



## City Research Online

### City, University of London Institutional Repository

---

**Citation:** Hwang, J., Lee, P., Mun, S., Karathanassis, I. K., Koukouvinis, P., Pickett, L. & Gavaises, M. (2021). Machine-learning enabled prediction of 3D spray under engine combustion network spray G conditions. *Fuel*, 293, 120444. doi: 10.1016/j.fuel.2021.120444

This is the published version of the paper.

This version of the publication may differ from the final published version.

---

**Permanent repository link:** <https://openaccess.city.ac.uk/id/eprint/25790/>

**Link to published version:** <https://doi.org/10.1016/j.fuel.2021.120444>

**Copyright:** City Research Online aims to make research outputs of City, University of London available to a wider audience. Copyright and Moral Rights remain with the author(s) and/or copyright holders. URLs from City Research Online may be freely distributed and linked to.

**Reuse:** Copies of full items can be used for personal research or study, educational, or not-for-profit purposes without prior permission or charge. Provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way.

---

---

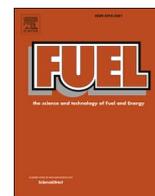
---

City Research Online:

<http://openaccess.city.ac.uk/>

[publications@city.ac.uk](mailto:publications@city.ac.uk)

---



## Full Length Article

# Machine-learning enabled prediction of 3D spray under engine combustion network spray G conditions



Joonsik Hwang<sup>a,b,\*</sup>, Philku Lee<sup>b</sup>, Sungkwang Mun<sup>b</sup>, Ioannis K. Karathanassis<sup>c</sup>,  
Phoevos Koukouvinis<sup>c</sup>, Lyle M. Pickett<sup>d</sup>, Manolis Gavaises<sup>c</sup>

<sup>a</sup> Department of Mechanical Engineering, Mississippi State University, Mississippi State, MS 39762, USA

<sup>b</sup> Center for Advanced Vehicular Systems (CAVS), Mississippi State University, Starkville, MS 39759, USA

<sup>c</sup> School of Mathematics, Computer Science and Engineering, City, University of London, UK

<sup>d</sup> Combustion Research Facility, Sandia National Laboratories, Livermore, CA 94550, USA

## ARTICLE INFO

## Keywords:

Machine-learning  
Linear regression  
Engine Combustion Network (ECN)  
Spray G  
Flash-boiling  
Projected liquid volume  
Tomographic reconstruction

## ABSTRACT

Spray and air–fuel mixing in gasoline direct-injection (GDI) engines play a crucial role in combustion and emission characteristics. While a variety of phenomenological spray models and computational fluid dynamics (CFD) simulations have been applied to identify air–fuel mixture distribution, most research efforts so far were concentrated on single axial-nozzle injectors and limited range of ambient conditions. Especially, the prediction of flash-boiling sprays in multi-hole injectors remains a great challenge due to the lack of understanding of the complicated two-phase flow dynamics. For the specific conditions, the question can arise concerning the capability of machine-learning algorithms to predict complex flash-boiling sprays. We developed a machine-learning algorithm, as a simple variant of linear regression, that is capable of predicting the spray 3D topology for various fuels and ambient conditions. A series of spray experiments were carried out in a constant-flow spray vessel coupled with high-speed diffused back-illumination extinction imaging to produce a data set for algorithm training. Nine different test fuels, including single component iso-octane (ic8) and multi-component EEE gasoline, that cover a wide range of fuel properties were injected using Engine Combustion Network (ECN) Spray G injector under ECN G2 (50 kPa absolute), G3 (100 kPa absolute), and G3HT (G3 with 393 K ambient temperature) conditions. Among the test fuels, ic8ib2 (ic8 80%, iso-butanol 20% v/v) and EEE gasoline were specified as target fuels for spray prediction by the machine-learning algorithm, thus they were not included in the training data. The macroscopic spray analysis based on projected liquid volume (PLV) and computed tomographic (CT) reconstruction showed that the spray prediction by the machine-learning algorithm showed excellent agreement with true values from the experimental data. The maximum differences in liquid penetration for ic8ib2 and EEE fuel were 3.6 mm (7.3% error) and 1.3 mm (2.32% error), respectively. The 3D spray predicted had a consistent trend to experimental data showing slight plume movement for ic8ib2 but complete spray collapsing for EEE gasoline fuel. The plume direction angle enabled by the CT data showed differences up to 2° compared to true values during the injection period. The quantitative validation results showed that the machine-learning algorithm is capable of predicting spray performance with nine input features (fuel properties and ambient conditions), and is actually superior to CFD performance for these same number of spray parameters.

**Abbreviations:** ANN, Artificial neural network; CFD, Computational fluid dynamics; CT, Computed tomographic; DI, Direct-injection; ECN, Engine Combustion Network; EGR, Exhaust gas recirculation; FGM, Flamelet generated manifolds; GDI, Gasoline direct-injection; LED, Light-emitting diode; LES, Large eddy simulation; LTGC, Low-temperature gasoline combustion; MSE, Mean squared error; PDEs, Partial differential equations; PDI, Phase-doppler interferometry; PFI, Port fuel injection; PINN, Physics-informed neural networks; PIV, Particle image velocimetry; PLV, Projected liquid volume; PM, Particulate matter; SGD, Stochastic gradient descent; SMD, Sauter mean diameter; SPCCI, Spark-controlled compression ignition.

\* Corresponding author at: Mississippi State University, USA.

E-mail address: [jhwang@cavs.msstate.edu](mailto:jhwang@cavs.msstate.edu) (J. Hwang).

<https://doi.org/10.1016/j.fuel.2021.120444>

Received 21 December 2020; Received in revised form 3 February 2021; Accepted 7 February 2021

Available online 22 February 2021

0016-2361/© 2021 The Author(s). Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license

(<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

### Nomenclature

$C_{\text{ext}}$	Extinction coefficient [ $\text{mm}^2$ ]
$d$	Droplet diameter [mm]
$I$	Transmitted light intensity [a.u.]
$I_0$	Incident light intensity [a.u.]
$\hat{I}_{i,j}$	PLV at $(i,j)$ [ $\text{mm}^3(\text{liquid})/\text{mm}^2$ ]
$P_a$	Ambient pressure [Pa]
$P_v$	Vapor pressure [Pa]
PLV	Projected liquid volume [ $\text{mm}^3(\text{liquid})/\text{mm}^2$ ]
$T$	Distillation temperature [ $^{\circ}\text{C}$ ]
$\tau$	Optical thickness [a.u.]
$w_{i,j}$	Regression weight vector at $(i,j)$ [a.u.]
$x_k$	Feature vector for training [a.u.]
$\hat{y}_{i,j}$	PLV vector at $(i,j)$ for training [a.u.]

## 1. Introduction

Increasing environmental concerns and fossil fuel costs have led the automobile industry to develop highly efficient and clean internal combustion engines. Amongst advanced engine technologies, high-pressure direct-injection (DI) engines have been widely adopted by the automotive industry because of lower knocking tendency and higher engine efficiency than port fuel injection (PFI). Reasons for these improvements include using DI fueling include charge cooling by fuel evaporation and high level of turbulence in the combustion chamber [1,2]. The application of DI strategy in engines is developing with other technologies such as multiple injections, exhaust gas recirculation (EGR), and enhanced air utilization by swirl or tumble motion [3–5]. Advanced gasoline direct-injection (GDI) engines encompassing low-temperature gasoline combustion (LTGC) or spark-plug controlled compression ignition (SPCCI) concepts are using multiple fuel injections (early intake stroke and top dead center) with high swirl and EGR ratios [6,7]. In these engines, inappropriate injection parameters or undesired fuel properties cause fuel impingement on the piston top or cylinder liner leading to a high amount of particulate matter (PM) emissions [8,9]. Therefore, understanding of plume direction and liquid penetration length under various ambient conditions and fuels is essential to improve the combustion and emission characteristics. Ideally, these processes could be predicted for all operating conditions to assist in the design of the engine or fuel injector.

Spray and air–fuel mixing processes under engine-like conditions are complicated two-phase flow phenomena involving high turbulence, small temporal/spatial scale, droplet breakup, and evaporation. Liquid fuel from the injector nozzle experiences throttling, breakup, atomization, evaporation, and mixing with ambient air [10]. The physics behind the spray formation process are highly dependent on thermodynamics conditions and fuel properties. For instance, spray breakup and mixing processes under critical temperature is very different compared to that of atmospheric or flash-boiling conditions [11,12]. Spray characteristics can be classified as macroscopic, i.e. liquid/vapor penetration and width, and microscopic, i.e. breakup process and droplet size. Many of the previous studies attempted to develop phenomenological models to predict spray characteristics for single nozzles. With reference to macroscopic aspects, Hiroyasu and Arai proposed semi-empirical correlations for spray breakup and liquid penetration length [13], which was derived based on experimental data and jet disintegration theory from a previous study of Levich [14], and is capable of predicting liquid penetration before/after the breakup time. Based on the model, the liquid penetration length increases linearly with time before the breakup time, while it has a square root dependency on time after the breakup. This correlation has been widely used for computational fluid dynamics

(CFD) simulations. Siebers also developed a model for prediction of liquid penetration length under a wide range of ambient density and temperature conditions with the hypothesis that maximum liquid penetration occurs at a point where vaporized fuel is at a saturated condition in thermodynamic equilibrium with the ambient gas [15]. On the other hand, for microscopic spray characteristics, models for atomization and droplet behaviors have been developed based on nondimensional parameters such as Reynolds ( $Re$ ), Ohnesorge ( $Oh$ ), and Weber ( $We$ ) numbers [16]. Two types of breakup mechanisms in sprays have been reported: the bag breakup for low  $We$  numbers and the shear breakup for high  $We$  conditions. Bag breakup is typically associated with Kelvin Helmholtz (KH) instability, while the shear breakup is related to Rayleigh Taylor (RT) instability [17]. For fuel sprays in engine applications, the KH model is used for intact liquid-core breakup (a primary breakup) length prediction and the RT model is utilized for a secondary breakup after the liquid-core length [18]. Apart from these models, there are several different numerical models to predict droplet deformation, droplet collisions, droplet diameter, and modified forms of KH-RT models [19–21].

Spray models incorporating the physics discussed above are desired to predict the liquid distribution for engine design. The models may be phenomenological, usually for single nozzles without complicated gas flow, or they may incorporate CFD simulations, typically using Lagrangian parcels. The models were also based on conventional spray formation and combustion conditions that correspond to high ambient temperature and pressure. This aspect is becoming more critical under flash-boiling conditions where plume-to-plume interactions become more significant and can lead to full spray collapse into a single plume. The current state of Lagrangian parcel simulation cannot yet adequately capture plume interaction and liquid distribution [22,23]. There are specified spray models called homogeneous relaxation model (HRM) for the Eulerian approach, however, the CFD simulation results still remain unsatisfactory [24–26].

On the contrary to the model-based microscopic and macroscopic approaches, there is a recent trend to apply artificial neural network (ANN) and machine-learning (ML) algorithms in fuels and spray combustion research [27–30]. Advances in experimental and numerical analysis, such as high-speed imaging and high-performance computing, are constantly accelerating the massive production of data across all fields. The analysis of those data through ANN and ML is offering novel breakthroughs in a wide variety of disciplines. As one promising solution to overcome the limitations of the model-based spray research, ANN and ML have been utilized to predict spray characteristics. In the work of Ikeda and Mazurkiewicz, they utilized ANN techniques to predict droplet diameter in a spray burner system [28]. The experimental measurements on droplet diameter and velocity were performed using phase Doppler anemometry (PDA) and particle image velocimetry (PIV). The model was trained to predict droplet size at different axial and radial locations with PIV data such as droplet travel time, droplet arrival time, radial velocity, and axial velocity. The results from the machine-learning showed reasonably good agreement in capturing the trend of droplet diameter depending on the relevant location in the spray region; however, it failed to match the quantitative profile by exhibiting discrepancies up to 10  $\mu\text{m}$  error in droplet diameter. Zhang et al. applied ANN and flamelet generated manifolds (FGM) to predict ignition delay time and lift-off length of a diesel spray flame using an ECN Spray H injector [29]. In this study, the ANN library PyTorch was linked to the CFD library OpenFOAM and the validation was conducted based on large eddy simulation (LES) results. The predicted spray combustion using ANN showed good agreement with other simulation results in terms of spray breakup length, evaporation rate, and mass fraction of chemical species, despite using eight times less memory. Meanwhile, combining model and data-driven analyses, Raissi et al. introduced physics-informed neural networks (PINN), which were trained to solve machine-learning tasks while reflecting any given laws of physics described by partial differential equations (PDEs) [27]. The proposed framework

demonstrated the capability of solving classical problems in fluids, quantum mechanics, reaction–diffusion systems, and the propagation of nonlinear pressure waves.

