclidean distance to the closest centroid, i.e.,

$$\arg\min_{c_i \epsilon C} (x - c_i)^2 \tag{4}$$

where C is the collection of centroids, c_i is the i^{th} centroid and x is the data point to be assigned to a cluster.

(2) In the second step of the sequence, centroids are recalculated as the mean of data points assigned to its cluster until the convergence is achieved. S_i is the set of data points assigned to i^{th} cluster.

$$c_i = \frac{1}{|S_i|} * \Sigma_{x_i \epsilon S_i} x_i \tag{5}$$

Algorithm stops iteration when no more data points switch between the clusters and the sum of euclidean distance becomes minimum. Based on the operating conditions, each HRR trace resulted in different peak heat release which affect the clustering. Therefore, HRR traces were normalized to cluster them based on shape rather than magnitude. Normalized heat release traces are input to the algorithm. The algorithm identifies the centroids for the cluster through the complete length of the heat release rate vector. Centroids are chosen randomly at the beginning of the classification and the euclidean distance of each trace from the centroid is calculated. Traces with the least distance from the centroid are clustered in a bin. From the clustered traces, centroid is recalculated. The process is repeated until the centroid and clustered traces remain the same after consecutive iterations. K-means clustering approach was used to classify data into 5 bins.

3.2 Supervised Learning using CNN

In supervised learning approach, convolutional neural network is a subset of artificial neural networks. Convolutional neural network (CNN) has been proved to be effective for image recognition. 1D CNN is used for identifying heat release rate traces. It is built as a combination of series of layers to extract the prominent features of the inputs and assign them to corresponding output labels. The CNN takes the 1D vector of HRR trace and passes it across multiple layers of convolutional, pooling and a fully connected layer to obtain output. Output here is the probability of five different classification bins which best represent the HRR traces. First layer of 1D CNN is a convolutional layer with an activation function, in which elements from the data, as per kernel dimension are taken and multiplied with the filter weights. It is summed up as a single element in the feature vector. The kernel slides all through the input data and elements of the the feature vector are obtained. The number of filters depicts multiple combinations of weights of the filter to extract features from input data. Each of these combinations result in a feature vector. All the feature vectors together constitute the convolutional layer.

Pooling is used to reduce the spatial dimension of the feature vector to reduce the computational time. Since, pooling operates individually on each of the feature vectors, the dimensions of maps reduce but the number of maps remains the same. In the final layer, global average pooling is used, where it reduces the complete dimension of the feature vector in to a single value. A dense layer is a fully connected neural network layer where each node on the input is connected to a node on the output. A dropout layer is very similar to dense layer except that when the layer is used, the activation is set to zero for some of the nodes randomly to avoid over fitting.

Training of neural network is achieved by adjusting the filter values through back propagation process. During the training process, initially the weights of the filter are randomly assigned and the output probabilities also end up as random values in the forward pass. The error of the output layer is calculated based on Eq. (6), referred to as loss or total error (L). In order to have the predicted and actual label to be the same, the loss has to be minimum.

$$L = \sum \frac{1}{2}(T - O)^2, \tag{6}$$

where T refers to the target probability and O refers to the output probability. By using back propagation method, the gradients of the error to weights in the network are adjusted to minimize error. By using gradient descent, the filter weights are minimized. Weight update is carried out based on Eq. (7).

$$\mathrm{W} = W_i - \eta \frac{dL}{dW}, \tag{7}$$

where W is the weight, W_i is the initial weight and η is the learning rate of the network. If the learning rate is set too high, it results in large jumps and makes it difficult to reach the optimized point. The process of forward pass is followed by loss calculation and backward pass, respectively. A trained model is achieved by carrying out 500 iterations.

When the same HR shape is input to the trained model, the probability results of the predicted label are more aligned with the actual label. Thus, the model has learnt to process the particular heat release trace to the corresponding label. Through the process of training, only the weights of the filter and connection weights are updated. However, the structure of the network such as number of filters and filter size, remains the same. For supervised learning approach, 65% of the data is used for training the model and the rest 35% is used to evaluate the trained model.

