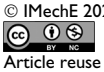
International Journal of ENGINE RESEARCH

Standard Article

International J of Engine Research
1–17


Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/14680874211020292
journals.sagepub.com/home/jer

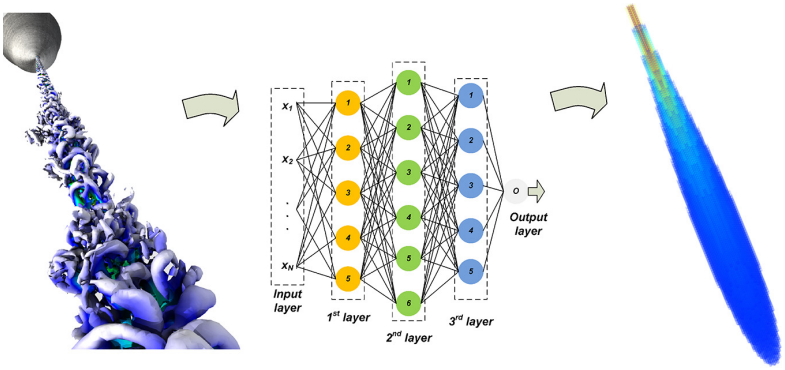

Machine Learning and transcritical sprays: A demonstration study of their potential in ECN Spray-A

Phoevos Koukouvinis[1], Carlos Rodriguez[1], Joonsik Hwang[2], Ioannis Karathanassis[1], Manolis Gavaises[1] and Lyle Pickett[3]

Abstract

The present work investigates the application of Machine Learning and Artificial Neural Networks for tackling the complex issue of transcritical sprays, which are relevant to modern compression-ignition engines. Such conditions imply the departure of the classical thermodynamic perspective of ideal gas or incompressible liquid, necessitating the use of costly and elaborate thermodynamic closures to describe property variation and simulation methods. Machine Learning can assist in several ways in speeding up such calculations, either as a compact, trained thermodynamic model that can be coupled to the flow solver, or as a surrogate predictive tool of spray characteristics. In this work, such applications are demonstrated and their performance is assessed against more traditional approaches. Such applications involve the prediction of macroscopic spray characteristics, for example, the spray penetration over time, or the spray distribution in space and time, and predictions of fluid properties for the thermodynamic states encountered in such applications. Macroscopic characteristics can be adequately predicted by relatively simple network structures, involving just a hidden layer of 3–4 neurons, whereas prediction of thermodynamic states requires several layers of 5–20 neurons each. The results of integrating Artificial Neural Networks in transcritical sprays are rather promising; prediction of thermodynamic properties at pressures greater than 1bar has effectively zero error, yielding simulations indistinguishable from standard tabulated approaches with minimal overhead. When used as a regression method for time-histories either of spray characteristics or spray distributions, the results are within experimental uncertainty of similar experiments, not included in the training dataset.

Graphical abstract



Accelerating atomisation predictions using Machine Learning

[1]School of Mathematics, Computer Science & Engineering, City, University of London, London, UK
[2]Center for Advanced Vehicular Systems (CAVS), Mississippi State University, MS, USA
[3]Combustion Research Facility, Sandia National Laboratory, Livermore, CA, USA

Corresponding author:
Phoevos Koukouvinis, School of Mathematics, Computer Science & Engineering, City, University of London, Northampton Square, London EC1V 0HB, UK.
Email: foivos.koukouvinis.1@city.ac.uk

Keywords
Fuel injection, real-fluid thermodynamics, transcritical mixing, Machine Learning, regression, Artificial Neural Network, multiphase flows

Date received: 8 February 2021; accepted: 1 May 2021

Introduction

Fuel atomisation and mixing is the core in every Internal Combustion Engine (ICE) operation, be it classical automotive engines,[1] aircraft turbines or aerospace propulsors,[2,3] affecting engine performance and emissions. More specifically to automotive and transportation sectors, the understanding and prediction of atomisation phenomena is of utmost importance to Fuel Injection Equipment (FIE) manufacturers, as they are effectively tied to vehicle regulations applicable globally, see the recent Euro-legislation in European Union[4] or the Clean Air Act Amendments in the US.[5] However, atomisation is a rather complex topic on its own,[6,7] both from an experimental and numerical standpoint, given the vast disparity of temporal and spatial scales in the flow field. To provide some context to this perspective, the spatial scales of injection in a modern vehicle may range from 10^{-2} m, corresponding to the cylinder bore-size,[8] to even the 10^{-6} m, corresponding to the miniscule droplets formed due to jet disintegration; typical jet velocities range from 200 m/s to even 700 m/s, for gasoline and diesel FIE respectively, consequently emerging time scales are in the order of 10^{-3}–10^{-6} s. Additional complexities arise from the transient nature of boundaries (moving geometries of the piston/cylinder or valves), phase change processes and finally combustion reactions and emission formation.

Admittedly, much progress has been made the last years to promote understanding of fuel atomisation, from both experimental and computational sides. Indicatively, on the experimental techniques for spray diagnostics, notable recent works involve, for example, the pioneering works of Pickett et al.[9] in an effort to quantify fuel mixing and penetration using Rayleigh Scattering investigations in high pressure and temperature Diesel injection. The works of Weiss et al.[10] and Hwang et al.[11] discussing the development and application of a novel optical 3D tomography technique for reconstructing the 3D liquid volume fraction of gasoline sprays, or Strek et al.[12] exploring the capabilities of X-ray radiography and tomography to determine density distribution of gasoline sprays. Computational models have evolved to capture the turbulent mixing of fuel sprays with the ambient gases, see, for example, the works of Yue et al.[13] in gasoline injectors and sprays, the fundamental work on highly resolved fuel break-up by Agarwal and Trujillo[14] investigating the effect of geometrical features of the injector in spray formation for the Engine Combustion Network (ECN) Spray-A injector,[15] or subsequent investigations of Battistoni et al.[16] in Diesel injector related geometries.

Additional complexities in Diesel fuel injection exist, due to strong property variations the fuel is subjected to. Indicatively, for modern Diesel engines, a pressure variation of more than 2000 bar (modern systems may even reach 3000 bar[17]) to effectively 0bar (cavitating regions in the fuel injector) and a temperature variation from 363 K (high pressure side of the fuel pump) to more than 1000 K (engine cylinder) are expected. Commonly used assumptions of, for example, incompressible fluids are no longer valid and non-ideal thermodynamic effects, such as viscous heating and Joule-Thomson cooling become important.[18,19] Such effects require accurate thermodynamic closures and have been progressively described with more advanced models. Most notable works are the pioneering work of Knudsen et al.[20] utilising cubic Equations of State (EoS) to model fuel properties, followed chronologically by Matheis and Hickel[21] and Yi et al.,[22] both introducing Vapor-Liquid Equilibrium (VLE) calculations for the fuel/gas mixture, and Koukouvinis et al.[23,24] examining higher order thermodynamic models, such as NIST REFPROP[25] and Perturbed Chain Statistical Associating Theory (PC-SAFT).[26] All the aforementioned works, represent rather detailed examples, involving a high computational cost and requiring large computing resources, which are not always readily available on the industrial level. Moreover, such investigations are likely to become more pronounced in the future, given the interest on investigating even more extreme injection pressures, see for example, Vera-Tudela et al.[27] where the development and testing of an experimental injector operating at 5000 bar is discussed.

Even if electrification is a promising candidate for alternative propulsion in vehicles, the interest on ICEs is unlikely to diminish, especially for heavy duty and transportation applications, where high energy density and vehicle autonomy are required,[28,29] hence the need of development of new and better engine concepts will undoubtedly persist.[30] An additional parameter that has to be considered is that in an effort to mitigate the immediate and inevitable environmental consequences, hydrogen-derived, CO_2-neutral synthetic fuels produced using renewable energy sources (*e-fuels*) are increasingly considered for future application.[31] Governments at the European level are investigating and are pushing for H_2 technological leadership,[32] while significant private sector investments,[33,34] have presented their plans to establish zero net emissions e-fuel production plants in Europe. Hence, it is expected

that in the future there will be a need of successfully designing versatile energy production and propulsion systems, operating with a variety of fuels derived from different sources, either synthetic, bio-derived, fossil or mixtures of all the aforementioned. The modelling of such fuels, of which measured properties may not be a-priori known, will require adaptive methods that can describe fundamental characteristics of operation with minimal input; to this goal, the use of Machine Learning is proposed here, as a means of providing quantitative predictions at unknown conditions using pre-existing investigations either from experiments or simulations.

Machine Learning (ML) as a concept has emerged in the 1960s,[35] in the form of learning automata, initially applied on games. Since then it has expanded and evolved greatly to cover a wide array of applications relevant to pattern recognition, feature extraction, classification and regression, by offering the capability of processing enormous amount of data and deriving meaningful relations to be subsequently used, hence their link to data mining. Nowadays it is being applied in the form of Artificial Neural Networks (ANNs), among others, on a vast array of fields spanning across all human activities, from telecommunications and medicine to even social studies and arts.[36] In the field of fluid mechanics, ANNs are currently being investigated as a complementary tool to Computational Fluid Dynamics (CFD) for accelerating design processes. A recent thorough review[37] discusses a variety of applications in fluid mechanics, involving flow feature extraction, turbulence modelling, optimisation and flow control. A variety of recent works demonstrates the capability of Neural Networks in modelling various types of flows, for example, Bhatnagar et al.,[38] Sekar and Khoo[39] the flow around different aerodynamic shapes, although application in multiphase flows is yet limited; indeed, only a handful of published works exist, with notable examples the works of Ansari et al.[40] and Sanchez-Gonzalez et al.[41] who simulated 'dam-break' and liquid sloshing scenarios using Machine Learning, and the work of Chaussonnet et al.[42] who employed Machine Learning for feature recognition of sprays.