While machine-learning approaches have been applied for spray combustion prediction on a limited basis, the current state of the art is for simplified cases with single-hole injectors or are with limited fuels or operating conditions that cover little of the operating envelope for GDI engine. In particular, injection from multi-hole nozzles often occurs where heated multi-component fuel is injected into a low-pressure gas to create flash-boiling sprays. Therefore, in this study, we aim to establish a framework to predict detailed quantitative information for flash-boiling and non-flashing multi-hole sprays using a machine-learning algorithm. We develop and apply a pixel regression model for time-resolved spray data as well as a variety of time-dependent spray characteristics without any time-marching simulations. For the training, experimental data using seven different fuels (single component *iso*-octane to multi-component EEE gasoline fuel) under ECN G2, G3, and G3HT conditions were utilized. Input features were composed of nine parameters including fuel and ambient properties. Spray profiles were predicted in three different viewing angles by the machine-learning algorithm and the data were fed into a separate 3D computed tomographic (CT) algorithm to generate quantitative spray information in a three-dimensional domain. To the best of the authors' knowledge, this is the first study in the literature to introduce a machine-learning based methodology suitable for predicting and quantifying the 3D spray topology of a multi-plume GDI injector under realistic operating conditions.

## 2. Test procedure and condition

### 2.1. Injector and fuels

An ECN Spray G injector (AV67-028) that has axisymmetrically drilled eight nozzles was utilized for fuel injection. Detailed geometric dimensions are shown in Fig. 1. The inner orifice and counterbore were designed to have diameters of 165 μm and 388 μm, respectively, however, recent x-ray measurement by Duke et al., indicated actual diameters of 173 μm and 394 μm [31,32]. The nozzles were drilled with an angle of 37° from the device center axis. It is noted that the plume direction angle can deviate from the nozzle drill angle due to in-nozzle two-phase phenomena and the plume dynamics under different

ambient conditions. The injector has specifically assigned *x*, *y*, and *z* axes as shown in the inset of the figure so that injector orientation coincides with ECN guidelines, in order to derive spray macroscopic features [33].

A major objective of this study is to develop a tool to predict spray behavior for various fuels, including renewable alternatives. Nine different fuels with significantly different fuel properties were included in the test matrix as are shown in Table 1. The test fuels were to maintain a high octane number (for high engine efficiency) and to have a wide range of fuel properties, especially with regards to vapor pressure and distillation temperature. The fuels include: single component *iso*-octane (ic8), multi-component surrogate di-isobutylene (1-hexene 4%, n-heptane 12.1%, *iso*-octane 44.2%, toluene 20.1%, and di-isobutylene 19.6% by volume), multi-component fuel with olefin molecular structure, e30 blend (gasoline 70%, ethanol 30% by volume), three-component e00 (ic8 46%, n-pentane 36%, n-undecane 18% by volume), high-cycloalkane, alkylate, two-component ic8ib2 (ib8 80%, *iso*-butanol 20% by volume), and EEE certification gasoline. The influence of compositional difference on distillation curves is presented in Fig. 2 [34]. The contrast in distillation profiles and boiling temperature is illustrated, for instance, with ic8 as a single component fuel with a boiling temperature of 98 °C in comparison to e30 blends that have not only light but also heavy components contributing to a wide spectrum of distillation temperatures from 34.3 °C to 204.4 °C. The large amount of light species in fuels attribute to the high vapor pressure as confirmed in Table 1. In order to obtain a format that can be easily integrated in the machine-learning algorithm, the distillation curve was processed to have a single value by summing up the differences in distillation temperature between a test fuel and ic8 (reference fuel) from distilled volumes of 0% to 100% with a 20% step size as presented in the following equation.

$$\text{Distillation parameter} = \frac{\sum_{i=0}^5 |T_{\text{testfuel@20-i}} - T_{\text{ic8@20-i}}|}{318.5} \quad (1)$$

Here, *T* is distillation temperature at a specific distilled volume, and 318.5 is an arbitrary constant to normalize the distillation parameter. Fig. 2(b) shows an example of EEE fuel that has a wide range of distillation temperatures. The distillation parameters are also reported in Table 1. In addition to the distillation parameters, fuel density [kg/m<sup>3</sup>], viscosity [mm<sup>2</sup>/s], vapor pressure [kPa], ambient to vapor pressure ratio (*P<sub>a</sub>/P<sub>v</sub>*) [n.a.], heat of vaporization [kJ/kg], ambient temperature

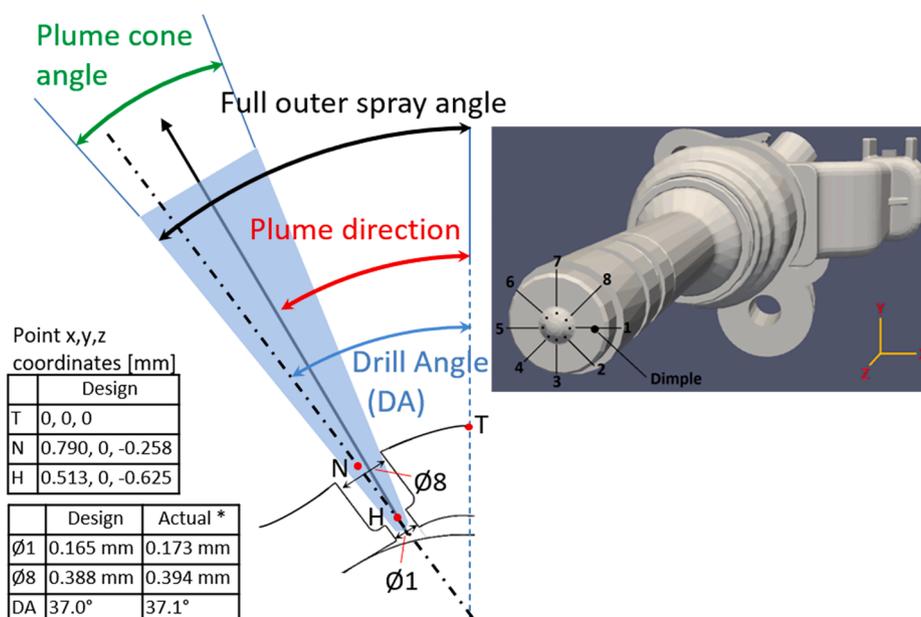
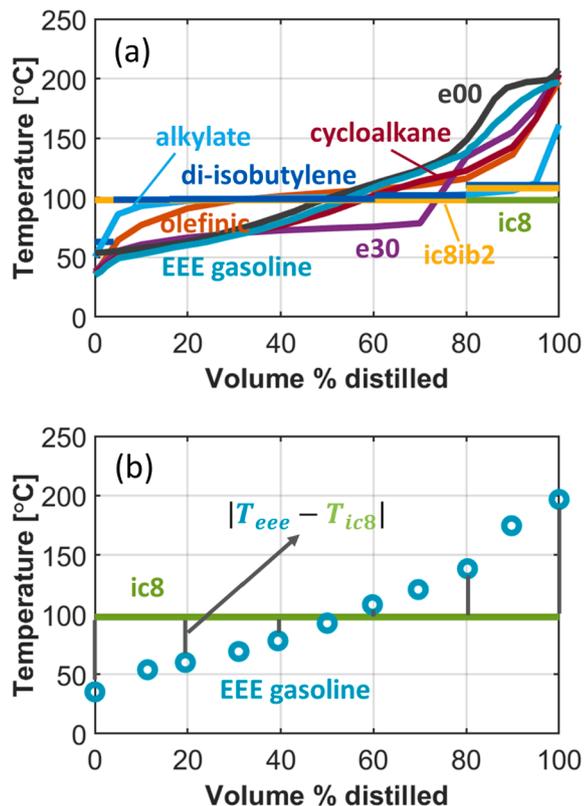


Fig. 1. Cut plane image of ECN spray G injector with dimensions (inset figure shows the 3-D rendering of the injector with primary orientation) [31,32].

**Table 1**  
Fuel properties [34].

Parameters	ic8	di-isobutylene	olefinic	e30	e00	alkylate	cycloalkane	ic8ib2	EEE gasoline
Density @15 °C [kg/m <sup>3</sup> ]	698.7	736.2	722.9	752.7	674.2	686.8	755.5	719.4	744
Viscosity ( $\nu$ ) @40 °C [mm <sup>2</sup> /s]	0.574	0.541	0.477	0.695	0.493	0.580	0.430	0.859	0.429
Vapor pressure @90 °C [kPa]	70.9	74.2	170.6	286.8	246	128.7	237.6	64.28	287
Heat of vaporization [kJ/kg]	271	295	337	565	339	309	393	333	349
Distillation parameter [a.u.]	0	0.171	0.626	0.883	0.848	0.398	0.792	0.079	0.848



**Fig. 2.** (a) Distillation curves of tested fuels and (b) example of distillation-parameter calculation [34].

[K], ambient density [kg/m<sup>3</sup>], and ambient pressure [kPa], which are the actual measurement data from [34], were set as input features for machine-learning algorithm.

## 2.2. Experimental setup

A high-speed imaging campaign was conducted in a constant flow vessel at Combustion Research Facility, Sandia National Laboratories. Constant nitrogen flow in the vessel enabled fast scavenging of residual air–fuel mixture so statistically converged data from 300 injections could be obtained with an injection frequency of 0.5 Hz. Two optical windows made of quartz were installed in parallel to allow for the optical access necessary for high-speed extinction imaging. The test injector was mounted at a port placed in the mid-distance of the windows and its temperature was regulated at 90 °C by circulating hot water in the injector jacket. Fuel spray was expelled by the injector into the ambient of constant pressure having nitrogen flow rate regulated by an electro-pneumatic flow controller. Low pressure flash-boiling condition was achieved by a custom-built vacuum pump system. The vessel pressure was measured at three different points by pressure transducers. The nitrogen flow entered the vessel through its lower side and flowed through a heating coil located at the bottom of the spray vessel surrounded by an insulator piece. Afterward, the flow passed through a

diffuser to enhance uniformity in the velocity and temperature field in the spray region. Using 24 thermocouples from the heater to the upper region of the spray, the temperature distribution was monitored during the experiment. The temperature at the target spray region was well controlled within 1 °C in the entire experiment.

The liquid spray was identified by the diffused back illumination extinction imaging. A high-speed green light-emitting diode (LED), Fresnel lens (150 mm,  $f = 150$  mm), engineered diffuser (20°), and bandpass filter (center wavelength: 527 nm, bandwidth: 20 nm, full width-half max: 22 nm) were utilized. A high-speed digital video camera (Photron, SA-Z) equipped with a prime lens (Nikkor, 50 mm f/1.8) was used to capture images of spray development in the vessel. The green LED was operated with a 24 ns command signal ( $\sim 220$  ns LED flash time) duration to freeze the spray in the visualized frame. The imaging was performed at a shutter speed of 67,200 frames per second (fps) with an image resolution of 512 by 512. The aperture of the lens and exposure time of the high-speed camera was set to 2.8 and 13.27  $\mu$ s, respectively. The engineered diffuser supplied a homogeneous light field and suppressed beam steering by evaporation or temperature field in the vessel [35]. This imaging technique is designed to collect extinction only by the fuel in its liquid phase, yet not from the respective vapor. Normalized by incident light intensity (and other optical parameters), the side-view extinction imaging can become quantitative for projected liquid volume (PLV) as explained in the next section. Further details of the spray vessel and optical setup can be found in [36].

