

Data-driven modeling of self-ignition properties of the renewable fuel PODE (Polyoxymethylen Dimethyl ether) using methods of machine learning

Daten-basierte Modellierung des Selbstzündungsverhaltens des erneuerbaren Kraftstoffs OME (Oxymethylether) mit Methoden des maschinellen Lernens

Bachelor thesis by Pascal Roth

Date of submission: August 26, 2020

1. Review: Prof. Dr.-Ing. Christian Hasse
2. Review: M. Sc. Philip Haspel
3. Review: M. Sc. Julian Bissantz

Darmstadt



TECHNISCHE
UNIVERSITÄT
DARMSTADT



Erklärung zur Abschlussarbeit gemäß §22 Abs. 7 und §23 Abs. 7 APB der TU Darmstadt

Hiermit versichere ich, Pascal Roth, die vorliegende Bachelorarbeit ohne Hilfe Dritter und nur mit den angegebenen Quellen und Hilfsmitteln angefertigt zu haben. Alle Stellen, die Quellen entnommen wurden, sind als solche kenntlich gemacht worden. Diese Arbeit hat in gleicher oder ähnlicher Form noch keiner Prüfungsbehörde vorgelegen.

Mir ist bekannt, dass im Fall eines Plagiats (§38 Abs. 2 APB) ein Täuschungsversuch vorliegt, der dazu führt, dass die Arbeit mit 5,0 bewertet und damit ein Prüfungsversuch verbraucht wird. Abschlussarbeiten dürfen nur einmal wiederholt werden.

Bei der abgegebenen Thesis stimmen die schriftliche und die zur Archivierung eingereichte elektronische Fassung gemäß §23 Abs. 7 APB überein.

Bei einer Thesis des Fachbereichs Architektur entspricht die eingereichte elektronische Fassung dem vorgestellten Modell und den vorgelegten Plänen.

Darmstadt, August 26, 2020



P. Roth

Aufgabenstellung für eine Bachelorarbeit am Fachgebiet STFS



TECHNISCHE
UNIVERSITÄT
DARMSTADT

1. Gutachten: Prof. Dr.-Ing. Christian Hasse, STFS
Studierender: Pascal Roth
Mat-Nr.: 2541363
Studiengang: Maschinenbau – Mechanical and Process Engineering,
B.Sc.
Betreuende: M.Sc. Philip Haspel, M.Sc Julian Bissantz

Titel der Arbeit:

Daten-basierte Modellierung des Selbstzündungsverhalten des erneuerbaren Kraftstoffs OME (Oxymethylether) mit Methoden des maschinellen Lernens
Data-driven modeling of self-ignition properties of the renewable fuel PODE (Polyoxymethylene Dimethyl ether) using methods of machine learning

Simulation reaktiver
Thermo-Fluid-Systeme

Simulation of reactive
Thermo-Fluid Systems



M.Sc. Philip Haspel
M.Sc. Julian Bissantz

Otto-Berndt-Straße 2
64287 Darmstadt
L1|01 286
Tel. +49 615116 - 24151
haspel@stfs.tu-darmstadt.de
bissantz@stfs.tu-darmstadt.de

Aufgabenstellung:

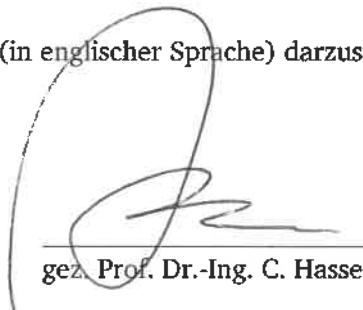
Schreiben Sie eine wissenschaftliche Arbeit zum Thema " Daten-basierte Modellierung des Selbstzündungsverhalten des erneuerbaren Kraftstoffs OME (Oxymethylether) mit Methoden des maschinellen Lernens".

Simulieren Sie zu Beginn der Arbeit charakteristische Zündverzugskurven für relevante Randbedingungen (Äquivalenzverhältnisse und Drücke). Nutzen Sie dazu detaillierte Reaktionsmechanismen und das sogenannte homogene Reaktormodell. Nutzen Sie anschließend die Ergebnisse dieser chemisch detaillierten Lösung um ein daten-basiertes Modell (z.B. ein neuronales Netz) zu trainieren. Mit Hilfe dieses Modells soll es möglich sein das Selbstzündungsverhalten, auf Basis von wenigen Zugriffsgrößen, qualitativ und quantitativ wiederzugeben.

Validieren Sie abschließend ihr Modell mit den Ergebnissen der detaillierten Chemie und vergleichen Sie es mit dem Modell eines globalen Reaktionsmechanismus.

Alle Ergebnisse sind in geeigneter Form (in englischer Sprache) darzustellen und zu diskutieren.

Darmstadt, den 26.03.2020,



gez. Prof. Dr.-Ing. C. Hasse

Abstract

Polyoxymethylene dimethyl ether (*PODE*) is a renewable fuel and a promising additive for diesel engines, owing a high cetane number and the capability to reduce soot emissions. The performance of the engine development using this alternative fuel is highly dependent on the possibility of three-dimensional computational fluid dynamics (CFD) engine simulations. However, using the direct chemistry approach with detailed chemical mechanisms, increases the computational effort to an extent where effective simulations are not feasible. In this thesis, a model is derived that employs artificial neural networks to model the detailed chemistry in the simulation while reducing the computational costs. As neural networks, multi-layer perceptrons (MLPs) are selected.

The MLP models learn the intended thermochemical properties for a representation of the entire thermochemical state-space and are able to interpolate states not covered in the representation. Therefore, the thermochemical states are precomputed by simulating the auto-ignition of an initialized *PODE/air* mixture and discretizing the results. Auto-ignition simulations are performed in a zero-dimensional homogeneous reactor (HR) model and use the detailed mechanism of Cai et al. [13]. In order to reduce complexity and optimize accuracy, the thermochemical properties are grouped according to their temporal evolution within combustion. Four regimes have been identified and for each one, a separate MLP model is constructed and trained. Training is performed by applying the backpropagation algorithm, rectifier activation function and mean squared error loss function. Look-up variables, including reaction progress variable, mixture fraction and absolute enthalpy, are introduced in order to characterize the thermochemical state in the CFD simulation and function as input variables of the MLP models.

In order to validate the derived method, the learning and interpolation capability of the MLPs are evaluated regarding the absolute and relative deviation between predicted curves and curves generated by the HR model. Besides, the reproduction of ignition delays is used to evaluate model performance. Furthermore, a comparison between the method developed in this work and the Global Reaction Mechanism by Haspel [27] is carried out. A good agreement was achieved for all thermochemical parameters as well as ignition delays. The non-linearities in the evolution progress have been reproduced accurately. Thus, the derived model has the precision of the detailed mechanism for every thermochemical parameter. Besides, the presented framework can be adjusted for other fuels.

Table of Contents

1	Introduction	1
1.1	Computational Fluid Dynamics Simulation	2
1.2	Objective of this Thesis	2
2	Homogeneous Reactor Simulation	4
2.1	Polyoxymethylene Dimethyl Ethers	4
2.2	Combustion Properties	6
2.2.1	Major Species and Equivalence Ratio	6
2.2.2	Two-Stage Ignition Behavior	7
2.3	Reaction Mechanism	8
2.3.1	Detailed Mechanism	9
2.3.2	Global Reaction Mechanism	13
2.3.3	Tabulated Chemistry Approach	14
2.4	Homogeneous Reactor Model	15
2.4.1	Mixture Fraction of Combustion Elements	17
2.4.2	Absolute Enthalpy	18
2.4.3	Progress Variable	19
3	Data-Driven Modeling of Chemistry	22
3.1	Neural Networks	22
3.1.1	Perceptron	23
3.1.2	Multi-Layer-Perceptron	24
3.2	Training Process of Neural Networks	26
3.2.1	Backpropagation Learning Algorithm	27
3.2.2	Training Optimization	30
3.3	Network Adjustments for Thermochemical Modeling	33
4	Results and Discussion	36
4.1	Data-driven Modeling of Chemistry	36
4.2	Comparison with a Global Reaction Mechanism	42
5	Conclusion and Outlook	47
Bibliography		vii
List of Figures		xii
List of Tables		xiii
Nomenclature		xiv

1. Introduction

UN chief António Guterres warned in the WMO Statement on the State of the Global Climate in 2019 that the world is "currently way off track to meeting either the 1.5°C or 2°C targets that the Paris Agreement calls for" [53]. The consequences of this development are already visible since the past five years are the five warmest on record and the past decade, 2010 to 2019, also is the warmest on record [53]. Other examples as the ocean warming, the increase in extreme weather phenomena or the related loss of biodiversity, as well as the increase in health risks haven't yet provoked the world to decrease or even reach a constant level of greenhouse gas emissions, instead, a further increase was monitored over the last years [21].

The transportation sector including aircraft, ships, and automobiles is responsible for around 14% of global emissions [14]. All types of carriers in this sector are mainly powered by internal combustion (IC) engines and hence produce emissions by burning fossil fuels like diesel, gasoline, or kerosene. Focussing on automobile transportation the energy demand for around 1.2 billion light-duty vehicles and around 380 million heavy-duty vehicles in 2019 with an increasing tendency has to be satisfied [60]. Other aspects such as the availability of affordable energy especially for developing countries and the importance of a reliable transportation system for a globalized world have to be taken into account when planning future transportation systems. Reitz et al. [60] concluded that there are still no real cost-efficient alternatives that are able to compete with IC engines over the entire range of applications. Furthermore, they noted that due to the continuous improvement of IC engines, possible alternatives face significant barriers to a fast adoption [60].

Taking into consideration this state of the art, further innovations of IC engines and the resulting decrease in pollutants as well as greenhouse gas emissions are necessary to fight climate change and secure the transportation system of the world until alternatives are found which are able to overcome IC engines. One improvement approach are E-Fuels [31]. This concept uses excessive renewable electricity in combination with different electrolysis technologies, CO_2 from exhaust gases, and Fischer-Tropsch synthesis to produce carbon-neutral fuels, which have similar combustion properties than their fossil counterparts. The accumulation of renewable energy in chemical compounds makes E-Fuels effective energy storage systems and allows the use of existing infrastructure such as gas stations. Furthermore, those carbon-neutral fuels tackle the problem of reduction in available petroleum reserves and have become more widely available over the last decade. Currently, E-Fuels are not an effective way to reduce greenhouse gas emissions in terms of financial aspects due to electricity, carbon dioxide and necessary investment costs. However, further improvements and the price increase of conventional fuels will compensate for this economic disadvantage over time[33]. Polyoxymethylene Dimethyl Ether (*PODE*) is one of the most promising representants of E-Fuels and has been proven to reduce soot emissions by up to 40% with a 20% *PODE* diesel blend [44]. As a consequence, this work will focus on *PODE* and will give a detailed introduction of its chemical properties in section 2.1.

1.1. Computational Fluid Dynamics Simulation

Necessary improvements are heavily dependent on the possibility to optimize IC engines powered by E-Fuels and blends. The general ability to test IC engines with the new fuel and use this information for the optimization process cannot exclusively be done experimentally but rather in three-dimensional Computational Fluid Dynamics (CFD) simulations. Those simulations discretize the given geometry using techniques such as finite element or finite volume method. The involved fluid flow and the interaction of the fluid with given surfaces are solved using turbulence models like Large Eddy Simulation or Reynolds-averaged Navier–Stokes. In addition to the Navier-Stokes equations and the equation of mass conservation, balance equations are solved for each species in the combustion process. During the simulation, energy and species content of each cell are not only changed by convective and diffusive flow over the cell boundaries but also due to chemical reactions. To account for those reactions, the chemical source term for each species is calculated. The calculation is carried out by including detailed chemistry in the simulation which is unique for the used fuel and summarized in a *reaction mechanism*. Even for simple fuels, a detailed reaction mechanism involves several hundred species and several thousand reactions. To avoid reducing the time step to the smallest chemical time scale, the chemical source terms are integrated over a time step Δt . This, however, implies that the source terms are the result of an n-species large, stiff, nonlinear ordinary differential equation (ODE) system. Regarding the number of species, solving this system is the most computationally intensive part of the simulation. Zirwes et al. [76] measured that for small reaction mechanisms about 60% of the complete simulation time is used for chemical calculations. For fuel reaction mechanisms the percentage can rise to 90%. Therefore, the performance of a 3D CFD engine simulation with the use of a detailed mechanism is too expensive to be applied effectively.

1.2. Objective of this Thesis

This work aims to introduce a novel method to model the chemistry of the renewable fuel *PODE* in 3D CFD simulations. The method is based on the tabulated chemistry approach, meaning that the solution of the detailed reaction mechanism is pre-computed for a representation of the thermochemical state-space and stored in a look-up table which is searched during runtime. This work, however, uses the pre-computed states to train an artificial neural network (ANN) which then is applied in the simulation. The novelty of the presented method is the use of an ANN that is chosen due to its ability to efficiently learn training states and interpolate thermochemical states it has not been trained with. A major advantage is a decrease in storage usage compared to an actual look-up table and the adaptability regarding the extension of the covered states. All of it while keeping the computational effort constantly low.

In order to generate the necessary training data, the auto-ignition of a *PODE/air* mixture has to be investigated. Therefore, a zero-dimensional homogeneous reactor (HR) model is applied in which the detailed mechanism is used to solve the combustion of the mixture. By identifying the major combustion species that are used to evaluate the combustion process in the CFD, the detailed chemistry is reduced to only a limited number of parameters. The ANN then learns to predict the values of the parameters given the thermochemical state. Consequently, the balance equations of the species are replaced and thus the need to calculate the species source terms. The inputs of the networks are look-up variables that identify the thermochemical state resulting in the fact that only their conservation equations have to be solved in the simulation. After training the networks with all results taken from the HR, an application in the CFD is possible and the speed of the optimization process of IC engines powered with E-Fuels is increased.

To present the workflow and the results of the method, this thesis is structured as follows: in chapter 2 the training data is generated by the homogeneous reactor simulation of the auto-ignition. Furthermore, the used fuel and the variables to access the thermochemical state-space are addressed. Chapter 3 presents neural networks, their learning algorithm and the adjustments made for chemistry modeling. Chapter 4 evaluates the ANNs regarding learning and interpolation capability. Furthermore, a comparison between the derived method and the global reaction mechanism of Haspel [27] is given. In the final chapter 5 the conclusion of this thesis is presented and an outlook for future research is provided.

2. Homogeneous Reactor Simulation

In the following, the simulation of the auto-ignition behavior of Polyoxymethylene Dimethyl Ether is performed and used to generate the temporal evolution of the major species and the necessary thermodynamic properties. For this purpose an understanding of the used fuel and the combustion process is necessary. Hence, this chapter presents the main properties of $PODE_n$ as well as the reason why it is claimed to be a promising renewable fuel. It discusses the major species for the combustion analysis in compression ignition (CI) engines as well as ignition delays (IDs). In order to mathematically describe the chemical kinetics, reaction mechanisms are introduced which include the chemical reactions, the species involved and the rate constants of each reaction. By extracting experimental data from the work of Jacobs et al. [35] and He et al. [29] the reaction mechanism with the best description of the auto-ignition behavior is selected. A chemical kinetic model is then presented by solving this mechanism in a zero-dimensional homogeneous reactor model.

2.1. Polyoxymethylene Dimethyl Ethers

Polyoxymethylene Dimethyl Ether, also known as Oxymethylene Ether (OME_n) , is a synthetic fuel with diesel-like combustion properties. Due to its capability to reduce soot emissions, $PODE_n$ is claimed as one of the most promising representatives of E-fuels. This section introduces the state of the art solution to synthesize $PODE_n$. Furthermore, the different degrees of polymerization are investigated in order to find the best suitable for IC engines. Lastly, the theoretical emissions reduction capabilities by using the selected fuel are introduced and experimental research is presented.

$PODE_n$ can be produced from synthesis gas on various routes. The state of the art is to start with hydrogen as synthesis gas and convert it in the first step to methanol. This usually also produces dimethyl ether (DME) as a by-product. In a second step, methanol is oxidized to formaldehyde (FA). Using DME and FA the synthesis of methylal becomes possible. Furthermore, trioxane is synthesized from FA in aqueous solution and then purified to an anhydrous product. In a final step $PODE_n$ is generated from methylal and trioxane [12, 25, 44, 51]. By using biological syngas and renewable electricity $PODE_n$ can be produced as quasi carbon-neutral fuel [13].

The $PODE_n$ condensed structural formula is $CH_3O(CH_2O)_nCH_3$ where the index n denotes the degree of polymerization. This degree influences the chemical properties of $PODE_n$ as shown in Tab. 2.1. Since the properties affect the combustion process, certain degrees of polymerization are better suitable for CI engines. For all $PODE_n$, Zheng et al. [74] investigated the connection between the ignitability of $PODE_n$ and its cetane number (CN). CN is an inverse function of the fuel's ignition delay time which is described in section 2.2.2 and thereby an indicator of the combustion speed as well as compression needed for ignition. In general, the ignitability of the fuel increases with a larger CN [9]. In the case of $PODE_1$, the excessive saturated vapor pressure and the low boiling point lead to vaporization before entering the combustion chamber and can cause the so-called "vapor lock". This behavior is more common for fuels used

Tab. 2.1.: Main combustion properties of $PODE_n$ [42]

	$PODE_1$	$PODE_2$	$PODE_3$	$PODE_4$	$PODE_5$	$PODE_6$
Molecular formula	$C_3H_8O_2$	$C_4H_{10}O_3$	$C_5H_{12}O_4$	$C_6H_{14}O_5$	$C_7H_{16}O_6$	$C_8H_{18}O_7$
Density at 25°C (g/cm³)	0.86	0.96	1.02	1.06	1.10	1.13
Melting point (°C)	-105	-65	-41	-7	18.5	58
Boiling point (°C)	42	105	156	202	242	280
Cetane number	29	63	78	90	100	104
Oxygen content (%)	42.1	45.3	47.1	48.2	49.0	49.6
Lower heating value (MJ/kg)	22.4	20.3	19.3	19.1	17.9	17.5

in spark-ignition engines [75]. This property combined with the low CN compared with higher degrees of polymerization make $PODE_1$ not suitable for CI engines. With a n larger than 1, the CN outranges classic diesel fuel as described in DIN EN 590 [3]. Thus, the CN of $PODE_2$ is sufficiently large. However, the flashpoint is too low to fulfill the security measures [74]. $PODE_{n>5}$ are not suitable due to their melting point which lies above room temperature and hence causes a tendency to crystallize at low temperatures. From $PODE_{3-5}$, $PODE_3$ and $PODE_4$ are the most relevant due to their lower melting and boiling point which indicates a lower low-temperature viscosity and a higher volatility [29], so that this work focuses on those two degrees of polymerization.

The condensed structural formula of $PODE_n$ shows an absence of carbon-carbon bonds and a relatively high oxygen content. The oxygen included in the fuel enhances the soot oxidation and hence diminish soot as well as reduce particulate matter (PM) emissions in CI engines [26, 42, 44]. Wang et al. [72] furthermore proved that these properties lead to a reduction of carbon oxide (CO) and unburned carbon (HC) emissions. In the case of $PODE_{3,4}$ the combustion temperature is higher compared to classic diesel fuel and thus issues a higher rate of thermal NO_x emissions [42]. In order to decrease these emissions, an effective approach is to reduce the peak combustion temperature, as done with an exhaust gas recirculation system [1]. However, the resulting lower temperature interferes with the oxidation of the fuel and therefore causes higher soot and PM emissions. Consequently, there is a tradeoff that has to be optimized during the engine development process [58].

The presented potential emission reduction capacities of $PODE_{3-4}$ have been studied in recent years. As mentioned in the introduction Lumpp et al. [44] have proved in 2011 that with a 10% to 20% $PODE_{3-4}$ diesel blend the PM and soot emissions can be reduced by approximately 40%. Further studies by Pellegrini et al. [55] showed that 50% $PODE_{3-5}$ expands the high-temperature region in the cylinder and increases the flame-propagation speed. In 2017, Liu et al. [43] stated that soot-free combustion can be achieved at near stoichiometric conditions by adding 20 to 30% $PODE_{3-4}$. Furthermore, they showed that CO emissions can be decreased by 90% while the NO_x emission reaches its minimum at 20% $PODE_{3-4}$. These studies underline the benefit in emission reduction and prove the possibilities of a $PODE_{3-4}$ diesel blends.

2.2. Combustion Properties

In order to investigate the ignition of $PODE_{3,4}$, the main properties of the thermochemical states within the combustion process have to be identified and calculated. The thermochemical state is defined by pressure, temperature as well as species composition. In the CFD, however, only major species are used to evaluate the combustion. Accordingly, the investigated composition is limited to the mass fractions of those species. Furthermore, thermodynamic parameters as temperature, pressure, etc. must be determined. Besides, ignition delay (IDs) as an important diesel characteristic have to be identified. In the first part, this section introduces the major species and the equivalence ratio. In the second part, IDs are defined and a method to measure them is presented.

2.2.1. Major Species and Equivalence Ratio

At the beginning of the combustion process, only the reactants, namely $PODE_{3-4}$ and the species in the oxidizer O_2 and N_2 , are present. Reactions between these substances, in which the fuel is branched, produce intermediate species and radicals. The following reactions between these substances take up most of the entire combustion process. Towards the end of the combustion, the intermediates and radicals are used to form the stable reaction products CO_2 , H_2O , and possible CO . In theory, all intermediates are consumed during the process. For the evaluation of the combustion in CFD simulation, only the educt and product species are relevant.

The fuel and the oxidizer mass fractions are obviously significant quantities and their ratio is described by the equivalence ratio Φ , which is commonly used to characterize the combustion. In the following the definition of Φ by Poinsot et al. [57] is provided. If not stated otherwise in the following indices F and O correspond to fuel and oxidizer. Considering a general global reaction of the type



where v'_F , v'_O are the coefficients corresponding to fuel and oxidizer, the mass fractions of fuel Y_F and oxydizer Y_O correspond to stoichiometric conditions when:

$$\left(\frac{Y_O}{Y_F} \right)_{st} = \frac{v'_O M_O}{v'_F M_F} = s, \quad (2.2)$$

where $M_{O/F}$ are the molecular masses of oxydizer and fuel and s the mass stoichiometric ratio. In an stoichiometric reaction fuel and oxydizer fully react to CO_2 , H_2O and N_2 . Φ of a certain mixture can then be calculated as follows:

$$\Phi = s \frac{Y_F}{Y_O} = \left(\frac{Y_F}{Y_O} \right) / \left(\frac{Y_F}{Y_O} \right)_{st}. \quad (2.3)$$

In case the fuel is in excess ($\Phi > 1$), rich combustion is obtained which results in an increase of CO as a combustion product because complete oxidation to CO_2 cannot be performed. Lean regimes ($\Phi < 1$) on the other side are achieved when the oxidizer exceeds. Under this condition, larger amounts of NO_x as a combustion product appears because N_2 starts to react with the unused O_2 . Hence, both conditions are not

Tab. 2.2.: Global reactions and mass stoichiometric ratio of the combustion of a $PODE_{3-4}/air$ mixture with $\Phi = 1.0$

Fuel	Global Reaction	s
$PODE_3$	$1CH_3O(CH_2O)_3CH_3 + 6(O_2 + 3.76N_2) \rightarrow 5CO_2 + 6H_2O + 22.56N_2$	1.41
$PODE_4$	$1CH_3O(CH_2O)_4CH_3 + 7(O_2 + 3.76N_2) \rightarrow 6CO_2 + 7H_2O + 26.32N_2$	1.35

optimal for CI engines due to an increase in pollutants. For a $PODE_{3-4}/air$ mixture the global reaction and mass stoichiometric ratio is given in Tab. 2.2.

2.2.2. Two-Stage Ignition Behavior

In CI engines, estimation of auto-ignition delays is of great importance due to their effect on the whole combustion process [56]. Thereby, the ignition delay time (IDT) is defined as the time interval between the start of injection and the ignition. In real applications, it can be divided into a physical delay, wherein atomization, vaporization, and mixing of air-fuel occur and a chemical delay attributed to pre-combustion reactions. Physical and chemical delays occur simultaneously [38]. For an accurate description of the auto-ignition, the IDTs have to be determined. Due to the high proportion of exothermic reactions happening at the ignition, the point in the combustion process can be identified as the global maximum of the heat release rate (HRR). The specific HRR is calculated using the production rates of all species $\dot{\omega}_i$ at a time t and their specific molar enthalpies h_i .

$$HRR = \sum_i \dot{\omega}_i * h_i \quad (2.4)$$

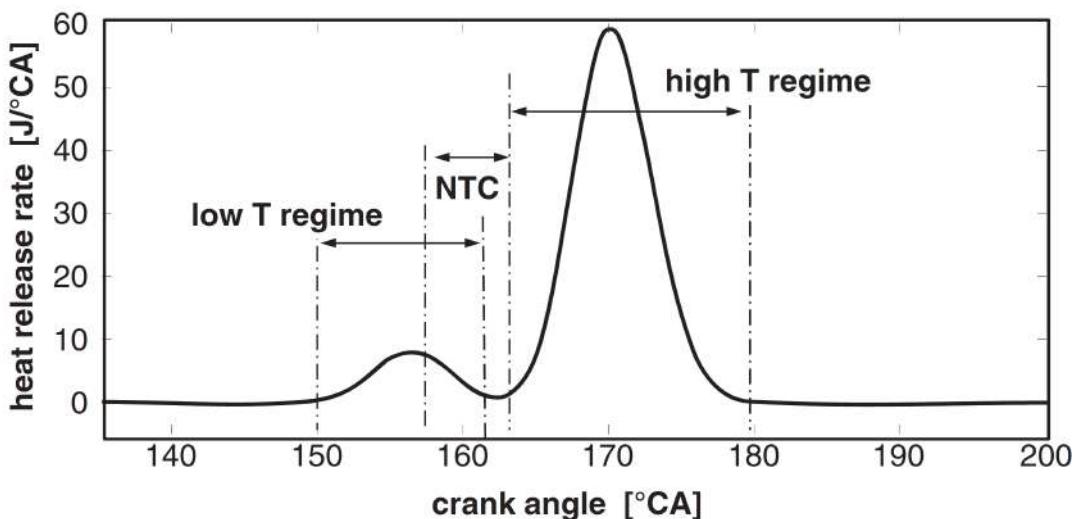


Figure 2.1.: Schematic heat release rate profile that shows the two stage ignition behavior. The low and high temperature ignition, as well as NTC area is clearly visible due to the peaks in the curve [71].