The motivation here is to discuss and present Machine Learning techniques that will significantly accelerate e-fuel and FIE development by predicting simultaneously the in-nozzle flow and its effect on the characteristics of vapourising liquid fuel sprays at time scales 3–4 orders of magnitude faster compared to today's state-of-the-art experimentation and CFD simulations. Predictions consider the fuel composition, FIE design and the varying P-T conditions realised in combustion systems. Central to this process is the method predicting the physical properties of the fuel. Despite that the existing fuel property libraries are limited to simplified hydrocarbon components, Equations of State (EoS) using the Perturbed Chain Statistical Associating Fluid Theory (PC-SAFT) have been efficiently applied to simultaneous predictions of nozzle flows and sprays for a wide range of fuel compositions.[23,43–45] They capture the effect of the variation of fuel density, viscosity, heat capacity and conductivity with respect to P-T realised in FIE and ICE (up to 30%, 10^4%, 40% and 60%, respectively), as well as the phase-change characteristics among different fuel components. Having these aspects in mind, the aim is to address a variety of applications relevant to high pressure/temperature fuel mixing sprays, in the following ways:

(a) Regression of thermodynamic properties: the behaviour of fluids at extreme variations of pressure and temperature departs classical assumptions, hence more complex Equations of State (EoS) are required to describe accurately property variation (the interested reader is addressed to, Koukouvinis et al.,[23] Vidal et al.[43] and Perez et al.[46]). Such models, while accurate, are rather time consuming to solve on the fly during simulation. On the other hand, precomputing the EoS in a tabular form and then using it through interpolations, whereas effective (see indicatively Koukouvinis et al.,[23] Kyriazis et al.[47] and Dumbser et al.[48]), can lead to cumbersome tables and interpolations, especially when considering multi-component mixtures. Here, ANNs can provide a tool for creating continuous interpolating functions with much smaller storage footprint than tables and faster calculation than the actual thermodynamic model.
(b) Regression of macroscopic spray characteristics: often, from an engineering perspective, it is of interest to know how certain parameters evolve, such as, for example, the spray penetration over time, and how they are affected by various factors. On this aspect, ANNs can be trained against results over a range of conditions predicted with CFD for estimating the behaviour of sprays at unknown conditions.
(c) Regression of spatio-temporal distributions of sprays: similar to the previous aspect, ANNs can also be trained against two-dimensional or three-dimensional (representing either planar or volumetric) distributions of a feature of interest, for example, fuel mass fraction, over time. In this way, new spray sequences can be generated at unknown conditions, at a fraction of the cost of the numerical simulation, as will be demonstrated later.

The above regression/fitting methods are combined with numerical modelling; however, it is highlighted that these operations can be applied either to experimental data (e.g. measurements of properties, high-speed videos of sprays) and numerical data alike, without loss of generality. The chosen case is Engine Combustion Network (ECN) Spray-A injector and conditions, mainly due to its simplicity (can relatively

accurately be approximated as axis-symmetric) and the fact that its geometry, conditions and data are open to the public (refer to Engine Combustion Network[15,49,50]).

Theoretical details

Short description of the ECN Spray-A

Here a short overview of the Spray-A configuration will be presented to familiarise the reader with the case of interest. The Spray-A injector is a single hole, tapered ($k_{\text{factor}} = 1.5$) Bosch solenoid-activated injector, with nominal orifice diameter $D_{\text{out}} = 90\,\mu\text{m}$ and orifice length of 1 mm (an illustration of the geometry is provided in the section 2.3: '*Fluid simulation model*'). The geometry of the Spray-A injector has been extensively studied with a variety of experimental techniques, including optical microscopy, X-ray tomography and X-ray phase contrast imaging and is publicly available.[15] The standardised spray-A conditions refer to the injection of dodecane fuel, at 1500 bar and 363 K to inert nitrogen atmosphere of 60 bar and 900 K, though parametric variations of this condition also exist, spanning 20.4–80 bar and 700–1200 K.[50]

Thermodynamic model

Accurate modelling of fuel/gas properties is achieved using the Perturbed Chain Statistical Associating Fluid Theory (PC-SAFT) EoS, which is a theoretically-derived model, based on perturbation theory, used to express relations between all the thermodynamic parameters of the fluid. PC-SAFT involves modelling of the intermolecular potential energy of a fluid using a reference term, for repulsive interactions and a perturbation term accounting for attractive interactions. The fluid is assumed to be composed of spherical segments that form molecular chains. The attractive interactions, perturbations to the reference system, are accounted for with the dispersion term. Intermolecular interaction terms accounting for segment self- or cross-associations are ignored. Hence, each component is characterised by three pure component parameters, which comprise a temperature-independent segment diameter, σ, a segment interaction energy, ε/k, and a number of segments per molecule, m; detailed databases for these parameters exist for non-associating fluids, such as hydrocarbons or gases, see Gross and Sadowski[26] and Polishuk.[51]

The PC-SAFT model provides an expression for the residual Helmholtz energy; once such an expression is obtained, all thermodynamic properties can be defined as functions of that expression. The derivation is rather lengthy and out of scope for the present work, however the interested reader can refer to Lemmon and Huber,[25] Gross and Sadowski[26] and Vidal et al.,[43] as an example of the necessary manipulations. The transport properties are estimated based on the residual entropy scaling method for dynamic viscosity[52] and thermal conductivity.[53] Identification of VLE is done through the minimisation of the molar Helmholtz free energy, defined in terms of density, temperature and mixture composition. This optimisation problem is solved via a combination of the successive substitution iteration (SSI) and the Newton minimisation method with a two-step line-search procedure, and the positive definiteness of the Hessian is guaranteed by a modified Cholesky decomposition. The algorithm consists of two stages: first, the mixture is assumed to be in a single-phase state and its stability is assessed via the minimisation of the Tangent Plane Distance (TPD).[25] The stability is tested by purposely dividing the homogeneous mixture in two phases, one of them in an infinitesimal amount and called 'trial phase'. For any feasible two-phase mixture, if a decrease in the Helmholtz free energy is not achieved, then the mixture is stable. In case the minimum of the TPD is found to be negative, the mixture is considered unstable and a second stage of phase splitting takes place consisting on the search for the global minimum of the Helmholtz Free Energy. As a result, the pressure of the fluid and the compositions of both the liquid and vapour phases are calculated.

It has to be highlighted here, that, the expression of Helmholtz energy is a rather lengthy polynomial usually expressed in terms of density, ρ, and temperature, T. Hence, manipulations may require numerical inversion, differentiation, non-linear equation system solution and are not trivial in terms of computational cost; direct evaluation of the thermodynamic model is often avoided during a numerical simulation (see Koukouvinis et al.,[23] Kyriazis et al.[47] and Dumbser et al.[48]). Instead, tabulation, done as a precursor step of the simulation, and interpolation of properties is preferred in practical applications. For example, in previous works of the authors, a thermodynamic table was used, with resolution of $100 \times 400 \times 101$ corresponding to $\log_{10}p$, T, y intervals, for a range of p:(10 Pa–2500 bar) $\times$ T:(280–2000 K) $\times$ y:(0–1), respectively. This results to a rather large file size, containing 4,040,000 entries, which, in ASCII, corresponds to a file size of 1.25 GB (or ~250 MB in binary); only for describing the properties of two components (fuel and ambient gas, an indicative phase diagram is provided in Figure 1). It is apparent that for more complex mixtures higher dimensional tables are necessary, with considerable storage overhead.

Fluid simulation model

The fluid simulation model presented here, is based on the Reynolds Averaged Navier Stokes (RANS) equations solved with Fluent v19.1.[54] The software, customised externally for the thermodynamic model described above, is used to resolve an axisymmetric representation of the Spray-A injector, using the published radial profile over the injector axis.[27] The computational domain is extended downstream by 50mm

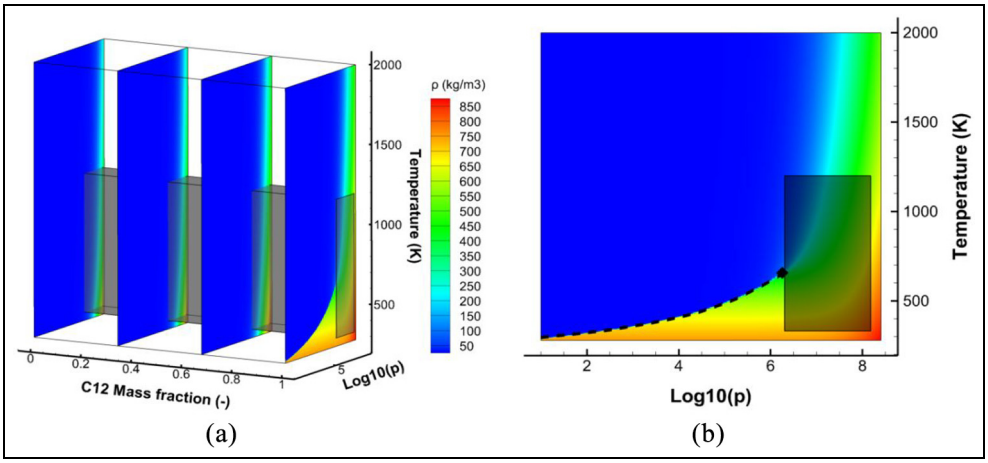


Figure 1. (a) Dodecane/Nitrogen phase diagram, showing density as function of $\log_{10}p$, T and y_{C12} and (b) detailed view of the dodecane phase diagram (slice at $y_{C12} = 1$). The saturation line is indicated with a dashed line and the critical point with a rhombus. In both diagrams, the dark shaded region represents the operating conditions of Spray-A.