## 2.3. Image processing method

### 2.3.1. Projected liquid volume (PLV) measurement

Extinction imaging is recommended by the ECN community for spray characterization because it can provide more quantitative information for liquid fuel concentration than conventional Mie-scattering imaging associated with lighting and scattering uncertainties [37]. Using the measured optical thickness, droplet size, and extinction coefficient, the projected liquid volume (PLV) along a line of sight can be derived for direct comparison with CFD results. The optical thickness in a spray region can be calculated based on Beer-Lambert law as follows:

$$\tau = -\ln(I/I_0) \quad (2)$$

where  $I$  is transmitted attenuated light intensity due to interaction with the liquid spray, and  $I_0$  is incident light intensity without any extinction [38]. This level of transmission intensity is reasonable for detection of the spray outline above the noise floor of the camera, but the vapor-phase beam steering needs to be considered and accounted for using engineered diffusers [35]. The measured optical thickness  $\tau$  is correlated to the PLV, which is the integral of liquid volume fraction (LVF) along the cross-stream direction  $y$ , as follows:

$$PLV = \tau \frac{\pi d^3 / 6}{C_{ext}} = \int_{-y_{\infty}}^{y_{\infty}} (LVF) \cdot dy \quad (3)$$

Mie scattering and extinction theories were applied in eq-3, along with assumptions that droplet diameter  $d$  and extinction coefficient  $C_{ext}$  (which depends upon  $d$ ) do not vary along the line of sight [37,38]. The PLV indicates how much liquid volume in a certain projected area, so it has a unit of mm<sup>3</sup>(liquid)/mm<sup>2</sup>. The PLV can easily be calculated from CFD simulations for direct comparison to experimental results.

However, the experimentalists must evaluate parameters such as  $d$  and  $C_{ext}$  to estimate PLV. In particular,  $C_{ext}$  is a function of droplet size, wavelength of light, and collection angle of the receiving optics. Fortunately, droplet diameter measurements have been performed using Spray G injector by General Motors and Shanghai Jiao Tong University by phase-doppler interferometry (PDI) [39]. The measurements show a Sauter mean diameter (SMD) near  $7 \mu\text{m}$  with fair uniformity across the plume during injection, and the  $C_{ext}$  was calculated as  $72.70 \cdot 10^{-6} \text{mm}^2$  with a droplet diameter of  $7 \mu\text{m}$  using MiePlot available at [40]. Fig. 3 shows an example of PLV at 0.45 ms after the onset of injection using ic8 fuel. The raw value of PLV is presented in Fig. 3(a). Overall the PLV profile

looks smooth; however, as it goes into a liquid dense region due to the uncertainty in the measurement, the jagged-edge noises appear in the PLV map, as seen in Fig. 3(b). Thus, data binning (averaging over corresponding binning size) using [5,5] size was utilized to remove the noise and provide with reduced noise dataset for machine-learning algorithm. The PLV map after data binning is presented in Fig. 3(c). As revealed by the PLV distribution at  $z = 19 \text{mm}$  indicated in Fig. 3(d), data binning did not affect the peak value of the PLV but smoothed the profile as shown in subplots from Fig. 3(e) to (g).

Fig. 4 depicts the image-processing techniques applied to quantify the spray liquid content and assess its macroscopic features. Starting

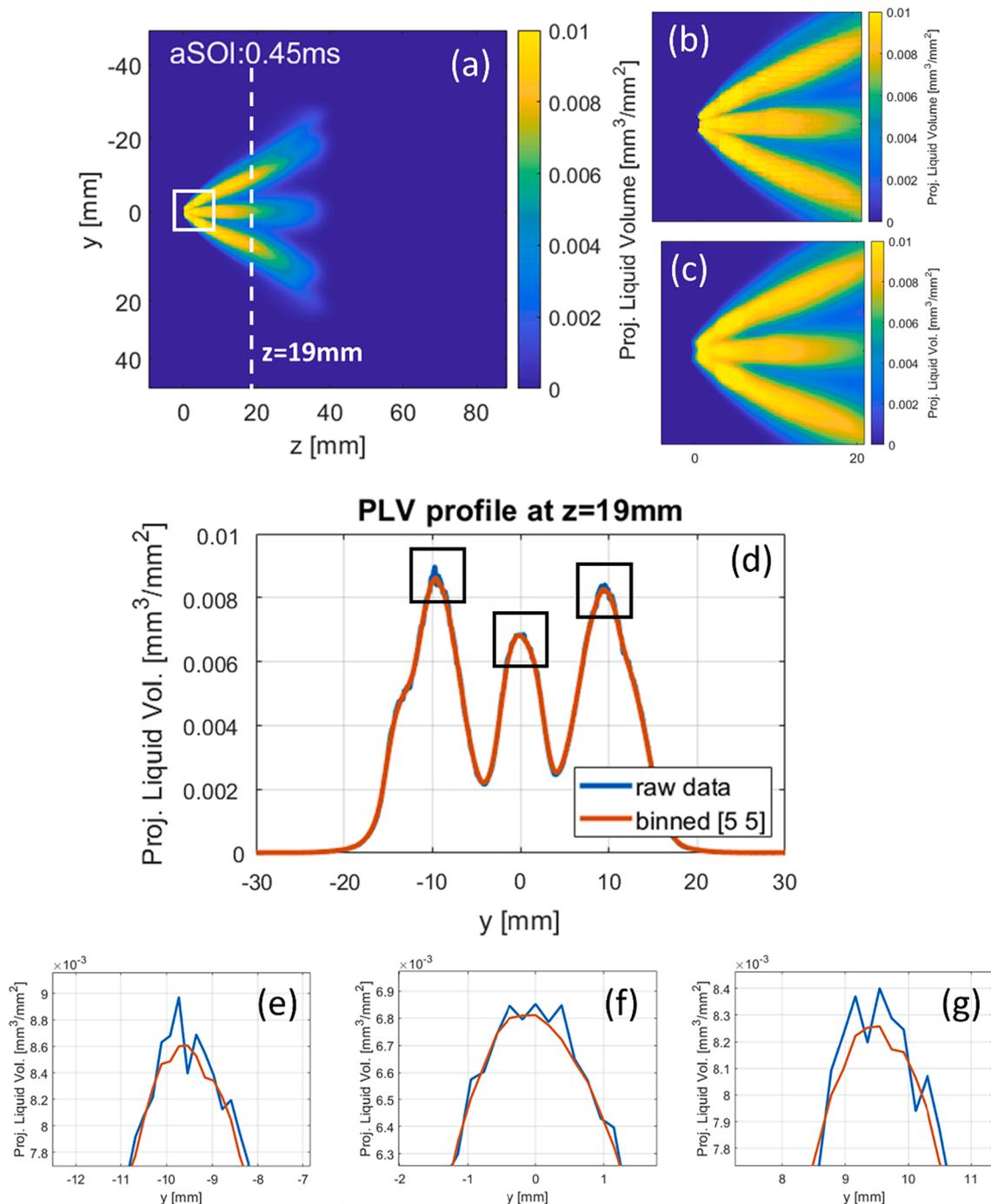
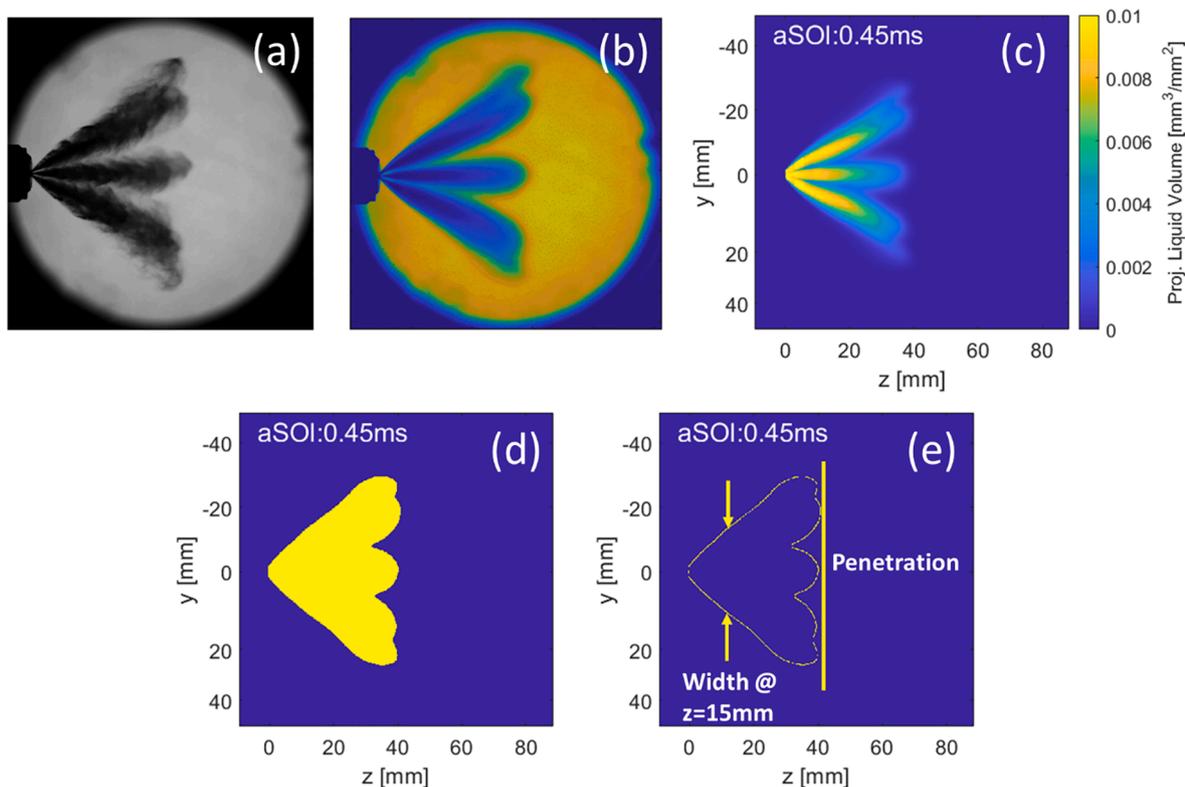


Fig. 3. Preprocessing of PLV map data for machine learning algorithm (a) raw PLV, (b) magnified view of raw PLV indicated white box in (a), (c) magnified view of PLV filtered by [5,5] data binning method, (d) PLV profile at  $z = 19 \text{mm}$  (blue raw data, orange binned data), comparisons of PLV at (e) first peak, (f) second peak, and (g) third peak of the graph shown in (d). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)



**Fig. 4.** Image processing procedure (a) a raw image from a single injection, (b) ensemble averaged image from 300 injections, (c) PLV map, (d) binarized image with PLV threshold of  $0.2 \cdot 10^{-3}$   $\text{mm}^3(\text{liquid})/\text{mm}^2$ , and (e) spray boundary for liquid penetration and liquid width measurements.

from the brightness of an image under evaluation (Fig. 4(a)), ensemble-averaged by 300 injections was normalized using a background image absent of spray (Fig. 4(b)). Subsequently, the PLV map (Fig. 4(c)) was calculated using Eq. (3). A single threshold for PLV was chosen to indicate the extent of liquid penetration and width. The ECN community recommends thresholds of  $0.2 \cdot 10^{-3}$  or  $2 \cdot 10^{-3}$   $\text{mm}^3(\text{liquid})/\text{mm}^2$  [31]. In this study, the lower threshold value of  $0.2 \cdot 10^{-3}$   $\text{mm}^3(\text{liquid})/\text{mm}^2$  was used to binarize PLV maps to get a filled binarized image as in Fig. 4 (d). In the binarized image, liquid penetration length was measured at the farthest axial distance from the nozzle at the primary viewing angle ( $0^\circ$  rotation angle), and the spray liquid width was measured at  $z = 15$  mm based on the coordinate system of Fig. 1. In this image set, Fig. 4(c) was used as input data for the machine-learning algorithm.

### 2.3.2. 3D computed tomography

PLV data were taken at three different viewing angles to construct a 3D spray by a CT algorithm known as inverse Radon transform. Fig. 5 indicates extinction images at three different views and corresponding PLV images. The reconstruction was carried out by using a built-in ‘iradon’ function in MATLAB. The reconstruction routine was applied from the nozzle tip to  $z = 60$  mm. Initially, a line of PLV data at a specific location was extracted and placed in a map shown at the bottom of the figure. Since the PLV data were available at three viewing angles, i.e. the data within  $0^\circ$ ,  $11.25^\circ$ , and  $22.5^\circ$ . The rest of the instances required to produce a  $180^\circ$  rotation matrix were derived through interpolation and mirroring as discussed in detail in [38]. Then the pattern was copied for the angled greater than  $22.5^\circ$  based on the symmetric assumption. Finally, after a projection map from  $0^\circ$  to  $180^\circ$  is generated, the CT algorithm was applied to build spray pattern at a certain location. More details of the 3D reconstruction routine can be found in [38], including confirmation of the process using synthetic model data for liquid volume fraction. It is noted that this is a separate procedure from the machine-learning algorithm. The CT routine was applied after the machine-learning algorithm predicted 2D PLV map.

