

Research Article

The Start of Combustion Prediction for Methane-Fueled HCCI Engines: Traditional vs. Machine Learning Methods

Mohammad Mostafa Namar ¹, Omid Jahanian ¹ and Hasan Koten ²

¹Faculty of Mechanical Engineering, Babol Noshirvani University of Technology, Babol, Iran

²Istanbul Medeniyet University, Mechanical Engineering Department, Faculty of Engineering and Pure Sciences, Istanbul, Turkey

Correspondence should be addressed to Omid Jahanian; omidjanavania@gmail.com and Hasan Koten; hasan.koten@medeniyet.edu.tr

Received 19 February 2022; Revised 26 April 2022; Accepted 30 April 2022; Published 30 May 2022

Academic Editor: Adriana Del Carmen Téllez-Anguiano

Copyright © 2022 Mohammad Mostafa Namar et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In this work, 11 regression models based on machine learning techniques were employed to provide a fast-response and accurate model for the prediction of the start of combustion in homogeneous charge compression ignition engines fueled with methane. These regression models are categorized into linear and nonlinear types. Although the robust random sample consensus (RANSAC) model is a nonlinear type as well as SAM (simple algebraic model), the prediction accuracy is enhanced from 89.3% to 98.4%. Such accuracy is also achieved for the linear models, namely, ordinary least squares, ridge, and Bayesian ridge models. Indeed, due to the linear hypothesis (the correlation for the start of combustion prediction), the presented models have an acceptable response time to be used in real-time control applications like the electronic control units of the engines.

1. Introduction

Using environmentally friendlier devices is one of the most important approaches to have a green atmosphere for the next decades. Homogeneous charge compression ignition (HCCI) engines are considered as more efficient and cleaner engines than both conventional types, spark ignition (SI) and compression ignition (CI), in the engine industry [1]. These engines release less NO_x emission due to the employed low temperature combustion (LTC) process and consume less fuel as they are designed in lean burnt mode [2]. Furthermore, the carbon-based emissions such as CO₂, CO, and HC may be decreased or even removed thanks to the ability of these engines in running by the light hydrocarbons such as methane or hydrogen [3, 4]. In consequence, the target of having a cleaner environment can be achieved by extending the applications of this type of engine in the automotive industry [5, 6] as well as the other approaches such as reactivity controlled compression ignition (RCCI) [7, 8], catalyst converter development [9, 10], and engine downsizing [11, 12].

One of the most challenging issues for the industrial applications of HCCI engines is the combustion stability, especially in high-speed operation conditions, and it is needed to be controlled by the start of combustion (SOC) timing [13]. There is no direct controller actuator for the SOC, such as the spark plug at the SI and injector (injection timing) at the CI engines, and the SOC will be affected by the engine inlet parameters like the pressure and temperature of the inlet charge, equivalence ratio, and engine speed [14]. Consequently, the real-time exact prediction of SOC may lead to extending the usage of these types of engines in movement applications, as it is considered as one of the key parameters in combustion stability [15, 16].

A wide range of SOC predictors has been presented in the literature. To study the effects of key parameters on the SOC, fundamentally, the three-dimensional computational fluid dynamics (3D-CFD) models are used [17, 18]. These models present high accuracy due to employing detailed chemical kinetics mechanism while they need a long run time. The 0D thermodynamic models are considered as the second type of SOC

predictors which have much better run time than 3D-CFD models [19, 20]. Although the run-time is noticeably declined by the 1D thermodynamic models, developing the predictor models leads the researcher to provide the third type called semiempirical models for the faster ones [21, 22]. These models generally use the engine performance data and provide an algebraic correlation to predict the SOC. The primary efforts caused the models which need to unmeasurable parameters as seen in literature [23, 24] and then they are developed for the models with measurable inputs [25, 26]. Although the accuracy of these models is acceptable for control application, the run time is still needed to be improved in the real-time applications.

Flourishing the machine learning (ML) techniques is a promising achievement to enhance the performance of SOC predictors. These techniques have been recently used by researchers to develop control models in the engine industry [27–29]. Using ML, it is possible to provide the numerical models to predict the SOC timing employing the engine performance dataset. These models may have better run times than the modified knock integral model (MKIM) [25] and simple algebraic model (SAM) [26] if the proposed correlations have less nonlinearity.

In this work, employing the dataset consists of the performance and operating conditions of 3 different HCCI engines fueled with Methane, adopted from a thermodynamic model, and using different regression ML techniques and several SOC prediction models are presented. The performance of them is compared with Simple Algebraic Model (SAM) model. The main novelty of this work is providing a simple linear model with higher accuracy thanks to the ML techniques. It is completely sufficient for control applications and to use in the Electronic Control Unit (ECU) via the linearity of the model.

2. Model Description

In general, three different types of numerical models are employed in this research, namely, SAM, virtual engine, and ML regressor models. In this study, 6 linear and 5 nonlinear supervised type regressor models were used. In this section, these models are described in detail.

2.1. SAM. The SOC time of methane-fueled HCCI engines will be predicted via a simple algebraic correlation due to the engine inlet parameters [26]:

$$\text{SOC} = C_1 \frac{N^A \phi^B (1 + \text{EGR})^D}{P_{\text{IVC}}^E T_{\text{IVC}}^F} + C_2, \quad (1)$$

where N is the engine rotary speed, ϕ is the equivalence ratio, EGR is the mass ratio of exhaust gas recirculated (EGR) to inlet charge, and P_{IVC} and T_{IVC} refer to the inlet charge pressure and temperature at the inlet valve closing (IVC) time, respectively. The constants A , B , D , E , and F relate to the used fuel, and C_1 and C_2 depend on the real compression ratio (CR).

2.2. Virtual Engine. Each control-oriented model uses the experimental dataset to provide its own semi-empirical correlations. In this research, the results of a stand-alone thermodynamic model of engine closed-cycle [3, 5, 26] are employed instead of the experimental data. This model includes a detailed chemical kinetics mechanism of methane oxidation called “GRI 3.0 [30]” which consists 325 reactions and 53 species. The validity of the employed model has been evaluated by several engines and different operation conditions, and SOC is defined at a crank angle where 5% of the fuel is consumed. The main excellence of this model is SOC detection with acceptable accuracy; therefore, it is sufficient to be used as the virtual engine for this study. The virtual engine is run for a wide range of engine inlet parameters, namely, engine speed, equivalence ratio, EGR, charge pressure and temperature at IVC, relative humidity (RH), and real CR.