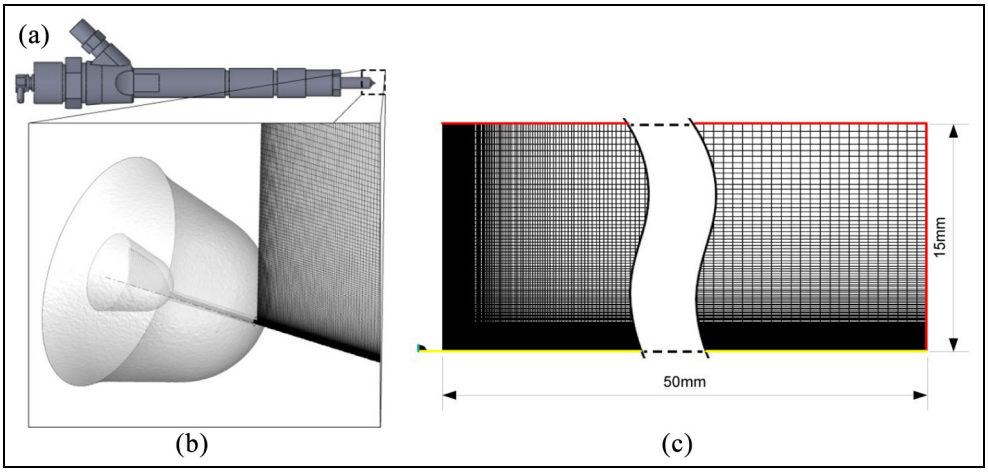


Figure 2. (a) Illustration of the Spray-A injector, (b) magnification at the tip of the injector; the meshed cutting plane shows the computational domain treated in axisymmetric configuration, with the dashed-dotted line being the axis of symmetry, and (c) illustration of the computational mesh, with boundary conditions: mass flow inlet (cyan), symmetry (yellow) and fixed pressure outlet (red).

in the axial and 15 mm in the radial directions (or by $\sim 56\, D_{\text{out}}$ and $\sim 17\, D_{\text{out}}$ respectively), to include part of the spray chamber over which measurements are available, see Figure 2. As shown in Figure 2(c), the computational mesh is block-structured quadrilateral and consists of 50,000 elements, with cell sizes ranging between 1 μm, inside the injector orifice, to 0.5 mm near the farfield, away from the injector axis. At the injector inlet (cyan boundary), a fixed mass flow rate, fuel mass fraction and temperature are imposed. At the fixed pressure outlet (red boundary), a fixed pressure is imposed, whereas temperature and fuel mass fraction are fixed in the case of inflow, or set as zero gradient in the case of outflow.

Briefly stated here, the CFD model corresponds to a multi-component diffuse interface approach, solving for mixture/species continuity, momentum and energy equations, as shown below, respectively:

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{u}) = 0 \tag{1}$$

$$\frac{\partial \rho y_{C12}}{\partial t} + \nabla \cdot (\rho \mathbf{u} y_{C12}) = -\nabla \cdot \mathbf{J} \tag{2}$$

$$\frac{\partial \rho \mathbf{u}}{\partial t} + \nabla \cdot (\rho \mathbf{u} \otimes \mathbf{u}) = -\nabla p + \nabla \cdot \tau \tag{3}$$

$$\frac{\partial \rho E}{\partial t} + \nabla \cdot (\mathbf{u}(\rho E + p)) = \nabla \cdot \left(\lambda_{eff} \nabla T\right) + \nabla \cdot (\tau \cdot \mathbf{u}) + \nabla \cdot (h\mathbf{J}) \tag{4}$$

In the aforementioned equation set, ρ is density, y_{C12} is dodecane mass fraction, $\mathbf{J}$ is the diffusion flux, $\mathbf{u}$ is the velocity vector field, p is pressure, τ is the viscous stress tensor, E is the total energy (as the sum of internal energy, e, and kinetic energy, $|\mathbf{u}|^2/2$), λ_{eff} is the effective heat conductivity and h is enthalpy (as $h = e + p/\rho$). Turbulence treatment is achieved with a RANS closure, in particular the standard k-ε model, which has

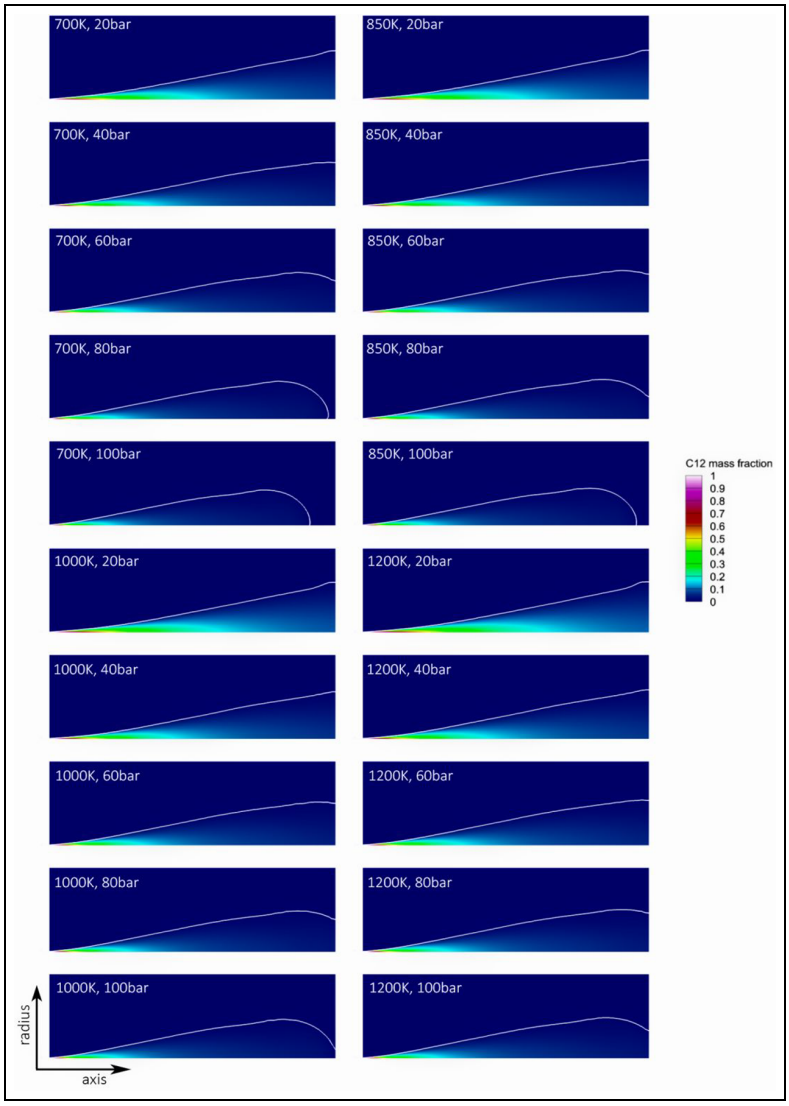


Figure 3. Indicative matrix for the cases examined, shown only for the ending of injection (1.5 ms). The continuous colouring represents the dodecane mass fraction as it mixes with nitrogen. The continuous white line is a representation of the dodecane jet, using a threshold value of 1% dodecane mass fraction. In all images, the horizontal axis represents the axis of symmetry.

demonstrated robust performance in sprays.[55,56] A more detailed discussion on the specific terms, as well as extensive validation of the model against mass fraction and temperature distributions from experimental data made available by ECN (see Engine Combustion Network[49]) is provided in detail in a previous publication by Koukouvinis et al.[24]

Here, the discussed methodology will be used to simulate an array of conditions covering the p/T range of Spray-A parametric investigations. Downstream ambient pressures that will be used are 20–100 bar at 20 bar intervals and temperatures of 700–1200 K at 150 K intervals forming an array of 20 different cases; each case requires ~10 h to compute (~100 k cells on six processes, time step of 0.5 μs, total simulated time 1.5 ms). The mass flow rate imposed at the injector inlet, is adjusted accordingly, based on the ECN tool.[57] An indicative instance list near the end of injection is provided below, in Figure 3. It can be clearly seen that (i) as downstream pressure increases, the jet propagation is retarded (ii) as ambient temperature increases, the jet propagation is slightly accelerated. Both are a direct consequence of the ambient density variation.

Machine learning and Artificial Neural Networks

In the present work, Machine Learning will be employed as a means of regression to approximate thermodynamic property variations of fluids, as they undergo injection in a modern fuel injection system, or to develop surrogate models of spray predictions. To achieve this purpose, an Artificial Neural Network will be defined for each case and trained against existing data, derived from dedicated tools, as outlined above (the thermodynamic model, or the fluid simulation model) and utilising Matlab Neural Network toolbox,[58] as the platform for training and exporting the created networks.

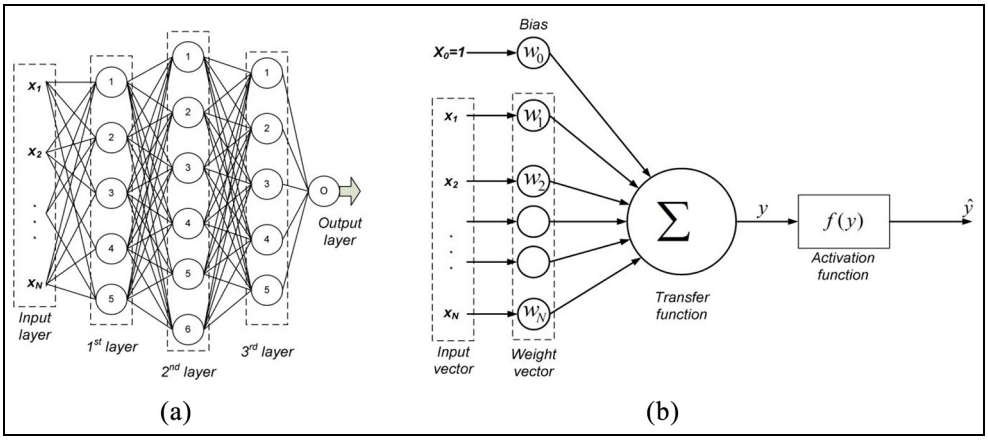


Figure 4. (a) Representation of a neural network consisting of one input layer, three hidden layers and one output layer and (b) typical structure of a neuron.

The building block of an Artificial Neural Network is the neuron or perceptron,[35,59] which represents the smallest processing element; a schematic representation of a network and a neuron is shown in Figure 4(a) and (b) respectively. In general, it can have one or more inputs x_i, that can be either from the environment or other neurons; these inputs are multiplied with a weight w_i and are shifted by a bias, w_0, to provide an intermediate value, y, as follows:

$$y = \mathbf{w}^T\mathbf{x} \tag{5}$$

where $\mathbf{w} = [w_0, w_1, w_2,\ldots w_N]$ and $\mathbf{x} = [x_0, x_1, x_2,\ldots x_N]$, assuming that the bias, w_0, is applied to $x_0 = 1$. Then, the intermediate value, y, is passed through an activation function, to provide the output, $\hat{y}$, of the neuron, to be passed to another neuron or is the output towards the environment. The activation function varies depending on the application; it can be linear, Gaussian, Heavyside, ramp, sigmoid etc. In the present work sigmoid will be used as activation function for all hidden layers and linear for the output layer, as this was found to provide best performance both in terms of training and predictions.