In comparison to fuels for spatial ignition engines, fuels for CI engines ignite in a low-temperature (LT) and high-temperature (HT) stage [15]. Fig. 2.1 shows a typical HRR profile of a fuel exhibiting two-stage ignition. The LT ignition is defined as the first local maximum in the HRR curve. In the following, the reaction kinetics of the LT and HT stage as derived in the work of Vandersickel et al. [71] are presented.

Low-temperature stage

- LT stage is initialized by the abstraction of H-atoms to form an alkyl radical R and a hydroperoxyl radical HO_2
- R reacts with O_2 in several steps to smaller species as ethers, aldehydes and to OH radicals.
- The OH radical then reacts with larger alkane molecules to regenerate R , with the consequence that the low-temperature reaction chain keeps propagating.
- Alternatively, one alkyl radical can combine with two O_2 atoms which lead over different reaction pathways to the formation of two OH radicals and therefore to an increase of the number of radicals. Consequently, this degenerate branching reaction leads to an acceleration of the reaction rate, which initializes an exothermic cycle.
- The exothermic cycle leads to an increase in temperature which causes different reactions because the reverse of the oxygen addition to the alkyl radical and hence the formation of HO_2 becomes important.
- This change is equivalent to a termination step since the new radical recombines to stable intermediates like H_2O or O_2 . This reduces the overall reaction rate and is the main reason for a negative temperature coefficient (NTC) area which concludes the low-temperature ignition.

High-temperatur stage

- During NTC the temperature only rises slowly until the reaction of H_2O_2 becomes important ($T > 1000K$), terminating the NTC regime and initiating a branched thermal explosion, indicating the start of the HT stage.
- Furthermore, the main ignition occurs due to the consumption of the fuel by reactions during NTC. Without the fuel as a radical sink, more OH radicals are available for rapid oxidation of the intermediate species, causing the main ignition at temperatures below 1000K.
- During HT ignition, large molecules and radicals decompose irreversible to H_2O and CO .
- Heat release is dominated by oxidation of the intermediates CH_2O to CO and successive oxidation of CO to CO_2 .

2.3. Reaction Mechanism

Reaction mechanisms are models that describe the chemical kinetics of a fuel in a range of validity [19]. Therefore, the mechanisms define the course of events from educts to products and include the elementary chemical reactions, the species involved and the rate parameters of each reaction. Thus, they can be used in order to calculate the chemical source terms of the species, which are also used in the calculation of the HRR, Eq. 2.4. In order to analyze and explain the experimental observations of $PODE_n/air$ mixtures with a numerical simulation, the parameters of the corresponding reaction mechanism have to be identified and determined.

The mechanism describes the chemical system of n_s species which react through n_r reactions [65]:

$$\sum_{i=1}^{n_s} v'_{ij} \mathcal{M}_i \rightleftharpoons \sum_{i=1}^{n_s} v''_{ij} \mathcal{M}_i, \quad (2.5)$$

where \mathcal{M}_i represents the symbol of species i and v'_{ij} , v''_{ij} indicate the molar stoichiometric coefficients of species i in reaction j . Due to the mass conservation in chemical reactions, the following mass relation has to be satisfied for every reaction j :

$$\sum_{i=1}^{n_s} v'_{ij} M_i \rightleftharpoons \sum_{i=1}^{n_s} v''_{ij} M_k, \quad (2.6)$$

where M_i denotes the molecular weight of species i . $\dot{\omega}_i$ is defined as the summation over the progress rate R_j for all reactions:

$$\dot{\omega}_i = \sum_{j=1}^{n_r} \dot{\omega}_{ij} = M_k \sum_{j=1}^{n_r} v_{ij} R_j, \quad (2.7)$$

where $v_{ij} = v''_{ij} - v'_{ij}$ is the net stoichiometric coefficient. R_j is defined by

$$R_j = k_{fj} \prod_{i=1}^{n_s} [X_i]^{v'_{ij}} - k_{bj} \prod_{i=1}^{n_s} [X_i]^{v''_{ij}}. \quad (2.8)$$

In this expression, k_{fj} and k_{bj} indicate the forward and the backward rate of reaction j which are defined with respect to the molar concentrations $[X_i]$. The Arrhenius law is usually used to model k_{fj} and k_{bj} . It is defined as follows:

$$k_{fj} = A_{fi} T^{\beta_j} \exp\left(-\frac{E_j}{RT}\right) \quad (2.9)$$

where A_{fi} is the preexponential factor, β_j the temperature exponent and E_j the activation energy which are defined for each reaction j . The symbol \mathcal{R} represents the ideal gas constant and T the temperature in Kelvin. For more complex reactions such as third-body-reactions or unimolecular reactions, the reaction rate cannot be determined using the Arrhenius law. Instead, different techniques like the Lindemann mechanism are employed. An overview of those methods is given in the book by Law [40].

2.3.1. Detailed Mechanism

Mechanism Presentation

In recent years, research has been focused on developing and validating new detailed chemical kinetic models of $PODE_n$. Detailed mechanisms are the most complex, non-simplified models, which aim to precisely reproduce the chemical kinetic behavior. Sun et al. [67] measured the laminar burning velocities

of $PODE_3$ /air mixtures in an electrically-heated constant-volume cylindrical combustion vessel for low-temperatures and Φ in a range from 0.7 to 1.6. By using this information several intermediate species have been investigated and a preliminary kinetic mechanism to simulate the reaction pathways in high-temperature regions has been introduced. He et al. [29] improved this detailed reaction mechanism (225 species, 1082 reactions) for low- and intermediate-temperature regions by investigating the IDTs of $PODE_3$ /oxygen/nitrogen mixtures in a rapid compression machine (RCM). Those experiments are conducted at pressures of 10/ 15 bar and equivalence ratios of 0.5, 1.0, and 1.5 with nitrogen diluted air. The mechanism itself is based on the assumption that reaction classes for $PODE_n$ are similar to the ones of regular alkanes, while in $PODE_n$ due to the absence of C-C bonds the pathways leading to the formation of alkanes do not exist. Therefore, the rate constants of the mechanism for $PODE_3$ are not only taken from $PODE_1$ but also from a comparison with dimethyl ether and diethyl ether. While the mechanism parameters of Sun et al. have been optimized to capture accurately the measured laminar flame speeds, He et al. used the determined IDTs to fit their mechanism.

Using a novel model development process Cai et al. [13] derived a detailed mechanism that includes low- and high-temperature regions, as well as $PODE_4$. In order to evaluate the mechanism the IDT investigation of Jacobs et al. [35] is used. Thereby, the IDTs are determined for high-temperatures in a shock tube (ST) and for low-temperatures, similar to He et al., in a RCM. Jacobs used non-diluted, stoichiometric $PODE_n$ /air mixtures and a pressure range from 10bar to 20bar. The measurements have been performed for different degrees of polymerization. Similar to He et al. the mechanism parameters are optimized on the available IDTs.

With the aim to develop a mechanism specialized in the mixture of diesel fuel with $PODE_3$ Lv et al. [45] coupled a simplified version of the mechanism of He et al. [29] with a simplified n-heptane and n-butylbenzene mechanism. The simplification was developed in line with the two-step iterative strategy of Ren et al. [61] that is continued until the deviations in IDTs are larger than a threshold. The first step of the method is to remove negligible species and reactions, while the second step is to group isomers with similar thermochemical properties. As a consequence, the IDTs are shorter due to more direct and effective reaction pathways.

Mechanism Selection

Since this thesis aims to model the detailed chemistry of $PODE_{3,4}$ in the simulation and thus give a method to reduce a detailed mechanism, no specialized mechanisms are used. The performance of the detailed mechanisms regarding the IDT reproduction capability is different due to the individual development process and whether they have been fitted to laminar burning velocities or to IDTs. The presented detailed mechanisms are evaluated using the experimental IDTs of He et al. [29] and Jacobs et al. [35]. The temperature range of the experimental IDTs starts at approximately 550K and ends at 1150K which leaves uncertainty for combustions with a higher initial temperature. Due to the close temporal distance of low temperature and high-temperature ignition, the comparison is performed solely for the high-temperature IDTs.

The experimental IDT investigation of He et al. is taken into consideration first. Using $PODE_3$ the comparison of the different mechanisms for Φ of 0.5, a pressure of 10 bar and a highly diluted oxidizer is shown in Fig. 2.2. Comparisons for different conditions are found in Fig. A.1. Obviously, He et al. and Cai et al. are able to reproduce the experimental IDTs accurately. Both achieve an almost identical compensation curve, which lies in the inaccuracies of the individual measuring points. Sun et al., however, largely overestimates the IDTs. The reason for this different performance is that Cai et al. and He et al. have been fitted to these experimental data points while Sun's mechanism is fitted to laminar burning velocities. Furthermore, the measurements of Sun et al. have been carried out with an atmospheric oxidizer which leads to uncertainties with diluted conditions.

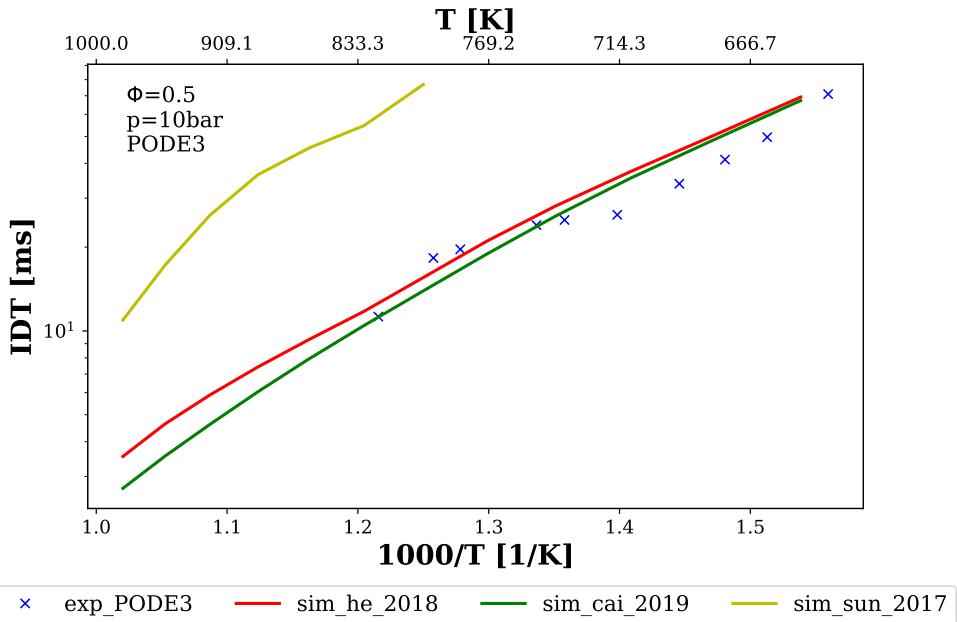


Figure 2.2.: Extracted experimental IDTs from He et al. [29] and simulated IDTs generated by the HR model with mechanisms of Sun et al. [67], He et al.[29] and Cai et al.[13]. Used oxydizer is diluted with the ratio: $O_2 : N_2 = 1 : 8$

On the other hand, the IDT investigation by Jacobs et al. [35] is shown for $PODE_3$ and $PODE_4$ under stoichiometric conditions and air as the oxidizer in Fig. 2.3 (other degrees of polymerization are shown in Fig. A.2). The undiluted conditions lead to significantly shorter IDTs since a higher rate of oxygen and fuel is included in the mixture. A correlation that has been reproduced by all mechanisms. Furthermore, the S-shape of the IDT curve caused by the difference between low-temperature and high-temperature reactions is captured. However, the mechanism of Sun et al [67] estimates the temperature range where the transition occurs too short. Besides, all mechanisms capture the correlation between IDT and degree of polymerization as well as between IDT and initial enthalpy which is determined by pressure and temperature.

The performances of the different mechanisms in order to reproduce the experimental IDTs are more evident in this study. While the mechanism of Sun et al. [67] still overestimates the IDT curve, the relative distance is closer than in the case of the experiments of He et al. [29]. Responsible is the undiluted air. The mechanism of He et al. shows a better approximation but also overestimates the IDTs. However, the outcome of the mechanism is surprising, since the IDTs deviate only slightly. The reason for this is that He's mechanism is also based on Sun's experiments and is therefore not just influenced by the measurement with diluted oxidizers. The mechanism of Cai et al. [13] generates compensation curves that fit all degrees of polymerization with the best accuracy. Only in the case of $PODE_4$, the mechanism tends to overestimate in high-temperature regions. Hence, the mechanism delivers the best results, while being applicable for the most fuels. Consequently, the mechanism of Cai et al. is chosen to generate the training data for the neural network.

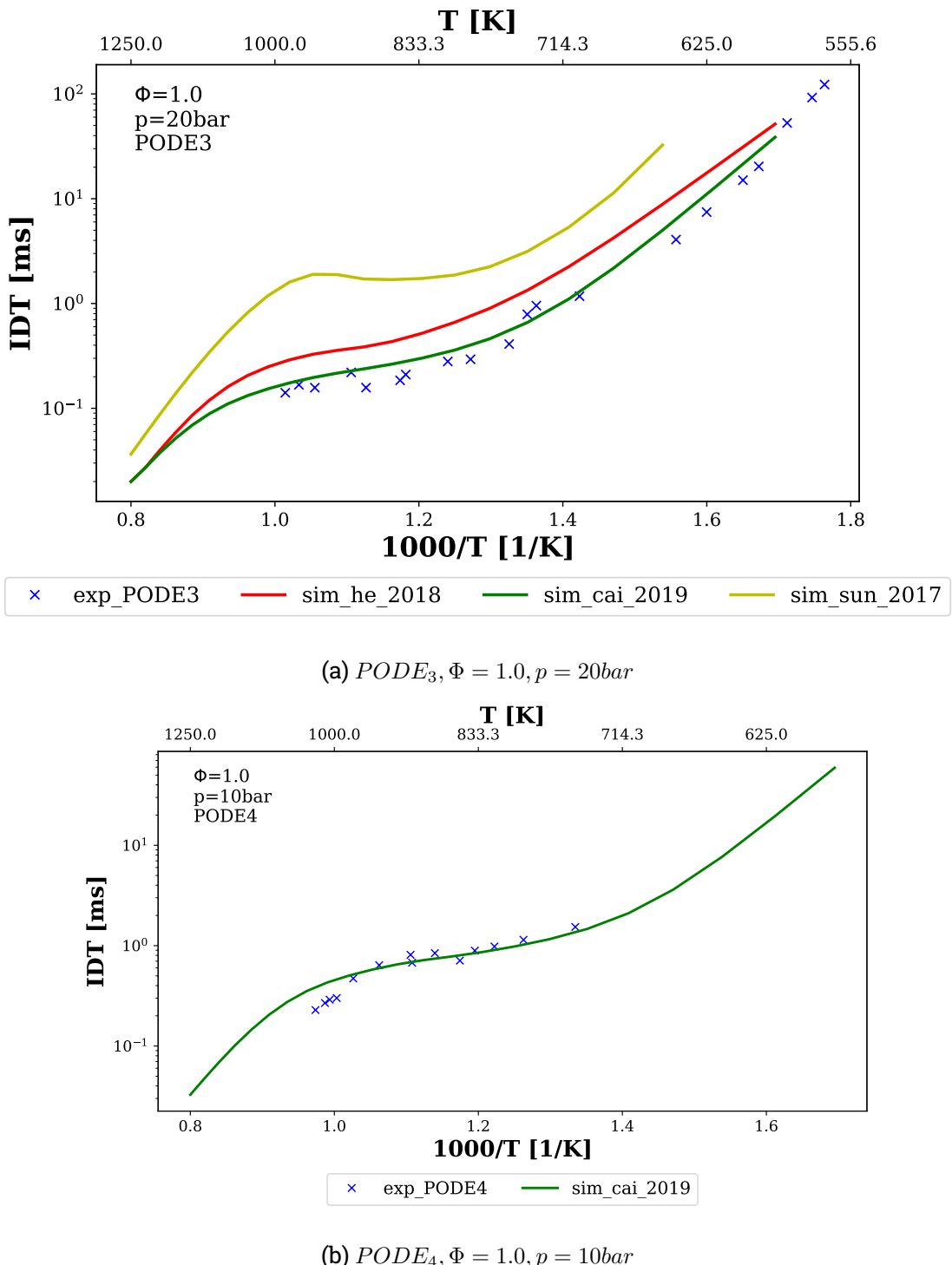


Figure 2.3.: Extracted experimental IDTs from Jacobs et al. [35] compared with IDTs generated by the HR model with mechanisms of Sun et al. [67], He et. al [29] and Cai et. al [13]

2.3.2. Global Reaction Mechanism

3D CFD engine simulations using a detailed mechanism are not feasible with regard to the computational effort. Consequently, in order to reduce the computational cost, the mechanism has to be reduced to a limited number of reactions that are relevant for the combustion evaluation [7]. While techniques as the one of Ren et al. [61] are able to reduce the complexity and keep nearly all information the resulting mechanism is still too large to be directly applied. This indicates that only the removal of similar reactions and species is not enough. In the literature, the development of a global reaction mechanism (GRM) exists as a strategy to ensure the necessary simplification and to solve the trade-off between information loss and computational effort

The global reaction mechanism reduces the detailed mechanism to only a limited number of species and hence reactions that summarize reaction classes. The resulting pathways can be separated into ones for low- and high-temperature regions. Vandersickel et al. [71] presented such a model that can be applied to any hydrocarbon fuel as long as the parameters are adjusted. Thereby, it covers a wide range of fuels and operating conditions while limiting the computational effort to seven reaction equations with eight species. An overview of these reactions and the generalized rate expressions fitted for $PODE_n$ are provided in Tab. 2.3. The species F , I_1 , I_2 and Y represent groups of species, more precisely represents F the fuel, alkyl radicals, and high-temperature intermediates, I_1 represents alkyl peroxy radicals and I_2 mainly covers H_2O_2 , while Y mainly correspond to HO_2 and OH . This mechanism uses its first two equations to describe high-temperature chemistry and equations three to five to include low-temperature phenomena.

In order to make sure that ignition delays, as well as major combustion products, are captured accurately by the model, Vandersickel et al. introduced four optimization functions. The first two to optimize the error in the IDTs for low-temperature and high-temperature regions:

$$f_{LT} = \frac{1}{N_{LT}} \sum_{i=1}^{N_{LT}} |\tau_{1i}^{exp} - \tau_{1i}^{sim}| \quad (2.10)$$

$$f_{HT} = \frac{1}{N} \sum_{i=1}^N |\tau_{2i}^{exp} - \tau_{2i}^{sim}|, \quad (2.11)$$

Tab. 2.3.: Reactions and reaction rate expressions of the GRM by Hasepl [27] for $PODE_n$ in reference to Vandersickel et al. [71]

Reaction	Reaction rate [$mol/cm^3 s$]
1. $F + \frac{2(x-z)+y}{4}O_2 \rightarrow \frac{y}{2}H_2O + xCO$	$R_1 = k_1[F]^r[O_2]^{1.5}(p/p_0)^c$
2. $CO + 0.5O_2 \leftrightarrow CO_2$	$R_{2+} = k_{2+}[CO][H_2O]^{0.5}[O_2]^{0.25}$ $R_2 = k_2[CO_2]$
3. $F + 2O_2 \leftrightarrow I_1$	$R_{3+} = C_{3+}k_{3+}[F][O_2][M_1](p/p_0)^a$ $R_{3-} = k_{3-}[I_1](p/p_0)^b$
4. $I_1 \rightarrow 2Y \frac{z}{2}$	$R_4 = k_4[I_1]$
5. $Y + 0.5F + (\frac{2(x-z)+y}{4} - 1)O_2 \rightarrow \frac{y}{2}H_2O + xCO$	$R_5 = k_5[F][Y]$
6. $I_1 \rightarrow I_2$	$R_6 = k_6[I_1]$
7. $I_2 \rightarrow 2Y$	$R_7 = k_7[I_2][M_2]^s$
$k_i = A_i \exp(-E_i/RT), p_0 = 1 MPa, [M_1] = p/RT$ and $[M_2] = [F] + [O_2] + [I_1] + [I_2] + [Y] + [CO]$	

where τ^{exp} indicates the experimental IDT, τ^{sim} the simulation IDT and N the number of initial conditions that caused the corresponding IDTs. The third objective function concerns the pressure at the end of the LT ignition and has been implemented to increase accuracy in the evolution of heat release and species profiles

$$f_{p_{LT}} = \frac{1}{N_{LT}} \sum_{i=1}^{N_{LT}} \frac{|p_{1i}^{det} - p_{1i}^{sim}|}{|p_{1i}^{det}|}, \quad (2.12)$$

where p_{1i}^{det} is the pressure calculated by using the detailed mechanism and p_{1i}^{sim} the pressure when the GRM is used. In order to ensure realistic species profiles, the fourth objective function compares the maximum in the reaction rate of the branching reaction 7 $t_{max_{R7i}}^{sim}$ and the simulated HT IDT τ_{2i}^{sim} :

$$f_{max_{R7}} = \frac{1}{N} \sum_{i=1}^N |t_{max_{R7i}}^{sim} - \tau_{2i}^{sim}|, \quad (2.13)$$

The final parameter setting depends on the trade-off between the objective functions and hence of the importance of each function for the desired fuel. The adjustment of this GRM model for $PODE_n$ has been performed by Haspel [27] and should be taken into account in section 4.2 to compare the performance with the method derived in this thesis.

2.3.3. Tabulated Chemistry Approach

An alternative to the global reaction mechanism is the tabulated chemistry approach which has been specially developed to avoid solving the stiff, non-linear ODE system in the simulation runtime. It thereby combines the high prediction accuracy of the chemical state of the detailed reaction mechanism with low computational costs within the simulation. Tabulated manifolds are based on the pre-calculation of a representation of the thermochemical state-space which is stored in a table that is accessed by look-up parameters. During the simulation, only the conservation equations of these look-up parameters need to be solved to account for the chemical reaction and not the ODE system [6]. In order to generate the look-up table, different approaches have been developed including FGM [52], FPI [22], and ILDM [46].

It is mentioned before that the approach presented in this work takes the tabulated manifold as a basis. More precisely the idea to eliminate the need to solve an ODE system in the simulation runtime is similar but instead of tabulating the thermochemical states, the same data is used to train an ANN. This network is able to predict the values of the selected thermodynamic property given look-up variables as input that identify the thermochemical state of the system. Computation of the output of an ANN is a task with a low computational effort which makes it suitable to be applied in a CFD simulation. Consequently, only the conservation equations for the input variables have to be solved in the simulation. Prominent variables are the reaction progress variable (section: 2.4.3), mixture fraction (section: 2.4.1) and absolute enthalpy (section: 2.4.2). Since the table content is only used once, the storage usage of the presented method is low compared to the tabulated chemistry approach.

Tab. 2.4.: Properties of the most common homogeneous reactor models

	Constant pressure & fixed mass	Constant volume & fixed mass	Well stirred	Plug Flow
time dependency	time dependencies	time dependencies	steady-state	steady-state
mass dependency	mass constant	mass constant	mass flow constant	mass flow constant
homogeneity	completely homogeneous	completely homogeneous	completely homogeneous	geometric dependencies
performance	volume changing work	no work done	no work done	no work done
mixing	perfectly mixed	perfectly mixed	perfectly mixed	assumption of no mixing in axial direction

2.4. Homogeneous Reactor Model

In order to train the ANN, a representation of the entire range of thermochemical states defined by temperature, pressure and species composition has to be pre-computed. Therefore, the auto-ignition of a $PODE_{3/4}/air$ mixtures are solved using a constant-volume zero-dimensional homogeneous reactor model. Each state in this model can be accessed by the initial thermochemical state and the time as a measurement for the reaction progress. By parameter variation of the initial state, all thermochemical states within defined boundaries can be covered. In general, HR models are used to couple chemical kinetics included in the detailed mechanism with fundamental conservation principles, like mass and energy conservation. Hence, they enable calculation of thermochemical properties such as the temporal evolution of species, pressure and temperature. All HR models are identical in their assumption of a perfectly mixed system [70], which results in an auto-ignition delay equal to the chemical ignition delay. Furthermore, the reactors are assumed to be adiabatic and a simplified assumption of air, consisting of 21% oxygen and 79% nitrogen, is used as oxidizer. In this work, the HR implementation of the Cantera software package [24] is employed and all relation codes can be found in the GitLab repository¹.

Tab. 2.4 shows an overview of the most common HR models and their properties. The ignition in a CI engine occurs around the moment when the piston reaches the top dead center (TDC). Since at this time the volume is assumed as constant while the pressure rises, this part of the combustion can be described with the constant-volume and fixed-mass reactor model.