2.3. ML Regressor Model. ML regressor models generally use a hypothesis to define the target value (SOC). For the multivariate linear regression, the hypothesis is defined as [31]

$$h_{\theta}(X) = \sum_{i=0}^n x_i \theta_i, \quad (2)$$

where x is the independent variable and θ is the coefficient of the related variable. The subscript n refers to the number of independent variables. It should be noted that the x_0 is considered equal to 1.00. Using the training dataset, the target of ML techniques is to find the best vector of θ which has the least residual for the test dataset. So, the type of considered cost function, which should be minimized, is the base of the definition of different ML techniques; as an example, the cost function of ordinary least squares (OLS) method is defined as [32]

$$J(\theta) = \frac{1}{2m} \sum_{i=1}^m (h_{\theta}(X^{(i)}) - Y^{(i)})^2, \quad (3)$$

where Y is the vector of dependent variable of the problem (target value of each case) and the superscript m is the number of train cases. A variety of approaches are used to minimize the cost function, but generally, they can be divided into two categories, namely, noniterative and iterative. In noniterative approach, the best vector of θ is calculated by $\partial/\partial\theta_j J(\theta) = 0$ in the train dataset domain as [33]

$$\theta = (X^T X)^{-1} X^T Y. \quad (4)$$

For the iterative approach, the vector of θ is simultaneously updated based on the learning rate (α) in each iteration [34].

$$\theta_j := \theta_j - \alpha \frac{\partial}{\partial \theta_j} J(\theta). \quad (5)$$

3. Results and Discussion

Considering the validated virtual engine [3, 5, 26], demanded dataset is constructed by running the virtual engine for 3 engines defined in Table 1. This dataset

TABLE 1: Studied engines' specifications.

Parameters	Values		
	Yanmar L70AE [35]	Caterpillar 3500 [36]	Volvo TD100 [37]
Bore (mm)	78	170	120
Stroke (mm)	62	190	140
CR	19.5	17	19.8
Displacement volume (L)	0.296	4.3	1.6
IVC [aBDC]	59°	20°	13°
EVO [bBDC]	59°	40°	39°

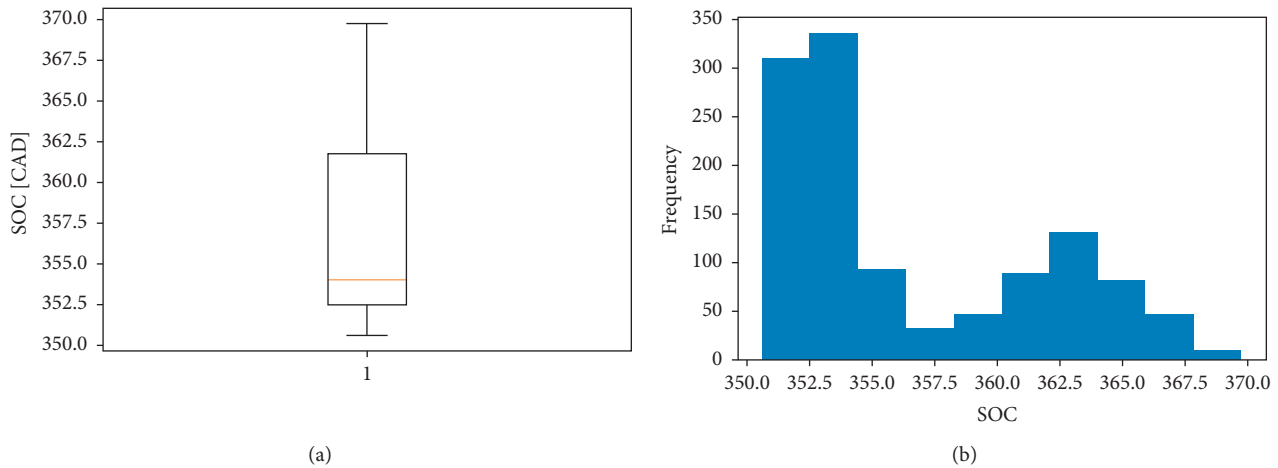


FIGURE 1: The dispersion of SOC in train dataset. (a) Box plot. (b) Histogram plot.

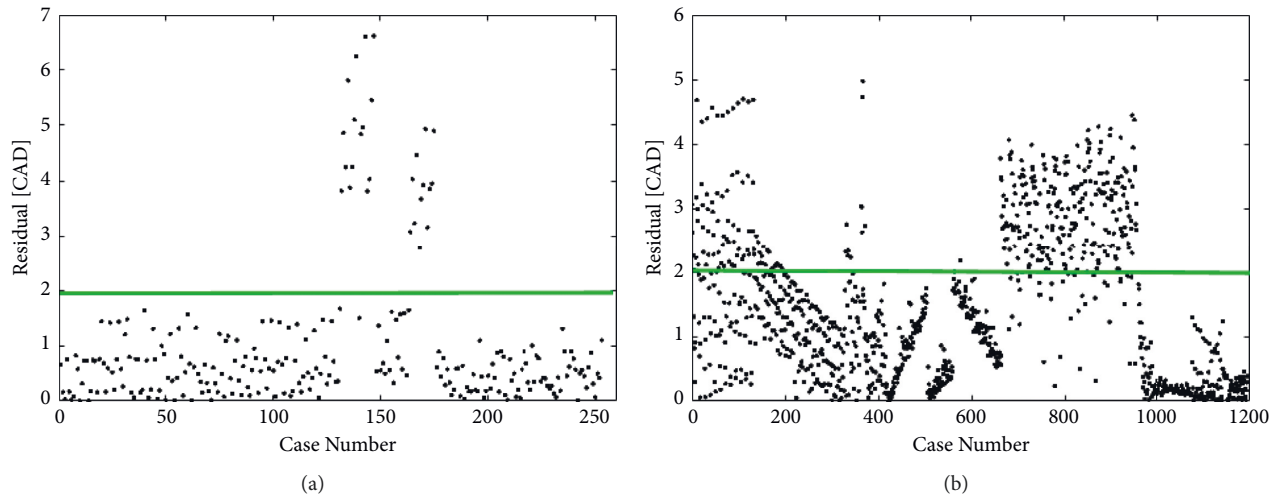


FIGURE 2: The SAM performance. (a) Test samples. (b) Train samples.

consists of 1177 sample train cases and 253 sample test cases which are adopted from the results of running the virtual engine for each engine at different operating conditions. Considered independent parameters are engine speed, equivalence ratio, EGR, charge pressure and temperature at IVC, RH, and real CR, and the dependent parameter is SOC. For the training dataset, the dispersion of SOC is between 10 CAD before Top Dead Center (TDC) and 10 CAD after TDC, as shown in Figure 1, however, it is mainly (more than 67%) occurred before TDC.