Before using the neural network, it is essential to perform the so-called training where the weights and biases of all involved neurons are determined. Often this is done as an optimisation process, to minimise the error in predicting the desired output for the input vector of a known dataset. In the present work, Bayesian Regularisation backpropagation is used as a method for training the created neural networks. This method offers good performance, is capable of handling complex datasets and minimises the likelihood of overfitting.[60]

Applications and results

Predicting fuel properties

A particular complexity of advanced thermodynamic models is their time-consuming nature when being evaluated. Indeed, considerable overhead arises from the way thermodynamic properties are derived; commonly, thermodynamic properties are expressed as functions of the Helmholtz energy, which traditionally is defined in terms of density, ρ, and temperature, T. However, fluid simulation algorithms will require properties defined with a different variable set, pressure, p, and temperature, T, or density, ρ, and internal energy, e. This necessitates the use of numerical inversion, as the complex thermodynamic model cannot be analytically inverted (multiple roots), adding further to the computational cost. Moreover, calculation of VLE of two-phase mixtures involves iterative solution of multiple non-linear equations. The interested reader is addressed to Vidal et al.,[61] for a more detailed discussion associated to the cost of evaluating complex thermodynamic models, such as PC-SAFT.

Whereas the aforementioned factors may not be that important for a single evaluation, the additional computational overhead becomes considerable in a simulation, where the thermodynamic model may be evaluated several times at each time step. Indicatively for simulations of sprays, the injection duration is ~1–2 ms and the computational time step size may vary from 10^{-6} s (implicit solvers) to 10^{-9} s (explicit solvers), showing that there will be at least 10^3–10^6 function calls to the thermodynamic subroutines. In practice, even the lowest estimation will be significantly higher, as in implicit solvers the iterative process necessitates multiple evaluations per time step.

A solution to this issue, which is employed in the literature is the pre-tabulation of properties in a structure that can be interpolated efficiently. Such examples involve structured tables with adaptive refinement,[48] unstructured tables that conform to saturation curve,[47] or very fine structured tables that can be accessed efficiently through indexing.[23] Admittedly, such methods lead to a dramatic reduction of computational cost, offering the capability of highly accurate simulations, however the storage of the tabulated properties can be

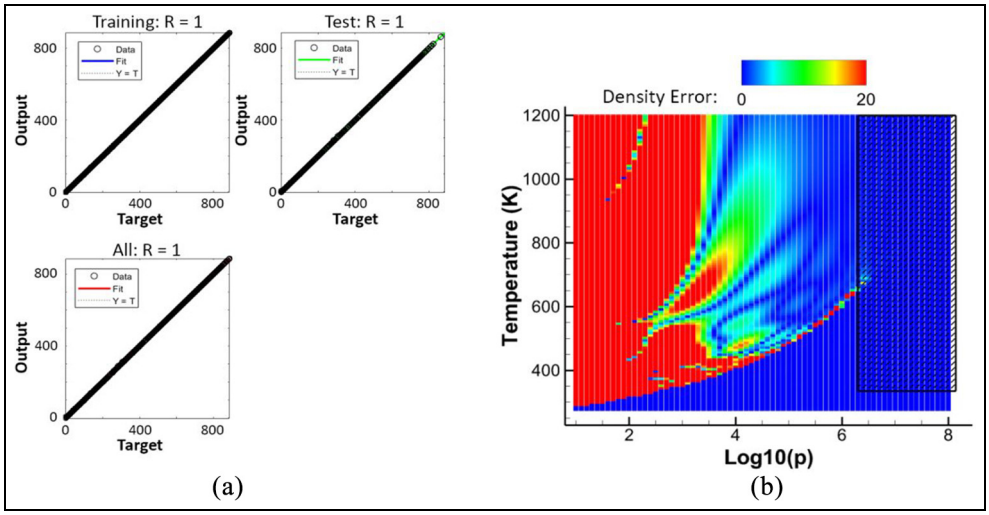


Figure 5. (a) Regression quality for reproducing dodecane density with ANN as function of log10 (pressure) and temperature. 'Training' refers to the regression of the dataset used for training (90% of data points), 'test' refers to the dataset used for testing the ANN performance (5%). As shown, the regression is practically perfect for both training and testing data and (b) relative error (%) of the ANN predicted dodecane density, for a range of temperatures and pressures. The hatched region represents Spray-A operating conditions.

problematic. Adequate resolution of the table implies a large table size with many sampling points; this can easily reach file sizes of several GBs. Inclusion of multiple species for practical spray applications will further increase the dimensionality of the tables, rendering their file size even larger. For example, the tables used in two-component Spray-A simulations[24] had a file size of ~1.25 GB; inclusion of only one additional specie with similar resolution would imply a table size of ~126 GB. It becomes apparent that tabulation will be very cumbersome when considering practical applications of actual fuels or fuel surrogates,[43] containing multiple gaseous or fuel species.

Artifical Neural Networks offer a very attractive alternative here, as a regression model for thermodynamic properties. In particular, tables can be used for the training of a sufficiently complex network, that is much more versatile in terms of file size (several KB) and minimal computational overhead. As an example, two cases will be presented here, one for the prediction of dodecane properties for different p, T conditions and for dodecane/nitrogen mixtures at different p, T, y combinations.

For dodecane modelling, a property table of 40,000 elements was used, spanning over a pressure range of (10 Pa–2500 bar) with 100 elements at regular $\log_{10}p$ intervals and temperature range of (280–2000 K) with 400 elements at constant T intervals. The Artificial Neural Network was structured to receive as an input combinations of ($\log_{10}p$, T) and output ρ, though application is straightforward for other thermodynamic variables as well. The minimal neural network structure that was found to have a very good performance in capturing both sharp density variations near the saturation curve and the smooth density transition beyond the critical point was a three layer structure with (4, 10, 20) hidden neurons respectively. Similarly to the above, for dodecane/nitrogen mixture modelling a property table of ~4.10^6 elements, was used, with resolution of 100 × 400 × 101 corresponding to regular $\log_{10}p$, T, y intervals, for a range of p:(10 Pa–2500 bar) × T:(280–2000K) × y:(0–1). In that case, the Artificial Neural Network was structured to receive as an input combinations of ($\log_{10}p$, T, y) and output ρ. The minimal neural network structure that was found to have a very good performance in capturing both sharp density variations near the saturation curve of pure dodecane and the smooth density transition in the multi-component mixture or pure nitrogen was a four layer structure with (20, 5, 10, 20) hidden neurons respectively. From these tables, 90% of the data were used for training, 5% for validation and 5% for testing.

In Figure 5(a), an indicative representation of the obtained regression is shown. It is clear that the ANN can almost perfectly capture the details of the dodecane phase diagram (pressure and temperature dependence only) in just 500 training iterations, done in less than 10 min as a single process, with a regression validation index of effectively unity. As a further test, the quality of the phase diagram obtained from ANN is assessed against the more classical method of interpolation and tabulation. Here, a sampling grid of $\log_{10}p$:(1–8) and T:(280–1200) with 71 and 123 intervals respectively, is used to evaluate the performance of both methods. The resolution is chosen different from the data table, so as to avoid sampling data points used for training. As shown in Figure 5(b), error is effectively zero at liquid

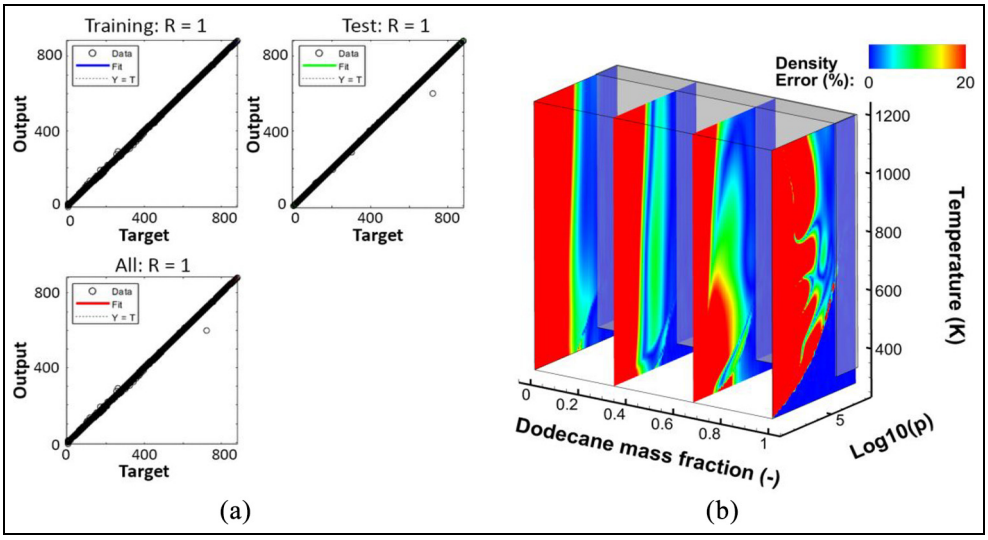


Figure 6. (a) Regression quality for reproducing dodecane/nitrogen density with ANN as function of log10 (pressure), temperature and composition (dodecane mass fraction). 'Training' refers to the regression of the dataset used for training (90% of data points), 'test' refers to the dataset used for testing the ANN performance (5%). As shown, the regression is practically perfect for both training and testing data, but some scatter can be observed and (b) relative error (%) of the ANN predicted mixture density, for a range of temperatures and pressures. The shaded region represents Spray-A operating conditions.

and supercritical states; it becomes considerable in the vaporous phase, at pressures below 10^5 Pa; this is mainly because at such conditions dodecane vapour density is very low (below 0.05 kg/m^3), hence small absolute errors are translated to large relative errors.

In Figure 6(a), a similar comparison is shown, this time for a dodecane/nitrogen mixture, which is admittedly more demanding, due to the larger data set and higher dimensionality. Indeed, this time 700 iterations were required and almost 6hrs to compute. However, despite the added complexity, the agreement is still very good, shown for a sampling grid of $\log_{10}p$:(1–8), T:(280–1200) and y_{C12}:(0–1) with 71, 123 and 21 intervals, respectively. Effectively at medium-high pressures (>1 bar) the error is zero, with issues arising at low pressures/low densities. ANN performance for training other thermodynamic variables is similar to density; in fact, for other variables, such as enthalpy, representation is even better. The only exception is thermodynamic derivatives and especially partial derivative of density in respect to temperature and pressure; irrespectively of attempts to scale the data, so as to assist the optimiser in finding the proper weights/biases, a good regression was not possible to be obtained with this network structure. In any further applications to be discussed hereafter, numerical differentiation of the trained network output for density and enthalpy will be used instead.