For the constant-volume and fixed-mass reactor under the above-mentioned assumption of an ideal homogeneous mixture, a system of first-order ODEs describing the temperature and pressure evolution, has been derived by Turns [70]. Starting with the first law of thermodynamics, which takes due to the absence of work the following form:

$$\frac{du}{dt} = \frac{\dot{Q}}{m}, \quad (2.14)$$

where u is the specific internal energy, \dot{Q} the heat transfer rate, and m the mass. The internal energy can then be expressed in terms of the systems species composition as

$$u = \frac{U}{m} = \frac{\sum_{i=1}^N N_i \bar{u}_i}{m}, \quad (2.15)$$

¹Codes used in this thesis can be found under: <https://git.rwth-aachen.de/pascal.roth.18/ba-pascal-roth>

where N_i are the number of moles for species i and \bar{u}_i the corresponding molar internal energy. Temporal differentiation of Eq. 2.15 leads to

$$\frac{du}{dt} = \frac{1}{m} \left[\sum_i (\bar{u}_i \frac{dN_i}{dt}) + \sum_i (N_i \frac{d\bar{u}_i}{dt}) \right] \quad (2.16)$$

Assuming ideal gas behavior and a constant volume the derivative of the internal energy $\bar{u}_i(v, T)$ is

$$\frac{d\bar{u}_i}{dt} = \frac{\partial \bar{u}_i}{\partial T} \frac{dT}{dt} + \frac{\partial \bar{u}_i}{\partial v} \frac{dv}{dt} = \frac{\partial \bar{u}_i}{\partial T} \frac{dT}{dt} = \bar{c}_{v,i} \frac{dT}{dt}, \quad (2.17)$$

where $\bar{c}_{v,i}$ is the molar specific heat capacity of species i . Hence, the internal energy is only dependent on the temperature $\bar{u}_i(T)$. This expression connects the internal energy with the temperature of the system. By introducing the definition of the molar concentrations $[X_i]$, as well as the species production rate $\dot{\omega}_i$

$$N_i = V[X_i] \quad (2.18)$$

$$\frac{dN_i}{dt} \equiv V\dot{\omega}_i, \quad (2.19)$$

where V denotes the volume, temperature and internal energy are coupled as follows:

$$\frac{dT}{dt} = \frac{\frac{\dot{Q}}{V} - \sum_i \bar{u}_i \dot{\omega}_i}{\sum_i [X_i] \bar{c}_{v,i}}. \quad (2.20)$$

By employing the ideal gas law, Eq. 2.20 can be expressed using enthalpies and specific heat capacity at constant pressure $\bar{c}_{p,i}$. Hence, it can be rewritten as follows:

$$\frac{dT}{dt} = \frac{\frac{\dot{Q}}{V} + RT \sum_i \dot{\omega}_i - \sum_i \bar{h}_i \dot{\omega}_i}{\sum_i [X_i] (\bar{c}_{p,i} - R)} \quad (2.21)$$

In this work, an adiabatic system is assumed, thereby the equation can be simplified to

$$\frac{dT}{dt} = \frac{RT \sum_i \dot{\omega}_i - \sum_i \bar{h}_i \dot{\omega}_i}{\sum_i [X_i] (\bar{c}_{p,i} - R)}. \quad (2.22)$$

In constant-volume reactors the time derivative of the pressure is of interest too and can be determined by differentiating the ideal-gas law, which leads to the following expression:

$$\frac{dp}{dt} = RT \sum_i \dot{\omega}_i + R \sum_i [X_i] \frac{dT}{dt} \quad (2.23)$$

The initial conditions of the HR are the equivalence ratio, temperature and pressure. Those conditions form, together with Eq. 2.22 and 2.23 an initial-value problem. The mass included in the HR is constant for the combustion process, due to the closed system. Also, the energy of the system has to be constant due to the assumption of an adiabatic system. In a CFD simulation, however, convective and diffusive mass and heat transport occur across the boundaries of each cell. Accordingly, the thermochemical state changes over time due to the combustion reactions and transport across the system boundaries. Therefore, the initial conditions and the physical time are unknown, which makes a correlation between the thermochemical state of the HR and the CFD cell via these parameters not possible. Instead, look-up parameters are employed. Hence, the mixture fraction variable and absolute enthalpy replace the equivalence ratio and the temperature. The physical time in the HR has to be non-dimensionalized in order to be utilized in the CFD, therefore the progress is determined by species mass fractions. A linear combination of these species forms the reaction progress variable. In CFD simulations, conservation equations are used to include the flow of the look-up variables between the different cells and calculate their progress for a variable time step.

2.4.1. Mixture Fraction of Combustion Elements

To identify the chemical state, Φ is replaced by the mixture fraction (Z). Both variables identify the ratio of fuel and oxidizer, but while Φ takes the initial amounts of both species, Z is defined by the ratio of elements involved in combustion. Under all conditions, this definition of Z remains representative of the local mixture of fuel and oxidant. Due to the assumption of a closed system and the usage of element ratios, Z is time-invariant in the HR model. For a system with two inlet streams, Bilger [8] introduced a general definition of a mixture fraction

$$Z = \frac{\beta - \beta_2}{\beta_1 - \beta_2} \quad (2.24)$$

where β is an arbitrary coupling function. In the case of $PODE_n$ diesel blend, the combustion process is mostly non-premixed and hence has an oxidizer and fuel stream. In the following, the fuel and oxidizer streams are denoted by 1 and 2 respectively, with the consequence that the mixture is in range of 0 to 1 or pure oxidizer to pure fuel. In order to find a suitable coupling function the mixture fraction for the chemical element j is defined as follows:

$$Z_j = \sum_{i=1}^{n_s} \frac{a_{ij} M_j}{M_i} X_i, \quad (2.25)$$

where n_s is equal to the number of species, M_i the molecular weight of species i , M_j the molecular weight of the element j and a_{ij} the number of elements j in species i . X_i indicates the mole fraction of species i in the stream. The coupling function β is then defined as the weighted sum of elemental mixture fractions

$$\beta = \sum_{j=1}^{n_e} \gamma_j Z_j = \sum_{j=1}^{n_e} \gamma_j \sum_{i=1}^{n_s} \frac{a_{ij} M_j}{M_i} X_i, \quad (2.26)$$

where n_e denotes the number of elements and γ_j the weight for a certain element. In the case of the combustion of $PODE_n$, the only elements are hydrogen, oxygen, and carbon. The weights are usually chosen that β is 0 for the stoichiometric condition. In that way the mixture fraction is calculated as follows:

$$Z = 2 \cdot Z_c + 0.5 \cdot Z_H - Z_O \quad (2.27)$$

Combining Eqs. 2.24, 2.25 and 2.26, the conservation equation solved in the CFD simulation for Z can be written as [65, 68]:

$$\rho \frac{\partial Z}{\partial t} + \rho \mathbf{u} \cdot \Delta Z = \frac{-1}{\beta_1 - \beta_2} \sum_{j=1}^{n_e} \gamma_j \sum_{i=1}^{n_s} \frac{a_{ij} M_j}{M_i} \Delta \cdot (\rho Y_i \mathbf{V}_i), \quad (2.28)$$

where ρ denotes the density, \mathbf{u} the velocity vector \mathbf{V}_i the diffusion velocity of species i . This general conservation equation makes no assumption concerning the diffusive flux. Furthermore, chemical source terms are not included since they vanish throughout the summation because elements are neither destroyed nor created.

2.4.2. Absolute Enthalpy

Since the model derived is valid for a certain initial pressure and due to the characterization of the initial composition by the mixture fraction, a look-up variable for the initial temperature has to be defined in order to determine the initial thermochemical state. Therefore, this thesis employs the absolute enthalpy h of the system. The conservation equation for h is denoted as follows [57]:

$$\rho \frac{Dh}{Dt} = \frac{\partial \rho h}{\partial t} + \frac{\partial}{\partial x_i} (\rho u_i h) = \frac{Dp}{Dt} - \frac{\partial q_i}{\partial x_i} + \tau_{ij} \frac{\partial u_i}{\partial x_j} + \dot{\mathcal{Q}} + \rho \sum_{k=1}^N Y_k f_{k,i} V_{k,i} \quad (2.29)$$

where τ_{ij} is the viscous tensor, q the heat flux and $\dot{\mathcal{Q}}$ the heat source term. Thereby, $\dot{\mathcal{Q}}$ does not include the heat released by combustion inside the control volume, but instead, heat transfer attempted through an external source like an electric spark. $\rho \sum_{k=1}^N Y_k f_{k,i} V_{k,i}$ is the power produced by volume forces f_k where $V_{k,i}$ denotes the i -component of the diffusion velocity V_k of species k . The above expression includes chemical terms in addition to sensible enthalpy which makes it not trivial to implement in 3D CFD simulations.

In the case of the HR model, the defined absolute enthalpy can be simplified due to different assumptions. The system is assumed to be closed and adiabatic, which eliminates $\dot{\mathcal{Q}}$. Furthermore, the reactor is supposed to be a perfectly mixed system and hence no diffusion takes place in it. Thus, Eq. 2.29 is changed as follows

$$\rho \frac{Dh}{Dt} = \frac{\partial \rho h}{\partial t} + \frac{\partial}{\partial x_i} (\rho u_i h) = \frac{Dp}{Dt} - \frac{\partial q_i^0}{\partial x_i} + \tau_{ij} \frac{\partial u_i}{\partial x_j} + \dot{\mathcal{Q}}^0 + \rho \sum_{k=1}^N Y_k f_{k,i} V_{k,i}^0. \quad (2.30)$$

The total derivative of the pressure can be rewritten, resulting in Eq. 2.30 to be expressed as:

$$\frac{\partial \rho h}{\partial t} + \frac{\partial}{\partial x_i} (\rho u_i h) = \frac{\partial p}{\partial t} + u_i \frac{\partial p}{\partial x_i} + \tau_{ij} \frac{\partial u_i}{\partial x_j}. \quad (2.31)$$

Since the HR reactor is defined as a zero-dimensional system, all spatial derivatives vanish. This leads to:

$$\frac{\partial \rho h}{\partial t} = \frac{\partial p}{\partial t}. \quad (2.32)$$

Accordingly, the temporal change of the enthalpy is only dependent on the work, which is performed by the increase of the pressure. In a regular CFD simulation, this enthalpy is used in order to identify the thermochemical state in the HR model. At the initialization of the mixture, no pressure increase happens and the enthalpy can be reduced to sensible enthalpy $h_{s,i}$ and enthalpy of formation $\Delta h_{f,i}^\circ$ under standard conditions ($T = 298.15K$ and $P = 1.0135bar$) [57]:

$$h_{init}(T) = \sum_i [(\Delta h_{f,i}^\circ(T_{ref}) * Y_i) + \Delta h_{s,i}(T)] = \sum_i \left[(\Delta h_{f,i}^\circ(T_{ref}) * Y_i) + \int_{T_0}^T c_{p,i} dT \right], \quad (2.33)$$

As the reactions occur, sensible enthalpy increases due to the rise in temperature while the chemical bonded energy decreases. The deviation between both is the work performed by pressure. For simplicity, in this thesis, the enthalpy h_{init} characterizes each state in the HR process with a certain initial thermochemical state and thus is used as a look-up variable to access the thermochemical state-space.

2.4.3. Progress Variable

With Z and h_{init} the initial thermochemical state is calculable at any time in HR. Using the conservation equation, the variables can also be adjusted in the CFD simulation in the case of enthalpy or mass transport. To describe the reaction progress and thus uniquely identify each further thermochemical state, time is used in the HR model. In CFD simulations, however, there is no possibility to determine this physical time. For this reason, the reaction progress variable (Y_c) is introduced as a dimensionless time substitute. In this subsection the theoretically properties are discussed and Y_c used in this work is derived.

Y_c is defined by a weighted linear combination of certain species mass fractions:

$$Y_c = \sum_i \alpha_i Y_i, \quad (2.34)$$

where α_i is a weight factor for each species i . Accordingly, each point of the reaction is characterized by the amount of certain species it contains. The corresponding conservation equation can be expressed as [65]:

$$\rho \frac{\partial Y_c}{\partial t} + \rho \mathbf{u} * \Delta Y_c = -\Delta \cdot (\rho Y_c \mathbf{V}_c) + \dot{\omega}_c, \quad (2.35)$$

where the diffusive flux of the process is defined as $Y_c \mathbf{V}_c = \sum_i \alpha_i Y_i \mathbf{V}_i$ and the chemical source term as $\dot{\omega}_c = \sum_i \alpha_i \dot{\omega}_i$. Since the conservation equations of Z and h_{init} do not include chemical source terms and the species mass fractions are predicted directly, $\dot{\omega}_c$ is the only source term that has to be predicted in the derived model. The combination of species that form Y_c has to fulfill certain criteria. Firstly, the set of species should uniquely characterize each point in the thermochemical state-space [11]. Thus, Y_c has to be monotonically increasing and therefore indicating the equilibrium state of the reaction as its maximum value [5]. Besides, Y_c should be zero at the beginning of the reaction. However, in CFD simulations Y_c is

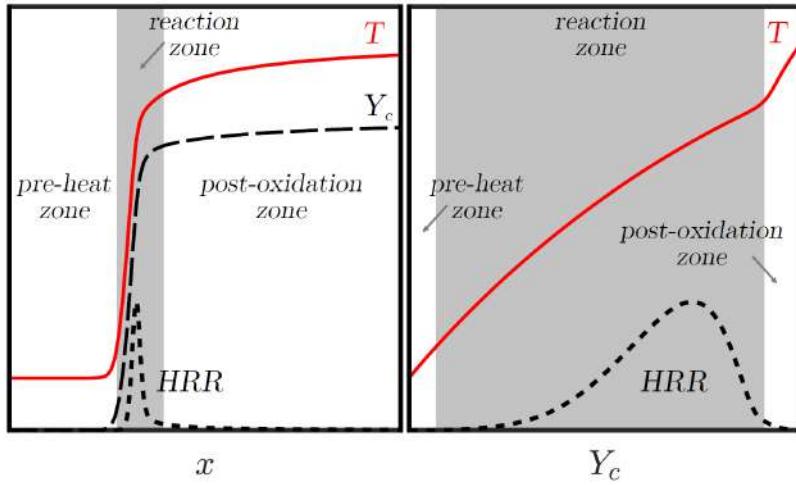


Figure 2.4.: Influence of the Progress Variable Space - left: Temperature, HRR and Y_c in pyhiscal time space. Right: HRR and temperature in Y_c space [65]

initialized with a small, non-zero value since otherwise the corresponding gradient is zero and thus no ignition occurs.

The theoretical influence of a progress variable for premixed flamelets is shown in Fig. 2.4 where a comparison between temperature and HRR evolution in physical space and Y_c space is given. Thereby, the reaction zone almost spans the whole computational domain in Y_c space, while it is tiny in physical space. Consequently, gradient resolution is increased by transforming into Y_c space. Optimization of the gradient resolution is achieved when Y_c increases in a linear shape.

In order to approximate the ideal linear shape, this thesis employs a combination of negative fuel consump-

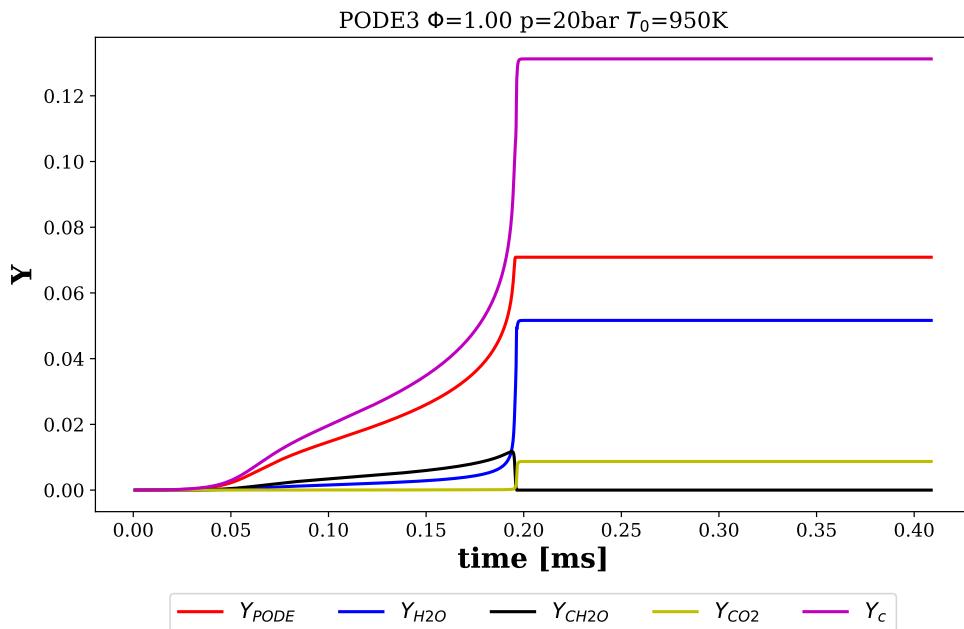


Figure 2.5.: Progress variable and the weighted species mass fractions that form Y_c

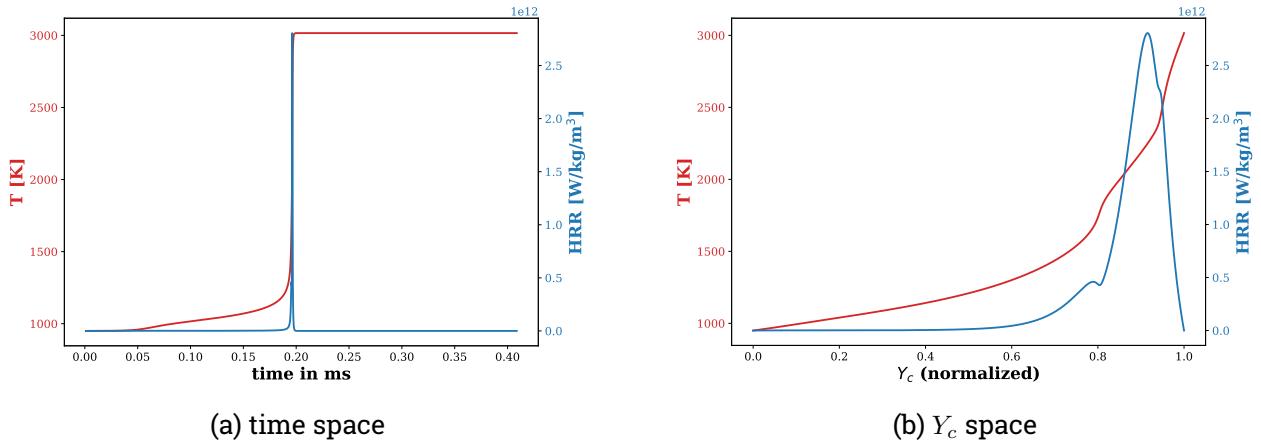


Figure 2.6.: HR executed with the mechanism by Cai et al. [13]. Monitored is temperature and HRR in physical time and Y_c space.

tion, CH_2O as intermediate species and the reaction products H_2O and CO_2 . $PODE_{3,4}$ is used to cover the start of the reaction. However, to ensure a monotonic increasing behavior of Y_c , its progress has to be inverted since the reactions branch out. The stable intermediate species CH_2O is directly generated by the branching of $PODE_{3,4}$. It is used to highlight the reaction period prior to ignition and since it vanishes in further reactions, to limit the steep gradient at the HT ignition. The non-monotonous behavior of CH_2O must under no circumstances jeopardize the monotony of Y_c , as otherwise, it is no longer possible to clearly assign the thermochemical state-space. Therefore, the product species CO_2 and H_2O have to be selected in such a way that they compensate for this behavior. Furthermore, these species are used to cover the end of the combustion process since the fuel is completely used beforehand. The weighted species mass fractions and the final Y_c are displayed in Fig. 2.5. Thereby, Y_c is calculated as follows:

$$Y_c = 0.5 \cdot Y_{H_2O} + 0.05 \cdot Y_{CO_2} + 0.5 \cdot Y_{CH_2O} + 0.5 \cdot (Y_{PODE_n}(t = 0) - Y_{PODE_n}) \quad (2.36)$$

The weights of Y_{CH_2O} and Y_{H_2O} are determined equally since H_2O is the earliest generated reaction product and defined to compensate for the decrease of CH_2O in order to ensure the monotony of Y_c . The weight of Y_{CO_2} is chosen small to limit the gradient at the main ignition. However, it has to be large enough to balance a small non-monotonic part in the H_2O evolution. For the comparison with the GRM of Haspel [27] the derived Y_c is reduced to H_2O and CO_2 since the other species are not available in the GRM model. A further explanation and the computation formula is given in subsection 4.2.

The gradient resolution capability and the stretch of the reaction zone are shown for the temperature and HRR progress in Fig. 2.6. Thereby, the derived two-stage auto-ignition behavior is visible in Y_c space and therefore confirms the expected combustion characteristic of fuel with large CN. The evolution of the species in progress variable space is shown in Fig. A.3. Thereby, it is visible that despite an equivalence ratio of 1.0 and hence stoichiometric conditions, the educts do not fully react to H_2O and CO_2 , instead CO and O_2 are part of the reaction products.

3. Data-Driven Modeling of Chemistry

This work aims to provide a combination of artificial neural networks that replace the ODE system in the simulation and can accurately predict the relevant species and other thermochemical parameters throughout the entire reaction process. Those networks are used by reason of their ability to memorize and interpolate unknown parameter values given a thermochemical state as input while having a low computational effort. In contrast to the tabulated chemistry approach discussed in subsection 2.3.3, ANNs can handle the increased complexity in combustion simulations as well as high dimensional manifolds without performance loss nor using gigabytes of storage [7]. The general capability of ANN to predict chemical kinetic properties has been shown in Lapeyre et al. [39].

This chapter aims to introduce Neural Networks, their central idea and their architecture. The basic training algorithm of NNs is discussed, as well as possibilities to optimize learning are presented. Furthermore, network adjustments for chemical modeling are given.

3.1. Neural Networks

Machine learning (ML) is a study field of artificial intelligence (AI) that enables systems to automatically learn and improve from experience without or with little explicit human interference [48]. It focuses on the development of algorithms that build mathematical models which recognize patterns in the given data and thereby are able to make predictions or decisions. Hence, the core objective of the model is to generalize from its experience and create the ability to perform accurately on new, unseen tasks that are similar to the ones it has been trained with [10]. ML methods are usually categorized as being either supervised or unsupervised. In Fig. 3.1 the schematic structure of both categories is shown. Thereby, in supervised learning, a model at hand is learned on certain data along with its respective outputs, so-called labels. Therefore, the model parameters are adjusted by a direct comparison between the actual network output and the desired output. Accordingly, supervised learning is a closed-loop feedback system. Once a model is learned on known data, it can be further fed with another set of data whose labels are unknown. In unsupervised learning, those prior labels are inaccessible or accessible but unimportant for the application being addressed. Here the model is trained to describe the associations and patterns among the set of inputs [28]. Due to the nature of the task in the thesis, a supervised approach is taken.

Artificial Neural Networks or simply Neural Networks (NNs) are one of the most prominent and most powerful ML methods. Their central idea is to extract linear combinations of the inputs as derived features and then model the target as a nonlinear function of these features [28]. The name neural network is chosen because NNs have first been developed as models for the human brain and hence are structured similarly even if much simpler. They consist of processing elements (called neurons) and connections between them with coefficients (weights) bound to the connections. NNs learn by processing samples, each of which contains a known set of inputs and a set of labels. Thus, learning from a given sample is the difference in the state of the network before and after processing the sample and updating the model parameters.



Figure 3.1.: Supervised and unsupervised learning method, where x_p and \hat{y}_p denote the input and output of the model. In supervised learning the weight \mathbf{W} are adjusted according to the error e_p which is the difference between \hat{y}_p and the desired output y_p . In unsupervised learning solely x_p influences the weights [17]

Thereby, the model parameters are the weights and bias terms of each neuron in the network. Being given a sufficient number and variety of samples, the NN becomes capable of predicting results from inputs, using the associations built from the training set. Besides, NN can be used for classification and regression tasks. In this section, the NN used in this thesis is introduced by explaining its general working principle and architecture.

3.1.1. Perceptron

The perceptron [62], also referred to as a McCulloch–Pitts neuron or linear threshold gate, is the earliest and simplest NN. Rosenblatt used a single-layer perceptron for the classification of linearly separable patterns, where the two decision regions are separated by a linear hyperplane. Each neuron in this kind of neural networks is a basic processing unit which takes the weighted sum of its continuous valued inputs to form its (scalar) *net activation* or simply *net*:

$$net_j = \sum_{i=1}^d x_i w_{ji} + w_{j0} = \sum_{i=0}^d x_i w_{ji} \equiv \mathbf{w}_j^t \mathbf{x}, \quad (3.1)$$

where x_i denotes the input i of unit j and w_{ji} the corresponding weight. The bias term w_{j0} is an additional parameter in order to adjust the inputs of the neuron. It can be interpreted as the intercept added in a linear equation. The basic structure is shown in Fig. 3.2. As analogy to neurobiology the expression "synapse" and "synaptic weight" are used for such connections and their weights. Each unit uses its net to generate a continuous-valued output according to a transfer function called the *activation function*, denoted by ϕ . This function represents a linear or nonlinear mapping from the input to the output [17]. Popular activation functions are the sigmoid function

$$\phi(net_j) = \frac{1}{1 + e^{net_j}}, \quad (3.2)$$

the hyperbolic tangent function

$$\phi(net_j) = \tanh(net_j) \quad (3.3)$$

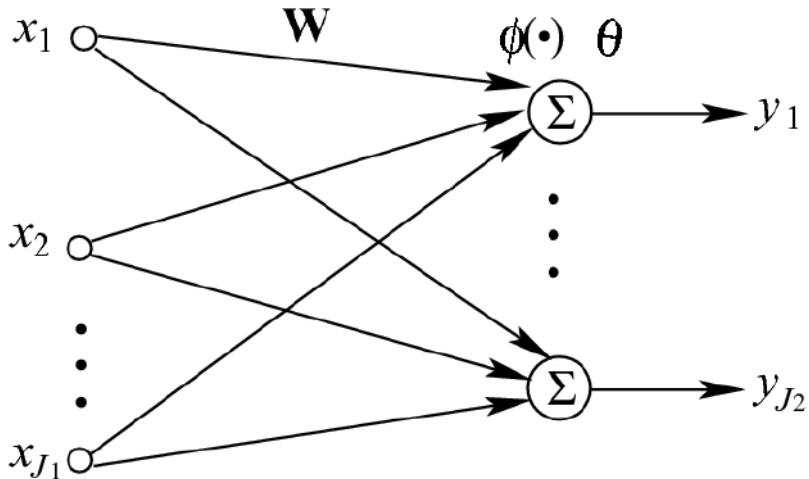


Figure 3.2.: The mathematical model of a single layer perceptron with the inputs x_j and their corresponding weights w_{ji} , summarized to W . Together with the bias $\theta = w_{j0}$ these inputs are summed up and the activation function ϕ generated the output y_i

and the rectifier activation function

$$\phi(\text{net}_j) = \max(0, \text{net}_j). \quad (3.4)$$

Already in 1962, Rosenblatt was able to prove that the presented structural combination of one or more single processing units and an activation function finds a hyperplane that can perfectly separate each linearly separable set of input patterns $x \in X$. Starting from an arbitrary initial state, the perceptron learning procedure always converges and yields a decision hyperplane between the two classes in limited time. Thus, the weights of the perceptron converge to a fixed point within a finite number of updates [17, 63]

3.1.2. Multi-Layer-Perceptron

Since most classification and regression tasks are too complex that a linear hyperplane is appropriate, layers of units are added between the input units x and the output units z , making the network a Multi-Layer-Perceptron (MLP). Due to the non-appearance of their mode of operation and the non-representants of real parameters or facts, they are called hidden layers with hidden units y . The nonlinearities in the nodes combined with the additional layers are responsible for the MLP to be a universal approximator and thus suitable for non-linearly separable cases [17]. Geometrically seen is the MLP a combination of linear hyperplanes that approximate the underlying function of the given data. It has been mathematically proved that a three-layer MLP using the sigmoidal activation function is able to approximate any continuous multivariate function to any accuracy, subject to a sufficient number of hidden nodes [32]. In practice, however, complex problems can often be adapted more quickly and effectively with two or more hidden layers due to the greater flexibility of the network and the fewer connection weights used [20, 50].