The performance of the SAM in SOC prediction for both the train and test samples is studied in the first step, as shown in Figure 2. The accepted range of SOC prediction in control applications is reported as the residual ($|SOC_{\text{real}} - SOC_{\text{model}}|$) with less than 2 CAD [26]. For the test samples, SAM predicted the SOC with 89.3% in the acceptable range as well as 70.6% for the train samples reported in Table 2. In addition, the root mean squared error (RMSE) is reported as 1.64 and 1.84 for the test and train samples, respectively.

TABLE 2: The performance of SAM.

	Train data		Test data	
Total cases	1177	100.00%	253	100.00%
Residual less than 2 CAD	831	70.60%	226	89.33%
Residual less than 3 CAD	1016	86.32%	227	89.72%
Residual less than 4 CAD	1152	97.88%	237	93.68%
Mean absolute error (MAE)	1.4279		0.9897	
Mean squared error (MSE)	3.3918		2.6933	
RMSE	1.8417		1.6411	

TABLE 3: The employed ML regression methods.

Num.	Name	Linear hypothesis
1	OLS	Yes
2	Lasso	Yes
3	Ridge	Yes
4	SVR with linear kernel	Yes
5	BR	Yes
6	Huber	Yes
7	Robust RANSAC	No
8	SVR with RBF kernel	No
9	Nearest neighbors with uniform weights	No
10	Decision tree	No
11	MLP neural network	No

TABLE 4: The coefficient vector of ML linear regression methods.

Model	N (rpm)	T (K)	P (kPa)	$\emptyset$	EGR (%)	RH (%)	Real CR	Intercept
OLS	$4.2962e-03$	$-2.493e-01$	$-4.0262e-02$	$2.1140e+01$	$3.0853e-02$	$-5.5125e-03$	-2.2814	499.1991
Lasso	0.0037	-0.2292	-0.0416	0.00	0.03259	-0.0061	-1.9549	491.9509
Ridge	$4.2417e-03$	$-2.4664e-01$	$-4.0385e-02$	$1.7200e+01$	$3.2108e-02$	$-5.7550e-03$	-2.2502	498.7426
SVR	$5.2693e-01$	$1.3572e+01$	1.7384	$1.3029e+03$	-3.1742	$8.9697e-01$	6.8226	-6227.37
BR	$4.2954e-03$	$-2.4926e-01$	$-4.0264e-02$	$2.1079e+01$	$3.0872e-02$	$-5.5163e-03$	-2.2809	499.1922
Huber	$-7.5840e-03$	$6.0737e-01$	$-2.3513e-01$	$1.7234e-02$	$-1.2149e-01$	$-1.2149e-01$	8.0966	0.6408

In the next step, both linear and nonlinear ML regression methods are applied to the training dataset and the results are illustrated in detail. The built-in Python code has been used for such an investigation, and the employed models are reported in Table 3. Achieved coefficient vector (vector of θ) for linear methods is presented in Table 4. The most effective parameter on SOC variations is reported the equivalence ratio by the OLS, Ridge, Support Vector Regression (SVR), and Bayesian Ridge (BR) models while the CR is considered with a more important role by Lasso and Huber models, as shown in Figure 3.

In Figure 3, the impact factors of independent parameters on SOC for the ML linear regressors; OLS, Lasso, ridge, SVR, BR, and Huber were shown. Looking more detail in Figure 3 and Table 4, and considering the out of range intercepts of SVR and Huber models, it seems, these models are not converged, and the learning is failed. In addition, the considered impact factor for equivalence ratio (0.00) in the Lasso model shows that the effect of equivalence ratio can be

ignored which is not acceptable in real applications. However, the main parameters of comparing these models are considered the acceptable residual (less than 2 CAD) percentage and the score of the model (the best score is equal to 1.00) in the test dataset. Table 5 reports the performance of defined models in Table 3 on the test dataset. Considering reported scores and the acceptable residual percentage, it can be concluded that among the nonlinear models, the Robust RANSAC model is greatly learned and has outstanding accuracy by predicting up to 98.4% in the acceptable range. The same noticeable performance is achieved by the OLS, ridge, and BR models from the linear models; however, the performance of the Lasso model is still acceptable by predicting up to 99.6% with less than 3 CAD errors. The scatter plots of achieved residuals by the models are shown in Figures 4 and 5 for the test and train datasets, respectively. Due to these plots, it can be concluded that the main reason for found errors in the Lasso model is approaching the learning process to overfitting. The same trend has existed