Combining machine learning and flow simulation

In this section, the p, T, y regression model developed for dodecane/nitrogen mixture is integrated as a User Defined Function to the flow solver to provide predictions for the Spray-A case. This is to assess the ANN thermodynamic functions' behaviour in a practical case, examine solver stability and evaluate solution quality against the more standardised tabulation and interpolation approach.

In Figure 7 an indicative comparison of flow instances is provided between the two methods. All images show the difference between the tabulated approach (which is considered as the reference) against the ANN implementation; the top part of the images shows % difference in mass fraction and the bottom in temperature distribution. It is clearly visible that results are effectively identical, with deviations of 1% or less in terms of mass fraction and temperature. For a more practical comparison, estimations of the liquid and vapour penetration are provided and are compared with reference experimental data (see Engine Combustion Network[62,63]) in Figure 8. Both vapour and liquid are estimated based on the ECN guidelines (for vapour $y_{C12} = 0.1\%$[64] and for liquid 0.15% liquid volume fraction[65]). Again here, it is clearly demonstrated that the ANN fitting and the tabular interpolation give identical results to the point of being indiscernible between each other. The results are very close and/or within experimental uncertainties. In terms of computational overhead, the use of ANN is relatively small, increasing computational time during execution by ~10%–15% compared to the table interpolation and the main overhead being training the network. Even in that case though, training took ~6 h, whereas the calculation of the input thermodynamic dataset required a couple of days, so as an additional step it does not constitute a large increase in computational cost.

For the interested reader, the regression networks used in this section are provided as C functions with an example calling function in the appendix.

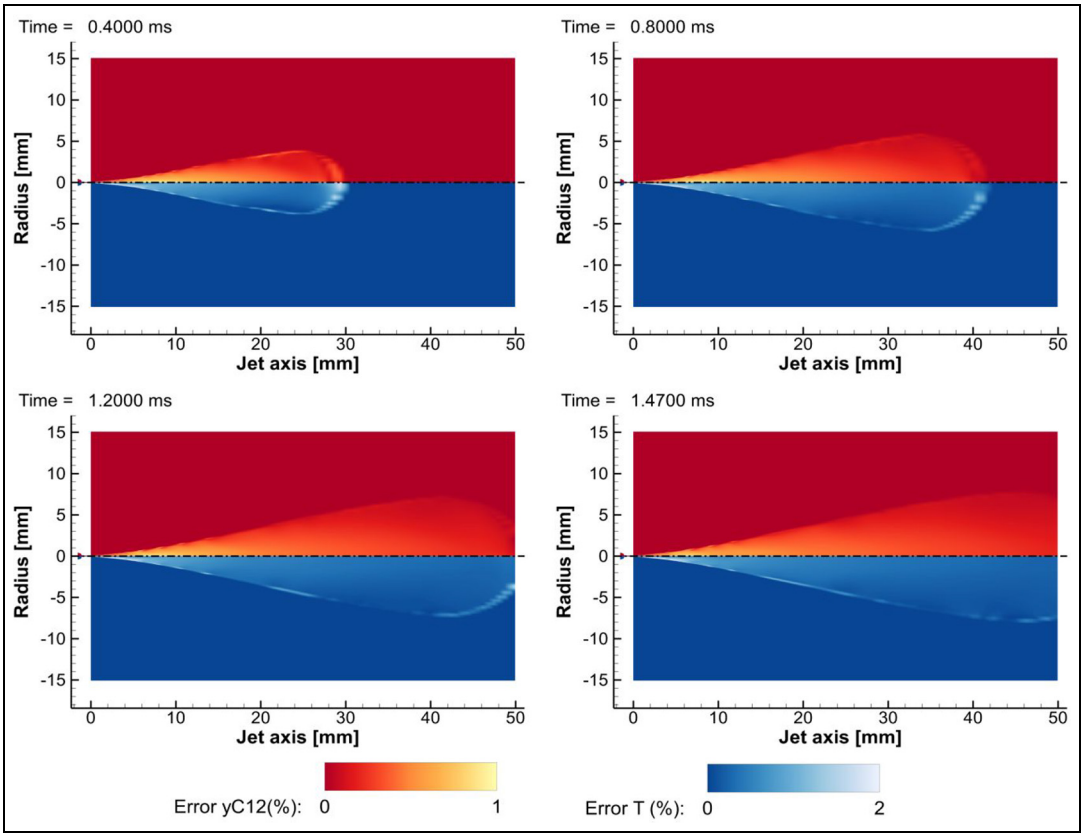


Figure 7. The difference between CFD calculations using the tabulated approach and ANN for thermodynamic properties, for different time instances. Each figure shows at the top side the difference in mass fraction distribution and at the bottom the respective difference in temperature distribution. The axis of the spray is indicated with the dashed dotted line. Standard Spray-A case, dodecane injected at 1500 bar/363 K to nitrogen 60 bar/900 K.

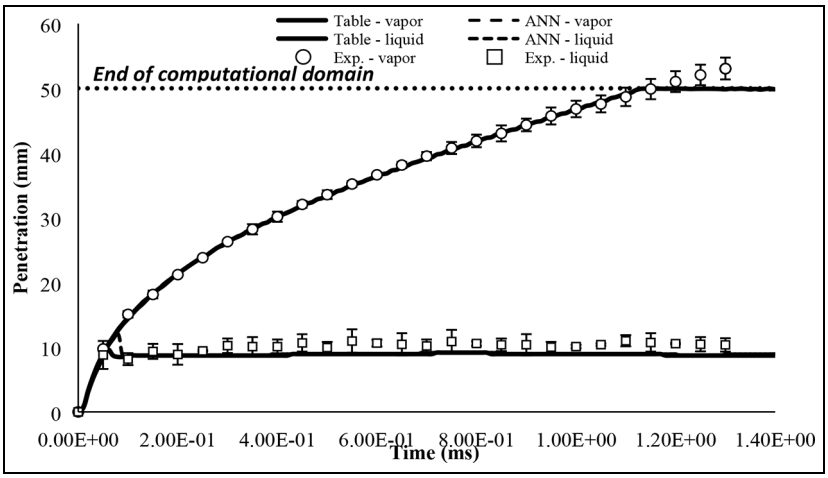


Figure 8. Comparison of the spray (both liquid and vapour) penetration, using tabulated interpolation and ANN, for the standard Spray-A case, dodecane injected at 1500 bar/363 K to nitrogen 60 bar/900 K. Experimental results are also provided for reference with associated error bars, as obtained from here: vapour,[63] liquid.[62]

Predicting macroscopic spray characteristics

Another aspect of ML and ANN that can be exploited, is regression/fitting of results, computational or experimental, to derive predictions at unknown conditions. For example, ANNs can be trained against a characteristic of interest for a specific case, say the penetration length, width of a spray, etc against different parameters, such as ambient conditions, fuel type, injection pressure, etc.

In this section, an example is provided for predicting spray penetration over time, as derived from the cases described in Figure 3. Spray penetration is estimated for all cases over time, using the 0.1% dodecane mass

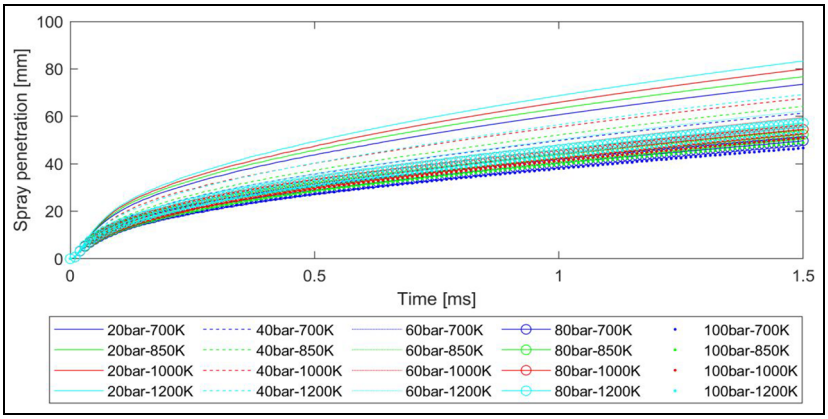


Figure 9. Summary of spray penetration for all parametric investigations of Spray-A cases, based on the ambient conditions.

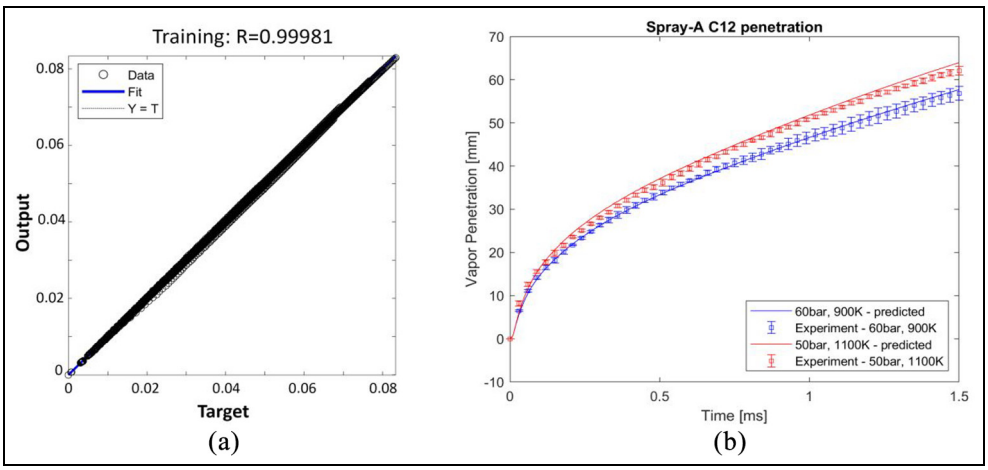


Figure 10. (a) Regression plot of the trained ANN output for predicting spray penetration as function of ambient conditions and (b) example predictions of the trained ANN for two different conditions where experiments are available.

fraction criterion, as shown in Figure 9. The time dependent penetration can be used as a vector (here 20 different cases by 151 time instances) to act as the output target of a specifically trained ANN, with input the different p, T conditions.