### 2.4. Test conditions

A series of spray imaging was carried out under ECN Spray G2 (flash-boiling), G3 (early injection), and G3HT (G3 with elevated temperature conditions) conditions employing the nine different fuels mentioned above. The duration of electronic injection command and injection pressure were fixed at  $680 \mu\text{s}$  ( $780 \mu\text{s}$  actual hydraulic duration) and 200 bar, respectively, as per ECN Spray G standard. The imaging was conducted at three different viewing angles by rotating the injector from  $0^\circ$  to  $22.5^\circ$ . The details of the experimental conditions are summarized in Table 2.

### 2.5. Machine-learning methodology and computational setup

An inference algorithm is a process using data analysis to infer underlying properties of data, for example, by testing hypotheses and deriving estimates. A machine-learning algorithm takes many pieces of data, each pre-labeled with a decision, and outputs a mapping from data space to decision space. Together with recent advances in multi-processor computing, machine-learning offers novel breakthroughs in a wide variety of disciplines. Many researchers apply machine-learning algorithms to regression analysis: given a set of vectors, what (low dimensional) subspace comes closest to containing them? In this study, a pixel regression model is considered to predict an image-based fuel spray profile.

#### 2.5.1. Pixel regression model

In the framework of the pixel regression model, instead of simulating the fuel spray from a physical equation-based model, the PLV value  $I_{ij}$  at a pixel position at  $(i, j)$  is predicted pixel-by-pixel from a pre-trained regression model with a given set  $\{x^l\}_{1 \leq l \leq L}$  of fuel, injection, and ambient properties:

$$I_{ij} = f_{ij}(x_1, x_2, \dots, x_L) \quad (4)$$

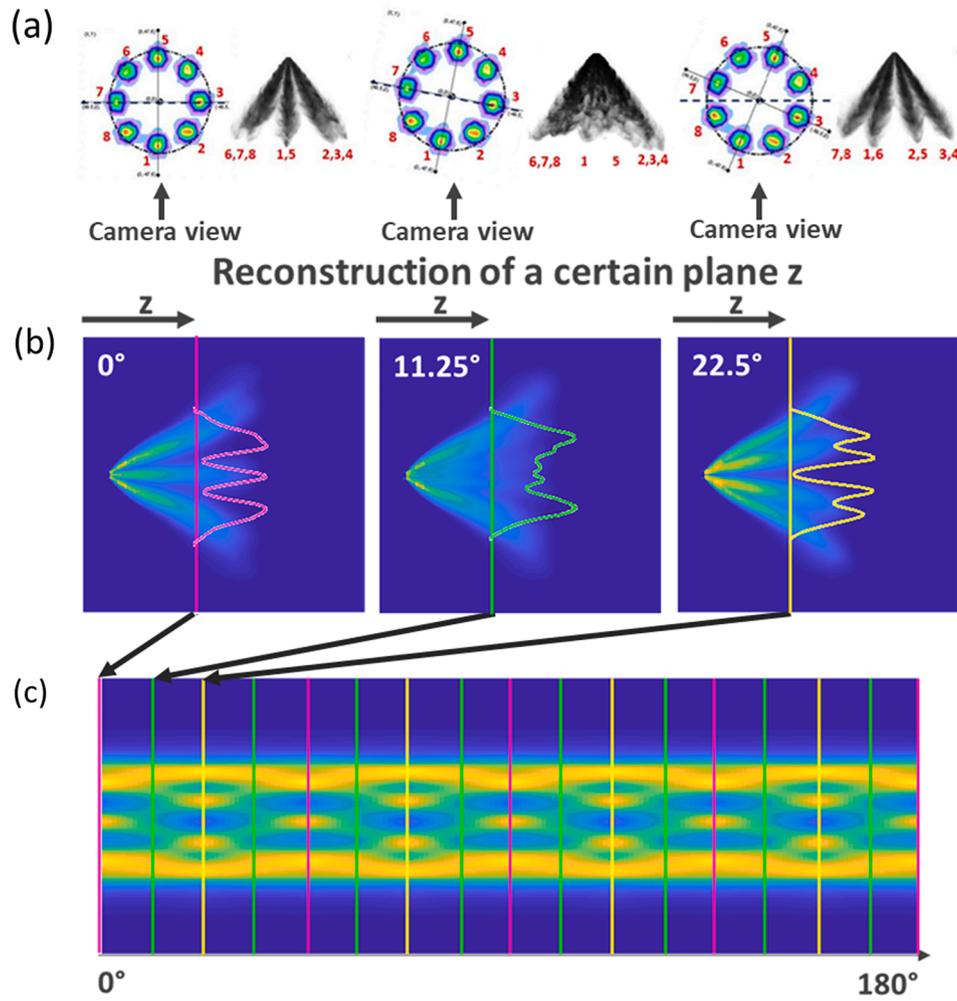


Fig. 5. (a) Raw extinction image at the top, (b) converted ensemble averaged PLV map in the middle, and (c) line data at certain axial location with linearly weighted interpolated map at the bottom [36].

Table 2  
Experimental conditions.

Item	G2	G3	G3HT
Injector viewing angle	3 views (0°, 11.25°, 22.5°)		
Repetition number	300		
Fuel	ic8, di-isobutylene, olefinic, e30, e00, cycloalkane, alkylate, ic8ib2, EEE gasoline		
Injection pressure [bar]	200		
Fuel temperature [°C]	90		
Injection duration (hydraulic duration) [μs]	680 (780) [32]		
Ambient pressure [bar]	0.5	1.0	1.18
Ambient temperature [°C]	60	60	120
Ambient density [kg/m <sup>3</sup> ]	0.5	1.01	1.01

where  $f_{ij}$  is a trained regression model for each pixel at  $(i, j)$ . For training, we utilized  $K$ -number of  $L$ -featured vectors  $\mathbf{x}_k = (x_k^1, x_k^2, \dots, x_k^L)$ ,  $(1 \leq k \leq K)$  with each observed PLV  $\hat{I}_{ij}^k$  via experiments. A simple linear regression model without bias is considered from the assumption that PLV value vanishes when every input feature is zero ( $x_i^k = 0, \forall i$ ). Moreover, since we focused on PLV value itself instead of building a classifier, our model was trained without any activation functions; however, it would be a meaningful investigation to introduce nonlinearity into the model through an activation function. The diagram of our simple pixel linear regression model is depicted in Fig. 6, and the formula is as follows:

$$I_{ij}^k = \sum_{l=1}^L w_{ij}^l x_k^l \quad (5)$$

where  $\mathbf{w}_{ij} = (w_{ij}^1, w_{ij}^2, \dots, w_{ij}^L)$  is a regression weight vector for each pixel at  $(i, j)$ . Moreover, our loss function for training is the most commonly used mean squared error (MSE):

$$MSE_{ij} = \frac{1}{K} \sum_{k=1}^K \left( \hat{I}_{ij}^k - \hat{I}_{ij} \right)^2 \quad (6)$$

where  $\hat{I}_{ij}^k$  is an observed PLV from experimental results. Finally, for each pixel at  $(i, j)$ , the regression weight vector  $\mathbf{w}_{ij}$  minimizing  $MSE_{ij}$  is iteratively updated through an optimizer such as gradient descent algorithm.

However, it is not always necessary to run an iterative optimization algorithm to obtain regression weights. We can solve a specific algebraic equation, the normal equation, to obtain the optimized regression weight vector  $\mathbf{w}_{ij}$  directly: let  $X$  be the matrix  $X = [\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_K]^T$  and  $\hat{\mathbf{y}}_{ij}$  be a vector  $\hat{\mathbf{y}}_{ij} = [\hat{I}_{ij}^1, \hat{I}_{ij}^2, \dots, \hat{I}_{ij}^K]^T$  for each pixel. Then, for each pixel, we can obtain regression weights by

$$\mathbf{w}_{ij} = (X^T X)^{-1} X^T \hat{\mathbf{y}}_{ij} \quad (7)$$

This can be regarded as a one-step learning algorithm, as opposed to

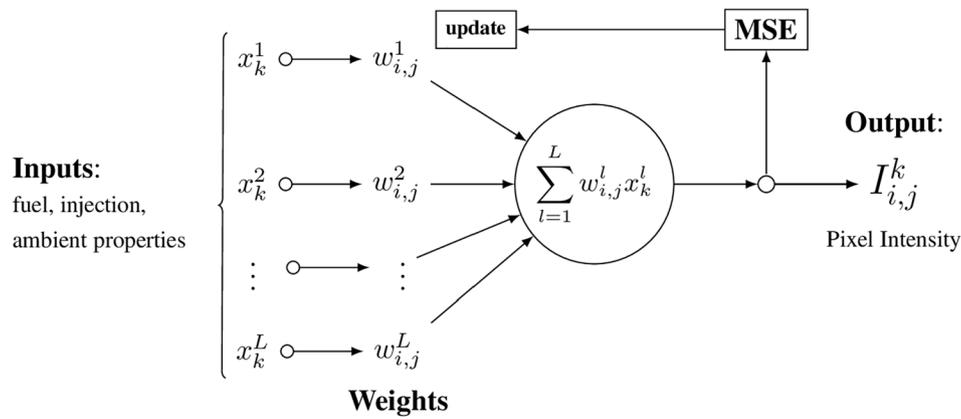


Fig. 6. Diagram of pixel linear regression model (MSE: mean squared error).

iterative machine-learning algorithms. The normal equation is relatively faster than iterative algorithms due to its one-step calculation, but it can be computationally expensive when it deals with a large number of features (large  $L$ ) since we need to take the  $L \times L$  inverse of  $X^T X$ . Moreover, the matrix  $X^T X$  can be singular (non-invertible) if some feature vectors of  $X$  are linearly independent.

2.5.2. Time-resolved image regression model

For one  $M \times N$  spray image, utilizing the matrix  $X = [x_1, x_2, \dots, x_K]^T$  of input feature vectors and each vector  $\hat{y}_{i,j} = [\hat{I}_{i,j}^1, \hat{I}_{i,j}^2, \dots, \hat{I}_{i,j}^K]^T$  of the corresponding PLV values, we trained the regression weight tensor  $W = [w_{i,j}]_{\substack{1 \leq i \leq M \\ 1 \leq j \leq N}}$  of an image. In particular, the training process was

accelerated by applying a parallel node computation. More precisely, we generated  $p$ -threads for  $p$ -number of pixel calculations (specifically row direction) in an image and spread them to a multi-processor pool at the same time so that  $p$ -number of pixel weight vectors are trained simultaneously in a parallel iteration. Here, an appropriate number  $p$  was determined from the number of available threads in the computing processor. The training schematic of an image is demonstrated in Fig. 7.

Independently repeating this procedure for each time-step, we trained a time series regression model for the prediction of spray dynamics. Thus, once the training is completed and the trained model is obtained, unlike CFD simulations, we can directly predict the spray image at any desired time-step without an accumulating time-stepping simulation.