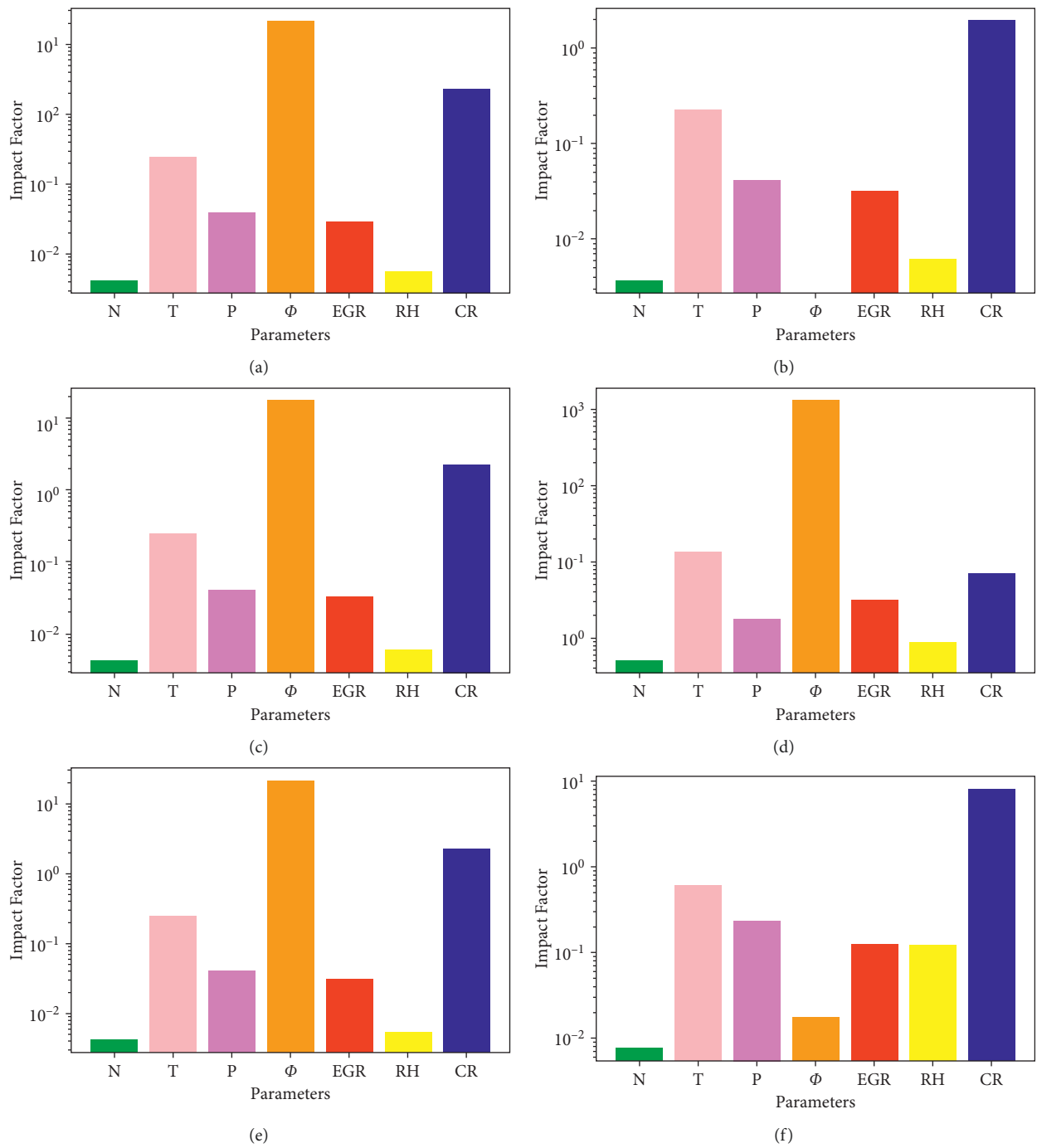


FIGURE 3: The impact factors of independent parameters on SOC for the ML linear regressors. (a) LOS. (b) Lasso. (c) Ridge. (d) SVR. (e) BR. (f) Huber.

TABLE 5: The performance of employed models on the test dataset.

Parameter	Model										
	1	2	3	4	5	6	7	8	9	10	11
Residual less than 2 CAD (%)	98.81423	92.49012	98.81423	0.395257	98.81423	9.486166	98.41897	26.08696	44.26877	77.0751	15.41502
Residual less than 3 CAD (%)	100	99.60474	100	0.790514	100	16.60079	100	34.78261	55.73123	89.72332	26.48221
Residual less than 4 CAD (%)	100	100	100	0.790514	100	25.29644	100	46.24506	73.12253	95.25692	33.99209
MAE	0.8233	0.9926	0.7878	484.0682	0.8226	7.8129	0.7766	4.3073	2.7197	1.3818	7.7372
MSE	0.9541	1.4513	0.8547	352987.6	0.9519	87.2833	0.8889	25.8072	11.5298	3.4353	92.1935
RMSE	0.9768	1.2047	0.9245	594.1276	0.9756	9.3425	0.9428	5.08	3.3955	1.8534	9.6017
Score	0.9039	0.8539	0.9139	-35526.2	0.9041	-7.7848	0.9105	-1.5974	-0.1604	0.6542	-8.279