MLPs are feedforward neural networks, which means that the connections between neurons are in one direction only. In layered feedforward networks, there is no connection between the neurons in the same layer and there is furthermore no feedback between layers. In general, MLPs are fully connected feedforward networks, with the consequence that every node in any layer is connected to every node in its adjacent

forward layer. The structure of a three-layer fully connected MLP is given in Fig. 3.3. During feedforward operation, a d -dimensional input is presented to the network and each input unit then emits its corresponding component x_i . Each hidden unit of the hidden layers computes its net activation using Eq. 3.1 and emits an output depending its activation function. Repeated for the number of hidden layers, the final output signals are the inputs of the output layer. Each output unit functions in the same manner as the hidden units do, however, the activation function can be different.

MLPs prove to be efficient for function approximation in high-dimensional spaces since there is no dependency between the error convergence rate and the input dimensionality. Conventional linear regression methods on the other side suffer from the curse of dimensionality, which results in a decrease of the convergence rate with an increase of the input dimensionality [4]. For the best performance in the context of generalization, the complexity of the NN architecture should match the complexity of the function underlying the data. If the model's number of adjustable parameters is smaller than the underlying function requires, the model has little flexibility and generalization is poor. Thus it yields underfitting with a low bias but high variance. In contrast, a network with too many adjustable parameters also gives a poor generalization performance, since it is too flexible and fits too much of the noise of the training data, thus yielding overfitting with high bias and low variance. By balancing bias and variance, the complexity of the model can be optimized [2]. Using a geometrical interpretation the MLP achieves the best generalization ability if the minimal number of line segments or hyperplanes that can construct the basic geometrical shape of the target function is chosen. Therefore, the approximation only depends upon the basic geometrical shape of the target function and not on the dimensionality of the input space [73].

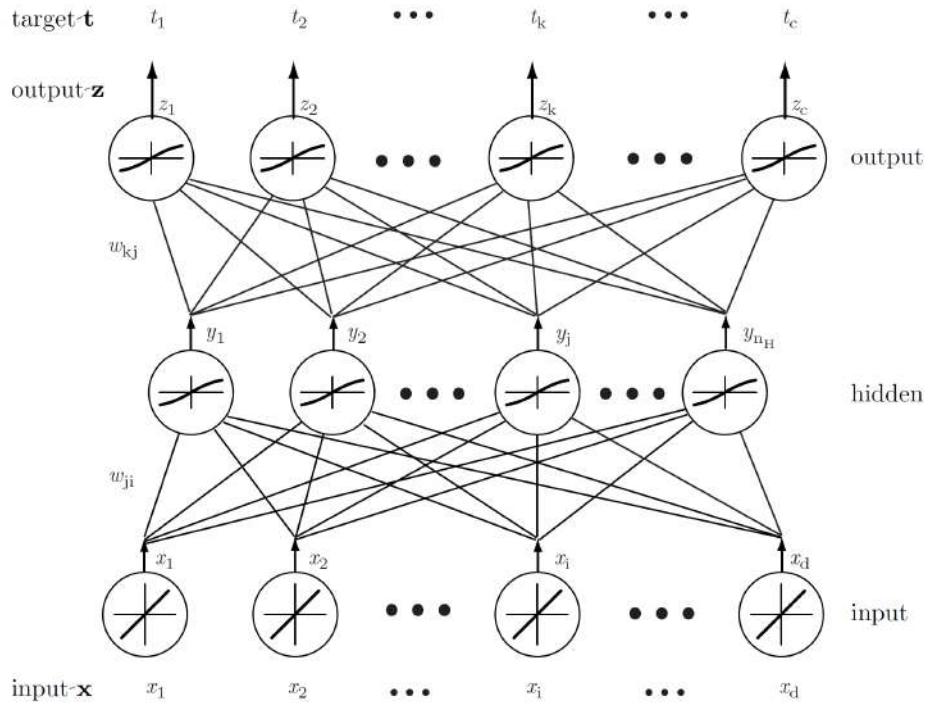


Figure 3.3.: Three layer fully connected network with sigmoidal activation function [18]

3.2. Training Process of Neural Networks

The basic approach in learning is to randomly initialize the weights of a given network. D-dimensional input patterns are presented to the network and the corresponding output is determined. The performance of the model is evaluated by applying an error or loss function that compares the network output with the intended output. Thereby, the loss function is a scalar function of the model parameters that is minimal when the output of the network matches the intended output. In this thesis, the training loss E is determined by the mean-squared-error (MSE) loss. Therefore, E is calculated as the mean sum over the squared difference between the desired output t_k and the actual output z_k for each output unit k [28]:

$$E(\mathbf{w}) \equiv \frac{1}{2n_c} \sum_{k=1}^{n_c} (t_k - z_k)^2 = \frac{1}{2n_c} (\mathbf{t} - \mathbf{z})^2, \quad (3.5)$$

where \mathbf{t} and \mathbf{z} are the target and the network output vector of length n_c . Notice that the factor $\frac{1}{2}$ is used in E for the convenience of derivation. The model parameters of the network are represented by \mathbf{w} . Training now aims to minimize the loss function and thus capture the intended output as accurately as possible. Accordingly, training a NN is an optimization problem that is solved by using gradient descent. Thereby, the error is propagated through the network and the randomly initialized weights are changed in the direction that reduces the error [18]:

$$\Delta\mathbf{w} = -\eta \frac{\partial E}{\partial \mathbf{w}}, \quad (3.6)$$

where η is the learning rate and determines the relative size of change in weights. This iterative algorithm requires taking a weight vector at iteration m and updating it as follows:

$$\mathbf{w}(m+1) = \mathbf{w}(m) + \Delta\mathbf{w}(m). \quad (3.7)$$

The best learning results are achieved if the global minimum of E is found. Local minimums, on the other hand, impact a loss in training and generalization ability. The number of local minimums depends on the underlying function as well as on the number of hidden layers and hidden units [59].

Additionally, during training an optimization of the hyperparameters is carried out. Hyperparameters are used to control the training process as well as model architecture and should not be confused with the model parameters, the weights and biases of the NN. Two classes of hyperparameters can be identified: model and algorithm hyperparameters. The model hyperparameters are used in the model selection task and thus are not inferred while fitting the network. However, these parameters influence the learning capability and have to be adjusted in case the model complexity is chosen differently compared to the complexity of the underlying function. Examples of the model hyperparameters are the number of hidden layers and hidden units. Algorithm hyperparameters, in theory, do not influence the model performance, instead, they impact the speed and quality of the learning process. In practice, however, algorithm hyperparameters do influence the capability and they have to be optimized. Typical algorithm hyperparameters are the learning rate and the batch size. This section continues by deriving the backpropagation learning algorithm and then presenting methods to optimize the learning and to avoid learning failure.

3.2.1. Backpropagation Learning Algorithm

Though it is biologically implausible, backpropagation (BP) learning is the most popular learning rule for performing supervised learning tasks [64]. It is not only used to train feedforward networks such as MLP but also adapted to Recurrent Neural Networks. The working principle is considered as the generalization of the delta rule, also called the Least-Mean-Squares (LMS) algorithm. The LMS algorithm is applied for one-layer perceptrons, as described in subsection 3.1.1 and minimizes the MSE loss between the desired and actual network output. Thus, BP is also called the generalized delta rule [17]. However, BP can be implemented with different loss functions. This loss or cost function is used to evaluate the performance of the network since the optimum is reached when the cost function is minimized. By applying a gradient-search technique, the weights are adjusted in order to reduce the loss function. Due to the BP algorithm, MLPs can be extended to many layers [28].

The key to BP is the chain rule of differentiation. With the aim to present the working principle of BP, the simple network shown in Fig. 3.4a is used. It consists of a single neuron perceptron with two input values and a linear activation function. For simplicity reasons, no bias terms are included and according to the use in this thesis, the MSE loss is employed. Since each network consists of a directed sequence of variables and operations, it can be represented as a computational graph. However, due to the complexity of NNs, this is only possible for small networks such as the one used here, whose graph is shown in Fig. 3.4b. The output of the example network is calculated as follows:

$$z = w_c \cdot (w_{x_1} \cdot x_1 + w_{x_2} \cdot x_2) \quad (3.8)$$

$$= w_c \cdot (a + b) \quad (3.9)$$

$$= w_c \cdot c, \quad (3.10)$$

where the abbreviations a, b, c are introduced for simplicity reasons. By applying the MSE-loss as defined in

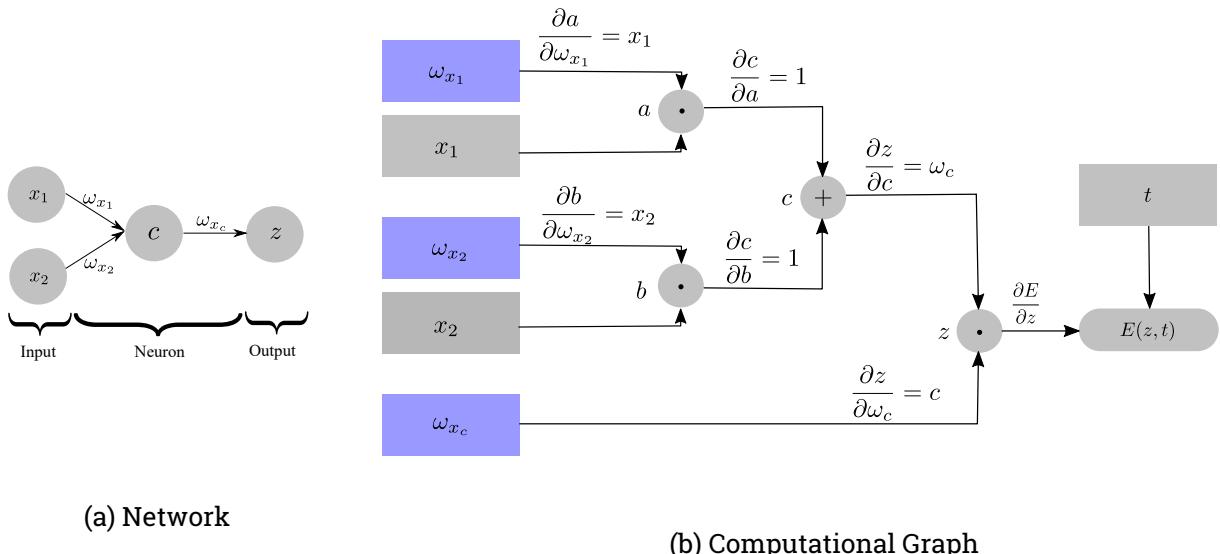


Figure 3.4.: The network structure and the corresponding computational graph for simplified BP explanation

Eq. 3.5, E is determined as:

$$E(z, t) = \frac{1}{2}(t - z)^2. \quad (3.11)$$

Since the weight w_i should be adjusted according to the negative gradient of the loss function regarding w_i , as seen in Eq. 3.7, a connection has to be found between E and w_i . However, E is only dependent on the output of the neuron z . At this point the chain rule of differentiation is used to propagate the error through the network. Hence, in order to reach w_i the derivatives of the quantities on the path between weight and loss function, as shown in the computational graph, are multiplied with each other. As an example, the chain for w_{x_1} is denoted as follows:

$$\frac{\partial E}{\partial w_{x_1}} = \frac{\partial E}{\partial z} \frac{\partial z}{\partial c} \frac{\partial c}{\partial a} \frac{\partial a}{\partial w_{x_1}} \quad (3.12)$$

In the case of an MLP with its multiple layers and units, BP is performed similarly. However, its implementation requires a nonlinear, monotonically increasing activation function. Prominent functions are the logistic function (Eq. 3.2), the hyperbolic tangent function (Eq. 3.3) and the rectifier activation function (Eq. 3.4). In the following, the BP algorithm for a fully connected three-layer MLP using the MSE loss as done in the work of Duda et al. [18] is presented. The architecture of the network is shown in Fig. 3.3. BP algorithms for NNs with other activation functions or structures are derived in a similar manner.

Training using BP is based on gradient descent which is presented in Eq. 3.6 and here used in component form for simplicity reasons. Accordingly to the introduced symbols the inputs are represented by x_i , the hidden units by y_j and the output units by z_k . First considering the hidden-to-output weights w_{jk} . Because the error is not explicitly dependent upon w_{jk} , the chain rule of differentiation is used:

$$\frac{\partial E}{\partial w_{jk}} = \frac{\partial E}{\partial \text{net}_k} \frac{\partial \text{net}_k}{\partial w_{jk}} = \delta_k \frac{\partial \text{net}_k}{\partial w_{jk}} \quad (3.13)$$

where the sensitivity δ_k of unit k is defined as

$$\delta_k \equiv -\frac{\partial E}{\partial \text{net}_k} \quad (3.14)$$

and describes the error change with the unit's activation. Using the chain rule and the differentiation of Eq. 3.5, the sensitivity is expressed as:

$$\delta_k \equiv -\frac{\partial E}{\partial \text{net}_k} = -\frac{\partial E}{\partial z_k} \frac{\partial z_k}{\partial \text{net}_k} = \frac{1}{n_c} (t_k - z_k) \phi'(\text{net}_k) \quad (3.15)$$

By applying the definition of the net activation from Eq. 3.1, the second derivative of the Eq. 3.13 can be simplified to:

$$\frac{\partial \text{net}_k}{\partial w_{jk}} = y_j \quad (3.16)$$

Taken together, the learning rule for the hidden-to-output weights is defined as:

$$\Delta w_{kj} = \eta \delta_k y_j = \frac{\eta}{n_c} (t_k - z_k) \phi'(\text{net}_k) y_j \quad (3.17)$$

The learning rule for the input-to-hidden weights w_{ji} is defined in a similar way, first starting by applying the chain rule of differentiation

$$\frac{\partial E}{\partial w_{ji}} = \frac{\partial E}{\partial y_j} \frac{\partial y_j}{\partial \text{net}_j} \frac{\partial \text{net}_j}{\partial w_{ji}}. \quad (3.18)$$

Evaluating the first term on the right hand side by using the chain rule again leads to:

$$\frac{\partial E}{\partial y_j} = -\frac{1}{n_c} \sum_{k=1}^{n_c} (t_k - z_k) \frac{\partial z_k}{\partial y_j} \quad (3.19)$$

$$= -\frac{1}{n_c} \sum_{k=1}^{n_c} (t_k - z_k) \frac{\partial z_k}{\partial \text{net}_k} \frac{\partial \text{net}_k}{\partial y_j} \quad (3.20)$$

$$= -\frac{1}{n_c} \sum_{k=1}^{n_c} (t_k - z_k) \phi'(\text{net}_k) w_{jk} \quad (3.21)$$

In analogy with Eq. 3.14 the sensitivity for a hidden unit is defined as:

$$\delta_j \equiv \phi'(\text{net}_j) \sum_{k=1}^c w_{kj} \delta_k \quad (3.22)$$

This expression is the key of the backpropagation algorithm since the sensitivity of a hidden unit is the sum of the individual sensitivities at the output units weighted by the hidden-to-output weights, all multiplied by $\phi'(\text{net}_j)$. Hence, the error is propagated from the output layer back to the hidden layer in order to perform the learning of w_{kj} . Thus, the learning rule for the input-to-hidden weights is:

$$\Delta w_{kj} = \eta x_i \delta_j = \eta x_i \phi'(\text{net}_j) \sum_{k=1}^c w_{kj} \delta_k \quad (3.23)$$

In practice, BP is not performed after each sample, in order to reduce computational effort and speed up the training procedure. Instead, batch learning is implemented, where the weight increments for each example are averaged over a defined number of training samples before the weights are adapted. Thus, batch learning is a deterministic optimization method. As the maximum batch size, all samples can be chosen. In this case, the weight update is performed at the end of one epoch, where epoch means, that the network has been trained once with all available training samples [64]. The total number of learning epochs, as well as the batch size, are part of the hyperparameters.

3.2.2. Training Optimization

Within the optimization process of the loss function using BP, training can get stuck in a local minimum or the learning rate can be too large, leading to the incapability to further reduce the error. Thus, BP does not ensure that the network converges to the global minimum. As more layers are added to the network as better it can approximate the target function due to increased nonlinearity and more precise feature representation. However, the increased network complexity and the larger number of adjustable parameters increase the number of local minimums and lead to further difficulties in training. With the aim to optimize training, the following techniques are introduced that preprocess the data, change the structure of the network, split the training data and adjust the learning rate.

Data pre-processing

Data pre-processing is performed once, prior to the training. It includes data cleaning, normalization, transformation, feature extraction and selection, etc. and results in the final training set. If there is much irrelevant and redundant information present or noisy and unreliable data included, data pre-processing is crucial in order to allow and simplify knowledge discovery [37].

With the aim to reduce the number of possible values of continuous features and thus avoid slow and ineffective learning, temporal discretization is applied. Since the solver of the HR can execute time steps equivalent to the smallest chemical time scale, the underlying function can be assumed as continuous and has to be discretized. For this purpose, adaptive time-stepping is applied which adjusts the time-step between two samples according to the gradient of Y_c . This gradient is chosen due to the extension of the reaction zone in Y_c space and the fact that it is constant when no reactions occur. Consequently, the majority of samples are taken in the reaction zone. In order to increase the resolution of the non-linearities, the time-step is decreased when the gradient is steeper. In this thesis, a four-staged time-stepping is implemented with the smallest time step of 10^{-10} s. Fig. 3.5 shows the ratio between the time passed and the samples taken. Thus, the connection is visible of how the amount of samples increases towards the ignition and especially

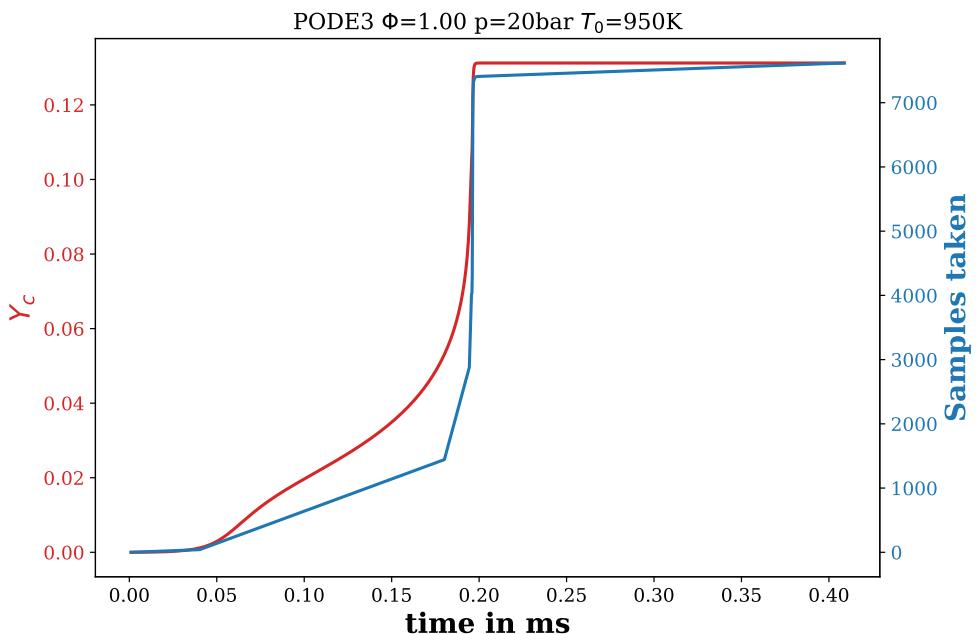


Figure 3.5.: Ratio between physical time and samples taken as well as Y_c evolution

while the ignition occurs. Furthermore, the Y_c progress is shown with the consequence that the correlation between its gradient and the step-size can be made.

Feature subset selection is the process of identifying features that cover the necessary information. This reduces the dimensionality of the data and enables the learning algorithms to operate faster and more effectively [37]. Feature selection in this work is part of section 3.3. In general, the resulting input layer contains all those features which, as independent variables, influence the result variable. Naturally, the network should prefer neither of these features over the other, as they differ only in their arbitrary presentation. Difficulties arise when values of the features have different orders of magnitude since then a relatively small change affects the network differently. Thus, the full data set should be scaled to have the same variance in each feature component. This can be ensured by the normalization of the single features to a range of zero to one [18] with the following rule:

$$\bar{x}_{ij} = \frac{x_{ij} - x_{j,min}}{x_{j,max} - x_{j,min}} \quad (3.24)$$

where x_{ij} is the element i of the feature j of the training data and $x_{j,max}, x_{j,min}$ the maximum and minimum of the feature for all samples. The resulting, normalized sample is denoted by \bar{x}_{ij} .

Training, validation and test data set

As mentioned in subsection 3.1.2, overfitting and underfitting occur if the number of adjustable parameters does not match the requirement of the underlying function. However, overfitting and underfitting cannot be avoided by only choosing the right number of parameters. Instead, also the number of training epochs influences the generalization ability. In order to optimize both, number of adjustable parameters and number of epochs, it is crucial to determine the performance on unseen data during training. As a consequence, the data set is split into training and validation set.

By splitting the data set, two kinds of loss can be monitored during training. Thereby, the validation set is passed through the network after a certain number of optimization steps and the error of the output is determined. However, the error is not used to adjust the weights. Therefore, the validation loss is an indicator if the model can predict unseen data. As long as it decreases, the model increases its generalization ability. If overfitting then occurs, the error in the validation set will increase while the error in the training

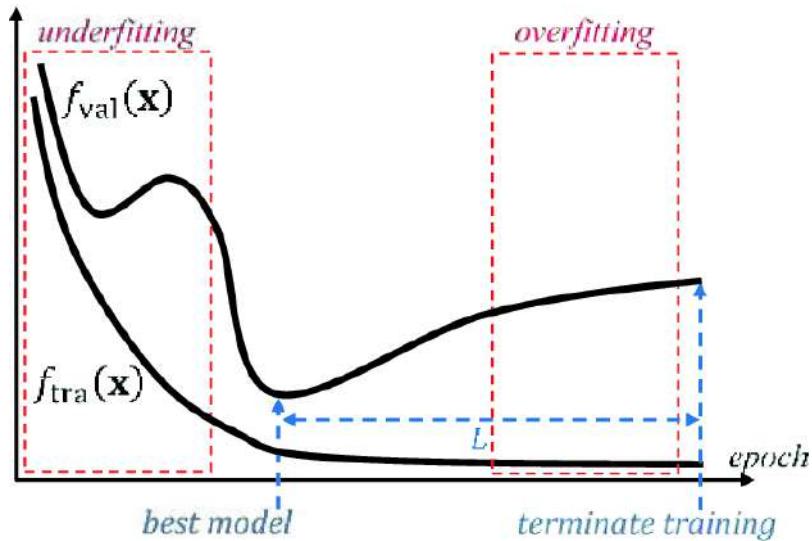


Figure 3.6.: Finding the best model by early stopping based on the validation performance [34]

data will continue to decrease, as seen in Fig. 3.6. When the validation error reaches its minimum, the training should be stopped to prevent the network from learning the high-frequency noise [17]. The training loss, on the other hand, is an indicator of whether training is taking place at all. Accordingly, it is used to adjust hyperparameters. Consequently, splitting the data is crucial in order to optimize training.

For optimal validation of the trained neural network, a third data set (test data) should be retained. This is neither used to adjust the connection weights nor to determine the end of the training and thus offers the possibility of an independent check of the network performance. When splitting the data sets, it should be ensured that all partial data sets statistically represent the same basic variety. In small data sets, there may not be enough data items for both training and validation. In this case, the whole data set is divided into n pieces, $n - 1$ pieces are used for training and the last piece is the validation set. This process of n -fold cross-validation builds n models; the numbers reported are the averages over all n validation sets. The extreme case of using only one data item for testing is known as leave-one-out cross-validation [36].

Dropouts

Standard BP learning builds up brittle co-adaptations that work for the training data but do not generalize to unseen data. Thereby, co-adaptation refers to when different hidden units in NNs have highly correlated behavior. Thus, the network is not using its full capacity efficiently since the dependant neurons detect the same feature and therefore give redundant information [30]. In order to break up these co-adaptations, dropout technique is used which randomly drops a certain percentage of units (along with their connections) from the neural network during training. Dropping means that the outputs of the nodes are set to be zero with the aim to deactivate the complete unit. Hence, the presence of any particular hidden unit becomes unreliable and the units are prevented from co-adapting too much. Therefore, it can be avoided that the network learns the training data noise. The dropout technique can be interpreted as a way of regularizing a NN by adding noise to its hidden units [66]. Fig. 3.7 shows the applied dropout technique for an MLP with two hidden layers. When the network is used for validation or in general applied to unseen data, all nodes are activated.

Dynamic Learning Rate

The learning rate is the step size of the optimization algorithm and therefore determines the amount the weights are updated during training. The performance of the BP algorithms is highly dependent upon a

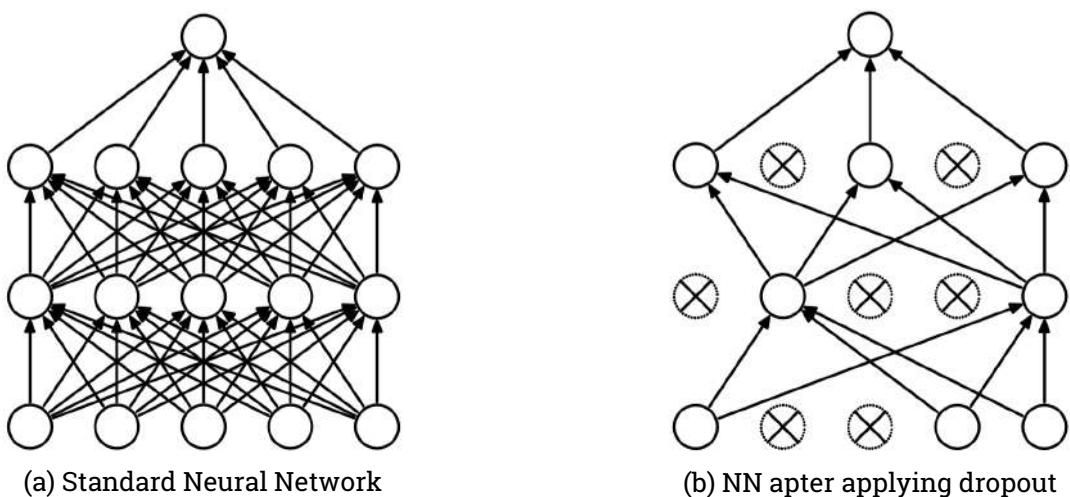


Figure 3.7.: Dropout Neural Net Model. Left: A standard neural net with 2 hidden layers. Right: An example of a thinned net produced by applying dropout to the network on the left. Crossed units have been dropped [66]

suitable selection of η . In principle, as long as η is sufficiently small to let the network converge to the mathematically closest minimum, η only defines the speed of the learning. Thus, the optimal learning rate η_{opt} is the one which leads to the local error minimum in one learning step and is determined by calculating the inverse of largest eigenvalue λ_{max} of the Hessian matrix of the error function [41]. The maximum η that is still converging is $\eta_{max} = 2\eta_{opt}$. If η is chosen small or close to η_{max} , the convergence speed of the network is slowed down.