2.6. Computational setup

Computations were carried out on a workstation computer with an AMD Ryzen Threadripper 3990X 64-Core 128-Thread 2.90 GHz processor, GeForce RTX 2080 SUPER (NVIDIA TU104; 8 GB RAM), and 128 GB DDR4 RAM. Pre-processing of spray images and their denoising processes were carried out in MATLAB, while the machine-learning processes were implemented using Python. To accelerate the calculations, we utilized PyTorch (stable 1.7.0), which is one of the widely used open-source machine-learning libraries that enables the training processes to cooperate with effective GPU array operation, and we parallelized whole pixel iterations in an image with a multiprocessing library. Concerning data structures, measured PLV values via experiments were assigned into pixel values of images, and selected  $L = 9$  features (mentioned in section 2.1) were utilized as training inputs. Also, we

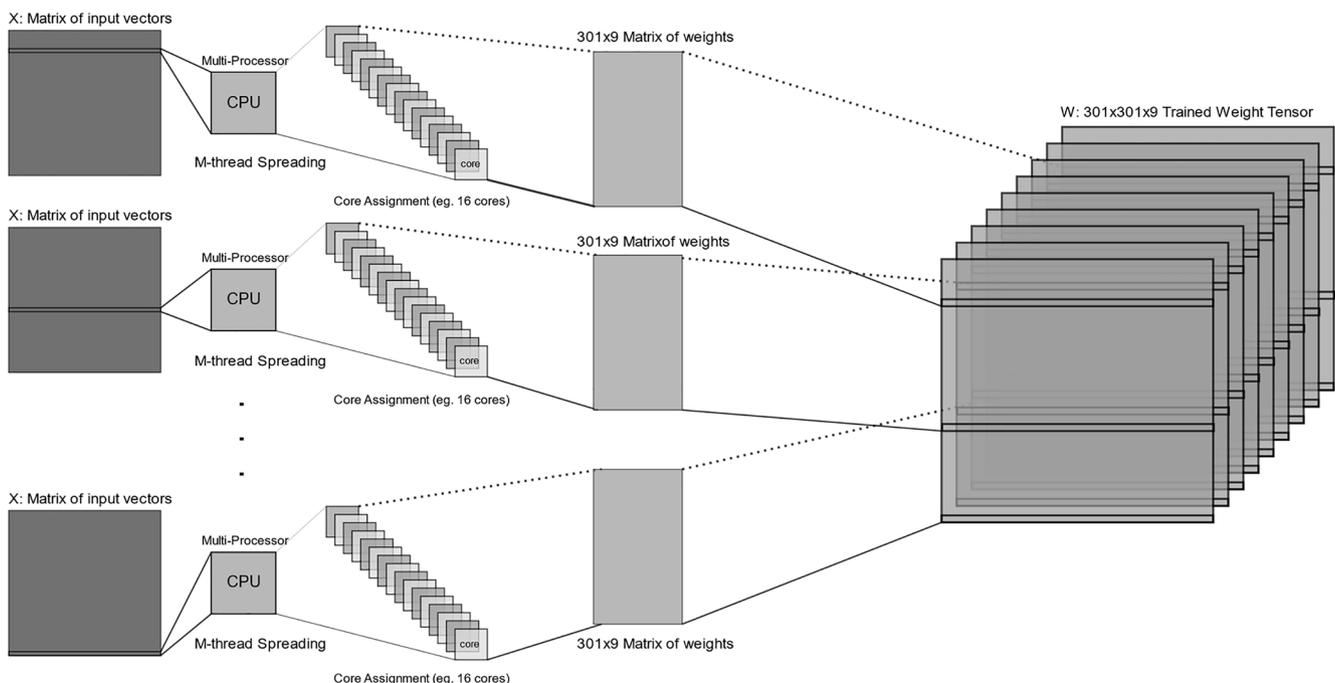


Fig. 7. Schematic of the algorithm training procedure to derive a predicted image of the spray topology.

prepared  $K = 21$  PLV at each time-step for the training. These feature data were not standardized since our regression model is a simple linear model without any interaction terms (eg.  $x_1x_2$ ) or power terms (eg.  $x^n$ ). Furthermore, it was revealed from a comparison between standardized and non-standardized data that non-standardization showed a much better accuracy for our model. In order to avoid unnecessary calculations of in pixels of zero PLV values, every image was cropped to the size of  $M \times N = 301 \times 301$  to cover entire spray regions with the smallest possible margins from zero-pixel regions. The training datasets had identical nozzle tip coordinates to prevent errors by the translational shifts. For the machine-learning process, stochastic gradient descent (SGD) optimizer as well as two adaptive optimizers namely Adam and DiffGrad, were considered for training, and we benchmarked the prediction performances of the three optimizers and the normal equation, which is discussed in Section 3.1. The details of computation, optimizers, and model hyperparameters are summarized in Table 3.

### 3. Results and discussion

#### 3.1. Validation of regression model

As aforementioned in Section 2.6, the regression model was trained by 21 different spray data consisting of 9-features and its PLV for each pixel. In order to validate the model, we depicted one of the spray images in the train set (ic8ib2 under G2, aSOI: 0.45ms) and its spatial heat map of the coefficient of determination ( $R^2$  score) with contour lines in Fig. 8, as an example, when we employed SGD as our optimizer for the heat map. The coefficient of determination was defined as the proportion of the variance in the dependent variable ( $I_{ij}^k$ ) that was predictable from the independent variable ( $x_k$ ) and it lies in the range of  $[-\infty, 1]$  (in the figure, we cut the values at  $-1$  i.e.  $R_{ij}^2 = -1$  if  $R_{ij}^2 < -1$ ):

$$R_{ij}^2 = 1 - \frac{\sum_{k=1}^K \left( I_{ij}^k - \bar{I}_{ij} \right)^2}{\sum_{k=1}^K \left( I_{ij}^k - \bar{I}_{ij} \right)^2} \quad (8)$$

where  $\bar{I}_{ij}$  is an observed pixel PLV via experiments and  $\bar{I}_{ij} = \frac{1}{K} \sum_{k=1}^K \hat{I}_{ij}^k$ , the mean of the desired PLV values.  $R^2$  score is often used to evaluate the relative variability of a regression model, if  $R_{ij}^2$  value is close to 1, then the two variables ( $I_{ij}^k$  and  $x_k$ ) are perfectly correlated and a low  $R_{ij}^2$  value would show a low level of correlation, meaning that a regression model is not valid. It can be seen from Fig. 8 that the region where  $R^2$  scores are above 0.5 contains the actual spray region, and the low scores, less than zero, are shown away from the periphery of the

**Table 3**  
Simulation details.

Specification	Description			
# of CPU threads	128			
Parallel pool size( $p$ )	301			
Image size( $M \times N$ )	$301 \times 301$			
Input data( $K$ )	21spray images			
Input features( $L$ )	9 features; Density [ $kg/m^3$ ], Viscosity [ $mm^2/s$ ], Ambient T [K], Ambient density [ $kg/m^3$ ], Ambient pressure [ $kPa$ ], Vapor pressure [ $kPa$ ], Pressure ratio [ $Pa/Pv$ ], Heat of vaporization [ $kJ/kg$ ], Distillation [n.a]			
Training epoch	50,000			
Optimizers	Learning rate	$\beta$	Tolerance	Momentum
SGD	$10^{-7}$	n.a	n.a	0
Adam	0.15	(0.9,0.999)	$10^{-8}$	n.a
DiffGrad	0.15	(0.9,0.999)	$10^{-8}$	n.a

actual spray region. These low scores may stem from the vanishing PLV values of the region. Since we are mainly interested in properly replicating the PLV values in the actual spray region, the low scores on the outer region can be ignored. Therefore, our model is still applicable to predict the spray profiles. It should be noted that more data and elaborated models are still necessary to enhance predictions and completely express the entire pixels.

Running losses (MES) versus epochs (iteration number) at pixel (60, 100) (marked in Fig. 8(a)) for Adam, DiffGrad, and SGD optimizers are presented in Fig. 9, with each terminating-loss value being reported in the legend. Note that the normal equation is not included here because it does not require training. From the terminating loss values, it seems that the adaptive optimizers are more effective than the SGD optimizer. However, unlike SGD in Fig. 9(b), both Adam and DiffGrad showed continuous spikes in their running losses plots in Fig. 9(a), which were an unavoidable consequence of the intended tactics for the adaptive optimizers to avoid the local minima. Although these spikes are soon damped out to stable value (see the magnified view at the right corner in Fig. 9(a)), it might be harmful once training is terminated at these spikes, details are as follows. To investigate the performance of each optimizer, in Fig. 10, we depicted the true cross-sectional line ( $z = 28$  mm; 151 column of image) of (ib20-G2) spray profiles from the experiment, the corresponding predicted results by the normal equation (NE), and the models trained with the different optimizers i.e., Adam, DiffGrad, and SGD. First, the NE results show a complete failure to predict the trend of the true values even though the algorithm is computationally much more efficient than the other optimizers (about 7.5 times faster). The reason for the inaccuracy of the normal equation is that the performance of our utilizing linear system solver might be suffered by the singularity of the matrix  $X^T X$ , which arises from linear dependencies between features. For the two adaptive optimizers, spike noises appear in their cross-sectional lines due to the premature termination at the unstable spike moment, as seen in Fig. 10(a). On the other hand, SGD has produced a smooth cross-sectional line without any noise. Applying the median filter to the noisy lines as postprocessing, we can obtain smooth lines in Fig. 10(b). As one can see from the figure, the adaptive optimizers show more accurate predictions in the center region ( $-10mm < y < 10mm$ ) than SGD optimizer, while SGD optimizer reveals a more accurate prediction in the outer regions ( $y < -10mm$  and  $y > 10mm$ ) than the adaptive optimizers. The possible reason for discrepancies depending on spray region is that this model is using global thermodynamic parameter for instance uniform ambient temperature and pressure not spatial values according to spray region.

From the observations in Figs. 9 and 10, we can conclude that SGD optimizer does not introduce unstable spikes into the prediction and can achieve better accuracy in the outer regions than the adaptive optimizers. Again, in this work, our main concern is the overall outer edge shape of spray. Moreover, noises in the spray profile and its filtered result can mislead our quantitative analysis. For these reasons, we employed SGD optimizer for our training and analysis of the results with 50,000 training epochs since it converges well enough at the epochs, as shown in Fig. 9.

#### 3.2. Liquid penetration and width

The liquid penetration length and liquid width can provide an overall insight into the spray morphology, so the line-of-sight data analysis was first carried out based on the PLV threshold of  $0.2 \cdot 10^{-3} \text{ mm}^3(\text{liquid})/\text{mm}^2$ . It has to be reminded that the liquid penetration length and liquid width were measured from an ensemble averaged image of 300 injections. The experimental data are presented in Fig. 11. It is noted that the spray results using ic8ib2 and EEE fuels were not included in the training data set but utilized only for comparison against the predictions of the machine-learning algorithm. True values from experiments are indicated with solid lines, while the predictions by the machine-learning

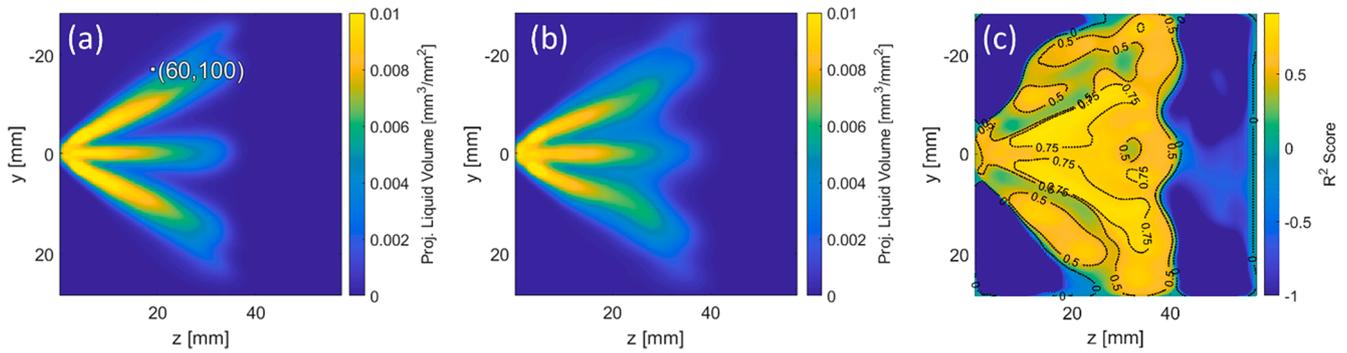


Fig. 8. (a) PLV map from experiment (ic8-G2), (b) predicted PLV map by machine learning algorithm (ic8-G2), and (c) its heat map with contour lines for the coefficient of determination ( $R^2$  score) at 0.45 ms after fuel injection.