Given the relatively straightforward relation between the ambient conditions and the spray penetration over time (see relevant discussion in section 2.3), only a single hidden layer of two neurons was found enough to obtain a decent model-regression quality and good predictions, using all the data set of 20 injection cases entirely for training. To demonstrate the capability of the trained ANN in predicting unknown conditions, two configurations were chosen for which experiments are available, namely at ambient conditions of 50 bar and 1100 K[66] and 60 bar and 900 K.[63] Note that such conditions have not been used for training, so this test acts as validation of the trained model. As shown in Figure 10(b), the predictions are very close to the actual experiments. In terms of computational cost, it is trivial, as training lasts less than a couple of seconds and predictions of penetration over time for unknown conditions are performed effectively instantaneously.

Predicting spatial distributions of quantities of interest

An extension of the previous technique is the prediction of 2D (or even 3D) distributions of sprays using ML. In the previous section, a single ANN was trained to predict the time history of a macroscopic characteristic of a spray. Here, a matrix of ANNs will be trained to predict all elements of (i, j)-arrays, effectively representing pixels of image sequences, coloured according to a quantity of interest (here mass fraction of dodecane), derived from simulations, representing the spray distribution. Rectangular images are sampled over a range of 0–50 mm and 0–10 mm at the axial and radial directions of the spray, with 401 and 81 elements respectively, every 10 μs, leading to a total of 150 frames for each case. The data for each condition are organised as a 3D array, with i-index and j-index representing axial and radial directions respectively and k-index the time instance (see Figure 11).

For each pixel, that is for each (i, j)-combination, a time vector of k-elements is constructed. For each pixel, training is performed using the time vectors corresponding to the associated input vector; in the present case the ambient pressure and temperature

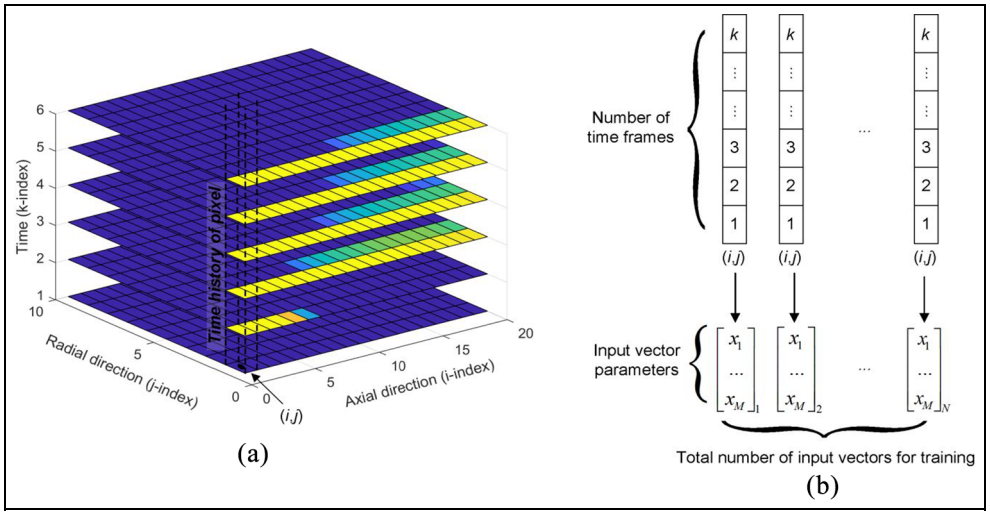


Figure 11. (a) Discrete pixel representation of the image sequence, obtained by resampling the simulation data for a single condition. The representation is in the form of a 3D array, with *i* and *j* indexes corresponding to spatial dimensions and *k* index to time and (b) the time history of each pixel is correlated to the associated input vector and is used to train a pixel-specific network.

conditions. Since the training of each pixel is independent of the other pixels, training is straightforward to be performed in parallel. In the end, an (*i*, *j*)-array of trained ANNs is produced, capable of reproducing the time history of each corresponding pixel, having the appropriate input vector (here the ambient conditions of the spray chamber). For the training of the networks, a single hidden layer with three neurons was employed. The training of the network required ~2 days with six processes. After the training, a new image sequence at untested conditions can be derived within ~5 min. To assess the quality of the reconstruction, a direct comparison with CFD results is performed for two conditions that have not been used for training (a) 50 bar and 1100 K and (b) 60 bar and 900 K, see Figures 12 and 13, respectively.

As shown in the figure, global agreement between the simulation and the ANN prediction is satisfactory; differences are in general well below 5%, with larger error values at very localised regions (front of the spray, or sides of the spray). Still, this applies for specific instances, whereas in general the spray core distribution is well predicted (effectively zero deviation).

To illustrate this effect even clearer, an indicative comparison of spray predictions using ANN, against experimentally quantified vapour mass fraction distributions is shown in Figure 14; as shown, even if a bit noisy, the agreement is within experimental uncertainties.

Discussion and conclusion

In this work, several options of implementing ML techniques, such as ANNs, for the prediction of high pressure/high temperature sprays are discussed. Several techniques are analysed for regression/fitting purposes that can be used as alternatives to tabulation of properties, or predicting quantities of interest in unknown/new conditions.

In terms of surrogate models for approximating thermodynamic functions, ANNs demonstrated a good performance in predicting properties both as function of pressure/temperature and pressure/temperature/mass fraction of a two-component mixture. For pressures higher than 1 bar, deviation from reference values, using tabular interpolation, is effectively zero. The time required for training ANNs against thermodynamic datasets is much smaller than the derivation of the datasets themselves, so there is only a small overhead there. When used in simulations, even though the ANN regression involves overhead during evaluation, the additional computational cost is relatively small. On the other hand, the only alternatives are either evaluating the complex thermodynamic model through the EoS, which is rather demanding, or performing interpolations from tables, the latter being the only practical way in complex cases. Tables have a large storage footprint and become very cumbersome as the interpolation dimensionality increases (interpolating pressure, temperature and the mass fractions of multiple components). ANN seems to be a very attractive option there, as the trained network has a size of several KBs, over a thousand or million times smaller than a table.

Further applications where ANN can be used, is the development of surrogate models that can be trained using existing, validated data, to derive either macroscopic spray characteristics, or even spatial distributions of the spray, over time. In the cases demonstrated, relatively simple networks (2–3 neurons with one hidden layer), can describe adequately the spray penetration and the mass fraction distribution for Spray-A. It is important to mention that the training for a single vector output network, as in the spray penetration prediction, takes a couple of seconds and its evaluation

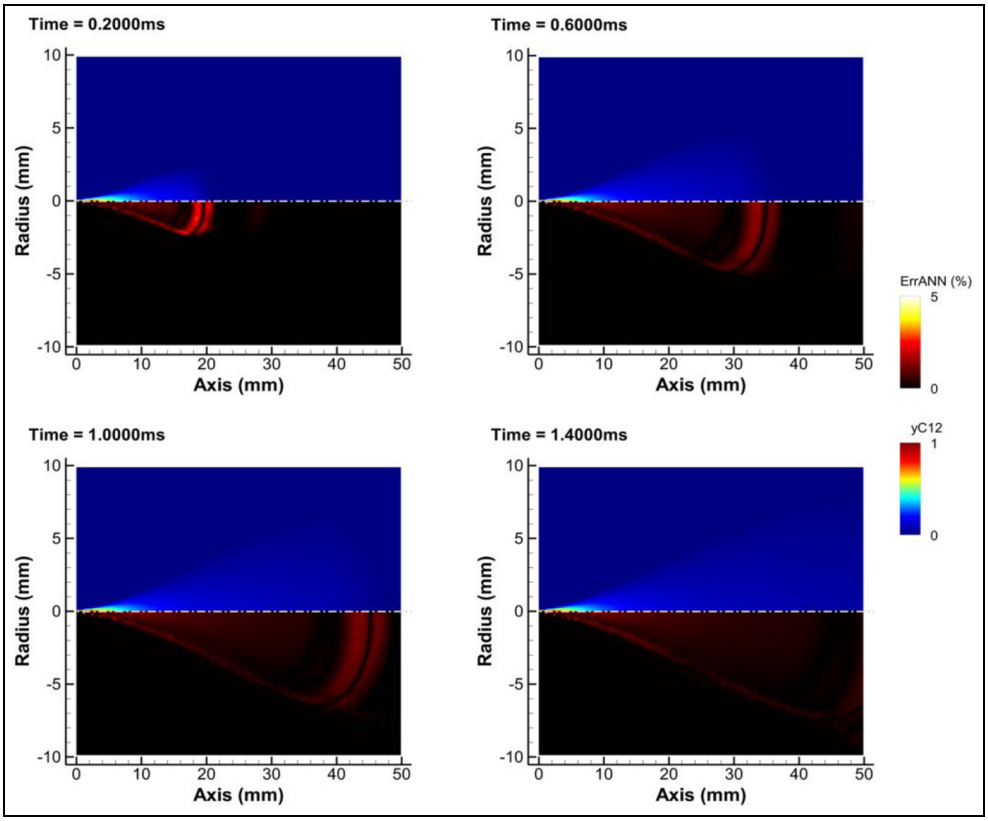


Figure 12. Evaluation of the performance of the ANN predictions; 900 K and 60 bar pressure. Top: Mass fraction distribution (y_{C12}– %), bottom: Absolute difference in mass fraction distribution (err_{ANN}– %). The grey dashed-dotted line indicates the axis of symmetry of the jet.