1D CNN model was built and tested using *keras* in python. In CNN approach, a filter of length 9 with 32 features is used. Exponential linear unit (ELU) is used as activation function. Maximum pooling is used in the CNN structure built for heat release trace identification. It helps to reduce dimension of feature map in patches. The layer at end is connected completely to its earlier activation layers.

Table 4. Data dimensions through layers of CNN

Layer (type)	Output Shape	Param #
Conv1D 1	(None, 292, 32)	320
Max pooling 1	(None, 97, 32)	0
Conv1D 2	(None, 91, 64)	14400
Max pooling 2	(None, 30, 64)	0
Conv1D 3	(None, 26, 128)	41068
Global avg. pooling	(None, 128)	0
Dropout 1	(None, 128)	0
Dense 1	(None, 5)	645
Lambda 1	(None, 5)	0

Total trainable parameters: 56,453

3.3 Supervised Learning using Decision Trees

Decision tree is one of the powerful supervised learning algorithms. High accuracy and interpretability are the important characteristics of the decision tree algorithm. Decision tree involves sequential hierarchical decisions which lead to final classification. The modelling process involves two main steps: (i) induction and (ii) pruning. Induction is a process in which a decision tree is built, but the nature of training process results in overfitting. Through the process of pruning, unnecessary structures from the decision tree are removed to prevent overfitting.

Decision tree consists of nodes, edges/branches and leaf nodes. Each node assesses an evaluation condition of a certain feature. Edges/Branch refers to the outcome of a node which connects with another node. Finally, leaf nodes refer to the final outcome resulting in the class labels. For classification of heat release rate traces, recursive binary splitting is used at every node. To calculate the accuracy of split at each node, cost of split is evaluated. For classification, a perspective of the goodness of the split is determined by evaluating the cost function (Gini Index Function) given by

$$G = 1 - \Sigma_k ({p_k}^2), \tag{8}$$

where p_k is the magnitude of the class inputs corresponding to a particular group. High level of purity (p_k) is achieved when the value of G is small. Information gain measures the concept of a single class segregation. Decision tree algorithm evaluates all the features for the highest value of information gain at every node which becomes the evaluation condition for each node. Equation (9) is used to calculate gain:

$$Gain(S,A) = Entropy(S) - \Sigma_{Values(A)} \frac{|S_v|}{|S|}.Entropy(S_v) \quad (9)$$

where S refers to set of occurrences, A represents the features. When A becomes equal to a particular classification value, then S_v denotes the subset of S. *Values(A)* represents the possible values of A in the training data set. Entropy is a measure of uncertainty in the random variable. It also depicts the impurity of the collection. At each node, the same step is evaluated till all the classes are achieved as leaf node.

To apply the decision tree method on HRR data, MATLAB predefined function *fitctree* is employed which uses binary recursive approach. In order to train the model, two major inputs are provided. One of the inputs is the features and other are the labels of the classification. In HRR classification, engine control inputs including engine speed, start of injection of DI fuel, total fuel quantity, premixed ratio and intake manifold temperature are selected as features. The output is the true labels for traces identified initially for training the model. The decision tree approach is prone to overfitting issue, hence the number of leaf nodes was restricted to a maximum of 12, to avoid overfitting.

4. RESULTS AND DISCUSSIONS

The models are trained using K-means clustering, CNN and decision tree algorithms. The performance of the trained models is evaluated by using the testing data. Figure 5 shows the results of HRR clustering in 5 different bins using K-means clustering algorithm. By comparing Fig. 4 and 5, we can not see any clear distinction between the shapes of heat release rates clustered by K-means algorithm. K-means clustering is an unsupervised ML approach which cluster the data points together without any prior knowledge of output labels. Furthermore, alignment of traces and centroid of bins changed with multiple iterations. This made it hard to identify the differences in the heat release patterns. Thus, it is difficult to justify the unique characteristics of each bin.

By evaluating with the testing data, the prediction accuracy of CNN mode is 70%. The prediction accuracy of the model is documented by using a confusion matrix, which provides a comparison between the actual and prediction values as shown in Fig. 6. Diagonal elements of the matrix depict the traces in which true label from the data and predicted labels of the model are the same. The higher the value of the diagonal elements, the better is the prediction accuracy of the model.