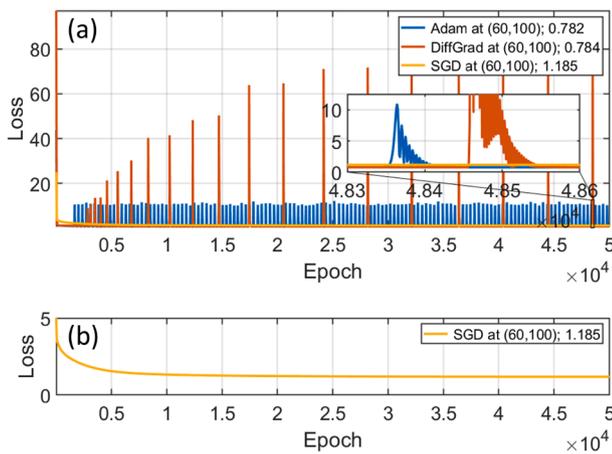


Fig. 9. (a) Running losses versus Epoch (iteration number) at the pixel position (60, 100) for each optimizer: Adam, Diffgrad, and SGD and (b) the magnified view of the losses for SGD; each terminating loss value is reported in legend. The small window at the right corner in (a) shows the magnified view of the losses at Epochs close to the termination.

are indicated with symbols. Under flash-boiling G2 condition presented in Fig. 11(a), the ratio of ambient pressure to vapor pressure ( $P_a/P_v$ ) dominates spray process, unlike sprays under high pressure ambient where nondimensional parameters such as Reynolds number  $Re$  and  $We$  typically determine spray characteristics [41,42]. Comparing liquid width at  $z = 15$  mm, both fuels showed a faster rate of decrease than G3 and G3HT conditions. This implies that a certain level of plume collapsing occurs for flash-boiling conditions. Especially for EEE fuel has a very low ambient to vapor pressure ratios ( $P_a/P_v$ ) equal to 0.17, 0.34, 0.41 for G2, G3, G3HT, respectively, spray collapse even under the relatively high-pressure G3HT condition can be detected. The lower ambient density and plume collapse into a single structure under G2 condition augmented the spray momentum at its leading edge so that the liquid penetration length increases compared to the G3 and G3HT conditions, despite the enhancement in vaporization for flash-boiling sprays. G3 and G3HT exhibited indistinguishable liquid penetration length because of identical ambient density, however, enhanced evaporation due to the higher ambient temperature under the G3HT condition resulted in the smallest liquid-outline width for the examined ambient conditions. The liquid width of ic8ib2 for the G3 condition did not show the drastic decrease as under G2 condition, since flash-boiling did not take place as suggested by the the pressure ratios ( $P_a/P_v$ ) of 1.55 and 1.83 for G3 and G3HT conditions, respectively. Details of plume direction derived from tomographic reconstruction will be discussed in the following section.

With regard to the predictive capability of the machine-learning

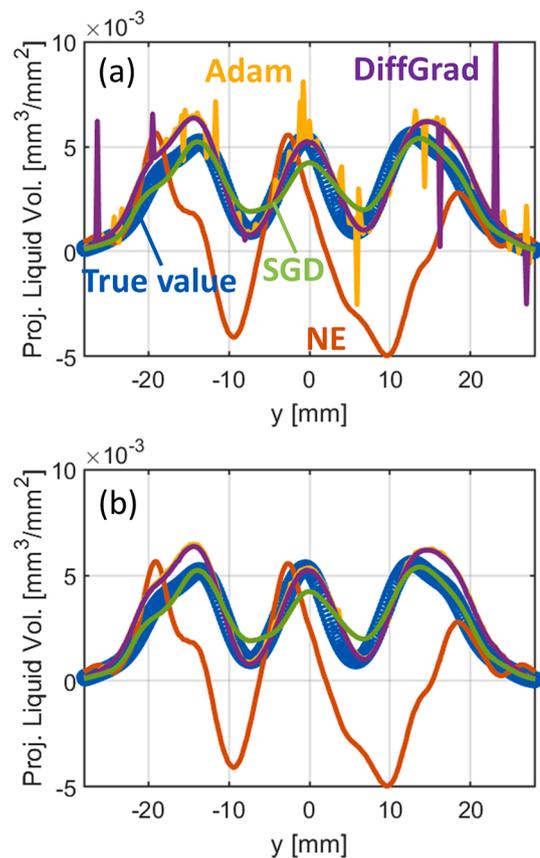
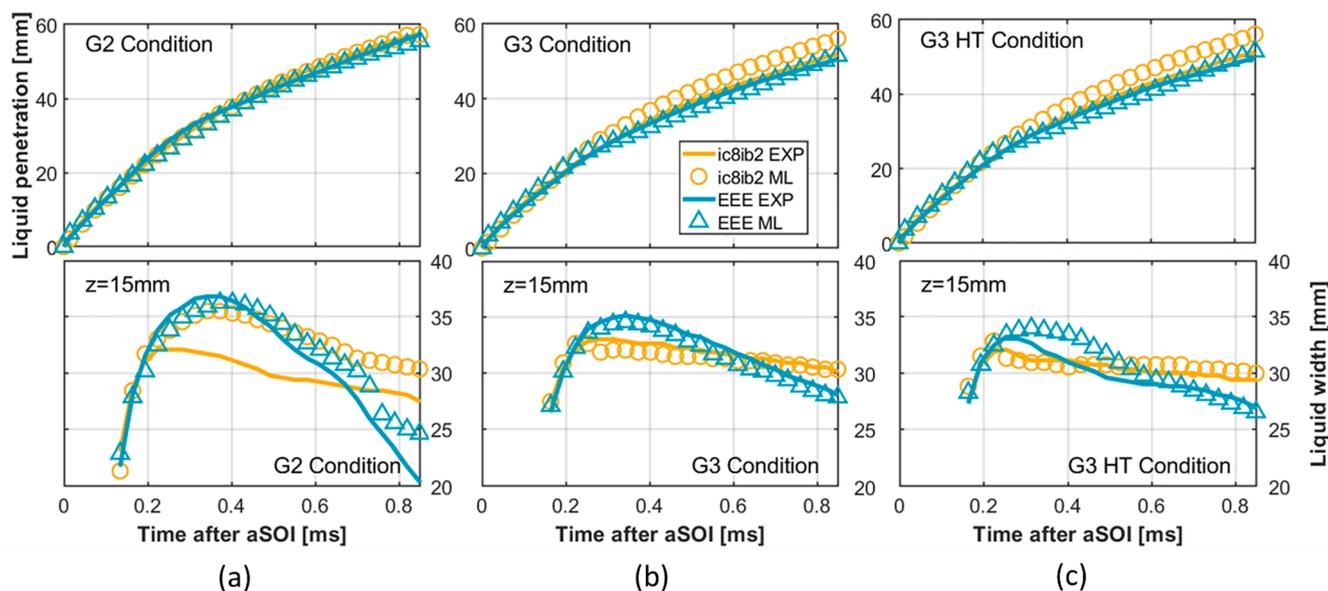


Fig. 10. (a) Cross-sectional line ( $z = 28$  mm; 151 column of image) of (ib20-G3) spray profiles obtained experiment and its predicted results by the models trained with different optimizers: normal equation (NE), Adam, DiffGrad, and SGD and (b) their filtered lines smoothed with a median filter.

algorithm, good overall agreement with experimental results was accomplished especially referring to the liquid penetration length for the EEE fuel. It is challenging to predict the spray topology under flare flash-boiling conditions even by CFD simulations [43,44], nevertheless, the machine-learning approach showed similar values of liquid penetration length and liquid width, consistent with the experimental results. The liquid width under the G2 condition and the decrease in liquid width due to spray collapse can also be captured by the machine-learning algorithm. For the EEE fuel, the maximum discrepancies between the machine-learning predictions and the experimental data regarding the liquid penetration length and liquid width during the injection period were 1.3 mm (2.32% error) and 2.4 mm (7.86%), respectively. The prediction of liquid penetration length of flash-boiling spray is very



**Fig. 11.** Comparison of liquid penetration length and liquid width (@ $z = 15$  mm) between experimental results and machine learning predictions under (a) G2, (b) G3, and (c) G3HT conditions.

crucial for engine design since the extent of liquid fuel impingement on the surface of the combustion chamber. On the other hand, for the ic8ib2 case, the machine-learning algorithm tends to have slightly higher values than the experimental results. The largest discrepancy occurred in the prediction of liquid width under G2 condition where the model predicted a slightly thicker spray plume on the upper part, as shown in the PLV distributions of Fig. 10. However, the algorithm predictions for G3 and G3HT conditions were satisfactory. The maximum differences in liquid penetration length and liquid width during the injection period were 3.6 mm (7.3% error) and 4.2 mm (13.3% error), respectively.

### 3.3. 3D spray prediction

The prediction of spray morphology using the machine-learning algorithm at different viewing angles was also carried out to reconstruct the 3D spray distribution. The LVF distribution in 3D spray according to time under G2 condition is shown in Fig. 12. It is noted that this is quantitative LVF data that can be utilized for CFD validation. The overall behavior of spray dynamics is well presented by the prediction result using the machine-learning algorithm. The ic8ib2 fuel showed slight plume collapse as suggested by the wide-open leading edge and curved shape but still maintaining eight separate plumes. The complete plume collapsing with EEE fuel also can be captured by the machine-learning algorithm. In terms of evaporation time, as the algorithm predicted a lower level of projected liquid volume, the overall level of liquid volume fraction is smaller, especially near the injector compared to the experimental results. The predicted 3D spray for ic8ib2 (at 0.88 aSOI) showed faster evaporation than the experiment.

To compare the detailed spray boundary at a certain plane, the 3D spray under G2 condition was cut along a yz plane with a LVF threshold of  $1 \cdot 10^{-4}$  as shown in Fig. 13. The nozzle drill angle is presented with grey dotted lines presenting  $37^\circ$  from the injector center axis. The spray boundaries obtained from the experiment and the machine-learning algorithm are indicated with solid and dotted lines, respectively. The spray boundary for ic8ib2 shows that it penetrates with a slightly smaller angle than the nozzle drill angle as it is under a mild flash-boiling regime. The decreasing liquid width in time is then confirmed in this cut plane data. Moreover, the consistent trend in liquid penetration length and liquid width also can be found in this plane image. The ic8ib2 showed longer penetration and wider liquid width than EEE fuel during the fuel injection. A previous comparative study for various

single- and multi-component fuels conducted in the same spray chamber had shown that strongly-collapsing sprays of multi-component fuels had increased momentum after the end of injection, leading to increased penetration [36]. The results for the EEE fuel in Fig. 13 also reveal vigorous spray collapsing from the beginning of the injection with a complete collapse of individual plumes. The strong plume-to-plume interaction finally left a significant amount of liquid phase fuel in the core region. The machine-learning algorithm demonstrated its capability to predict spray collapse for the EEE fuel under a flash-boiling condition.

The comparison of LVF predictions and experimental measurements at an xy plane 30 mm downstream of the injector tip is shown in Fig. 14. A contour plot of LVF prediction is presented in Fig. 14 (a), and the averaged LVF through the centerline of the plumes, in Fig. 14 (b). Discrete plumes can be seen for the ic8ib2 case, yet the plumes have converged towards the injector axis for the EEE fuel. As the machine-learning algorithm tends to retain original plumes, it exhibited star-shaped morphology in the bottom view, however, the collapsing of the plumes is well presented. The quantitative comparison by averaged LVF (Fig. 14 (b)) showed a lower peak with the machine-learning algorithm equipped with SGD optimizer, however, the prediction of plume dynamics was consistent regardless of fuel type. Overall, the machine-learning algorithm showed better agreement in the level of LVF and distribution for ic8ib2 than the EEE fuel. The possible reason for inaccurate prediction for flare flash-boiling is that the critical features may not have been fed into the machine-learning algorithm. In this study the features are limited in nine different parameters, however, one can expand them by applying vapor pressures at various temperatures, plume cone angle, and different method for distillation parameter.