Training is optimized when the global minimum of the loss function is reached. However, if η is chosen sufficiently small that training only optimizes towards the mathematically closest minimum, the risk increases that the found minimum is a local minimum. Consequently, the network loses a part of its capability to learn the underlying function. By choosing η large at the beginning of the training process, the BP algorithm can drop out of local minimums and the probability to reach the global minimum increases. Furthermore, in practice, networks are rarely fully trained to a training error minimum. Instead, η is chosen in a way that converges fast at the beginning of the training and then decreases in order to find the right epoch to stop the training in order to maximize generalization ability. A practical approach is to lower η if the loss function diverges and increase its value if learning seems unduly slow [18].

3.3. Network Adjustments for Thermochemical Modeling

This work aims to provide MLPs that model the thermochemical properties given the thermochemical state as input. However, the models are constructed and trained for a predefined initial pressure. Therefore, as initial conditions of the combustion process, only the temperature and equivalence ratio are parameterized. Adding the pressure as a variable is possible and results in increased complexity regarding data variance and required network architecture. This thesis refrains from introducing the pressure as a variable in order to compare it with the Global Reaction Mechanism developed by Haspel et al., as done in section 4.2. Hence, the evaluation of the models developed and a demonstration of the benefits of using ML techniques are feasible.

In this section, the adjustments of the general MLP structure given in subsection 3.1.2 are introduced. This includes the necessary input parameters, a structure that provides the ability to model the underlying function and the definition of the output parameters. The models are implemented using the PyTorch Deep Learning Library [54]. The training and evaluation codes, as well as the trained MLP models, can be found in the GitLab repository¹.

Input features

As mentioned before, the inputs characterize the thermochemical state in order to connect the CFD simulation to the data generated by the HR model and learned by the MLP. Furthermore, the used degree of polymerization of $PODE_n$ is part of the input vector in order to enable the network to be used for $PODE_3$ and $PODE_4$. The addition of the degree is possible due to the similar combustion of both fuels. However, if the framework is used for other tasks, the models are more likely trained separately for each fuel since the similarity of the underlying function can not be guaranteed and hence an unknown complexity is added which can prevent effective training.

As discussed in section 2.4, within the HR model each thermochemical state can be accessed with the initial state and the physical time. Considering a 3D CFD simulation, on the other hand, these variables cannot be determined using the information available in a single cell at each time. Consequently, look-up variables are employed which fully parameterized the entire thermochemical state-space and which can be

¹Codes used in this thesis can be found under: <https://git.rwth-aachen.de/pascal.roth.18/ba-pascal-roth>

computed using conservation equations. In accordance with section 2.4 the species composition defined by Φ is substituted by the mixture fraction (subsection 2.4.1) and the initial absolute enthalpy (subsection 2.4.2) replaces the initial temperature. Since the absolute enthalpy is dependent on the pressure increase, the selection of the initial value for all thermochemical states in the reaction process is a simplification of this thesis. In the case of a unitless time replacement, the progress variable defined in subsection 2.4.3 is employed. As a result, the input vector contains Z , h_{init} , Y_c and the degree of polymerization. Thereby, the pressure is not included since the networks are developed for constant initial pressure.

Activation function

In this work, the rectifier activation function is chosen. Accordingly, each unit employing the function is called a rectified linear unit (ReLU). The function is given in Eq. 3.4. It is a piecewise linear function which prunes the negative part to zero and retains the positive part. As a consequence, the network easily obtains sparse representation and thus results in the networks' non-linearity resulting from the path selection associated with individual neurons being active or not. Consequently, the computation of the output is a linear function of the network's input once the subset of active neurons is selected. Therefore, the model may be considered an exponential number of linear models that share parameters [49]. Due to this linearity and the constant zero gradient whenever a unit is inactive, the gradient used for learning (described in subsection 3.2.1) flows well through the network. Furthermore, the risk of the gradient vanishing effect which can happen as hyperbolic tangent or sigmoid activation are used, is decreased. Those functions have a gradient in range (0,1) and since in BP the learning update is computed by applying the chain rule, the gradients of each layer are multiplied. The gradient vanishing effect denotes the problem that if the gradients are small the corresponding weight update becomes vanishingly small, effectively preventing the weight from changing its value. Furthermore, ReLU function better transmits the error than sigmoid function, since sigmoid function limits the node outputs to the unity. Besides, the computational effort is reduced due to the absence of the exponential function in activation [23].

Output features and Network partitioning

Since only major combustion species are usually of interest for the evaluation of the process, the species mass fractions have to be determined for the species mentioned in subsection 2.2.1, namely the reaction educts O_2 , $PODE_n$ and products CO_2 , CO as well as H_2O . Due to the ability of the MLP models to directly predict the mass fractions, the balance equation of the individual reactions can be avoided in addition to the stiff ODE system. The only "species" for which a balance equation must be solved and accordingly a chemical source term calculated, is the process variable. This is necessary because as part of the look-up variable it determines the thermochemical state in the simulation. Therefore, its progress must be determined depending on the convective and diffusive flow as well as the chemical reactions. Since the conservation equation of the mixture fraction variable does not include any kind of source terms, no output quantity has to be defined for it. The conservation equation for the absolute enthalpy, on the other hand, includes thermodynamic properties that are changed due to the combustion. Thus, the MLP models must be able to predict quantities such as temperature, pressure, density, etc. Furthermore, the HRR has to be predicted in order to determine IDs. However, the output values of the models are chosen in order to perform the comparison with the GRM by Haspel [27]. Therefore, this work refrains from adding the source term of the progress variable as an output quantity and limits the thermodynamic quantities to temperature and pressure.

When comparing the curves of the species mass fractions and the other thermochemical parameters given in chapter 2, large differences are evident. Obviously, the mass fractions of the educts decrease with increasing reaction progress, while those of the products rise. The HRR shows a strongly pronounced maximum at the time of ignition. Temperature and pressure are similar to the products, but increase earlier and have a different order of magnitude. In order to reduce complexity and to facilitate as well as accelerate training, this work does not define all variables as outputs of one network. Instead, the quantities are grouped

Tab. 3.1.: Properties of MLPs used in this work to predict thermochemical properties

Input	Output	Architecture	Activation Function
$PODE_n, Y_c, Z, H$	T, P	64, 128, 128, 64	ReLU
$PODE_n, Y_c, Z, H$	$Y_{CO}, Y_{CO_2}, Y_{H_2O}$	64, 128, 256, 128 64	ReLU
$PODE_n, Y_c, Z, H$	Y_{PODE_n}, Y_{O_2}	64, 128, 128, 64	ReLU
$PODE_n, Y_c, Z, H$	HRR	64, 128, 256, 128 64	ReLU

according to their functional progress, thus the underlying functions are similar. This results in a total of four networks. The number of intermediate layers and hidden neurons must be tested individually for each output group. There are different empirical principles for the choice of this architecture, but in the end, the optimal constellation has to be found according to the principle of trial and error. This optimum represents the smallest possible network that can still map the problem with sufficient accuracy [47]. The properties of the networks can be taken from Tab. 3.1 where the activation function is used for both, hidden and output units. The architecture is denoted by the number of hidden units for the corresponding hidden layer. The number of input and output quantities, on the other hand, defines the number of units of the input and output layer. In the case of the reaction product species, the network has an additional layer compared to the other species mass fraction network and hence is able to capture more complex functions. This is necessary due to a different shape of the Y_{CO} evolution compared to evolutions of Y_{H_2O} and Y_{CO_2} . Y_{CO} has its maxima not at the end of the reaction, but rather at the low-temperature ignition. Therefore, while the other two curves are increasing within the main ignition, CO is used and the curve decreases. By adding the layer it is ensured that the network captures both curves types accurately. Since the underlying functions for all educts are similar, no additional layer is necessary and the same network architecture as used for temperature and pressure is chosen.

4. Results and Discussion

With the purpose of modeling chemistry in 3D CFD simulations with low computational effort and memory requirement, this thesis introduces a combination of Multi-Layer Perceptrons that are able to predict the species mass fraction and the thermodynamic properties directly, given the thermochemical state at a certain time as input. The thermochemical state defined by pressure, temperature and species composition is distinguished by look-up variables that can be calculated by conservation equations. Thereby, NNs are not only able to learn the thermochemical states with which they have been trained, but also to interpolate between them and thus to make accurate predictions for every possible state in the simulation. While training the MLPs can be computationally costly, their application within the simulation does not lead to any significant effort.

This chapter evaluates the derived MLP model regarding the capability to learn the training data and interpolate unseen thermochemical states. Furthermore, a comparison is done between the model introduced in this thesis and the Global Reaction Mechanism by Haspel [27] regarding the learning requirements and the prediction of the thermochemical properties as well as ignition delays.

4.1. Data-driven Modeling of Chemistry

MLPs are evaluated with regard to their ability to acquire training data, to interpolate this data and thus predict even unknown labels with high accuracy. Therefore, this section firstly presents the modifications done to allow and optimize learning and shows the accuracy of the best-trained model. In the second part, unknown data is predicted and the capability to generalize from the learned pattern discussed. Furthermore, the prediction of ignition delays is evaluated.

Training

The four MLPs presented in section 3.3 have been trained with a dataset including the results of the HR model for an initial pressure of 40 bar, an initial temperature in the range of 650 to 1250K with 30K step size, as well as an equivalence ratio in the range of 0.5 to 1.5 with step size 0.05. The network for the thermodynamic properties temperature and pressure as well as the network for the HRR have been trained for 300 epochs. Both species networks, on the other hand, for 400 epochs. The training has been executed on four Nvidia Tesla K20Xm with a training time of approximately one to two hours each. Since the models are initialized randomly and there is the risk of falling into a local minimum, the models have been trained multiple times and the one with the best results is used.

In order to optimize the training, the data set has been separated into training and validation set with the validation set being 20% of the whole samples. A further test data set is created separately and introduced when considering the interpolation ability of the networks. The samples for the validation set are not taken randomly, but instead, all samples for certain equivalence ratios and temperatures are filtered. Therefore, the ability to generalize between the initial conditions can be controlled instead of determining whether the progress between two training samples is accurate. As a consequence, over- and underfitting can be

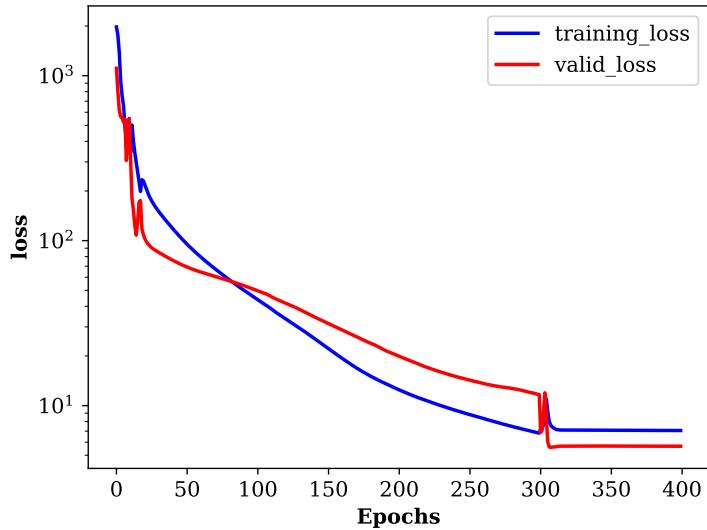


Figure 4.1.: Training and validation loss curves for the MLP with species products as outputs

monitored during training, allowing the parameters to be adjusted if necessary. Since the training data includes features with different orders of magnitude (H : approx. 10^5 , Z : 10^{-1}) the samples are normalized using Eq. 3.24. The risk of overfitting is partly caused by noise in the training data. However, in the case of the adiabatic, closed HR model, the reaction is assumed to be ideal and thus to produce training data without noise. If dropout as a method to prevent the network from learning the high-frequency noise, is applied to training data without noise, the network's loss is not decreasing monotonically. Instead, there are large fluctuations that make efficient learning not possible. Therefore, in this thesis regularization methods such as dropout or normalization layers are not used. Learning rate adjustment is performed when the validation loss is not decreasing in two successive epochs. This strict regulation regarding the number of epochs is due to the large fluctuations that occur when the learning rate is adjusted.

The loss curves of the MLP with the product species (CO , CO_2 , H_2O) as output are presented in Fig. 4.1 (the loss curves of the other networks are shown in Fig. A.4). The presented adjustment of the learning rate is visible due to the shape of the validation loss curve. The loss starts to increase at several points but due to the decrease of η by a factor of 10, a further reduction can be accomplished. Therefore, training can be stopped at the point where η has been reduced to such an extent that the change in the validation loss becomes minimal. In the case presented in Fig. 4.1, the initial η of 10^{-3} has been reduced three times to 10^{-6} until the validation loss stopped to decrease and the learning could be terminated.

The training data is learned by the NN with a certain absolute and relative accuracy. The tolerance ranges of the accuracies show differences for small and large values. As shown in Fig. 4.2, the relative accuracy defines a tolerance range that is strict for small values and widens to larger values. The absolute accuracy, on the other hand, is reversed in its progression, and therefore more strict for large values. In the case of the selected absolute and relative tolerance values, both accuracies have similar tolerance margins for Y_{H_2O} equals approx. 0.02. In order to evaluate the entire training performance, both accuracies have to be taken into consideration since predictions that fulfill both tolerance margins are considered as correctly reproduced.

For the output quantities temperature and Y_{CO} the results of the HR, of the MLP model as well as the absolute and relative deviation are shown in Fig. 4.3 (comparisons of the other output quantities are given

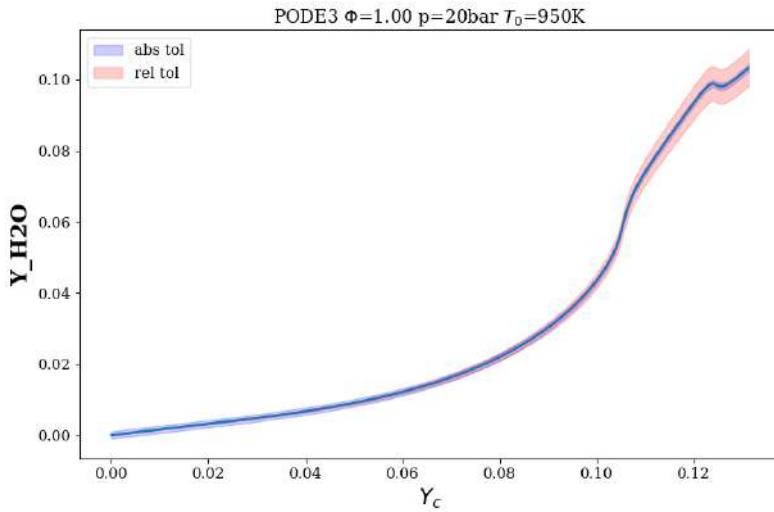


Figure 4.2.: Comparison of the absolute and relative tolerance margins in case of an Y_{H_2O} progress. The absolute tolerance is defined as $\pm 1\%$, the relative tolerance as $\pm 5\%$

in Fig. A.5 and Fig. A.6). Thereby, the training data used has a Φ of 1.0 and an initial pressure of 40bar. The difference in the strictness of relative and absolute accuracy indicated above is confirmed since large relative deviations are mainly found in the range of small values of the quantity and large absolute deviations in the range of large values. In general, the majority of the samples are learned with a deviation in the low single-digit percentage range which indicates successful learning. However, in the case of the species mass fractions, the relative deviation can be much larger. This is caused by the magnitude of resolution generated by the HR model which can be less than 10^{-5} . The MLP models are not able to predict such small values precisely, leading to small absolute but large relative deviations.

It can be concluded, that the marginal areas are the weak points of the NN. Thus, learning is difficult when the gradients of the underlying function are either steep or if the values of the output quantities are small. In the case of steep gradients, the adaptive time-step of the discretization is one method to improve learning ability by adding more samples and focus the learning on these particular parts of the reaction progress. Furthermore, the gradient resolution has been increased due to the projection in Y_c space. It can be assumed that these methods improved the learning since the values of the absolute deviations are small compared to the ones for the relative deviations. With the aim to further improve the training, Y_c can be adjusted to a more linear form. The large relative deviations are partly caused by the MSE loss function which focuses on the absolute deviations. Thus, an adjustment of the function in order to increase accuracy in marginal areas might be necessary. As an example, the relative MSE loss gives more attention to the range of small values with small absolute but large relative inaccuracies.

In summary, the MLP models were trained with the presented hyperparameters up to their maximum generalization ability, indicated by the minimum of the validation loss. This implies that the training data has not been learned to its maximum possible accuracy. However, most training samples are captured within a relative and absolute error of under $\pm 1\%$. Deviations are still too large for the relative error in the range of small values with flat gradients and for the absolute error in the range of larger values with steep gradients. Especially adjustments of the loss function are crucial in order to improve the relative accuracy of the MLP models.

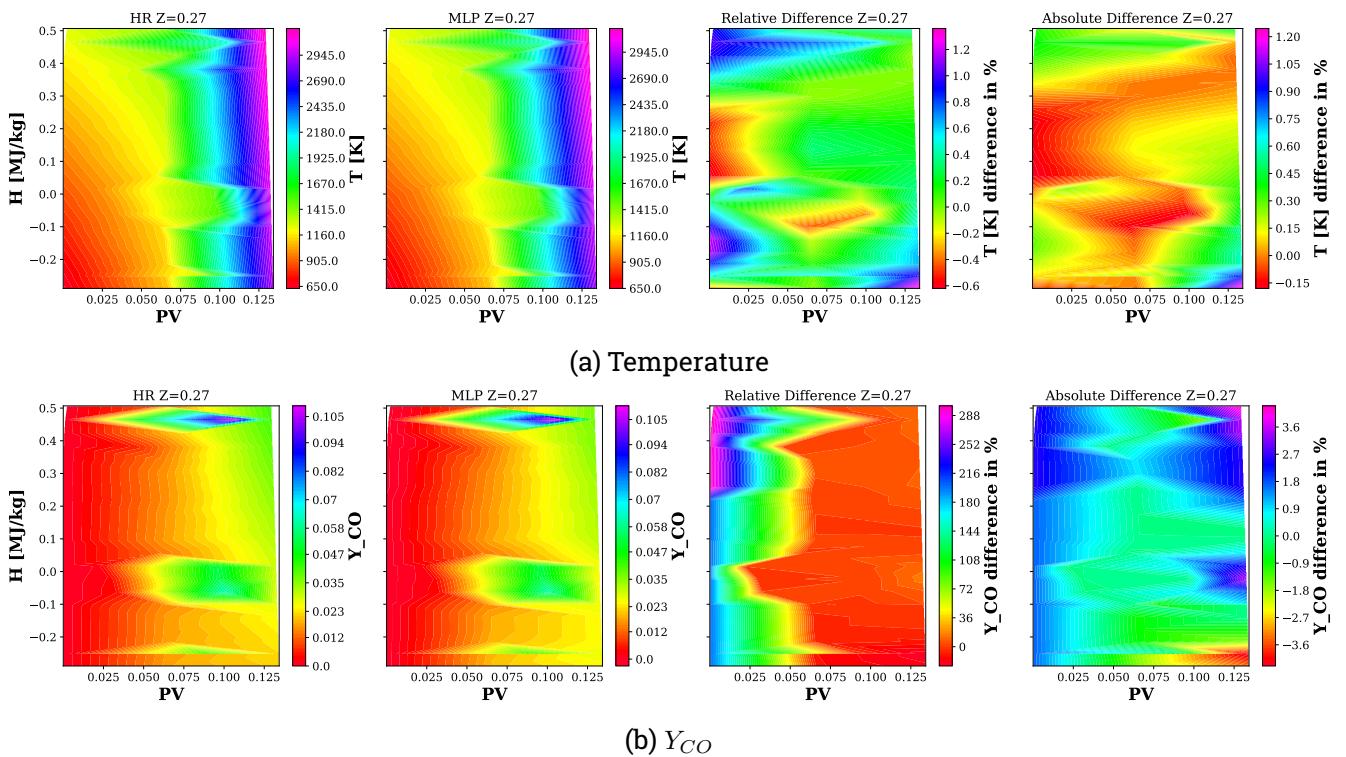


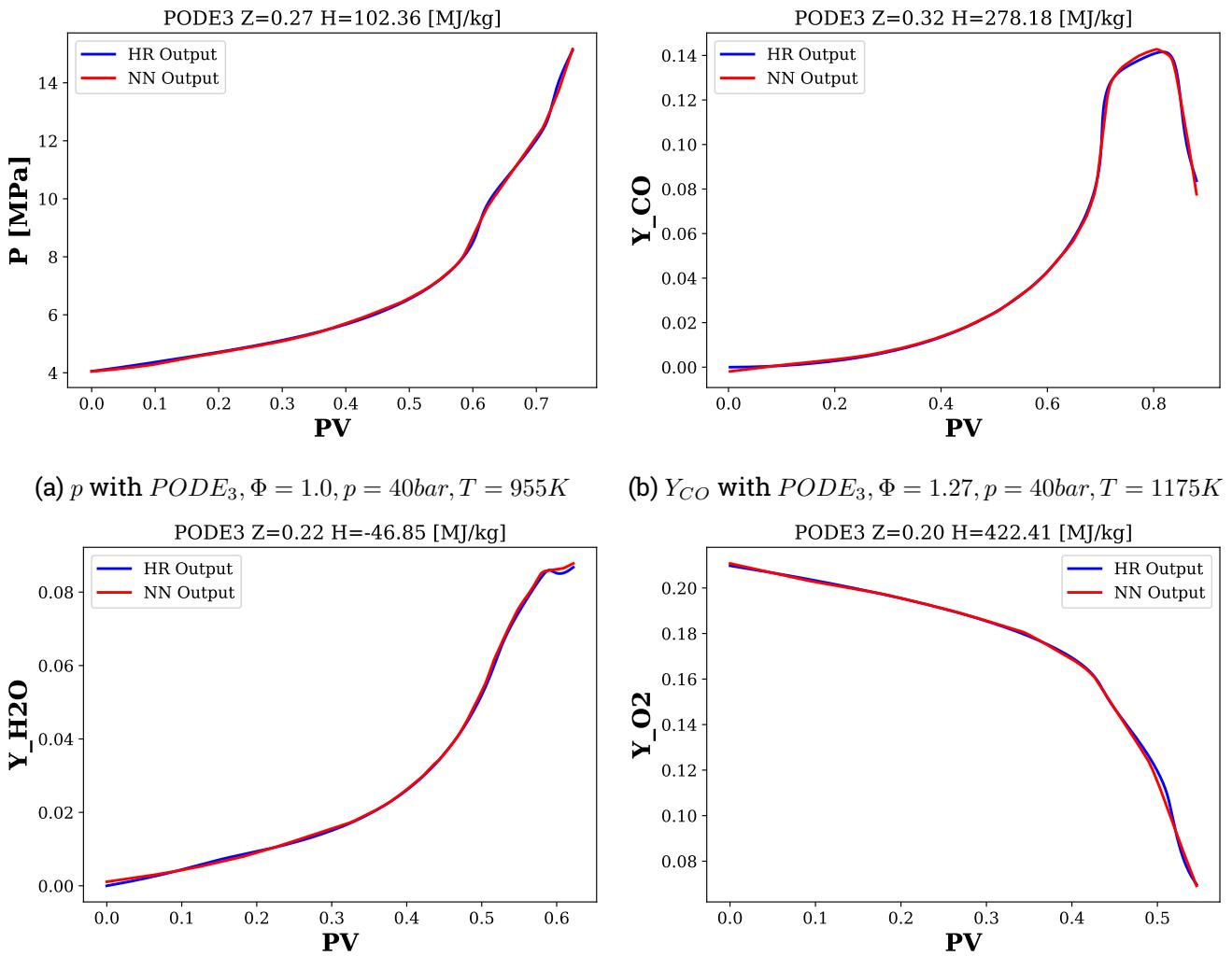
Figure 4.3.: Output of the HR model, corresponding output of the MLP as well as relative and absolute are shown. Comparison is performed for $\phi = 1.0$, a temperature range of 650 to 1250K and a pressure of 40bar.

Interpolation

Generalization for thermochemical states that are not captured in the training data, is the key task of the MLP in order to be applied in the 3D CFD. With the aim to control the abilities of the network to predict unseen data, a test set is created (Tab. A.1 shows the parameters of the set). This set includes initial thermochemical states where the two variable parameters, namely temperature and equivalence ratio, are defined with values that do not appear in the training data. Thereby, either one of the parameters is changed or both. The performance of the different networks on the test set is evaluated by a comparison of the curves for certain initial conditions and by an accuracy calculation.

In Fig. 4.4 selected output quantity curves are compared with the corresponding labels generated by the HR model (the comparison for the other output quantities is shown in Fig. A.7). In all cases, the non-linearities included in the evolution of the properties are captured well. Since temperature and pressure curves rise over the entire combustion process and thus are less steep than other output quantities, their progress is captured more accurately compared to the species mass fractions. The small area where Y_{H_2O} does not increase monotonically and thus has caused the addition of Y_{CO_2} to Y_c is not accurately mapped by the MLP. However, as this is a peculiarity of the course without any significant impact on the combustion process, this should not be of further concern. Considering the educt species, the corresponding MLP is able to capture the results sufficiently.

This work defines an MLP output as correct if the output is within a certain absolute and relative tolerance range. Both tolerances are differently affected by small and large values. Consequently, if a sample is in both ranges, a correct reproduction by the MLP is ensured. In this thesis, the relative tolerance is $\pm 3\%$ and the absolute tolerance $\pm 1\%$. The resulting mean accuracies for the output quantities in the test data set



(c) Y_{H_2O} with $PODE_3$, $\Phi = 0.77, p = 40bar, T = 740K$ (d) Y_{O_2} with $PODE_3$, $\Phi = 0.67, p = 40bar, T = 1065K$

Figure 4.4.: Comparison between the labels generated by the HR model and the output of the MLPs for data out of the Test data set.