Once the decision tree model is determined, the prediction accuracy of the trained model is evaluated using testing data. The summary of the true label and predicted value is shown in Fig. (7). The prediction accuracy of the model is 74.5%, with diagonal elements signifying the predictions tallying with the true label. By comparing the performance of CNN with K-means clustering it can be observed that the CNN algorithm showed better prediction accuracy. This can be explained by the fact that supervised learning algorithms have prior knowledge about the class labels incorporated which leads to a better prediction. By using

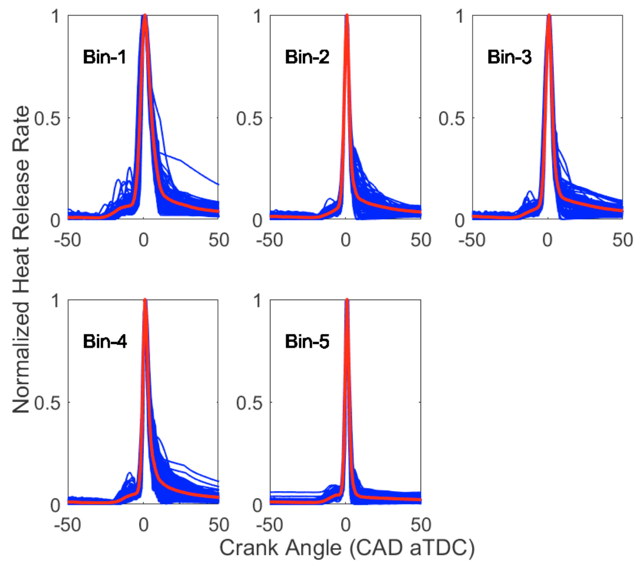

Fig. 5. K-means classification of heat release rate traces

CNN		Predicted Label				
	Bin #	1	2	3	4	5
True Label	1	36	0	0	2	1
	2	1	16	0	0	18
	3	3	0	97	29	5
	4	9	1	6	38	1
	5	5	2	0	1	9

Fig. 6. Prediction summary of the HRR shapes by the model trained using CNN algorithm

the CNN trained model, the type of heat release rate can be determined by providing the normalized heat release rate traces as inputs without calculating the fractions of early and late heat release rate. Furthermore, the decision tree model can determine the type of combustion event based on the control inputs to realize the different types of heat release rate in the RCCI mode.

5. SUMMARY AND CONCLUSIONS

In this work, heat release rate traces of an RCCI engine obtained at about 800 different operating conditions were classified using classification techniques. To develop classification models, supervised and unsupervised ML algorithms were adopted to identify different types of heat release traces. The main findings are:

- Based on the experimental data, five different types of heat release rate were identified using two parameters including fractions of early and late heat release. Type-1 is a single stage heat release with

Decision Tree		Predicted Label				
	Bin #	1	2	3	4	5
True Label	1	16	1	4	0	4
	2	0	23	6	2	4
	3	0	2	146	5	3
	4	2	0	11	12	1
	5	0	2	21	3	11

Fig. 7. Prediction summary of the HRR shapes by the model trained using decision tree algorithm

shorter burn duration. Type-2 HR shows predominant low temperature heat release followed by main stage combustion. In type-3, main stage heat release is followed by diffusion type combustion. Type-4 HR shows transition between type 1 and 3 heat release. Type-5 HR shows transition between type 1 and 2 heat release rates.

- K-means clustering algorithm could not identify distinct patterns of heat release rates. This can be explained by the fact that this algorithm does not have any information about the output labels. Therefore, supervised learning algorithms were preferred for HRR shape classification.
- Convolution neural networks (CNNs) were used to classify five different types of heat release rate. Normalized HRR traces were used as inputs to the model to determine the type of heat release rate. By using the trained model, the heat release patterns can be identified with a prediction accuracy of 70%.
- Decision trees were used to identify the heat release types on the basis of control inputs. The control inputs include start of injection of directly injected fuel, total fuel quantity, pre-mixed ratio and intake manifold temperature and engine speed. The model can predict five different HRR types with an overall prediction accuracy of 74.5%.

In our future work, selected classification algorithms will be used to develop the linear parameter varying (LPV) control oriented models to control RCCI engine operation based on early and late heat release.