As shown in the yz and xy cut plane information (Figs. 13 and 14), the spray plume does not penetrate along the nozzle drill angle but at a smaller angle. Understanding of the fuel direction is essential to assess air-fuel mixing in an engine cylinder. Thus, the plume direction angle was derived from 3D spray results and defined as an angle between the axis corresponding to the plume center and injector axis. The plume center was selected at the 99% peak LVF position to eliminate bias induced by the skewed shape of plumes. As shown in Fig. 15, the plume direction angle of ic8ib2 maintained a constant value of about  $35^\circ$  for the G3 and G3HT conditions where flash-boiling was not an influencing factor, while plume direction exhibited slightly lower yet also relatively stable values for the G2 condition. On the other hand, the plume

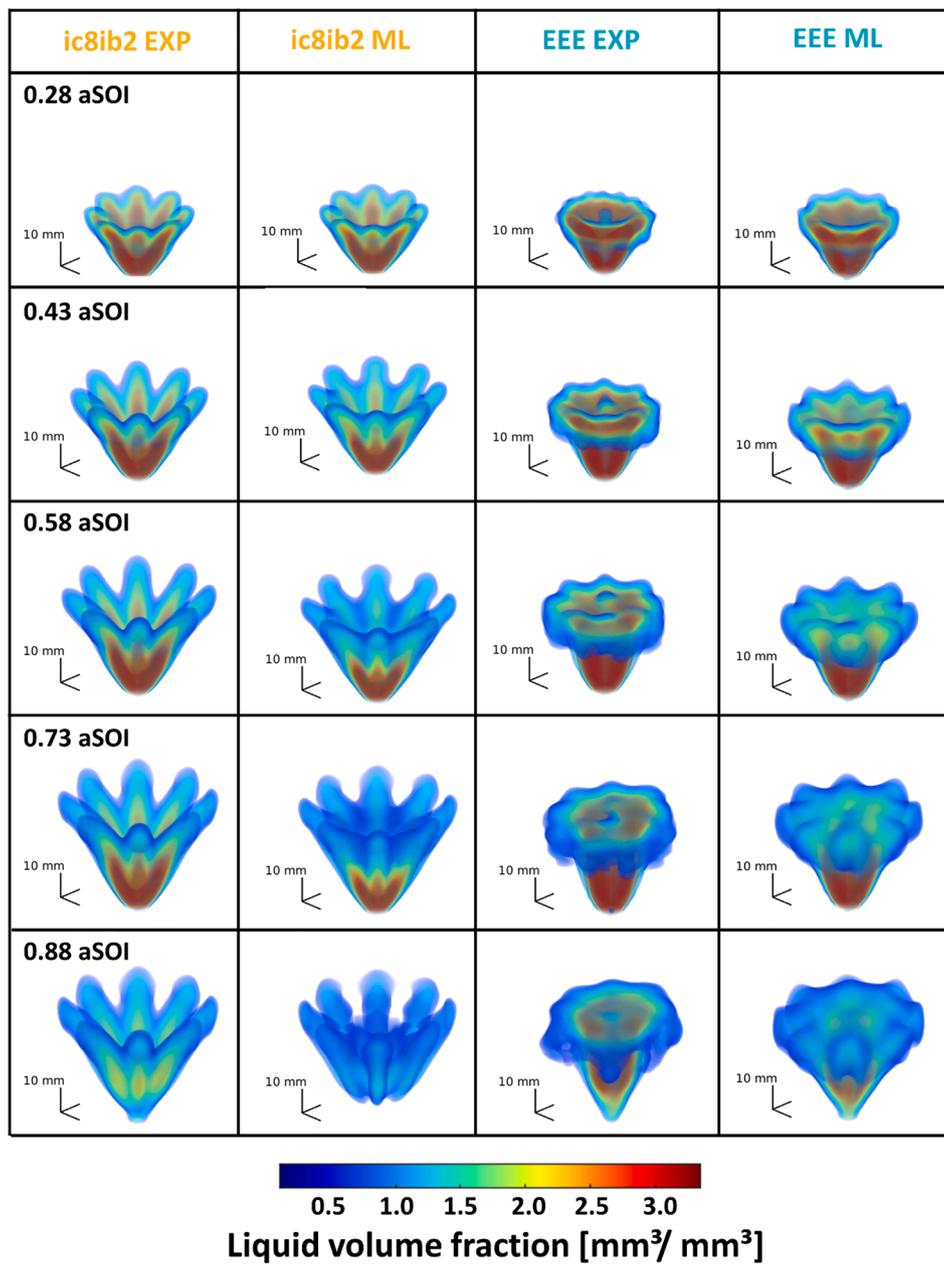


Fig. 12. Three-dimensional LVF distribution by experiment and machine learning prediction under G2 condition.

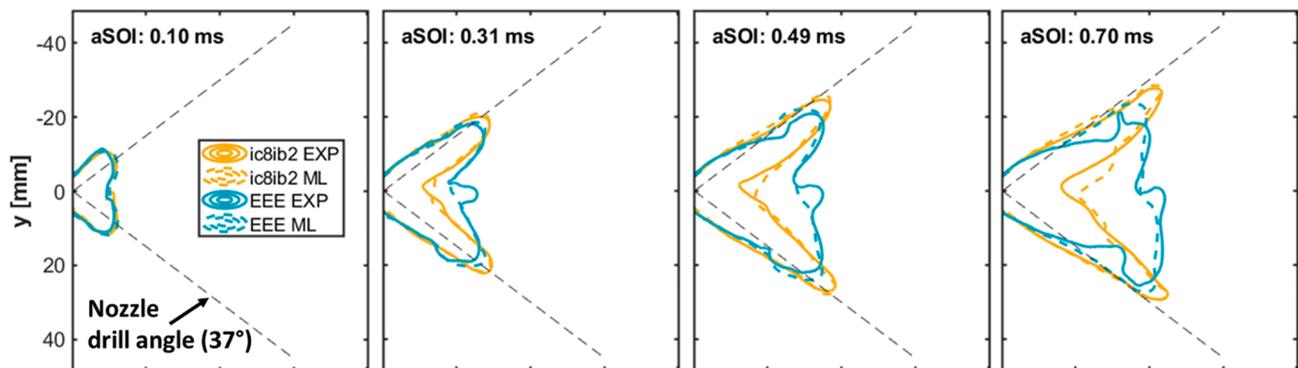


Fig. 13. Boundary of liquid spray at a yz plane enabled by 3D CT. The boundary was determined by an LVF threshold of  $1 \cdot 10^{-4}$  under G2 condition.

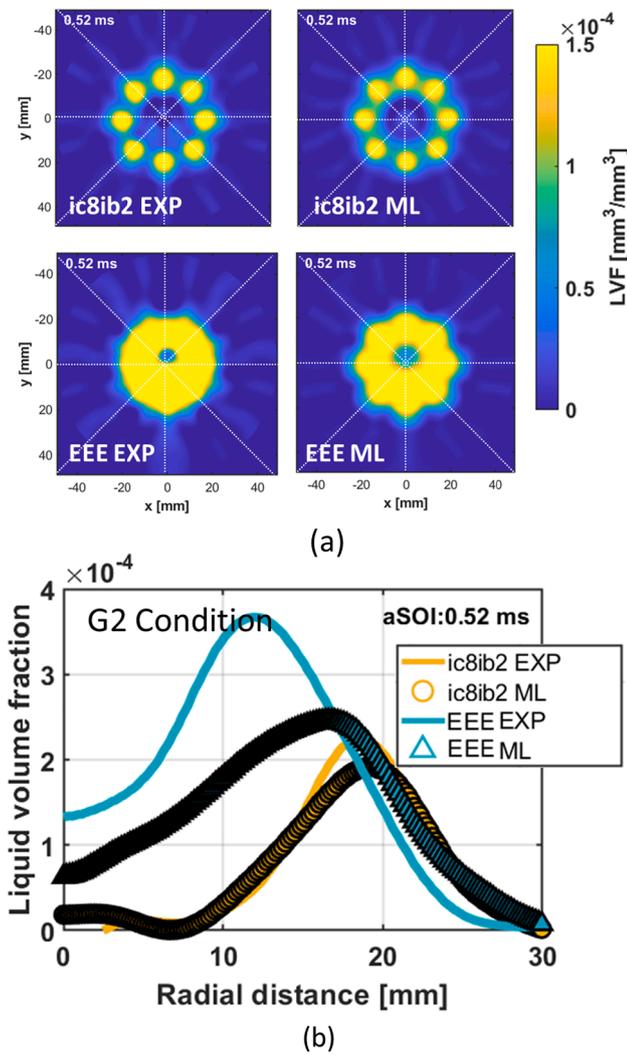


Fig. 14. (a) Liquid volume fraction contour plots, and (b) averaged liquid volume fraction distribution across the centerline (white dotted lines in (a)) at  $z = 30$  mm ( $xy$  plane) under G2 condition.

direction angle of the EEE fuel continuously decreased continuously during the injection event regardless of the ambient condition since a certain level of plume-to-plume interaction took place even under G3 and G3HT conditions. The plume direction angle from the machine-learning algorithm showed the best match under G3 and G3HT conditions with ic8ib2 fuel. The prediction for the flash-boiling condition and EEE fuel showed a slight discrepancy compared to the experimental

data, but still the difference was small up to  $2^\circ$ . This result implies that the machine-learning algorithm can effectively predict quantitative LVF and overall spray dynamics under different ambient and fuel conditions.

#### 4. Conclusion

In this study, a machine-learning algorithm was implemented to predict gasoline sprays under engine-like conditions. The non-deterministic numerical predictions were compared to high-speed extinction imaging performed in a constant-flow vessel under ECN defined G2, G3, G3HT conditions. Nine different fuels were injected using an ECN Spray G 8-hole injector. The machine-learning algorithm was capable of predicting not only line-of-sight measurement data but also 3D liquid volume fraction by computed tomographic reconstruction. Quantitative LVF and plume direction angle data of ic8ib2 and EEE fuels (not included in the training data set) were obtained by machine-learning prediction. The major findings from this study can be summarized as follows.

- 1) The introduced machine-learning technique is a simple regression model without a bias term that does not pose significant complexities regarding its implementation. Despite its simplicity, reliable results could be achieved. These model characteristics will enable future researchers to easily modify and enhance the model to accomplish better efficiency and accuracy in spray prediction. Here, we leave some suggestions for enhancement of the algorithm predictive capability: application of other regression models such as polynomial and logistic regression; application of neural network (NN), especially convolutional NN (CNN); a collaboration between the model-based and data-driven algorithm such as physics informed deep learning (PINN).
- 2) The strength of the developed regression model is its natural neural network configuration. Although we trained our regression model per each pixel and per time-step independently without considering any spatio-temporal neural networks, the predicted spray profiles and their time sequences naturally show a continuous spatio-temporal spray dynamics evolution with high accuracy (as verified by predicted images and their temporal changes). This constitutes proof of the interconnection between the underlying spray dynamics and fuel, injection, and ambient properties, and eventually implies that our pixel-independent model can well-preserve the fundamental dynamics underneath training data, thus realizing a virtual spatio-temporal neural network.
- 3) The machine-learning algorithm developed in this study predicts the spray patterns in an independent time-stepping manner. This indicates that this algorithm can be used to reduce a significant amount of computational time and cost compared to CFD simulations that needs sequentially accumulated data to predict spray evolution in the subsequent time step.

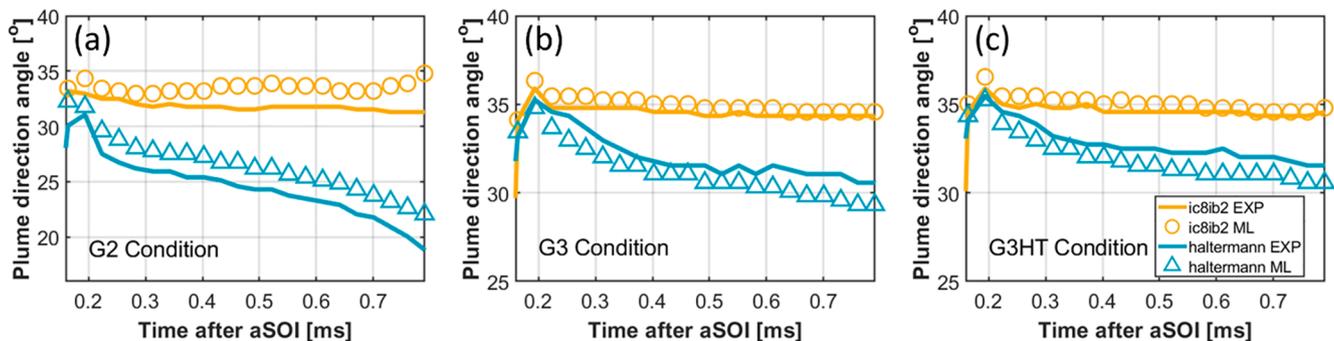


Fig. 15. Comparison of plume direction angle from experiment and machine learning algorithm under G2 condition. The plume direction angle was defined by an angle between the nozzle orifice axis and the center of plume width placed at 99% of the peak liquid volume fraction at  $z = 30$  mm downstream the injector tip.

- 4) Predicted macroscopic spray characteristics, including liquid penetration length and width, showed good agreement with experimental data, while the machine-learning based 3D computed tomography of the spray topology was demonstrated to be capable of predicting distinct features such as plume collapse and direction angle variation. The machine-learning data showed a slightly lower peak of LVF, however, the plume dynamics showed consistent trends with the experimental results.