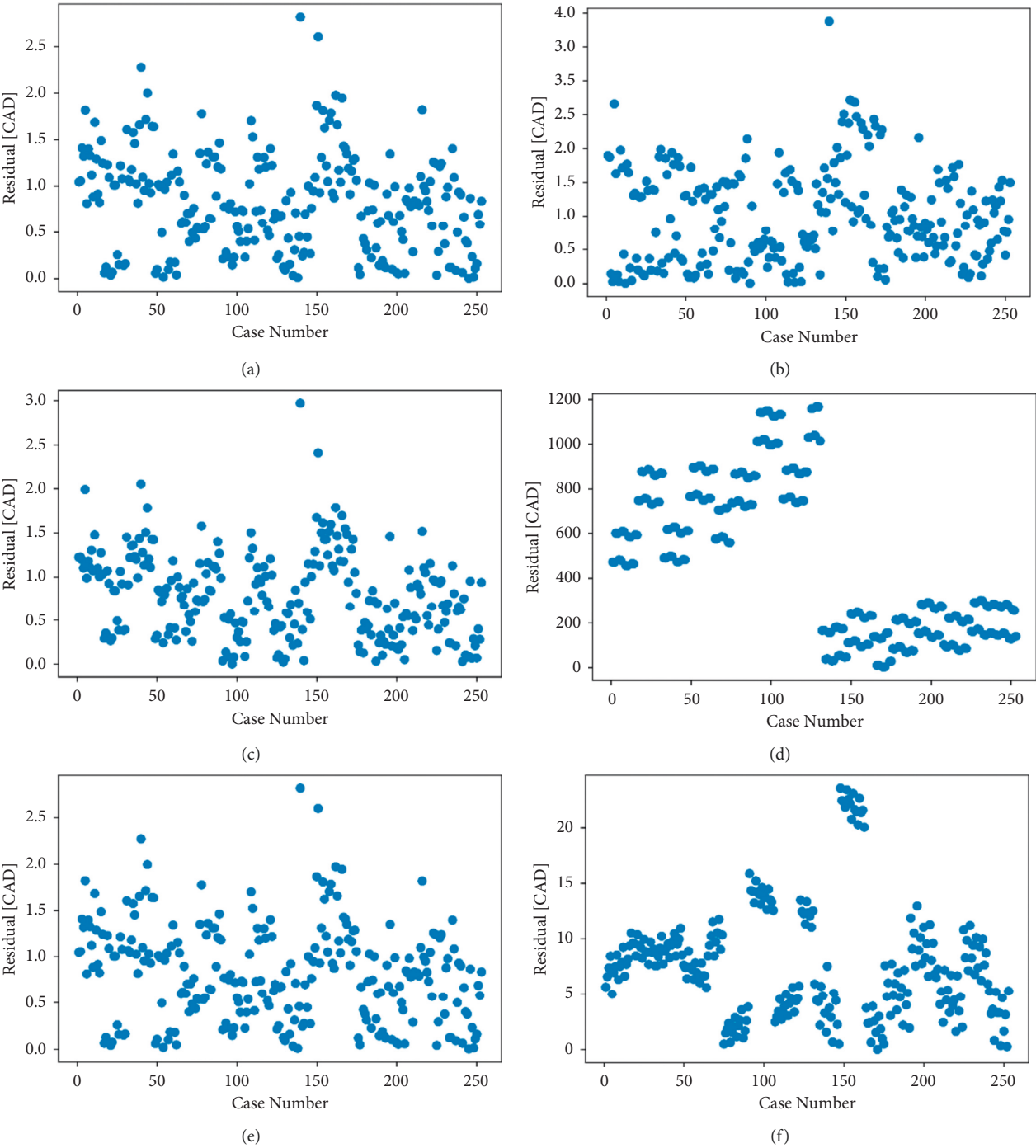


FIGURE 4: Continued.

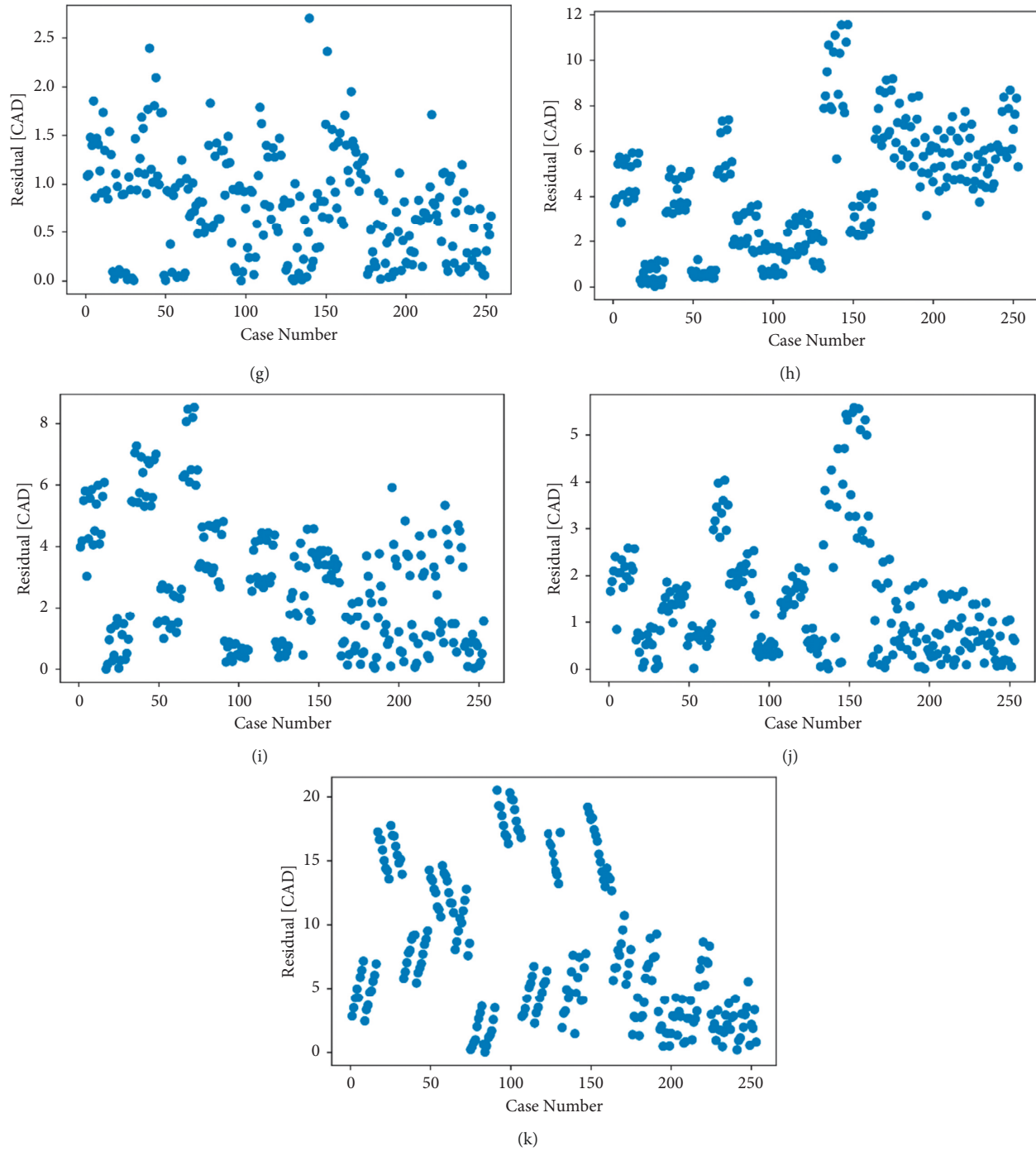


FIGURE 4: The residual of employed models in the test dataset. (a) LOS. (b) Lasso. (c) Ridge. (d) SVR_linear. (e) BR. (f) Huber. (g) Robust RANSAC. (h) SVR_rbf. (i) Nearest neighbors. (j) Decision tree. (k) MLP neural network.