REFERENCES

Agarwal, A., Singh, A., and Maurya, R. (2017). "Evolution, challenges and path forward for Low temperature combustion". *Progress in Energy and Combustion Science*, 61, 1–56.

Aliriamezani, M., Koch, C.R., and Shahbakhti, M. (2021). "Modeling, diagnostics, optimization, and control of internal combustion engines via modern machine learning techniques: A review and future directions". *Prog. in Ener. and Comb. Sci.*, 88, 100967.

Batool, S., Naber, J.D., and Shahbakhti, M. (2021). "Data-Driven Modeling and Control of Cyclic Variability of an Engine Operating in Low Temperature Combustion Modes". *IFAC Mod., Est. and Cont. Conf. (MECC)*.

Batool, S., Naber, J.D., and Shahbakhti, M. (2022a). "Closed-Loop Predictive Control of a Multi-Mode Engine Including Homogeneous Charge Compression Ignition, Partially Premixed Charge Compression Ignition and Reactivity Controlled Compression Ignition Modes". *SAE Int. J. of Fuels and Lub.*

Batool, S., Naber, J.D., and Shahbakhti, M. (2022b). "Multi-mode Low Temperature Combustion (LTC) and Mode Switching Control". In *Agarwal A.K., Martínez A.G., Kalwar A., Valera H. (eds). Advanced Combustion for Sustainable Transport, Energy, Environment, and Sustainability*, 43–93. Springer, Singapore.

Bidarvatan, M. and Shahbakhti, M. (2014). "Gray-box modeling for performance control of an HCCI engine with blended fuels". *Journal of Engineering for Gas Turbines and Power*, 136.

Dempsey, A. (2013). *"Dual Fuel Reactivity Controlled Compression Ignition(RCCI) with Alternative Fuels"*. Ph.D. thesis, University of Wisconsin-Madison.

Kanda, T., Hakozaki, T., Uchimoto, T., Hatano, J., Kitayama, N., and Sono, H. (2005). "PCCI Operation with Early Injection of Conventional Diesel Fuel". In *SAE Technical Paper 2005-01-0378.*

Kokjohn, S., Reitz, R., Splitter, D., and Musculus, M. (2012). "Investigation of Fuel Reactivity Stratification for Controlling PCI Heat-Release Rates Using High-Speed Chemiluminescence Imaging and Fuel Tracer Fluorescence". 5(2), 248–269.

Lee, J., Chu, S., Kang, J., Min, K., Jung, H., Kim, H., and Chi, Y. (2019). "The classification of gasoline/diesel dual-fuel combustion based on the heat release rate shapes and its application in a light-duty single-cylinder engine". *International Journal of Engine Research*, 20(1), 69–79.

Li, C., Yin, L., Shamun, S., Tuner, M., Johansson, B., Solsjo, R., and Bai, X. (2016). "Transition from HCCI to PPC: the Sensitivity of Combustion Phasing to the Intake Temperature and the Injection Timing with and without EGR". In *SAE Technical Paper 2016-01-0767.*

Ravi, N., Liao, H., Jungkunz, A.F., Widd, A., and et. al. (2012). "Model predictive control of HCCI using variable valve actuation and fuel injection". *Cont. Engg. Prac.*, 20, 421–430.

Rezaei, J., Shahbakhti, M., Bahri, B., and Aziz, A. (2015). "Performance prediction of HCCI engines with oxygenated fuels using artificial neural networks". *Applied Energy*, 138, 460 – 473.

Shaver, G.M., Roelle, M., and Gerdes, J.C. (2004). "Modeling Cycle-to-Cycle Coupling in HCCI Engines Utilizing Variable Valve Actuation". *IFAC Symp. on Adv. in Auto. Cont.*, 37(22), 227 – 232.

Sitaraman, R., Batool, S., Borhan, H., Velni, J.M., Naber, J.D., and Shahbakhti, M. (2022). "Data-Driven Model Learning and Control of RCCI Engines based on Heat Release Rate". *IFAC Mod., Est. and Cont. Conf. (MECC)*.

Wissink, M. and Reitz, R.D. (2015). "Direct Dual Fuel Stratification, a Path to Combine the Benefits of RCCI and PPC". *SAE Int. J. Engines 8(2):878-889.*