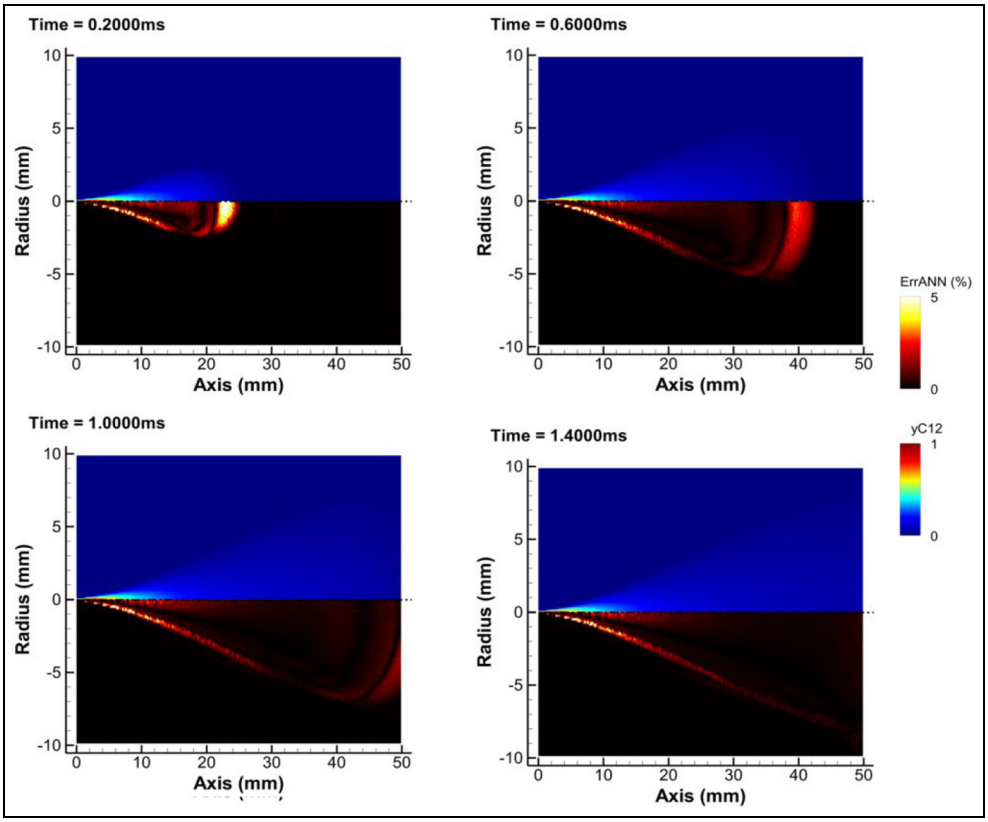


Figure 13. Evaluation of the performance of the ANN predictions; 1100 K and 50 bar pressure. Top: Mass fraction distribution (y_{C12}– %), bottom: Absolute difference in mass fraction distribution (err_{ANN}– %). The grey dashed-dotted line indicates the axis of symmetry of the jet.

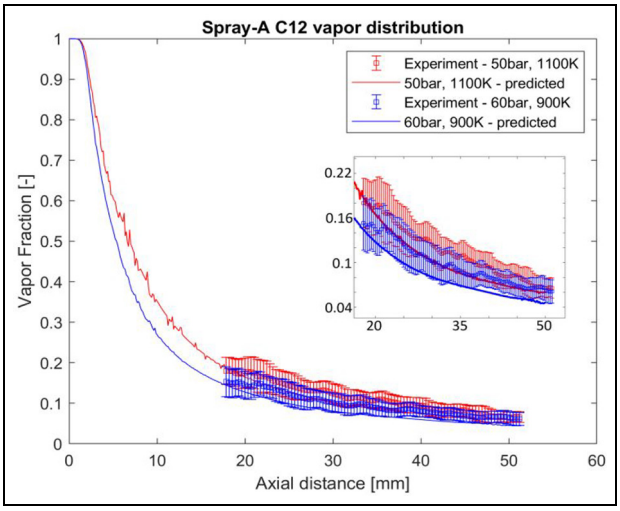


Figure 14. Reconstructed spatial distribution of dodecane mass fraction with ANN, compared against experimental data (available from[49]) at the end of injection. The comparison is shown for two cases, 900 K/60 bar and 1100 K/50 bar, not used in ANN training. The magnified inset shows with greater detail the comparison in the area of the experimental data.

happens instantaneously, whereas the training for two dimensional arrays of time series (three total dimensions) is much more demanding, however any subsequent evaluation takes only several minutes, much faster than any simulation could produce.

Another aspect worth mentioning is that a very attractive characteristic of ML is versatility; the discussed methods here, even if applied in conjunction with simulations, are not restricted by the input type. It can be any form of data in an appropriate format, as high-speed image sequences from optical techniques, X-ray radiographies, or even volumetric sequences (as in the work of Hwang et al.[11] and Weiss et al.[10]), although in the later case the computational cost of training will be considerably higher due to the much larger number of voxels to be trained.

Finally, it should be highlighted that this work can in no way claim that it has explored the full capability of ML in sprays; admittedly, there are many different ways where ML can be integrated, using classification techniques to identify spray characteristics, time series neural networks for predicting the evolution of sprays, or even convolutional neural networks and physics informed deep-learning for predicting flow fields. Indeed, ML constitute a vast array of techniques that can be adapted to a multitude of problems and only recently have such techniques started to be explored in multiphase flows.

Acknowledgements

The corresponding author would like to thank Dr. I Malgarinos for the helpful comments. The corresponding author would also like to thank Sandia National Laboratories, the US Department of Energy (DoE) and the staff at Combustion Research Facility for hosting him. 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.

Declaration of conflicting interests

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

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Sklodowska-Curie grant agreement No 748784 (project UNIFIED) and 794831 (project AHEAD).

ORCID iD

Phoevos Koukouvinis https://orcid.org/0000-0002-3945-3707

Supplemental material

Supplemental material for this article is available online.

References

1. Xia J, Huang Z, Xu L, Ju D and Lu X. Experimental study on spray and atomization characteristics under subcritical, transcritical and supercritical conditions of marine diesel engine. *Energy Convers Manag* 2019; 195: 958–971.
2. Smith JJ and Mayer WOH. Fundamentals of Supercritical Mixing and Combustion of Cryogenic Propellants. In: M Popp, J Hulka, V Yang and M Habiballah (eds) *Liquid rocket thrust chambers*. Reston,VA: American Institute of Aeronautics and Astronautics, 2004, pp.339–367.
3. Smith JJ, Schneider G, Suslov D, Oschwald M and Haidn O. Steady-state high pressure LOx/H2 rocket engine combustion. *Aerosp Sci Technol* 2007; 11(1): 39–47.
4. Hooftman N, Messagie M, Van Mierlo J and Coosemans T. A review of the European passenger car regulations – Real driving emissions vs local air quality. *Renew Sustain Energy Rev* 2018; 86: 1–21.
5. Gerard D and Lave LB. Implementing technology-forcing policies: the 1970 Clean Air Act Amendments and the introduction of advanced automotive emissions controls in the United States. *Technol Forecast Soc Change* 2005; 72(7): 761–778.
6. Gorokhovski M and Herrmann M. Modeling primary atomization. *Annu Rev Fluid Mech* 2008; 40(1): 343–366.
7. Benjamin MA, Jensen RJ and Arienti M. Review of atomization: current knowledge and future requirements for propulsion combustors. *Atomiz Sprays* 2010; 20(6): 485–512.
8. van Basshuysen R and Schäfer F. *Internal combustion engine handbook*. 2nd ed. Warrendale: SAE International, 2004.

9. Pickett LM, Manin J, Genzale CL, Siebers DL, Musculus MPB and Idicheria CA. Relationship Between Diesel Fuel Spray Vapor Penetration/Dispersion and Local Fuel Mixture Fraction. *SAE Int J Engines* 2011; 4(1): 764–799.
10. Weiss L, Wensing M, Hwang J, Pickett LM and Skeen SA. Development of limited-view tomography for measurement of Spray G plume direction and liquid volume fraction. *Exp Fluids* 2020; 61(2): 51.
11. Hwang J, Weiss L, Karathanassis IK, Koukouvinis P, Pickett LM and 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.
12. Strek P, Duke D, Swantek A, Kastengren A, Powell CF and Schmidt DP. X-ray radiography and CFD studies of the spray G injector. SAE technical paper 2016-01-0858, 2016.
13. Yue Z, Battistoni M and Som S. Spray characterization for engine combustion network Spray G injector using high-fidelity simulation with detailed injector geometry. *Int J Engine Res* 2020; 21(1): 226–238.
14. Agarwal A and Trujillo MF. The effect of nozzle internal flow on spray atomization. *Int J Engine Res* 2020; 21(1): 55–72.
15. Engine Combustion Network. Spray A nozzle geometry, https://ecn.sandia.gov/diesel-spray-combustion/target-condition/spray-a-nozzle-geometry/ (2012, accessed 18 May 2021).
16. Battistoni M, Som S and Powell CF. Highly resolved Eulerian simulations of fuel spray transients in single and multi-hole injectors: nozzle flow and near-exit dynamics. *Fuel* 2019; 251: 709–729.
17. Boccardo G, Millo F, Piano A, et al. Experimental investigation on a 3000 bar fuel injection system for a SCR-free non-road diesel engine. *Fuel* 2019; 243: 342–351.
18. Strotos G, Koukouvinis P, Theodorakakos A, Gavaises M and Bergeles G. Transient heating effects in high pressure diesel injector nozzles. *Int J Heat Fluid Flow* 2015; 51: 257–267.
19. Theodorakakos A, Strotos G, Mitroglou N, Atkin C and Gavaises M. Friction-induced heating in nozzle hole micro-channels under extreme fuel pressurisation. *Fuel* 2014; 123: 143–150.
20. Knudsen E, Doran EM, Mittal V, Meng J and Spurlock W. Compressible Eulerian needle-to-target large eddy simulations of a diesel fuel injector. *Proc Combust Inst* 2017; 36(2): 2459–2466.
21. Matheis J and Hickel S. Multi-component vapor-liquid equilibrium model for LES of high-pressure fuel injection and application to ECN Spray A. *Int J Multiph Flow* 2018; 99: 294–311.
22. Yi P, Yang S, Habchi C and Lugo R. A multicomponent real-fluid fully compressible four-equation model for two-phase flow with phase change. *Phys Fluids* 2019; 31(2): 026102.
23. Koukouvinis P, Vidal-Roncero A, Rodriguez C, Gavaises M and Pickett L. High pressure/high temperature multiphase simulations of dodecane injection to nitrogen: Application on ECN Spray-A. *Fuel* 2020; 275: 117871.
24. Koukouvinis P, Vidal-Roncero A, Rodriguez C, Gavaises M and Pickett L. Enhancing the predictive capabilities for high P/T fuel sprays; non-ideal thermodynamic modelling using PC-SAFT. *ERCOFTAC Bulletin* 2020; 124.
25. Lemmon EW and Huber ML. Thermodynamic properties of n-dodecane. *Energy Fuels* 2004; 18(4): 960–967.
26. Gross J and Sadowski G. Perturbed-chain SAFT: an equation of state based on a perturbation theory for chain molecules. *Ind Eng Chem Res* 2001; 40(4): 1244–1260.
27. Vera-Tudela W, Haefeli R, Barro C, Schneider B and Boulouchos K. An experimental study of a very high-pressure diesel injector (up to 5000 bar) by means of optical diagnostics. *Fuel* 2020; 275: 117933.
28. Reitz RD, Ogawa H and Payri R, x IJER editorial: the future of the internal combustion engine. *Int J Engine Res* 2020; 21(1): 3–10.
29. Kalghatgi G. Is it really the end of internal combustion engines and petroleum in transport?. *Appl Energy* 2018; 225: 965–974.
30. Bloomberg. Electric vehicle outlook 2019, https://about.bnef.com/electric-vehicle-outlook-table-of-contents/ (2019, accessed 6 August 2020).
31. European Union. Communication COM/2020/301: a hydrogen strategy for a climate-neutral Europe, [Online]. https://ec.europa.eu/knowledge4policy/publication/communication-com2020301-hydrogen-strategy-climate-neutral-europe_en (2020, accessed 18 May 2020).
32. Hoppe J. MAN energy solutions partners up with the Bavarian Hydrogen Council. *MAN Energy Solutions SE*, https://www.man-es.com/company/press-releases/press-details/2019/09/05/man-energy-solutions-partners-up-with-the-bavarian-hydrogen-council (2019, accessed 18 May 2020).
33. Norsk e-fuel. Supplying your renewable fuel. Unlimited. On the road to climate neutral transportation, https://www.norsk-e-fuel.com/en/ (2020, accessed 18 May 2020).
34. Repsol. Repsol to develop two major emissions-reductions projects in Spain, https://www.repsol.com/en/press-room/press-releases/2020/repsol-to-develop-two-major-emissions-reductions-projects-in-spain.cshtml (2020, accessed 18 May 2021).
35. Alpaydin E. *Introduction to machine learning*. Cambridge, MA: MIT Press, 2014.
36. Abiodun OI, Jantan A, Omolara AE, Dada KV, Mohamed NA and Arshad H. State-of-the-art in artificial neural network applications: a survey. *Heliyon* 2018; 4(11): e00938.
37. Brunton SL, Noack BR and Koumoutsakos P. Machine learning for fluid mechanics. *Annu Rev Fluid Mech* 2020; 52(1): 477–508.
38. Bhatnagar S, Afshar Y, Pan S, Duraisamy K and Kaushik S. Prediction of aerodynamic flow fields using convolutional neural networks. *Comput Mech* 2019; 64(2): 525–545.
39. Sekar V and Khoo BC. Fast flow field prediction over airfoils using deep learning approach. *Phys Fluids* 2019; 31(5): 057103.
40. Ansari A, Boosari SSH and Mohaghegh SD. Successful implementation of artificial intelligence and machine learning in multiphase flow smart proxy modeling: two case studies of gas-liquid and gas-solid CFD models. *J Pet Environ Biotechnol* 2020; 101(1): 1–8.
41. Sanchez-Gonzalez A, Godwin J, Pfaff T, Ying R, Leskovec J and Battaglia PW. Learning to simulate complex physics with graph networks [Online], http://arxiv.org/abs/2002.09405 (2020, accessed 18 May 2021).
42. Chaussonnet G, Lieber C, Yikang Y, et al. Towards deepspray: using convolutional neural network to post-process shadowgraphy images of liquid atomization.