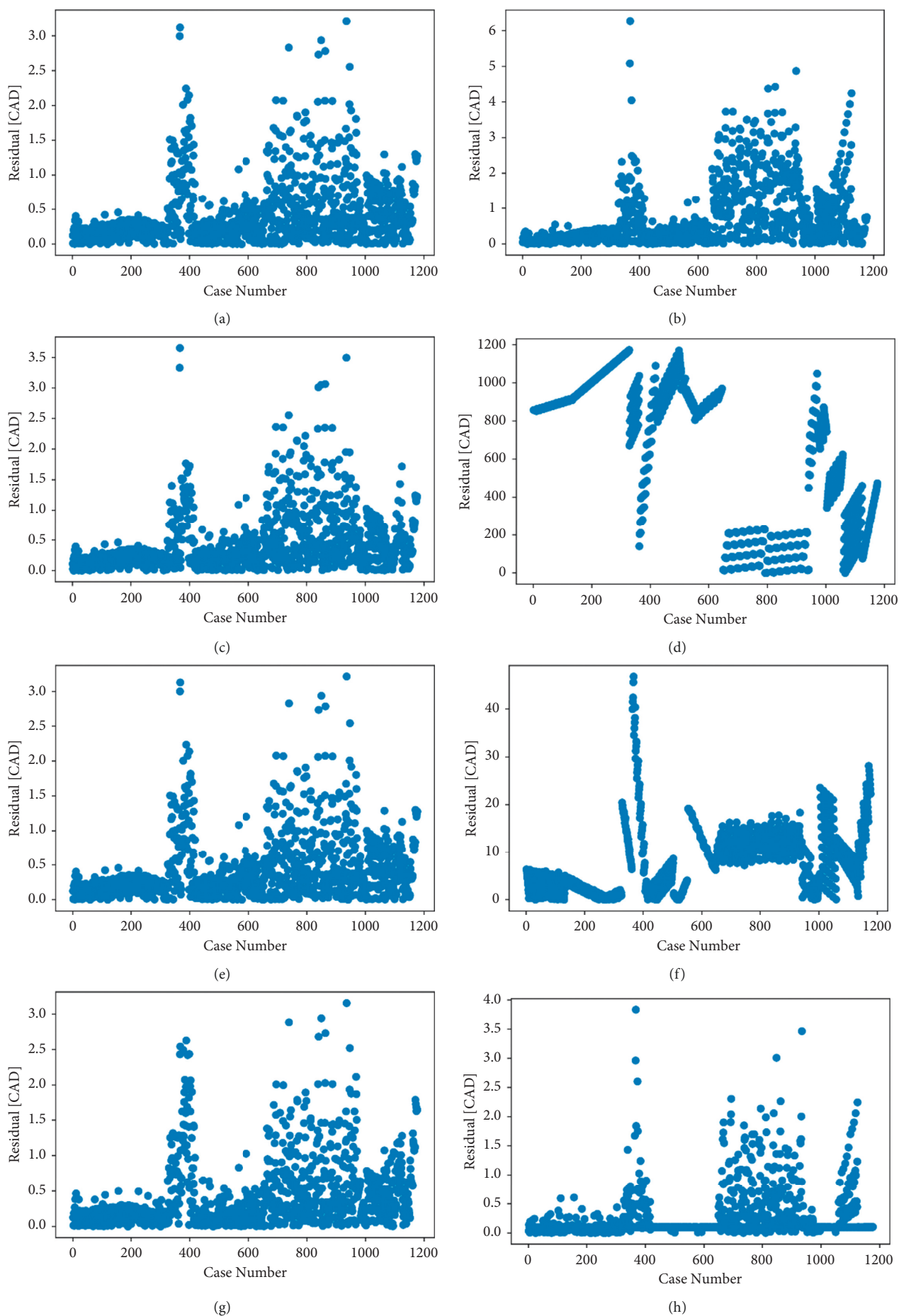


FIGURE 5: Continued.