Tab. 4.1.: The accuracy of the output quantities is given. A sample is considered correct if it is in the absolute tolerance (atol) margin with $\pm 1\%$ and/or the relative tolerance (rtol) margin with $\pm 3\%$

Output	atol	rtol	both	Output	atol	rtol	both
P	92.27%	98.25%	92.37%	Y_{H_2O}	79.46%	71.52%	58.83%
T	93.36%	99.99%	93.36%	Y_{PODE_3}	96.75%	52.75%	50.11%
Y_{CO}	72.62%	66.06%	49.42%	Y_{O_2}	87.12%	91.23%	85.79%
Y_{CO_2}	89.14%	22.25%	21.45%	HRR	61.48%	18.68%	13.57%

are given in Tab. 4.1. The best accuracies are achieved for the thermodynamic quantities, temperature and pressure, with the lowest combined accuracy of 92%. This performance is caused by an overall more flattened curve and thus an underlying function that can be learned more easily. While for those quantities the difference between absolute and relative tolerance is small, output quantities with a higher rate of small values show that the relative tolerance is more strict. Since the comparison between the output and label curves has shown a close approximation, these relative accuracies are lower than expected. However, the tolerance range for such small values cannot be recognized in the figures and thus the small absolute deviation leads to wrong conclusions. Therefore, the assumption from the training process is confirmed, stating that the loss function does not weight the relative deviations sufficiently. Besides, it can be noticed that the performance of the two quantities included in the educt MLP is different, more precisely the accuracies of Y_{PODE_3} compared to Y_{O_2} are much worse. This deviation is caused by a strong non-linearity close to ignition in the course of Y_{PODE_3} which is not captured within the margins. In further work, it has to be considered to split the two species if the difference is still evident after the loss function and other hyperparameters have been changed. Thereby, inaccuracies in range of ignition have a higher weight in the accuracy calculation since the time step is decreased and thus more samples are taken. However, due to more samples, the resolution in this part of the combustion is better which can be seen in the good absolute accuracies for all output quantities.

In summary, the overall discrepancy between predicted and intended output is small. However, the insight from the learning process that the marginal areas are particularly challenging and could not be learned to the same extent as areas with moderate gradients is consistent with the accuracy results. Possible changes to improve prediction capabilities in these areas are similar to the ones presented above. It is assumed that a change in the loss function is the most effective option. However, adjusting the learning rate, discretization and other methods such as weight-pruning can also be used to improve generalization.

Ignition Delay

The ignition delay as defined in subsection 2.2.2 is a crucial parameter for CI-Engines that impacts the overall combustion process. In order to determine IDs, the HRR rate is used since it has a distinct peak value at the point of ignition. Due to this shape and the importance of IDs an MLP model solely to predict the HRR is implemented. Therefore, the ID prediction is separated from the prediction of all other thermochemical properties. Capturing the HRR correctly is especially difficult since the strong maximum causes steep gradients. However, the discrepancy is not the critical criterion since only the location of the maximum value is of interest. Therefore, a structure similar to the other MLPs is chosen.

Fig. 4.5a shows the interpolation of a HRR curve for a variation in both, temperature and equivalence ratio. The shape of the low-temperature ignition is not captured accurately. However, the deviation between the maximum of the HR model and the MLP is small which indicates a good high-temperature ID prediction. Regarding the input variables of the MLP model, the ID for a certain initial thermochemical state is only depending on Y_c . Since there is no functional connection between Y_c and time, as well as the possibility that different times are projected on the same Y_c value, no IDTs can be generated using the MLP. The reason for projecting multiple time values on one Y_c value is that the time until the reaction significantly starts and the speed of the reaction are different depending on the initial conditions and hence Y_c levels can be reached at a different time. Therefore, a comparison between the IDs of the HR model and the NN in order to evaluate the capability to predict IDs can only be performed in the Y_c space. Such an ID curve for a temperature range from 665 to 1235K is given in Fig. 4.5b. It is shown that the curves are close and only small deviations are determined between the HR and the MLP model. Consequently, the MLP with HRR as output is able to predict the IDs with sufficient exactness and allows evaluation of the combustion process for certain tasks.

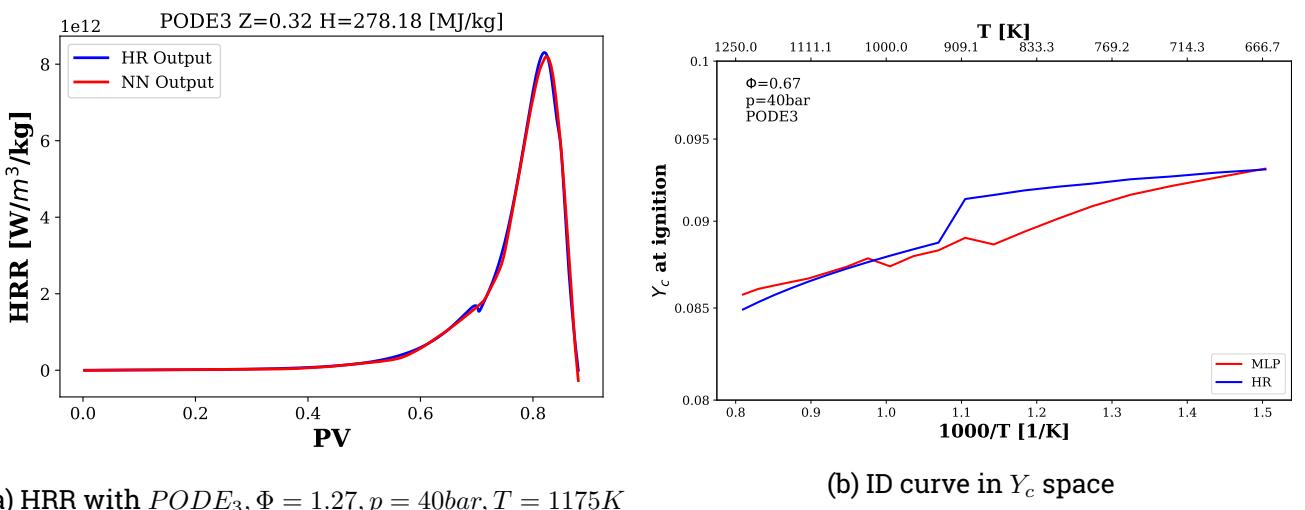


Figure 4.5.: Left: Interpolation of the HRR for a single thermochemical state. Right: Y_c values at the main ignition of the HR model and the MLP

4.2. Comparison with a Global Reaction Mechanism

Next to the tabulated chemistry approach, this work has presented the Global Reaction Mechanism (see subsection 2.3.2) as a method to reduce the computational effort of modeling detailed chemistry in a 3D CFD simulation. Instead of pre-computing the entire thermochemical state-space, the GRM reduces the mechanism by summarizing reaction classes. A GRM for $PODE_{3,4}$ has been developed by Haspel [27] following the example of Vandersickel et al. [71]. This mechanism includes seven reactions with eight species. In the first part, this section compares the general aspects of the derived method and the GRM. In the second part, a predicted ID and species curve are compared with regard to the ideal curve generated by the HR.

Using the GRM in a 3D CFD simulation does not remove the stiff ODE system, but instead, it is significantly reduced in size. Accordingly, a conservation equation that includes a chemical source term has to be solved for each of the eight species. The MLP model, on the other hand, uses conservation equations only for the look-up variables defined in section 2.4 whereas a chemical source term is exclusively included in the equation of Y_c . The computation effort in order to calculate the output of the MLPs is small since it only includes addition and multiplication operations. Therefore, it can be assumed that the method derived in this work, is faster in computation. In the case of the parameter fitting, the GRM has multiple objective functions, more precisely the four objectives presented in Eq. 2.10 - 2.13. In principle, this results in a set of optimal solutions, instead of a single optimal solution. In order to generate a solution that is suitable for the combustion of $PODE_{3,4}$ Haspel [27] uses a multiobjective genetic algorithm. Genetic algorithms are random-based classical evolutionary algorithms. They apply random changes to current solutions in order to generate new ones [16]. With each solution, a HR model similar to the one implemented in this work is executed with multiple initial thermodynamic states and the performance of the different GRM solutions are compared. The one which fulfills certain criteria regarding the accuracy of IDTs, species and thermodynamic properties is applied in the CFD simulation. Using this working principle, the process of finding a suitable solution takes about one to two days on 20 CPU cores. In contrast, the training procedure of the MLP models takes approx. 6 hours on 4 GPUs. Hence, the training procedure of the MLP models is faster which makes them more adaptable for different tasks.

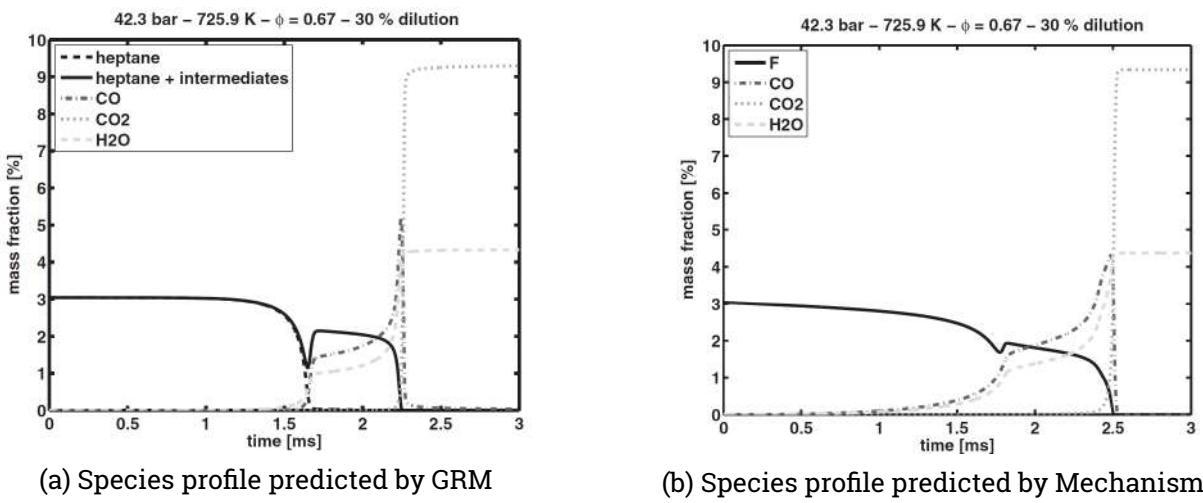


Figure 4.6.: Vandersickel et al. [71] performed a comparison of the main species profiles between the GRM derived and the detailed mechanism of Tsurushima [69] for n-heptane

Progress Variable Space

A comparison between any curve of the MLP model and the GRM has to be performed in Y_c space. The reason for it is that regarding the input variables of the MLP, the properties are only dependent on Y_c . Since there is no functional connection and no unique mapping between the physical time and Y_c , it is not possible to draw a conclusion from the Y_c space to the time-space. Furthermore, the usage of the Y_c of the derived model is not possible due to the non-appearance of CH_2O in the GRM. Therefore, a new Y_c has to be defined in order to compare the methods. Thereby, not all of the species included in the GRM are usable since some contain species groups. In Fig. 4.6 the predicted progress of the species included in the GRM of Vandersickel et al. [71] are compared with a detailed mechanism for the used fuel. Thereby, it is evident that the assigned fuel in the GRM furthermore contains high-temperature intermediates and alkyl radicals since instead of being used at approx. 1.7ms, its curve continues until the ignition takes place. Due to the requirement of linearity, the combination of Y_{CO_2} and Y_{H_2O} has been chosen as Y_c for the comparison. The process of the new Y_c is shown in Fig. A.8. It is calculated as follows:

$$Y_c = 0.5 \cdot Y_{H_2O} + 0.25 \cdot Y_{CO_2}. \quad (4.1)$$

According to the change in the input variable space, the MLP models defined in section 3.3 are retrained. Thereby, the architecture is equal to the one presented in Tab. 3.1. However, 400 training epochs are used for all networks. As a consequence of the new definition of Y_c , the models are not able to capture the start of the reaction since it only contains reaction products. Furthermore, the gradient at the high-temperature ignition is steeper, leading to a reduced resolution of the gradient. In Fig. A.10 the absolute and relative deviations of the training samples are shown. The impact of the steeper gradient is visible in larger areas of higher inaccuracies. Thus, as concluded in the previous section, the steeper gradient negatively influences the learning capability of the MLP model.

The GRM model, on the other hand, is more affected in its significance. This is due to the fact that the species progress of H_2O and CO_2 are also changed by the grouping of the other species, as shown in Fig. 4.6. An explanation for the connection is given in the first reaction of the GRM that represents the global reaction of the combustion process, as shown in the Tab. 2.3. Therefore, the mass fraction of the fuel directly

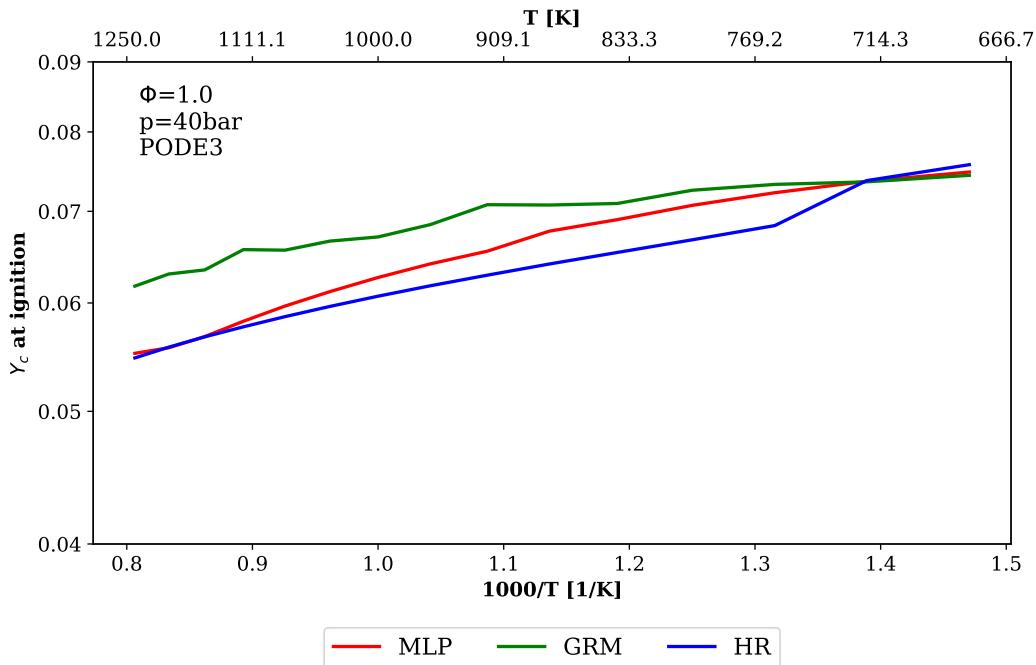


Figure 4.7.: Comparison of the ignition delay (ID) in Y_c space for an initial temperature range from 680 to 1240K at a constant initial pressure of 40bar. Ideal ID is taken from the HR model and then predicted by the MLP and the GRM.

influences the species used for Y_c . If the grouping results in too much or too little fuel being present in parts of the combustion, the production of CO_2 and H_2O is inaccurate.

Comparison of Ignition Delays

As discussed in subsection 2.2.2 ignition delays are an import property of CI-engines. Therefore, it is crucial for the derived models that a prediction of the ID is possible. Similar to the method used in the HR model, the distinct peak of the predicted HRR is used to identify the ID. The resulting ID curve in Y_c space for a $\Phi = 1.0$ and $p = 40bar$ is shown in Fig. 4.7 (the curves for different Φ 's is given in Fig. A.9). The deviation between the IDs generated by the MLP and GRM from the ideal IDs of the HR model is small. In the case of $\Phi = 1.0$, the MLP shows a slightly better approximation to the HR curve while in the case of the other comparisons, the GRM seems to reproduce the results slightly closer to the intended values. In summary, both models capture the IDs. The relatively small differences are almost negligible since the Y_c space causes disadvantages in accuracy for both models.

Comparison of Thermochemical Properties

In the case of the species progresses, MLP and GRM reveal large deviations. Fig. 4.8 shows the progression of selected species and the thermodynamic variable temperature (the remaining properties are shown in Fig. A.11). Thereby, each evolution curve is interpolated by the MLP and not included in its training data. By grouping the species as in the case of the fuel shown in Fig. 4.8a, their course is decoupled from the temporal evolution observed in the HR. The observed non-monotonic part in the fuel evolution is consistent with the temporal evolution presented in Fig. 4.8 and caused by the addition of the high-temperature intermediates to the fuel mass fraction. However, Fig. A.11d shows that for higher initial temperatures the non-monotonic part vanishes. The simplified reactions lead to missing formation reactions of CO and to premature formation of CO_2 . The course of the thermodynamic variables temperature and pressure, on the

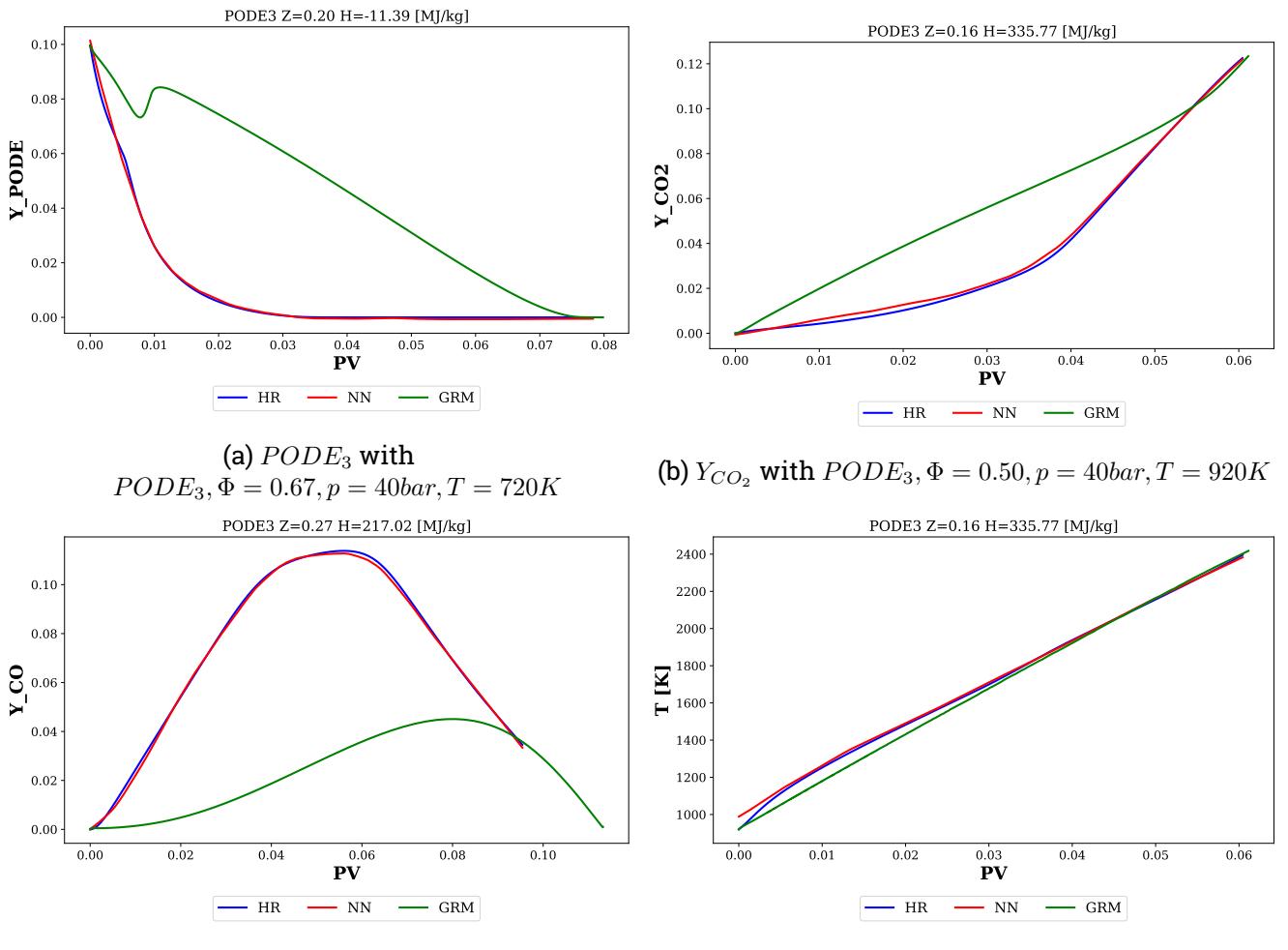


Figure 4.8.: Comparison between the species progress in Y_c between the predicted results of the derived MLP model, the GRM model by Haspel [27] and the HR output

other hand, is similar between the GRM and HR models. The difference in the species progress seen in the time space Fig. 4.6 are therefore even more evident in Y_c space. Thereby, a further difference between HR and GRM is evident since the maximum of Y_c under stoichiometric conditions is different. This is caused by the HR since the educts do not fully react to H_2O and CO_2 (as seen in Fig. A.3). In the GRM, on the other side, the fuel directly reacts to H_2O and CO_2 causing a complete use if enough oxygen is present. In summary, the comparison shows that the evolution of the species in the GRM does not correspond to the evolution generated by the HR, whereas the MLP model is able to reproduce them precisely. Hence, the structure of the GRM causes that the approach can only make qualitative statements for the species mass fractions.

Summary

The basic idea of the GRM and the MLP model is different. While the MLP model is trained to reproduce all thermochemical properties that are important for the evaluation with the precision of the detailed mechanism, the GRM is limited to IDTs and to the thermodynamic parameters temperature and pressure. Therefore, in the case of the species evolutions, large deviations between GRM and HR are observed. The higher level of detail of the derived method is caused by the possibility to train different models for different

tasks and thus avoid tradeoffs in optimization. As a consequence, the adaptability of the MLP model is an advantage compared to the GRM, because new output quantities can be added with high precision and without requiring chemical knowledge.

However, the level of detail and the adaptability are only possible because the output quantities are decoupled. In contrast, the thermochemical state of the GRM is always consistent since the parameters are mutually dependent. Thus, chemical implausible states can occur in the MLP model. Furthermore, the GRM model is to some extent applicable beyond the boundaries in which it has been validated. Thus, if unforeseen states occur in the simulation, the results can still be useful. For the MLP model, on the other hand, it is unknown how the prediction will behave if states outside the range of validity occur. Accordingly, those states are scaled back to the corresponding boundary state of the training data set.

5. Conclusion and Outlook

In this thesis, a model has been derived that employs a combination of MLPs in order to model the detailed chemistry in a 3D CFD simulation. Therefore, the thermochemical states included in the training data are pre-computed in a zero-dimensional homogeneous reactor model by using the detailed reaction mechanism of Cai et al. [13]. The results are discretized with an adaptive time step. The different thermochemical properties which should be included in the simulation are grouped according to their underlying function and for each group, a separate MLP is constructed and trained, an overview is given in Tab. 3.1. Each MLP is adjusted individually and equipped with look-up variables as input in order to access the thermochemical state in CFD and HR.

MLPs are chosen due to their capability to learn the evolution of the properties in the thermochemical state-space they have been trained with and interpolate between the discretized samples. Hence, the development included in the CFD has the precision of the detailed mechanism while the employment of MLPs avoids large memory usage and computational expensive calculation. Furthermore, the MLPs used are highly adaptable. New species or thermodynamic properties can easily be defined as the output of the corresponding network, without the need for knowledge about reactions and without using large amounts of memory. In order to extend the thermochemical state space by adding pressure as a variable quantity, the input variable space is extended by one dimension and the architecture of the networks is slightly modified. However, it can be assumed that computational effort and memory usage do not change significantly. Besides, the MLP framework is supposed to be applicable for other fuels if the networks are retrained with the corresponding thermochemical states. However, adjustments of the architecture depending on the complexity of the combustion evolution and of the look-up variables can be necessary. Consequently, the model derived has high adaptability, while it neither requires large computational effort nor large amounts of memory. Especially in contrast to the tabulated chemistry approach or the GRM, this applicability makes the technique interesting for further research in order to provide an alternative to traditional methods.

By evaluating the interpolation capabilities of the MLP models, it is evident that the non-linearities in the curves of the thermochemical properties are captured. Most samples have been fitted within an absolute deviation of $\pm 1\%$ and a relative deviation of $\pm 3\%$. However, in the high-temperature ignition with steep gradients and in range of small quantity values, a higher percentage of samples exceeded these tolerance margins. Since in a CFD simulation these differences can have an influence on the result, future research has to address this problem. Investigations regarding a progress variable that has higher gradient resolution or an adaptive time-stepping that automatically searches the best time-step are promising fields to increase accuracy. However, changing the loss function might be the most effective method since the focus of the learning can be shifted towards the marginal areas. Besides, a two-step training, firstly in order to reduce the absolute deviation and secondly in order to reduce the relative definition, each with different loss functions, is an option and has to be taken into account.

The MLP structure presented in this thesis enabled the learning of the thermochemical properties. The defined architecture is chosen to be the minimum size that is still able to capture the non-linearities with high accuracy. However, further optimization is necessary. By applying different activation functions, new

regularisation methods, or by adopting the architecture, presented inaccuracies can be reduced. Besides, also slight changes such as in the learning rate adjustment can be addressed for improvement. Furthermore, network-pruning as a weight-decay technique can be applied in order to increase generalization capability. In summary, the derived architecture should provide the basis which shows the possibilities of using ML techniques and which can now be optimized by further research.

While this work is only a comparison of the reproductive capabilities of the developed model, further research is needed to test the model in a 3D CFD simulation. This allows making realistic statements about the true requirements for the computing power, the resulting inaccuracies and the general performance. In any case, it can be concluded that neural networks open up new possibilities in this area of chemical simulation. Due to the great adaptability, the efficient computing power, the low requirements of memory as well as chemical understanding of the process, the presented model shows advantages which cannot be provided by any other known method. It remains open to what extent possible inaccuracies can still be reduced and how the method performs in the simulation of other fuels but we can look to the future with confidence.