https://arxiv.org/pdf/1910.11073.pdf (2019, accessed 18 May 2021).
43. Vidal A, Rodriguez C, Koukouvinis P, Gavaises M and McHugh MA. Modelling of diesel fuel properties through its surrogates using perturbed-chain, statistical associating fluid theory. *Int J Engine Res* 2020; 21(7): 1118–1133.
44. Rodriguez C, Rokni HB, Koukouvinis P, Gupta A and Gavaises M. Complex multicomponent real-fluid thermodynamic model for high-pressure diesel fuel injection. *Fuel* 2019; 257: 115888.
45. Vidal-Roncero A. Preferential cavitation and friction-induced heating of multi-component diesel fuel surrogates up to 450MPa. *Int J Heat Mass Transf* 2020; 166: 120744.
46. Perez AG, Coquelet C, Paricaud P and Chapoy A. Comparative study of vapour-liquid equilibrium and density modelling of mixtures related to carbon capture and storage with the SRK, PR, PC-SAFT and SAFT-VR Mie equations of state for industrial uses. *Fluid Phase Equilib* 2017; 440: 19–35.
47. Kyriazis N, Koukouvinis P and Gavaises M. Numerical investigation of bubble dynamics using tabulated data. *Int J Multiph Flow* 2017; 93: 158–177.
48. Dumbser M, Iben U and Munz C-D. Efficient implementation of high order unstructured WENO schemes for cavitating flows. *Comput Fluids* 2013; 86: 141–168.
49. Engine Combustion Network. Rayleigh scattering n-dodecane/ambient mixing images for Spray A and several other experimental conditions, https://ecn.sandia.gov/data/bkldaal4mixing/ (accessed 18 May 2021).
50. Engine Combustion Network. Diesel data search page, https://ecn.sandia.gov/ecn-data-search/ (accessed 18 May 2021).
51. Polishuk I. Standardized critical point-based numerical solution of statistical association fluid theory parameters: the perturbed chain-statistical association fluid theory equation of state revisited. *Ind Eng Chem Res* 2014; 53(36): 14127–14141.
52. Lötgering-Lin O and Gross J. Group contribution method for viscosities based on entropy scaling using the perturbed-chain polar statistical associating fluid theory. *Ind Eng Chem Res* 2015; 54(32): 7942–7952.
53. Hopp M and Gross J. Thermal conductivity of real substances from excess entropy scaling using PCP-SAFT. *Ind Eng Chem Res* 2017; 56(15): 4527–4538.
54. ANSYS. Ansys® Fluent, Release 19.1.
55. García-Oliver JM, Novella R, Pastor JM and Pachano L. Computational study of ECN Spray A and Spray D combustion at different ambient temperature conditions. *Transp Eng* 2020; 2: 100027.
56. Anez J, Ahmed A, Hecht N, Duret B, Reveillon J and Demoulin FX. Eulerian–Lagrangian spray atomization model coupled with interface capturing method for diesel injectors. *Int J Multiph Flow* 2019; 113: 325–342.
57. CMT. Virtual injection rate generator, https://www.cmt.upv.es/ECN03.aspx (accessed 18 May 2021).
58. The MathWorks Inc. *MATLAB*. Natick, MA: The MathWorks Inc, 2020.
59. Fernandes de Mello R and Antonelli Ponti M. *Machine learning*. Cham: Springer International Publishing, 2018.
60. MacKay DJC. A Practical Bayesian framework for backpropagation networks. *Neural Comput* 1992; 4(3): 448–472.
61. Vidal A, Koukouvinis P and Gavaises M. Vapor-liquid equilibrium calculations at specified composition, density and temperature with the perturbed chain statistical associating fluid theory (PC-SAFT) equation of state. *Fluid Phase Equilib* 2020; 521: 112661.
62. Engine Combustion Network. Liquid penetration, Spray-A, 900K, 60bar, https://ecn.sandia.gov/cvdata/assets/datafiles/liq/bkldaAL1-liqDBI.txt (accessed 18 May 2021).
63. Engine Combustion Network. Vapor penetration, Spray-A, 900K, 60bar, https://ecn.sandia.gov/cvdata/assets/datafiles/pen/bkldaAL1-pen.txt (accessed 18 May 2021).
64. Engine Combustion Network. Engine combustion network, modeling standards and recommendations. https://ecn.sandia.gov/diesel-spray-combustion/computational-method/modeling-standards/ (accessed 18 May 2021).
65. Pickett LM, Genzale CL and Manin J. Uncertainty quantification for liquid penetration of evaporating sprays at diesel-like conditions. *Atomiz Sprays* 2015; 25(5): 425–452.
66. Engine Combustion Network. Vapor penetration, Spray-A, 1100, 50bar, https://ecn.sandia.gov/cvdata/assets/datafiles/pen/bgcdaAL4-pen.txt (accessed 18 May 2021).

Appendix. Sample Neural Network C code for thermodynamic properties of dodecane/nitrogen mixtures

In the included Matlab files, ANN functions for the thermodynamic parameters of dodecane/nitrogen mixture are provided as supplementary material. The included functions are:

rhoFunction.m – density in kg/m^3
hFunction.m – enthalpy in J/kg
sFunction.m – entropy in J/kg/K
aFunction.m – speed of sound in m/s
cpFunction.m – heat capacity at p = ct in J/kg/K
mFunction.m – viscosity in Pa.s
kFunction.m – heat conductivity in W/m/K

The input of all these functions is a three-element vector, containing the decimal logarithm of pressure (in Pa), the temperature (in K) and the dodecane mass fraction (–). Thermodynamic derivatives $\frac{\partial \rho}{\partial \mathrm{T}}, \frac{\partial \rho}{\partial p}, \frac{\partial h}{\partial p}$ are estimated by differentiating the corresponding functions of ρ and h, since their values are very small and a good regression could not be obtained. A simple program calling all the relevant functions is also included.

Notation

Commonly used abbreviations

ICE	Internal Combustion Engines
FIE	Fuel Injection Equipment
EoS	Equation of State
ECN	Engine Combustion Network
NIST	National Institute of Standards and Technology

ML	Machine Learning
ANN	Artificial Neural Network
PC-SAFT	Perturbed Chain Statistical Associating Theory
CFD	Computational Fluid Dynamics
VLE	Vapor-Liquid-Equilibrium
RANS	Reynolds Averaged Navier Stokes

Commonly used symbols

ρ	Mixture density (kg/m^3)
$\mathbf{u}$	Mixture velocity (m/s)
T	Mixture temperature (K)
p	Mixture pressure (Pa)
τ	Stress tensor (Pa)
y	Mass fraction (–)
E	Total energy (J/kg)
e	Internal energy (J/kg)
h	Enthalpy (J/kg)
s	Entropy (J/kg/K)
λ	Thermal conductivity (W/m/K)
μ	Dynamic viscosity (Pa.s)
c_p	Heat capacity at constant pressure (J/kg/K)