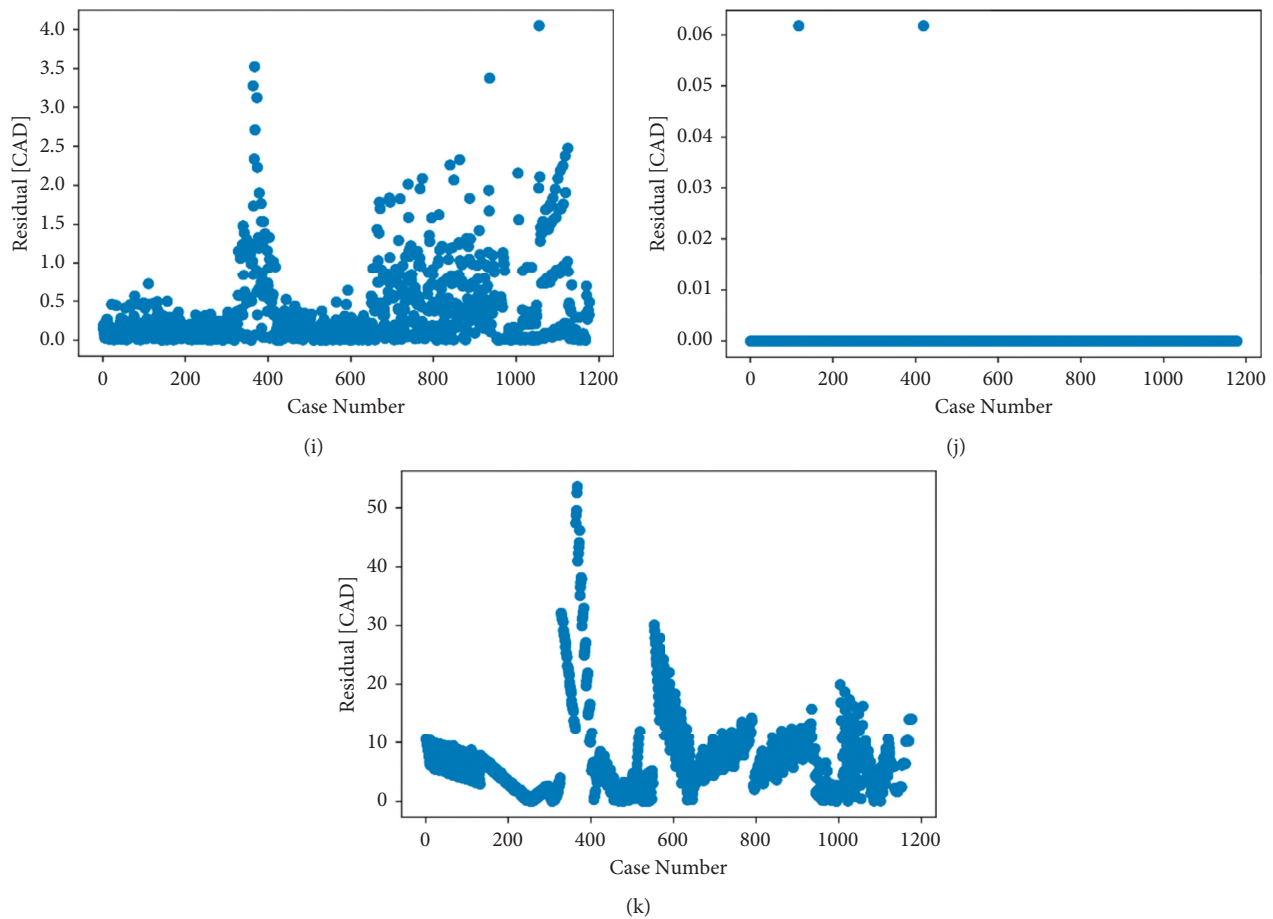


FIGURE 5: The residual of employed models in the train dataset. (a) LOS. (b) Lasso. (c) Ridge. (d) SVR_linear. (e) BR. (f) Huber. (g) Robust RANSAC. (h) SVR_rbf. (i) Nearest neighbors. (j) Decision tree. (k) MLP neural network.

for the nonlinear SVR and nearest neighbors models. Complete overfitting has occurred for the decision tree model, and the reason for the pretty weak performance of MPL neural network, Huber, and linear SVR models is the convergence issue.

4. Conclusion

In this work, 11 regression models based on machine learning techniques were employed to provide a fast-response and accurate model for the prediction of the start of combustion in homogeneous charge compression ignition engines fueled with methane. The dataset for such an investigation is adopted from validated virtual engines which used the detailed chemical kinetics of methane combustion. The main achievements of the work are listed in the following:

- (i) The robust RANSAC model as a nonlinear model enhanced the accuracy of SOC prediction to 98.4%.
- (ii) The linear models, namely, ordinary least squares, ridge, and Bayesian ridge models, presented

outstanding accuracy in SOC prediction (up to 98% correct predictions).

- (iii) For the linear hypothesis (the correlation for SOC prediction), the presented models have acceptable response time to be used in the real-time control applications as the engine ECU.

Abbreviations

0D:	0-dimensional
3D:	3-dimensional
BDC:	Bottom dead center
BR:	Bayesian ridge
CAD:	Crank angle degree
CFD:	Computational fluid dynamics
CI:	Compression ignition
CR:	Compression ratio
ECU:	Electronic control unit
EGR:	Exhaust gas recirculated
EVO:	Exhaust valve opening
HCCI:	Homogeneous charge compression ignition

IVC: Inlet valve closing
 LTC: Low temperature combustion
 MAE: Mean absolute error
 MKIM: Modified knock integral model
 ML: Machine Learning
 MSE: Mean squared error
 OLS: Ordinary least squares
 RCCI: Reactivity controlled compression ignition
 RH: Relative humidity
 RMSE: Root mean squared error
 SAM: Simple algebraic model
 SI: Spark Ignition
 SOC: Start of combustion
 SVR: Support vector regression
 TDC: Top dead center

N : Engine speed (rpm)

P : Pressure (kPa)

T : Temperature (K)

α : Learning rate

θ : Vector of coefficients

ϕ : Equivalence ratio.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

- [1] J. Moradi, A. Gharehghani, and M. Mirsalim, "Numerical comparison of combustion characteristics and cost between hydrogen, oxygen and their combinations addition on natural gas fueled hcci engine," *Energy Conversion and Management*, vol. 222, Article ID 113254, 2020.
- [2] H. Solmaz, "A comparative study on the usage of fusel oil and reference fuels in an hcci engine at different compression ratios," *Fuel*, vol. 273, Article ID 117775, 2020.
- [3] M. M. Namar and O. Jahanian, "Energy and exergy analysis of a hydrogen-fueled hcci engine," *Journal of Thermal Analysis and Calorimetry*, vol. 137, no. 1, pp. 205–215, 2019.
- [4] M. Rabeti, O. Jahanian, A. A. Ranjbar, S. M. Safieddin Ardebili, and H. Solmaz, "Potential of semi-empirical heat transfer models in predicting the effects of equivalence ratio on low temperature reaction and high temperature reaction heat release of an hcci engine," *Journal of Applied and Computational Mechanics*, 2021.
- [5] M. Namar and O. Jahanian, "Defining start and duration of combustion in hcci engines using mean-value method for control applications," *Amirkabir Journal of Mechanical Engineering*, vol. 50, no. 6, pp. 1265–1276, 2019.
- [6] M. Fathi, O. Jahanian, D. D. Ganji, S. Wang, and B. Somers, "Stand-alone single-and multi-zone modeling of direct injection homogeneous charge compression ignition (di-hcci) combustion engines," *Applied Thermal Engineering*, vol. 125, pp. 1181–1190, 2017.
- [7] A. P. Singh, V. Kumar, and A. K. Agarwal, "Evaluation of comparative engine combustion, performance and emission characteristics of low temperature combustion (pcci and rccl) modes," *Applied Energy*, vol. 278, Article ID 115644, 2020.
- [8] S. S. Motalebi Hasankola, R. Shafaghat, O. Jahanian, S. Talesh Amiri, and M. Shooghi, "Numerical investigation of the effects of inlet valve closing temperature and exhaust gas recirculation on the performance and emissions of an rccl engine," *Journal of Thermal Analysis and Calorimetry*, vol. 139, no. 4, pp. 2465–2474, 2020.
- [9] S. Pandiaraj, D. Subbaiyan, T. Ayyasamy, and S. Nagarajan, "Design, development and analysis of mullite catalytic converter for ci engines," *SAE Technical Paper*, 2019.
- [10] T. Agrawal, V. K. Banerjee, B. S. Sikarwar, and M. Bhandwal, "Optimizing the performance of catalytic converter using turbulence devices in the exhaust system," in *Advances in Interdisciplinary Engineering* Springer, Berlin, Germany, 2019.
- [11] M. M. Namar, O. Jahanian, R. Shafaghat, and K. Nikzadfar, "Numerical/experimental study on downsized iranian national engine (ef7) performance at low engine speeds," *International Journal of Engineering*, vol. 34, no. 9, pp. 2137–2147, 2021.
- [12] M. M. Namar, O. Jahanian, R. Shafaghat, and K. Nikzadfar, "Feasibility study for downsizing ef7 engine, numerical and experimental approach," *The Journal of Engine Research*, vol. 61, no. 61, pp. 73–85, 2021.
- [13] X. Duan, M.-C. Lai, M. Jansons, G. Guo, and J. Liu, "A review of controlling strategies of the ignition timing and combustion phase in homogeneous charge compression ignition (hcci) engine," *Fuel*, vol. 285, Article ID 119142, 2021.
- [14] S. Khandal, N. Banapurmath, and V. Gaitonde, "Performance studies on homogeneous charge compression ignition (hcci) engine powered with alternative fuels," *Renewable Energy*, vol. 132, pp. 683–693, 2019.
- [15] M. Taghavi, A. Gharehghani, F. B. Nejad, and M. Mirsalim, "Developing a model to predict the start of combustion in hcci engine using ann-ga approach," *Energy Conversion and Management*, vol. 195, pp. 57–69, 2019.
- [16] Z. Wang, G. Du, Z. Li, X. Wang, and D. Wang, "Study on the combustion characteristics of a high compression ratio hcci engine fueled with natural gas," *Fuel*, vol. 255, Article ID 115701, 2019.
- [17] H. Ezoji and S. S. M. Ajarostaghi, "Thermodynamic-cfd analysis of waste heat recovery from homogeneous charge compression ignition (hcci) engine by recuperative organic rankine cycle (rorc): effect of operational parameters," *Energy*, vol. 205, Article ID 117989, 2020.
- [18] J. Moradi, A. Gharehghani, and M. Mirsalim, "Numerical investigation on the effect of oxygen in combustion characteristics and to extend low load operating range of a natural-gas hcci engine," *Applied Energy*, vol. 276, Article ID 115516, 2020.
- [19] G. Du, Z. Wang, D. Wang, X. Wang, and X. Fu, "Study on the effect of water addition on combustion characteristics of a hcci engine fueled with natural gas," *Fuel*, vol. 270, Article ID 117547, 2020.
- [20] M. Fathi, O. Jahanian, and M. Shahbakhti, "Modeling and controller design architecture for cycle-by-cycle combustion control of homogeneous charge compression ignition (hcci) engines—a comprehensive review," *Energy Conversion and Management*, vol. 139, pp. 1–19, 2017.
- [21] X. Han, P. Divekar, M. Zheng, and J. Tjong, "Empirical investigation and control-oriented modeling of n-butanol hcci for improving combustion controllability," *Fuel*, vol. 280, Article ID 118551, 2020.