### CRedit authorship contribution statement

**Joonsik Hwang:** Conceptualization, Investigation, Writing - original draft. **Philku Lee:** Investigation, Writing - original draft. **Sungkwang Mun:** Methodology, Validation, Writing - original draft. **Ioannis K. Karathanassis:** Validation, Writing - review & editing. **Phoevos Koukouvinis:** Validation, Writing - review & editing. **Lyle M. Pickett:** Writing - review & editing, Supervision, Funding acquisition. **Manolis Gavaises:** Supervision, Funding acquisition.

### Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

### Acknowledgements

The authors would like to thank Tim Gilbertson, Alberto Garcia, Brian Patterson, Keith Penny, Nathan Harry, Aaron Czeszynski, Chris Ingwerson, Laurie Bell, Paul Abers, and Peter Cable for their dedicated support for the research facility. The spray experiment was performed at the Combustion Research Facility (CRF), Sandia National Laboratories, Livermore, CA. This research was conducted as part of the Co-Optimization of Fuels & Engines (Co-Optima) project sponsored by the U.S. Department of Energy (DOE) Office of Energy Efficiency and Renewable Energy (EERE), Bioenergy Technologies and Vehicle Technologies Offices. Sandia National Laboratories is a multi-mission laboratory managed and operated by National Technology and Engineering Solutions of Sandia, LLC., a wholly owned subsidiary of Honeywell International, Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525. We also gratefully acknowledge EU Horizon-2020 Marie Skłodowska-Curie Global Fellowships AHEAD (IK, Grant No. 794831) and UNIFIED (PK, Grant No. 748784) for supporting an international visiting program of Ioannis K. Karathanassis and Phoevos Koukouvinis at the Sandia National Laboratories. Additional funding has been received by the UK's Engineering and Physical Sciences Research Council (EPSRC) through grant EP/K020846/1. We also appreciate internal funding from Center for Advanced Vehicular Systems (CAVS) for the development of the machine-learning algorithm.

### References

- Harada J, Tomita T, Mizuno H, Mashiki Z, Ito Y. Development of direct injection gasoline engine. SAE Tech Pap 1997. <https://doi.org/10.4271/970540>.
- Duronio F, De Vita A, Montanaro A, Villante C. Gasoline direct injection engines – A review of latest technologies and trends. Part 2. Fuel 2020;265:116947. <https://doi.org/10.1016/j.fuel.2019.116947>.
- Costa M, Sorge U, Merola S, Irimescu A, La Villetta M, Rocco V. Split injection in a homogeneous stratified gasoline direct injection engine for high combustion efficiency and low pollutants emission. Energy 2016;117:405–15. <https://doi.org/10.1016/j.energy.2016.03.065>.
- Park C, Kim S, Kim H, Moriyoishi Y. Stratified lean combustion characteristics of a spray-guided combustion system in a gasoline direct injection engine. Energy 2012; 41:401–7. <https://doi.org/10.1016/j.energy.2012.02.060>.
- Yang J, Dong X, Wu Q, Xu M. Effects of enhanced tumble ratios on the in-cylinder performance of a gasoline direct injection optical engine. Appl Energy 2019;236: 137–46. <https://doi.org/10.1016/j.apenergy.2018.11.059>.
- Dec JE, Dernotte J, Ji C. Increasing the Load Range, Load-to-Boost Ratio, and Efficiency of Low-Temperature Gasoline Combustion (LTGC) Engines. SAE Int J Engines 2017;10. <https://doi.org/10.4271/2017-01-0731>.
- Stuhldreher M, Kargul J, Barba D, McDonald J, Bohac S, Dekraker P, et al. Benchmarking a 2016 Honda Civic 1.5-Liter L15B7 Turbocharged Engine and Evaluating the Future Efficiency Potential of Turbocharged Engines. SAE Int J Engines 2018;11:1273–305. <https://doi.org/10.4271/2018-01-0319>.
- Luijten B, Adomeit P, Brunn A, Somers B. Experimental investigation of in-cylinder wall wetting in GDI engines using a shadowgraphy method. SAE Tech Pap 2013;2. <https://doi.org/10.4271/2013-01-1604>.
- Tang Q, Liu H, Li M, Yao M. Optical study of spray-wall impingement impact on early-injection gasoline partially premixed combustion at low engine load. Appl Energy 2017;185:708–19. <https://doi.org/10.1016/j.apenergy.2016.10.108>.
- Arai M. Physics behind Diesel Sprays. ICLASS 2012, 12th Trienn Int Conf Liq At Spray Syst Heidelberg, Ger Sept 2-6, 2012 Phys 2012:1–18.
- Crua C, Manin J, Pickett LM. On the transcritical mixing of fuels at diesel engine conditions. Fuel 2017;208:535–48. <https://doi.org/10.1016/j.fuel.2017.06.091>.
- Wang Z, Jiang C, Xu H, Badawy T, Wang B, Jiang Y. The influence of flash boiling conditions on spray characteristics with closely coupled split injection strategy. Appl Energy 2017;187:523–33. <https://doi.org/10.1016/j.apenergy.2016.11.089>.
- SAE Technical Paper 900475 Structures of Fuel Sprays in Diesel Engines Hiro Hiroyasu and Masataka Arai Dept. of Mechanical Engineering University of Hiroshima Japan n.d.:900475.
- Levich VG, Seeger RJ. Physicochemical Hydrodynamics. 892–892 Am J Phys 1963; 31. <https://doi.org/10.1119/1.1969158>.
- Siebers DL. Scaling liquid-phase fuel penetration in diesel sprays based on mixing-limited vaporization. SAE Tech Pap 1999. <https://doi.org/10.4271/1999-01-0528>.
- Reitz RD, Bracco FV. Mechanisms of breakup of round liquid jets. Encycl Fluid Mech 1986:233–49.
- Marmottant PH, Villermaux E. On spray formation. J Fluid Mech 2004;498: 73–111. <https://doi.org/10.1017/S0022112003006529>.
- Reitz RD, Diwakar R. Effect of drop breakup on fuel sprays. SAE Tech Pap 1986. <https://doi.org/10.4271/860469>.
- Dahms RN, Oefelein JC. The significance of drop non-sphericity in sprays. Int J Multiphase Flow 2016;86:67–85. <https://doi.org/10.1016/j.ijmultiphaseflow.2016.07.010>.
- O'rouke PJ, Amsden AA. The Tab Method for Numerical Calculation of Spray Droplet Breakup International Fuels and Lubricants Meeting and Exposition Toronto, Ontario 2018.
- Som S, Aggarwal SK. Effects of primary breakup modeling on spray and combustion characteristics of compression ignition engines. Combust Flame 2010; 157:1179–93. <https://doi.org/10.1016/j.combustflame.2010.02.018>.
- Paredi D, Lucchini T, D'Errico G, Onorati A, Pickett L, Lacey J. Validation of a comprehensive computational fluid dynamics methodology to predict the direct injection process of gasoline sprays using Spray G experimental data. Int J Engine Res 2020;21:199–216. <https://doi.org/10.1177/1468087419868020>.
- Mohan B, Badra J, Sim J, Im HG. Coupled in-nozzle flow and spray simulation of the Engine Combustion Network Spray-G injector. Int J Engine Res 2020. <https://doi.org/10.1177/1468087420960612>.
- Hwang J, Yasutomi K, Arienti M, Pickett LM. Numerical Investigation of Near Nozzle Flash-Boiling Spray in an Axial-Hole Transparent Nozzle. SAE Tech Pap 2020;2020-April:1–17. [10.4271/2020-01-0828](https://doi.org/10.4271/2020-01-0828).
- Saha K, Som S, Battistoni M. Investigation of homogeneous relaxation model parameters and their implications for gasoline injectors. At Sprays 2017;27: 345–65. <https://doi.org/10.1615/AtomizSpr.2017016338>.
- Saha K, Som S, Battistoni M, Li Y, Pomraning E, Senecal PK. Numerical Investigation of Two-Phase Flow Evolution of In- and Near-Nozzle Regions of a Gasoline Direct Injection Engine During Needle Transients. SAE Int J Engines 2016; 9:1230–40. <https://doi.org/10.4271/2016-01-0870>.
- Raissi M, Perdikaris P, Karniadakis GE. Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. J Comput Phys 2019;378:686–707. <https://doi.org/10.1016/j.jcp.2018.10.045>.
- Ikedo Y, Mazurkiewicz D. Application of neural network technique to combustion spray dynamics analysis. Lect Notes Comput Sci (Including Subser Lect Notes Artif Intell Lect Notes Bioinformatics) 2002;2281:408–25. [https://doi.org/10.1007/3-540-45884-0\\_30](https://doi.org/10.1007/3-540-45884-0_30).
- Zhang Y, Xu S, Zhong S, Bai X-S, Wang H, Yao M. Large eddy simulation of spray combustion using flamelet generated manifolds combined with artificial neural networks. Energy AI 2020;2:100021. <https://doi.org/10.1016/j.egyai.2020.100021>.
- Han Z, Hossain MM, Wang Y, Li J, Xu C. Combustion stability monitoring through flame imaging and stacked sparse autoencoder based deep neural network. Appl Energy 2020;259:114159. <https://doi.org/10.1016/j.apenergy.2019.114159>.
- <https://ecn.sandia.gov/gasoline-spray-combustion/computational-method/mesh-and-geometry/>.
- Duke DJ, Kastengren AL, Matusik KE, Swantek AB, Powell CF, Payri R, et al. Internal and near nozzle measurements of Engine Combustion Network "Spray G" gasoline direct injectors. Exp Therm Fluid Sci 2017;88:608–21. <https://doi.org/10.1016/j.exptthermfluidsci.2017.07.015>.
- <https://ecn.sandia.gov/gasoline-spray-combustion/target-condition/spray-g-plume-orientation/>.
- McCormick RL, Fouts LA, Fioroni GM, Christensen ED, Ratcliff MA, Zigler BT, et al. Co-Optimization of Fuels & Engines. Properties of Co-Optima Core Research Gasolines 2018.

- [35] Westlye FR, Penney K, Ivarsson A, Pickett LM, Manin J, Skeen SA. Diffuse back-illumination setup for high temporally resolved extinction imaging. *Appl Opt* 2017; 56:5028. <https://doi.org/10.1364/ao.56.005028>.
- [36] Hwang J, Weiss L, Karathanassis IK, Koukouvinis P, Pickett LM, Skeen SA. Spatio-temporal identification of plume dynamics by 3D computed tomography using engine combustion network spray G injector and various fuels. *Fuel* 2020;280: 118359. <https://doi.org/10.1016/j.fuel.2020.118359>.
- [37] Pickett LM, Genzale CL, Manin J. Uncertainty quantification for liquid penetration of evaporating sprays at diesel-like conditions. *At Sprays* 2015;25:425–52. <https://doi.org/10.1615/AtomizSpr.2015010618>.
- [38] Weiss L, Wensing M, Hwang J, Pickett LM, Skeen SA. Development of limited-view tomography for measurement of Spray G plume direction and liquid volume fraction. *Exp Fluids* 2020;61:1–17. <https://doi.org/10.1007/s00348-020-2885-0>.
- [39] <https://ecn.sandia.gov/gasoline-spray-combustion/target-condition/primary-spray-g-datasets/>.
- [40] <http://www.philiplaven.com/mieplot.htm>.
- [41] Aleiferis PG, Van Romunde ZR. An analysis of spray development with iso-octane, n-pentane, gasoline, ethanol and n-butanol from a multi-hole injector under hot fuel conditions. *Fuel* 2013;105:143–68. <https://doi.org/10.1016/j.fuel.2012.07.044>.
- [42] Wang Z, Wang B, Jiang C, Xu H, Badawy T. Microscopic characterization of iso-octane spray in the near field under flash boiling condition. *Appl Energy* 2016; 180:598–606. <https://doi.org/10.1016/j.apenergy.2016.07.115>.
- [43] Marco A, Hwang J, Pickett L, Shekhawat Y. A thermally-limited bubble growth model for the relaxation time of superheated fuels. *Int. J. Heat Mass Transf* 2020; 159. <https://doi.org/10.1016/j.ijheatmasstransfer.2020.120089>.
- [44] Price C, Hamzehloo A, Aleiferis P, Richardson D. Numerical modelling of droplet breakup for flash-boiling fuel spray predictions. *Int J Multiph Flow* 2020;125. <https://doi.org/10.1016/j.ijmultiphaseflow.2019.103183>.