- [22] M. Nazoktabar, S. A. Jazayeri, M. Parsa, D. D. Ganji, and K. Arshatabar, "Controlling the optimal combustion phasing in an hcci engine based on load demand and minimum emissions," *Energy*, vol. 182, pp. 82–92, 2019.
- [23] M. Shahbakhti and C. R. Koch, "Dynamic modeling of hcci combustion timing in transient fueling operation," *SAE International Journal of Engines*, vol. 2, no. 1, 2009.
- [24] K. Swan, M. Shahbakhti, and C. R. Koch, "Predicting start of combustion using a modified knock integral method for an hcci engine," *SAE Transactions*, pp. 611–620, 2006.
- [25] M. Shahbakhti, R. Lupul, and C. R. Koch, "Predicting hcci auto-ignition timing by extending a modified knock-integral method," *SAE Technical Paper*, 2007.
- [26] M. M. Namar and O. Jahanian, "A simple algebraic model for predicting hcci auto-ignition timing according to control oriented models requirements," *Energy Conversion and Management*, vol. 154, pp. 38–45, 2017.
- [27] C. Mishra and P. M. V. Subbarao, "Machine learning integration with combustion physics to develop a composite predictive model for reactivity controlled compression ignition engine," *Journal of Energy Resources Technology*, vol. 144, no. 4, Article ID 042302, 2021.
- [28] C. Mishra and P. Subbarao, "Design, development and testing a hybrid control model for rcci engine using double wiebe function and random forest machine learning," *Control Engineering Practice*, vol. 113, Article ID 104857, 2021.
- [29] Z. H. Zheng, X. Lin, M. Yang et al., "Progress in the application of machine learning in combustion studies," *ES Energy & Environment*, vol. 9, pp. 1–14, 2020.
- [30] G. P. Smith, D. M. Golden, M. Frenklach et al., *Gri 3.0 mechanism*, Gas Research Institute, Chicago, IL, USA, 1999, <http://combustion.berkeley.edu/gri-mech/version30/text30.html>.
- [31] M. Alizamir, S. Kim, O. Kisi, and M. Zounemat-Kermani, "A comparative study of several machine learning based non-linear regression methods in estimating solar radiation: case studies of the USA and Turkey regions," *Energy*, vol. 197, Article ID 117239, 2020.
- [32] A. Almaghrebi, F. Aljuheshi, M. Rafaie, K. James, and M. Alahmad, "Data-driven charging demand prediction at public charging stations using supervised machine learning regression methods," *Energies*, vol. 13, no. 16, 2020.
- [33] R. Li, J. M. Herreros, A. Tsolakis, and W. Yang, "Machine learning regression based group contribution method for cetane and octane numbers prediction of pure fuel compounds and mixtures," *Fuel*, vol. 280, Article ID 118589, 2020.
- [34] G. Hackeling, *Mastering Machine Learning with Scikit-Learn*, Packt Publishing Ltd, Birmingham, United Kingdom, 2017.
- [35] M. Ishida, S. Jung, H. Ueki, and D. Sakaguchi, *Combustion Characteristics of Hcci Engines Fuelled with Natural Gas and Dme*, 2007.
- [36] S. B. Fiveland and D. N. Assanis, "Development and validation of a quasi-dimensional model for hcci engine performance and emissions studies under turbocharged conditions," *SAE Transactions*, pp. 842–860, 2002.
- [37] S. B. Fiveland, R. Agama, M. Christensen et al., "Experimental and simulated results detailing the sensitivity of natural gas hcci engines to fuel composition," *SAE Transactions*, pp. 2123–2134, 2001.