Bibliography

- [1] M.S. Abd-Elhady et al. “Influence of gas velocity on particulate fouling of exhaust gas recirculation coolers”. In: *International Journal of Heat and Mass Transfer* 54.4 (Jan. 2011), pp. 838–846.
- [2] Ethem Alpaydin. *Introduction to machine learning*. Cambridge, Mass: MIT Press, 2010. ISBN: 978-0262012430.
- [3] *Automotive fuels - Diesel - Requirements and test methods; German version EN 590:2013+A1:2017*. Norm. Oct. 2017.
- [4] A.R. Barron. “Universal approximation bounds for superpositions of a sigmoidal function”. In: *IEEE Transactions on Information Theory* 39.3 (May 1993), pp. 930–945.
- [5] C. Bekdemir et al. “Predicting diesel combustion characteristics with Large-Eddy Simulations including tabulated chemical kinetics”. In: *Proceedings of the Combustion Institute* 34.2 (Jan. 2013), pp. 3067–3074.
- [6] Cemil Bekdemir, Bart Somers, and Philip de Goey. “DNS with detailed and tabulated chemistry of engine relevant igniting systems”. In: *Combustion and Flame* 161.1 (Jan. 2014), pp. 210–221.
- [7] Sushrut Bhalla et al. “Compact Representation of a Multi-dimensional Combustion Manifold Using Deep Neural Networks”. In: Sept. 2019.
- [8] R. W. Bilger. “The Structure of Diffusion Flames”. In: *Combustion Science and Technology* 13.1-6 (July 1976), pp. 155–170.
- [9] Klaus B. Binder. “Diesel Engine Combustion”. In: *Handbook of Diesel Engines*. Springer Berlin Heidelberg, 2010, pp. 61–75.
- [10] Christopher Bishop. *Pattern recognition and machine learning*. New York: Springer, 2006. ISBN: 978-0-387-31073-2.
- [11] K BRAY, P DOMINGO, and L VERVISCH. “Role of the progress variable in models for partially premixed turbulent combustion”. In: *Combustion and Flame* 141.4 (June 2005), pp. 431–437.
- [12] Jakob Burger et al. “Poly(oxymethylene) dimethyl ethers as components of tailored diesel fuel: Properties, synthesis and purification concepts”. In: *Fuel* 89.11 (Nov. 2010), pp. 3315–3319.
- [13] Liming Cai et al. “Automatic mechanism generation for auto-ignition of the promising e-fuels oxymethylene ethers (OMEn, n = 2-4)”. In: Jan. 2019.
- [14] Intergovernmental Panel on Climate Change. *Climate Change 2014: Mitigation of Climate Change* -. Cambridge: Cambridge University Press, 2015. ISBN: 978-1-107-05821-7.
- [15] Jinhwa Chung et al. “Rapid-compression machine studies on two-stage ignition characteristics of hydrocarbon autoignition and an investigation of new gasoline surrogates”. In: *Energy* 93 (Dec. 2015), pp. 1505–1514.
- [16] K. Deb et al. “A fast and elitist multiobjective genetic algorithm: NSGA-II”. In: *IEEE Transactions on Evolutionary Computation* 6.2 (2002), pp. 182–197.

- [17] Ke-Lin Du and M. N. S. Swamy. *Neural Networks and Statistical Learning*. Springer London, 2019.
- [18] Richard Duda. *Pattern classification*. New York: Wiley, 2001. ISBN: 978-0-471-05669-0.
- [19] J.H. Espenson. *Chemical Kinetics and Reaction Mechanisms*. Advanced Chemistry Series. McGraw-Hill, 1995. ISBN: 9780070202603.
- [20] Ian Flood and Nabil Kartam. “Neural Networks in Civil Engineering. I: Principles and Understanding”. In: *Journal of Computing in Civil Engineering* 8.2 (Apr. 1994), pp. 131–148.
- [21] Pierre Friedlingstein et al. “Global Carbon Budget 2019”. In: *Earth System Science Data* 11.4 (Dec. 2019), pp. 1783–1838.
- [22] Olivier Gicquel, Nasser Darabiha, and Dominique Thévenin. “Liminar premixed hydrogen/air counter-flow flame simulations using flame prolongation of ILDM with differential diffusion”. In: *Proceedings of the Combustion Institute* 28.2 (Jan. 2000), pp. 1901–1908.
- [23] Xavier Glorot, Antoine Bordes, and Yoshua Bengio. “Deep Sparse Rectifier Neural Networks”. In: *Proceedings of the Fourteenth International Conference on Artificial Intelligence and Statistics*. Ed. by Geoffrey Gordon, David Dunson, and Miroslav Dudík. Vol. 15. Proceedings of Machine Learning Research. Fort Lauderdale, FL, USA: PMLR, Nov. 2011, pp. 315–323.
- [24] David G. Goodwin et al. *Cantera: An Object-oriented Software Toolkit for Chemical Kinetics, Thermodynamics, and Transport Processes*. <https://www.cantera.org>. Version 2.4.0. 2018.
- [25] Philipp Haltenort et al. “Heterogeneously catalyzed synthesis of oxymethylene dimethyl ethers (OME) from dimethyl ether and trioxane”. In: *Catalysis Communications* 109 (May 2018), pp. 80–84.
- [26] Martin Härtl et al. “Oxygenate screening on a heavy-duty diesel engine and emission characteristics of highly oxygenated oxymethylene ether fuelOME1”. In: *Fuel* 153 (Aug. 2015), pp. 328–335.
- [27] Philip Haspel. haspel@stfs.tu-darmstadt.de; Research assistant at STFS, TU-Darmstadt.
- [28] Trevor Hastie, Robert Tibshirani, and Jerome Friedman. *The Elements of Statistical Learning*. Springer New York, 2009.
- [29] Tanjin He et al. “A chemical kinetic mechanism for the low- and intermediate-temperature combustion of Polyoxymethylene Dimethyl Ether 3 (PODE3)”. In: *Fuel* 212 (Jan. 2018), pp. 223–235.
- [30] Geoffrey E. Hinton et al. “Improving neural networks by preventing co-adaptation of feature detectors”. In: *ArXiv* abs/1207.0580 (2012).
- [31] Laura E. Hombach et al. “Economic and environmental assessment of current (2015) and future (2030) use of E-fuels in light-duty vehicles in Germany”. In: *Journal of Cleaner Production* 207 (Jan. 2019), pp. 153–162.
- [32] Kurt Hornik, Maxwell Stinchcombe, and Halbert White. “Multilayer feedforward networks are universal approximators”. In: *Neural Networks* 2.5 (Jan. 1989), pp. 359–366.
- [33] Mingdi Huang et al. “Diesel engine CFD simulations: Influence of fuel variability on ignition delay”. In: *Fuel* 181 (Oct. 2016), pp. 170–177.
- [34] Serdar Iplikci et al. “A Novel Modification on the Levenberg-Marquardt Algorithm for Avoiding Overfitting in Neural Network Training”. In: *Lecture Notes in Computer Science*. Springer International Publishing, 2019, pp. 201–207.
- [35] S. Jacobs and K.A. Heufer. “An experimental and theoretical comparison of the novel e-fuel class oxymethylene ether”. In: 2019.
- [36] P. JANSEN et al. “Model structure selection for multivariable systems by cross-validation methods”. In: *International Journal of Control* 47.6 (June 1988), pp. 1737–1758.

- [37] Sotiris Kotsiantis, Dimitris Kanellopoulos, and P. Pintelas. “Data Preprocessing for Supervised Learning”. In: *International Journal of Computer Science* 1 (Jan. 2006), pp. 111–117.
- [38] P. A. Lakshminarayanan and Yogesh V. Aghav. “Ignition Delay in a Diesel Engine”. In: *Mechanical Engineering Series*. Springer Netherlands, Dec. 2009, pp. 59–78.
- [39] Corentin J. Lapeyre et al. “Training convolutional neural networks to estimate turbulent sub-grid scale reaction rates”. In: *Combustion and Flame* 203 (May 2019), pp. 255–264.
- [40] Chung K. Law. *Combustion Physics*. Cambridge University Press, 2006.
- [41] Yann LeCun, Patrice Y. Simard, and Barak Pearlmutter. “Automatic Learning Rate Maximization by On-Line Estimation of the Hessian’s Eigenvectors”. In: *Advances in Neural Information Processing Systems* 5. Ed. by S. J. Hanson, J. D. Cowan, and C. L. Giles. Morgan-Kaufmann, 1993, pp. 156–163.
- [42] Haoye Liu et al. “Improvement of emission characteristics and thermal efficiency in diesel engines by fueling gasoline/diesel/PODEn blends”. In: *Energy* 97 (Feb. 2016), pp. 105–112.
- [43] Haoye Liu et al. “Study on combustion and emission characteristics of Polyoxymethylene Dimethyl Ethers/diesel blends in light-duty and heavy-duty diesel engines”. In: *Applied Energy* 185 (Jan. 2017), pp. 1393–1402.
- [44] Björn Lumpp et al. “OXYMETHYLENE ETHERS AS DIESEL FUEL ADDITIVES OF THE FUTURE”. In: *MTZ worldwide* 72.3 (Feb. 2011), pp. 34–38.
- [45] Delin Lv et al. “Development of a reduced diesel/PODEn mechanism for diesel engine application”. In: *Energy Conversion and Management* 199 (Nov. 2019), p. 112070.
- [46] U. Maas and S.B. Pope. “Simplifying chemical kinetics: Intrinsic low-dimensional manifolds in composition space”. In: *Combustion and Flame* 88.3-4 (Mar. 1992), pp. 239–264.
- [47] Holger R. Maier and Graeme C. Dandy. “Neural networks for the prediction and forecasting of water resources variables: a review of modelling issues and applications”. In: *Environmental Modelling & Software* 15.1 (Jan. 2000), pp. 101–124.
- [48] Tom Mitchell. *Machine Learning*. New York: McGraw-Hill, 1997. ISBN: 978-0070428072.
- [49] Vinod Nair and Geoffrey E. Hinton. “Rectified Linear Units Improve Restricted Boltzmann Machines”. In: *ICML*. 2010.
- [50] Takehiko Nakama. “Comparisons of Single- and Multiple-Hidden-Layer Neural Networks”. In: *Advances in Neural Networks – ISNN 2011*. Springer Berlin Heidelberg, 2011, pp. 270–279.
- [51] Dorian Oestreich et al. “Reaction kinetics and equilibrium parameters for the production of oxymethylene dimethyl ethers (OME) from methanol and formaldehyde”. In: *Chemical Engineering Science* 163 (May 2017), pp. 92–104.
- [52] J.A. VAN OIJEN and L.P.H. DE GOEY. “Modelling of Premixed Laminar Flames using Flamelet-Generated Manifolds”. In: *Combustion Science and Technology* 161.1 (Dec. 2000), pp. 113–137.
- [53] World Meteorological Organisation. *Statement on the State of the Global Climate in 2019*. WMO, 2019. ISBN: 978-92-62-11248-5.
- [54] Adam Paszke et al. “PyTorch: An Imperative Style, High-Performance Deep Learning Library”. In: *Advances in Neural Information Processing Systems* 32. Ed. by H. Wallach et al. Curran Associates, Inc., 2019, pp. 8024–8035.
- [55] Leonardo Pellegrini, Renata Patrini, and Mario Marchionna. “Effect of POMDME Blend on PAH Emissions and Particulate Size Distribution from an In-Use Light-Duty Diesel Engine”. In: *SAE Technical Paper Series*. SAE International, Apr. 2014.

- [56] N. V. Petrukhin, N. N. Grishin, and S. M. Sergeev. "Ignition Delay Time - an Important Fuel Property". In: *Chemistry and Technology of Fuels and Oils* 51.6 (Jan. 2016), pp. 581–584.
- [57] Thierry Poinsot and Denis Veynante. *Theoretical and Numerical Combustion* -. T. Poinsot, D. Veynante, 2011. ISBN: 978-2-746-63990-4.
- [58] T. Popp et al. "Potentials of OME/diesel blends for stationary power production – Improving emission characteristics of a diesel CHP unit". In: *Applied Thermal Engineering* 153 (May 2019), pp. 483–492.
- [59] M.F. Redondo and C.H. Espinosa. "Generalization capability of one and two hidden layers". In: *IJCNN'99. International Joint Conference on Neural Networks. Proceedings (Cat. No.99CH36339)*. IEEE, 1999.
- [60] R D Reitz et al. "IJER editorial: The future of the internal combustion engine". In: *International Journal of Engine Research* 21.1 (Sept. 2019), pp. 3–10.
- [61] Shujin Ren et al. "Development of a reduced polyoxymethylene dimethyl ethers (PODEn) mechanism for engine applications". In: *Fuel* 238 (Feb. 2019), pp. 208–224.
- [62] F. Rosenblatt. "The perceptron: A probabilistic model for information storage and organization in the brain." In: *Psychological Review* 65.6 (1958), pp. 386–408.
- [63] Frank F. Rosenblatt. "PRINCIPLES OF NEURODYNAMICS. PERCEPTRONS AND THE THEORY OF BRAIN MECHANISMS". In: *American Journal of Psychology* 76 (1963), p. 705.
- [64] David E. Rumelhart, Geoffrey E. Hinton, and Ronald J. Williams. "Learning Internal Representations by Error Propagation". In: *Parallel Distributed Processing: Explorations in the Microstructure of Cognition, Volume 1: Foundations*. Ed. by David E. Rumelhart and James L. McClelland. Cambridge, MA: MIT Press, 1986, pp. 318–362.
- [65] Arne Scholtissek. *Flamelet modeling in composition space for premixed and non-premixed combustion*. München: Verlag Dr. Hut, 2019. ISBN: 978-3843939300.
- [66] Nitish Srivastava et al. "Dropout: A Simple Way to Prevent Neural Networks from Overfitting". In: *Journal of Machine Learning Research* 15.56 (2014), pp. 1929–1958.
- [67] Wenyu Sun et al. "Speciation and the laminar burning velocities of poly(oxymethylene) dimethyl ether 3 (POMDME3) flames: An experimental and modeling study". In: *Proceedings of the Combustion Institute* 36.1 (2017), pp. 1269–1278.
- [68] J. C. Sutherland, P. J. Smith, and J. H. Chen. "Quantification of differential diffusion in nonpremixed systems". In: *Combustion Theory and Modelling* 9.2 (May 2005), pp. 365–383.
- [69] T. Tsurushima. "A new skeletal PRF kinetic model for HCCI combustion". In: *Proceedings of the Combustion Institute* 32.2 (2009), pp. 2835–2841.
- [70] S.R. Turns. *An Introduction to Combustion: Concepts and Applications*. McGraw-Hill series in mechanical engineering. McGraw-Hill, 2012. ISBN: 9780071086875.
- [71] A. Vandersickel, Y. M. Wright, and K. Boulouchos. "Global reaction mechanism for the auto-ignition of full boiling range gasoline and kerosene fuels". In: *Combustion Theory and Modelling* 17.6 (Dec. 2013), pp. 1020–1052.
- [72] Zhi Wang et al. "Homogeneous charge compression ignition (HCCI) combustion of polyoxymethylene dimethyl ethers (PODE)". In: *Fuel* 183 (Nov. 2016), pp. 206–213.
- [73] C. Xiang, S. Ding, and T.H. Lee. "Geometrical Interpretation and Architecture Selection of MLP". In: *IEEE Transactions on Neural Networks* 16.1 (Jan. 2005), pp. 84–96.
- [74] Y. Zheng et al. "Synthesis of a Green Fuel Additive Over Cation Resins". In: *Chemical Engineering & Technology* 36.11 (Oct. 2013), pp. 1951–1956.

-
- [75] Ruijun Zhu et al. “Performance and Emission Characteristics of Diesel Engines Fueled with Diesel-Dimethoxymethane (DMM) Blends”. In: *Energy & Fuels* 23.1 (Jan. 2009), pp. 286–293.
- [76] Thorsten Zirwes et al. “Automated Code Generation for Maximizing Performance of Detailed Chemistry Calculations in OpenFOAM”. In: *High Performance Computing in Science and Engineering ’17*. Springer International Publishing, 2018, pp. 189–204.

List of Figures

2.1	Schematic heat release rate curve with two-stage ignition behavior	7
2.2	Exp. IDT of He et al. compared with simulations for $PODE_3$, $\Phi=0.5$ and $p=10\text{bar}$	11
2.3	Exp. IDT of Jacobs et al. compared with simulations for $PODE_{3/4}$, $\Phi=1.0$	12
2.4	Schematic Progress Variable	20
2.5	Y_c species	20
2.6	Temperautre and HRR in time and progress varibale space	21
3.1	Supervised and unsupervised learning methods	23
3.2	Single layer perceptron	24
3.3	Three layer fully connected MLP	25
3.4	Computational Graph for BP explanation	27
3.5	Adaptive time-stepping	30
3.6	Overfitting and underfitting	31
3.7	Working principle of Dropouts	32
4.1	Loss curves of MLP with species products as output	37
4.2	Comparison of absolute and relative tolerance margins	38
4.3	Contour plots to show the learning of the training samples	39
4.4	Interpolation Examples of the Test Data Set	40
4.5	Interpolation of HRR and ignition curve for $\Phi = 0.67$, $PODE_3$ and $p = 40\text{bar}$	42
4.6	Comparison species profiles GRM and detailed mechanism by Vandersickel [71]	43
4.7	ID comparison between HR, GRM and MLP	44
4.8	Comparison of the species progress between the MLP and GRM model	45
A.1	Exp. IDT of He et al. compared with simulations for $PODE_3$	xix
A.2	Exp. IDT of Jacobs et al. compared with simulations for $PODE_{1,2}$	xx
A.3	Species mass fraction evolution in Y_c space	xxi
A.4	Loss curves for different MLP models	xxii
A.5	Further contour plots to show the learning of the training samples (Part1)	xxiii
A.6	Further contour plots to show the learning of the training samples (Part2)	xxiv
A.7	Interpolation of output quantities	xxv
A.8	Y_c comparison	xxvi
A.9	ID comparison between HR, GRM and MLP	xxvii
A.10	Contour plots to show the learning of the training samples	xxviii
A.11	Comparison of the species progress between the MLP and GRM model	xxix

List of Tables

2.1	PODE properties	5
2.2	Global Reactions of $PODE_{3-4}$	7
2.3	Reactions and reaction rates global reaction mechanism	13
2.4	Overview of homogeneous reactor models	15
3.1	Properties of MLPs used in this work	35
4.1	MLP Accuracies	40
A.1	Parameters of Test data set	xviii

Nomenclature

Acronyms

AI	Artificial Intelligence
ANN	Artificial Neural Network
BP	Backpropagation
CFD	Computational Fluid Dynamics
CI	compression ignition
CN	cetane number
DME	dimethyl ether
FA	formaldehyde
GRM	Global Reaction Mechanism
HR	Homogeneous Reactor
HRR	heat release rate
HT	high temperature
IC	internal combustion
ID	ignition delay
IDT	ignition delay time
LMS	Least-Mean-Squares algorithm
LT	low temperature
ML	Machine Learning
MLP	Mulit-Layer Perceptron
MSE	Mean Squared Error
NN	Neural Network

NTC	negative temperature coefficient
ODE	ordinary differential equation
OME	Oxymethylene Ether
PM	particulate-matter emissions
PODE	Polyoxymethylene Dimethyl Ether
RCM	Rapid Compression Machine
ST	Shock Tube
TDC	top dead center

Greek Symbols

v''_{kj}	molar stoichiometric coefficients of species i in the backward reaction j
v'_{kj}	molar stoichiometric coefficients of species i in the forward reaction j
v_{kj}	molar stoichiometric coefficients of species i in reaction j
β_j	temperature exponent of reaction j
δ_k	sensitivity of unit k
ω_c	chemical source term of the progress variable
$\dot{\omega}_i$	production rate of species i
η	learning rate
Φ	equivalence ratio
ϕ	activation function
ρ	denisty
τ_{ij}	viscous tensor
E_j	activation energy of reaction j

Latin Symbols

\mathcal{M}_k	species symbol
$\Delta h_{f,i}^\circ$	enthalpy of formation of species i
\dot{Q}	heat source term
\dot{Q}	heat transfer rate

\mathcal{R}	ideal gas constant
\mathbf{t}	labels vector
\mathbf{u}	velocity vector
\mathbf{V}_i	diffusion velocity of species i
\mathbf{w}	model parameters
\mathbf{x}	input vector
\mathbf{y}	output vector of hidden layer
\mathbf{z}	output vector
A_{fi}	preexponential factor of forward reaction j
a_{ij}	number of elements j in species i
$c_{p,i}$	specific heat capacity of species i at constant pressure
$c_{v,i}$	specific heat capacity of species i at constant volume
E	loss / error / cost function
h	absolute enthalpy
h_i	molar enthalpies of species i
$h_{s,i}$	sensible enthalpy of species i
k_{bj}	backward rate of reaction j
k_{fj}	forward rate of reaction j
m	mass
M_i	molecular weight of species i
n	degree of polymerization
n_c	length of output vector
n_e	number of elements
N_i	number of moles of species i
n_r	number of reactions
n_s	number of species
p	pressure

q	heat flux
R_j	progress rate of reaction j
s	mass stoichiometric ratio
T	Temperature
u	specific internal energy
V	Volume
w_{j0}	bias term of unit j
w_{ji}	weight of input i of unit j
X_i	mole fraction of species i
x_i	input i of unit j
Y_c	reaction progress variable
Y_i	mass fraction of species i
Z	mixture fraction

Other Symbols

$[X_i]$	molar concentrations of species i
atol	absolute tolerance
rtol	relative tolerance

A. Appendix

Tab. A.1.: Initial conditions of the homogenous reactor simulation that have been used to generate the test data set

Fuel	Φ	Pressure [bar]	Temperature [K]
3, 4	0.77	40	740
3, 4	1.00	40	955
3, 4	0.67	40	1065
3, 4	1.27	40	1175

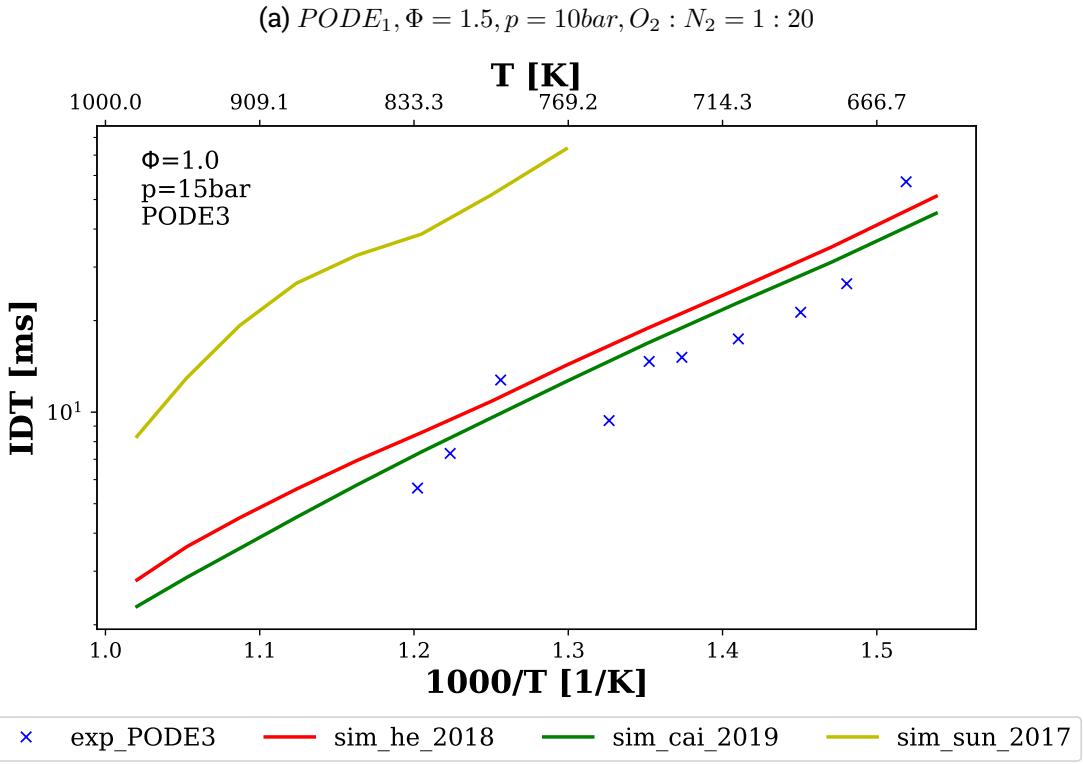
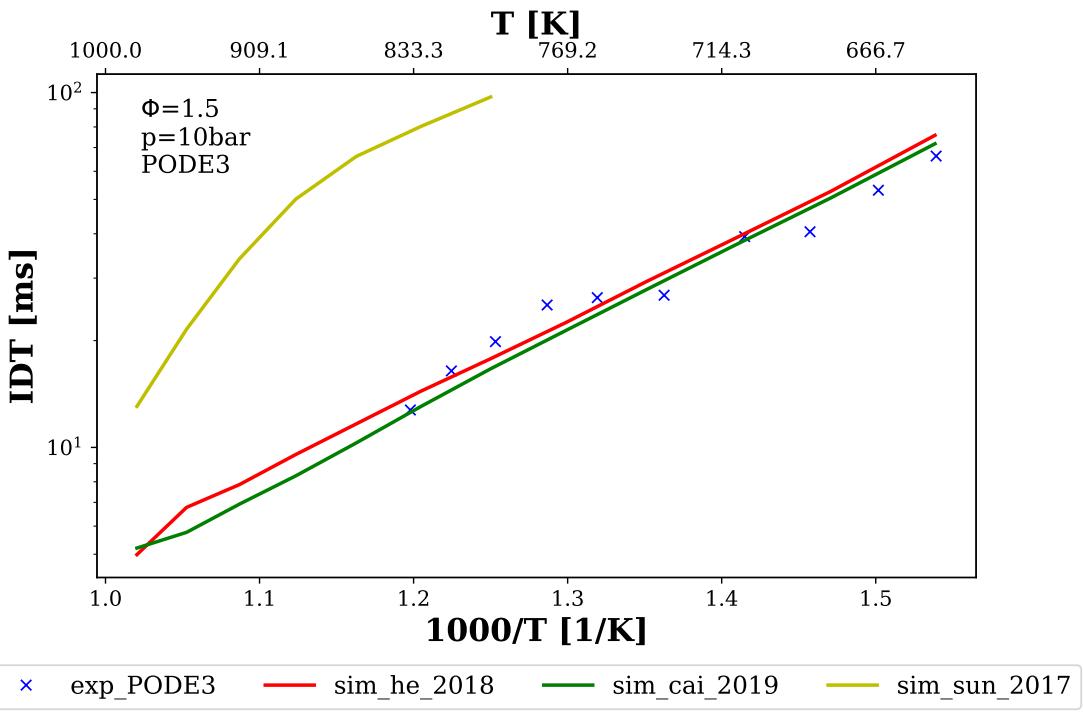


Figure A.1.: Extracted experimental IDTs from He et al. [29] for $PODE_3$. Simulated IDTs generated by HR model with mechanisms of Sun et al. [67], He et al [29] and Cai et al.[13]. Used oxydizer is diluted with the ratio:

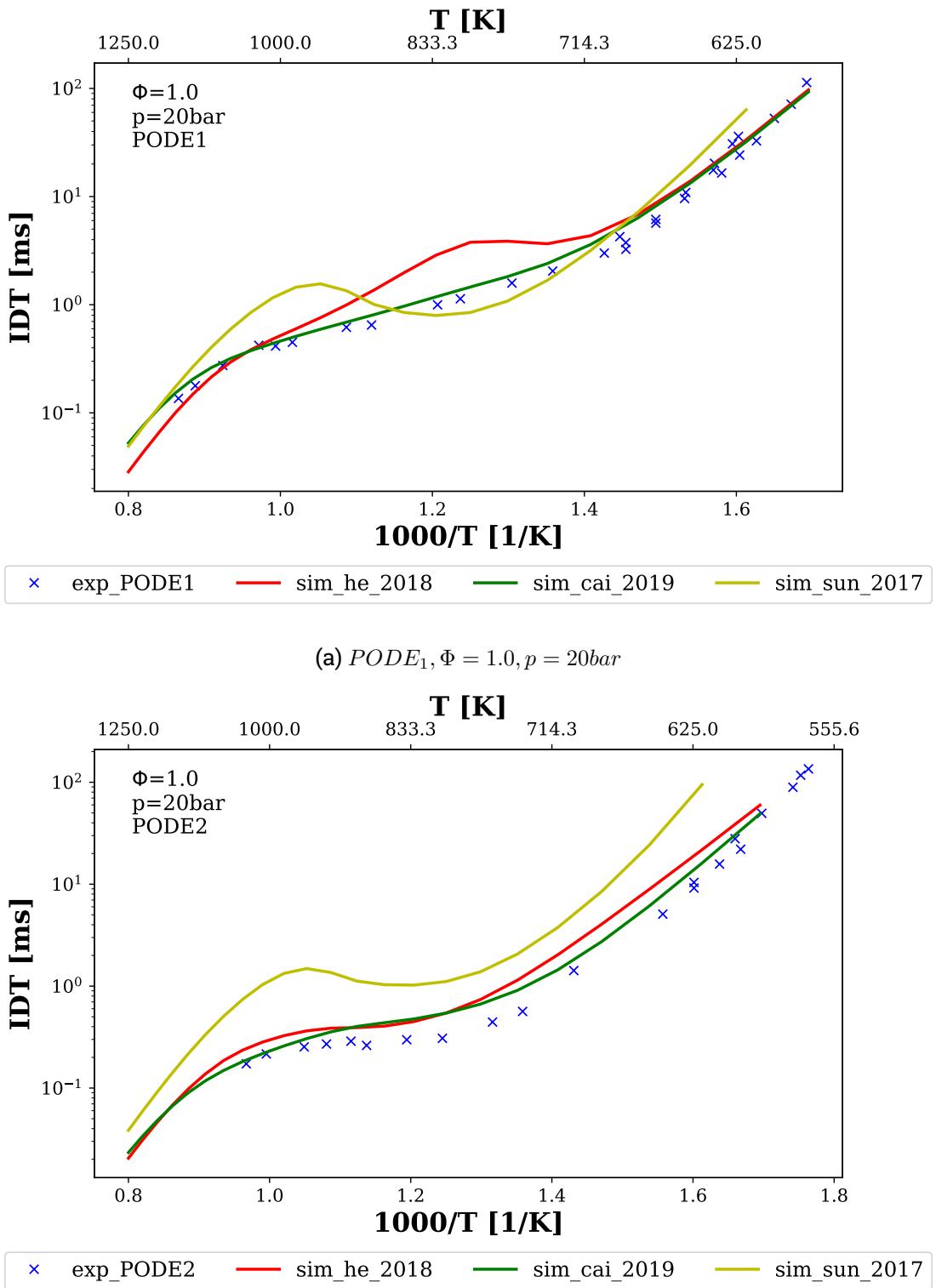


Figure A.2.: Extracted experimental IDTs from Jacobs et al. [35] compared with the delay times generated by the implemented reactor model with mechanisms of Sun et al. [67], He et. al [29] and Cai et. al [13]

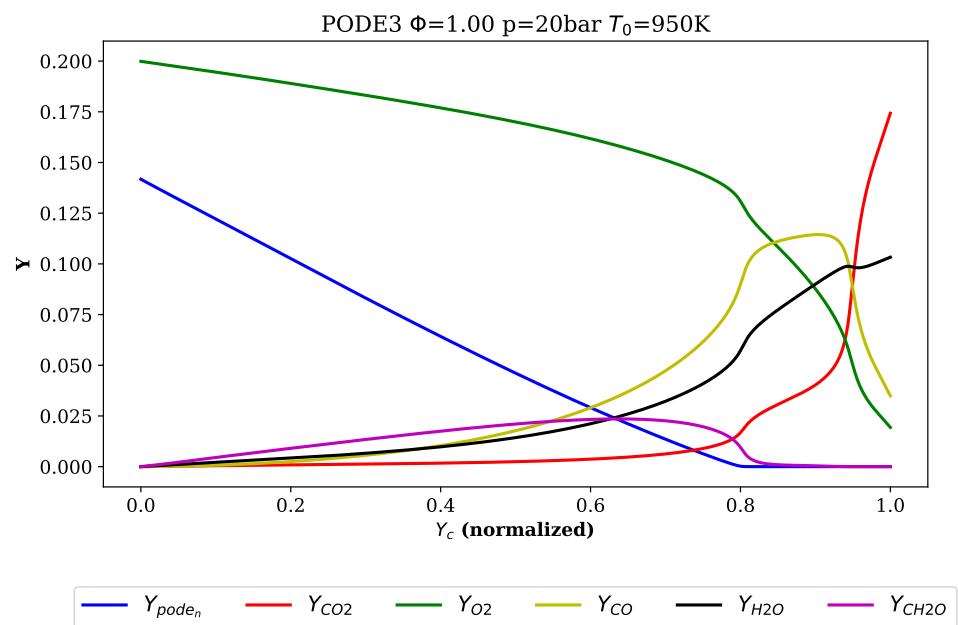


Figure A.3.: HR executed with the mechanism by Cai et al. [13]. Monitored is the species mass fraction development in Y_c space.

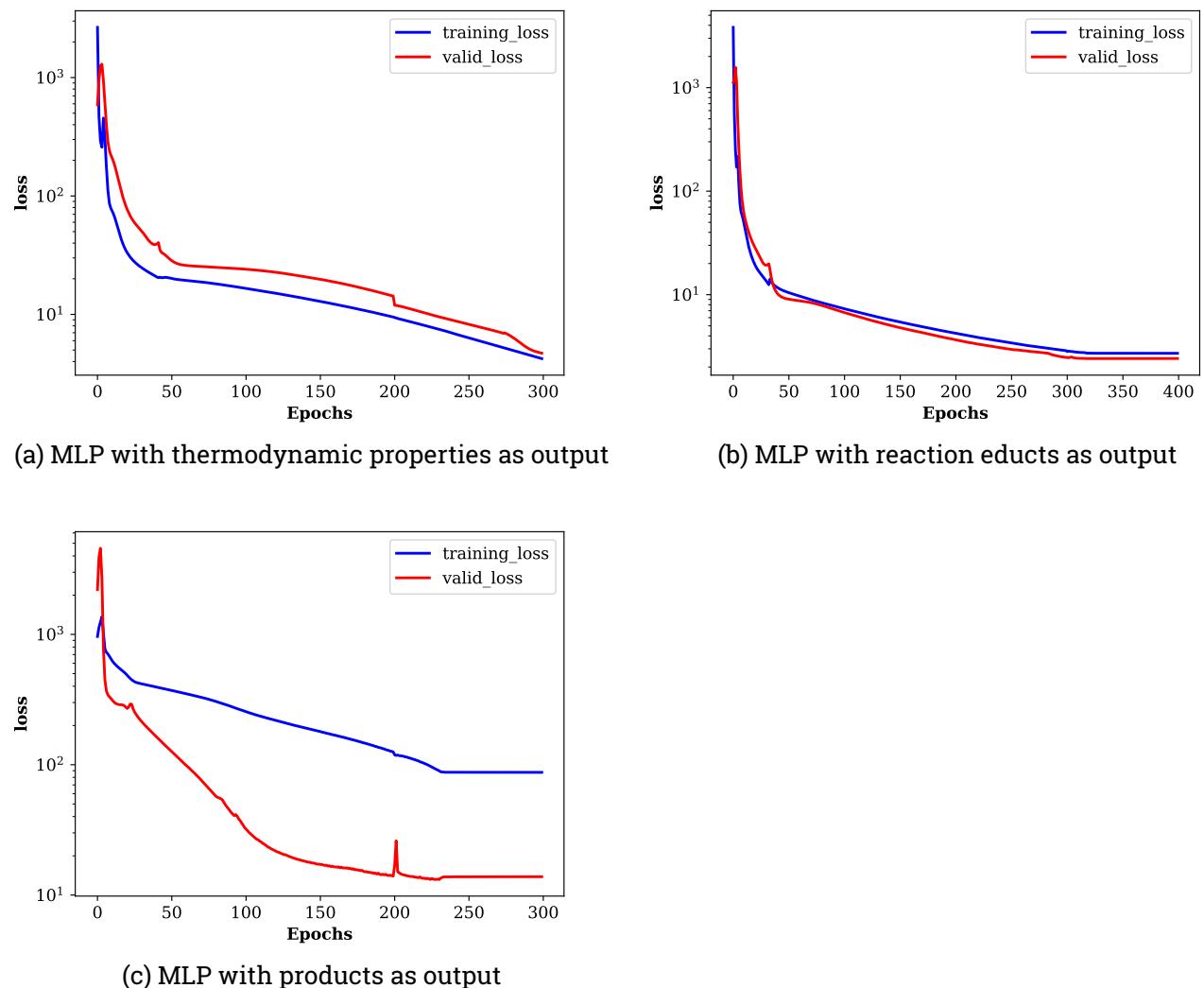


Figure A.4.: Training and validation loss curves for different MLPs. LR adjustments are performed with the validation loss starts to increase

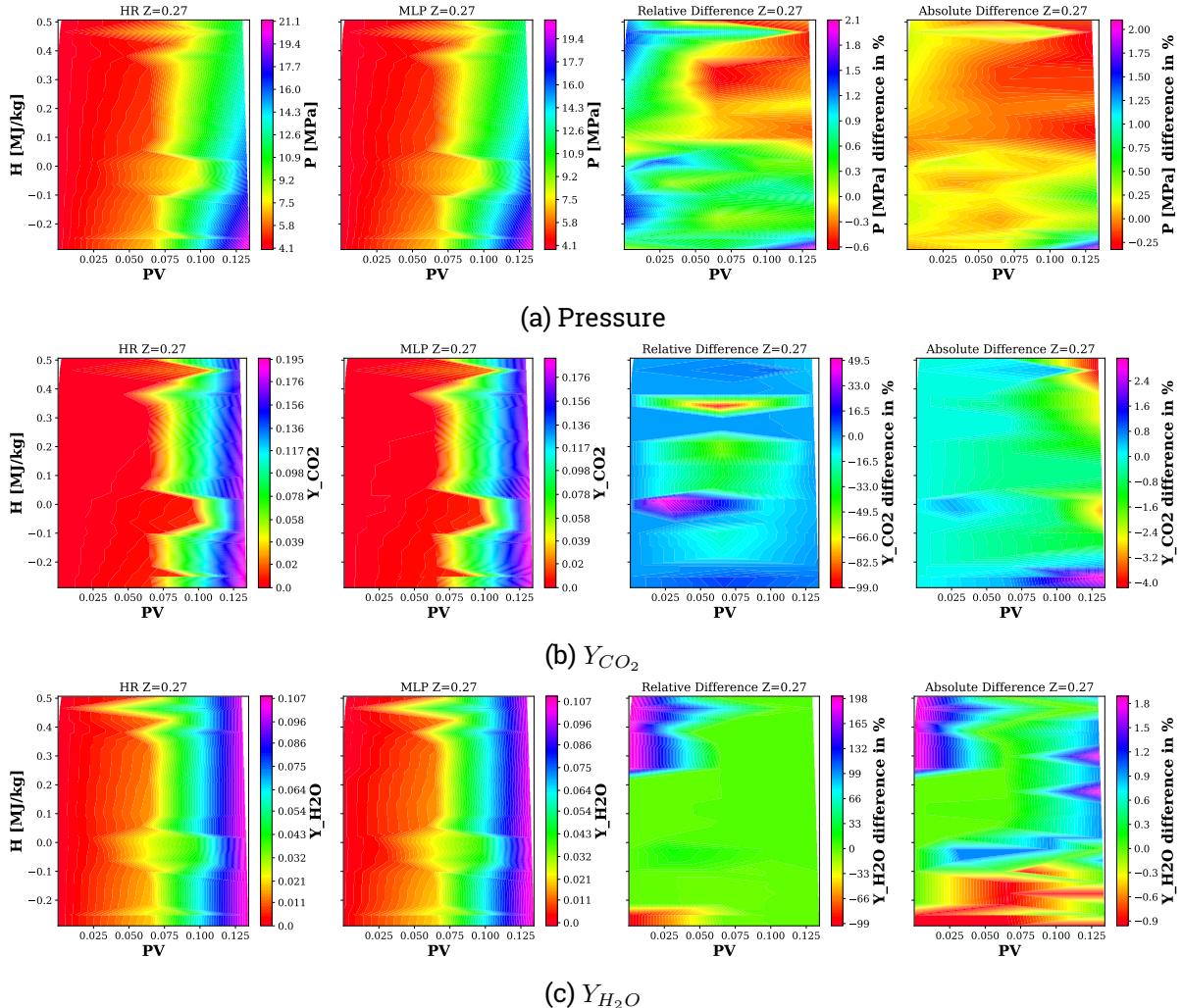


Figure A.5.: Output of the HR model, corresponding output of the MLP as well as relative and absolute deviation are shown. Comparison is performed for $\phi = 1.0$, a temperature range of 650 to 1250K and a pressure of 40bar.

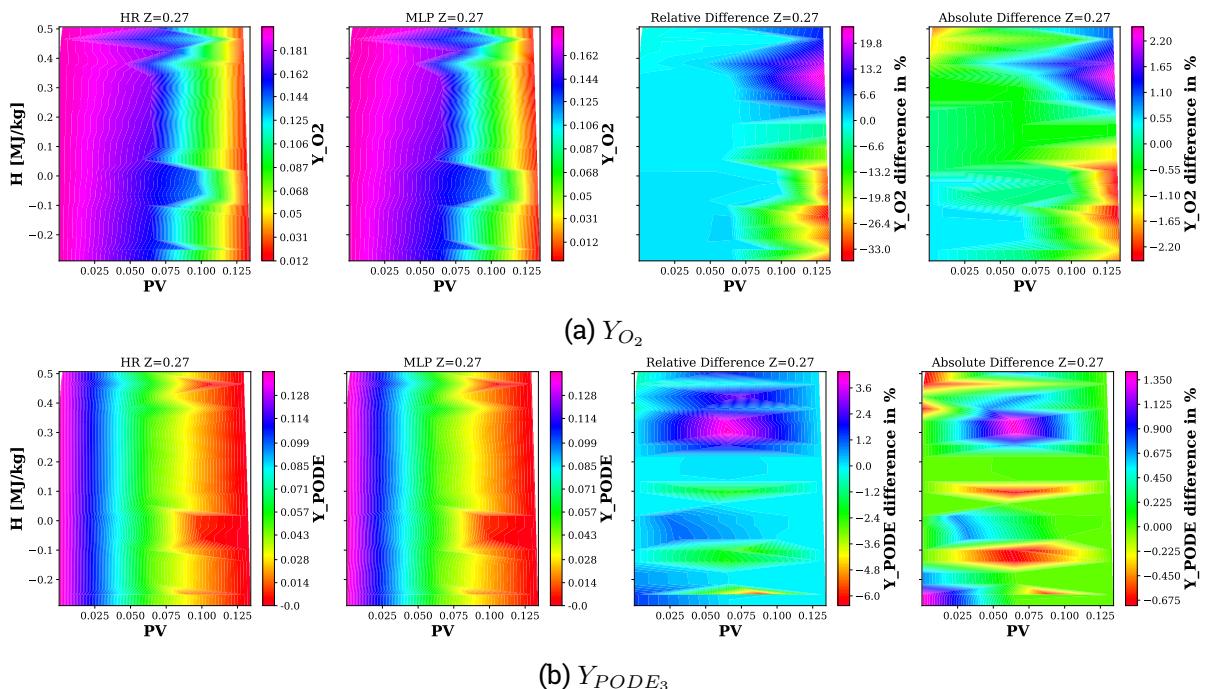


Figure A.6.: Output of the HR model, corresponding output of the MLP as well as relative and absolute deviation are shown. Comparison is performed for $\phi = 1.0$, a temperature range of 650 to 1250K and a pressure of 40bar.

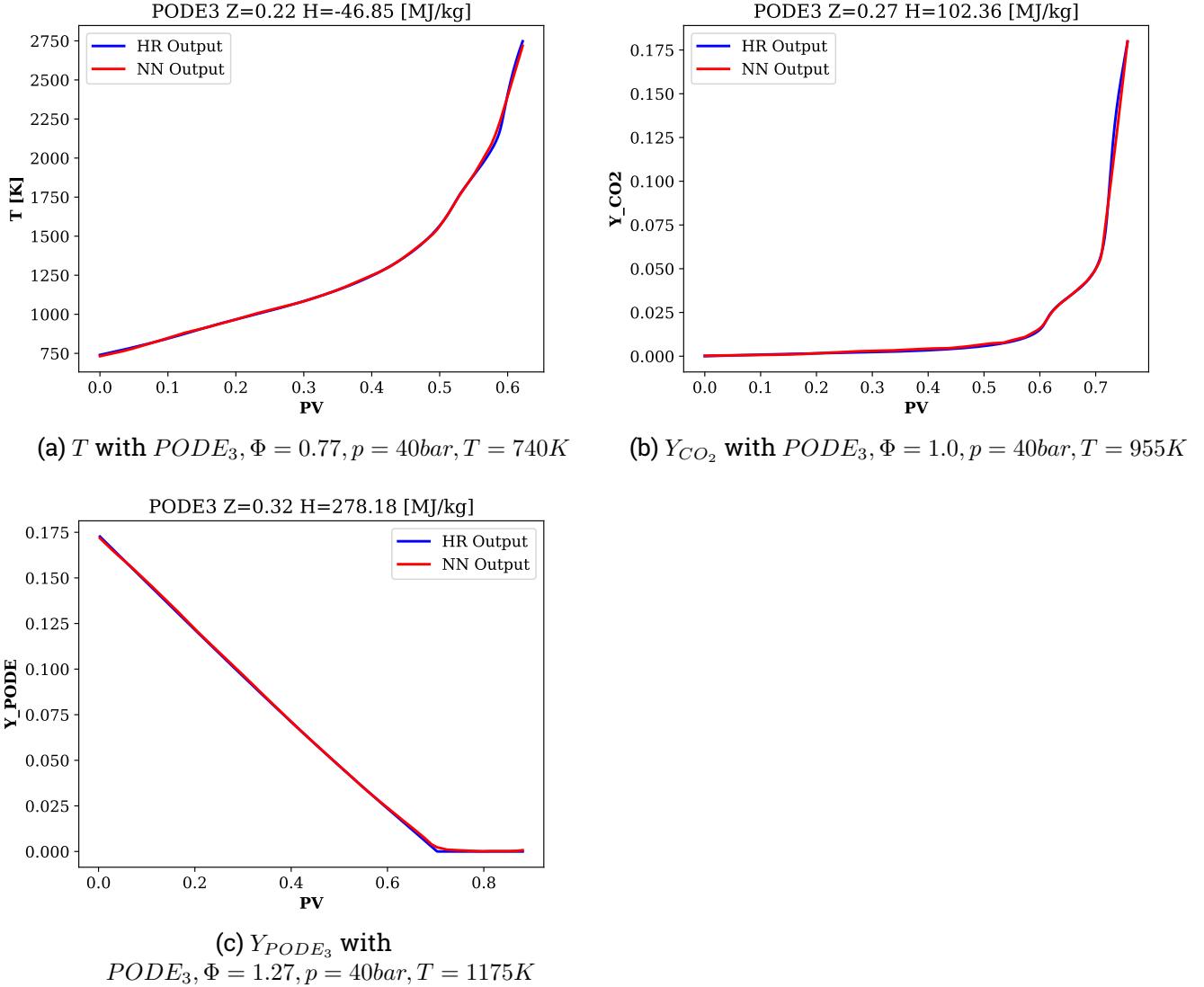


Figure A.7.: Comparison between the labels generated by the HR model and the output of the MLPs for data out of the Test data set. In each setting at least one parameter setting is not included in the training data and thus had to be interpolated.

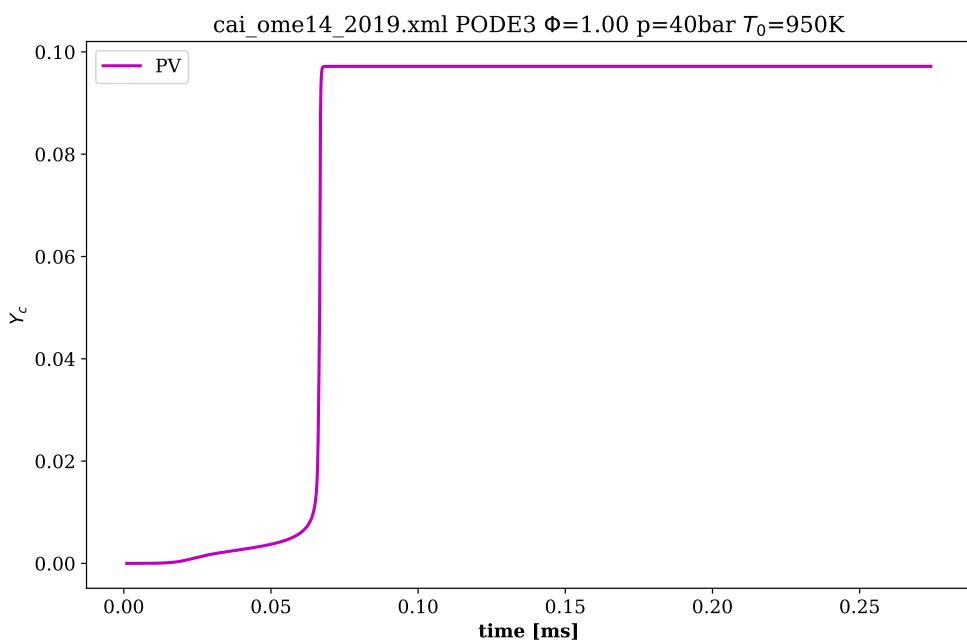


Figure A.8.: Progress variable Y_c of the comparison between MLP and GRM model

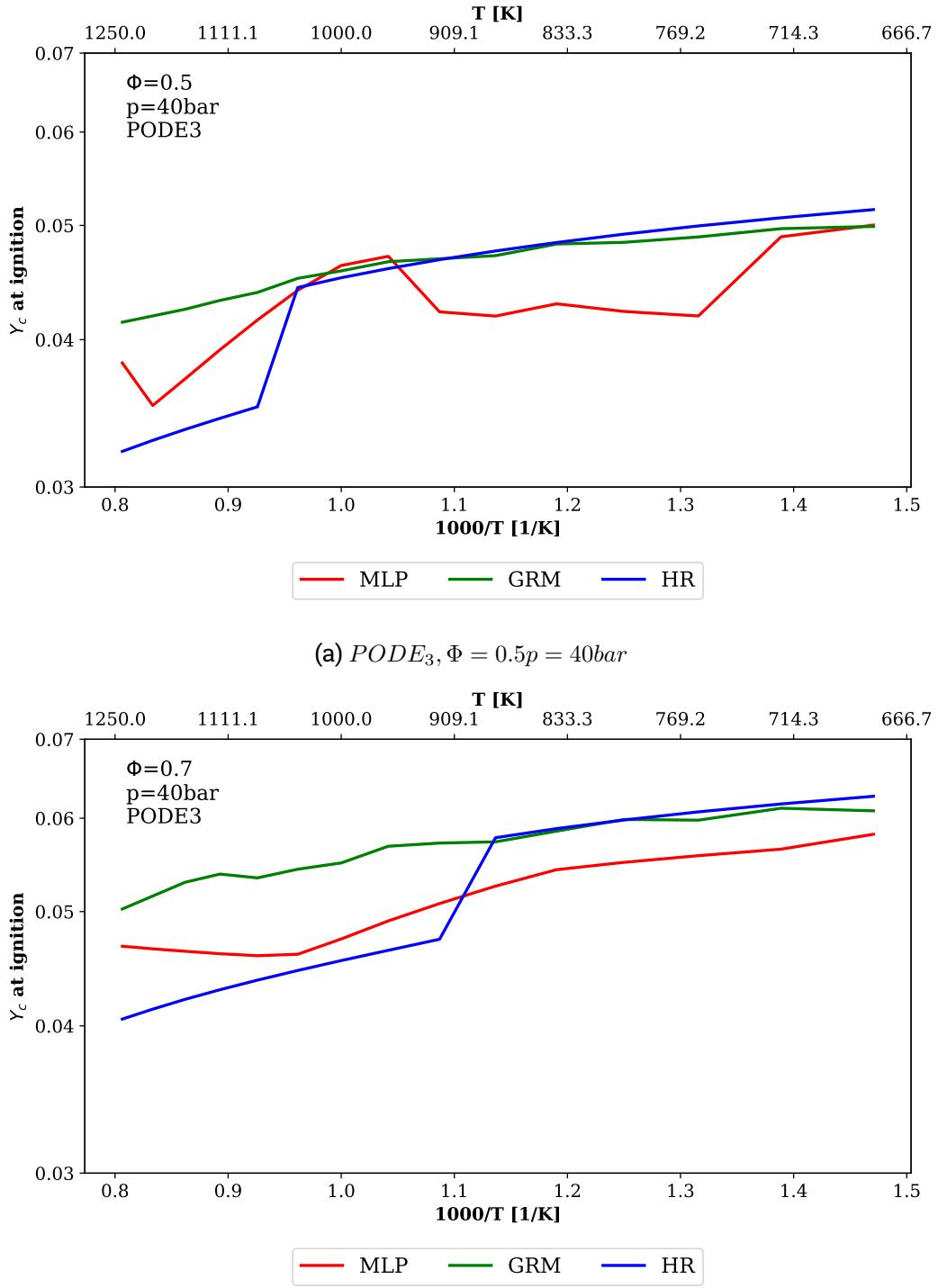


Figure A.9.: Comparison of the ignition delay (ID) in PV space for an initial temperature range from 680 to 1240K at a constant initial pressure of 40bar. Ideal ID is taken form the HR model and then predicted by the MLP and the GRM.

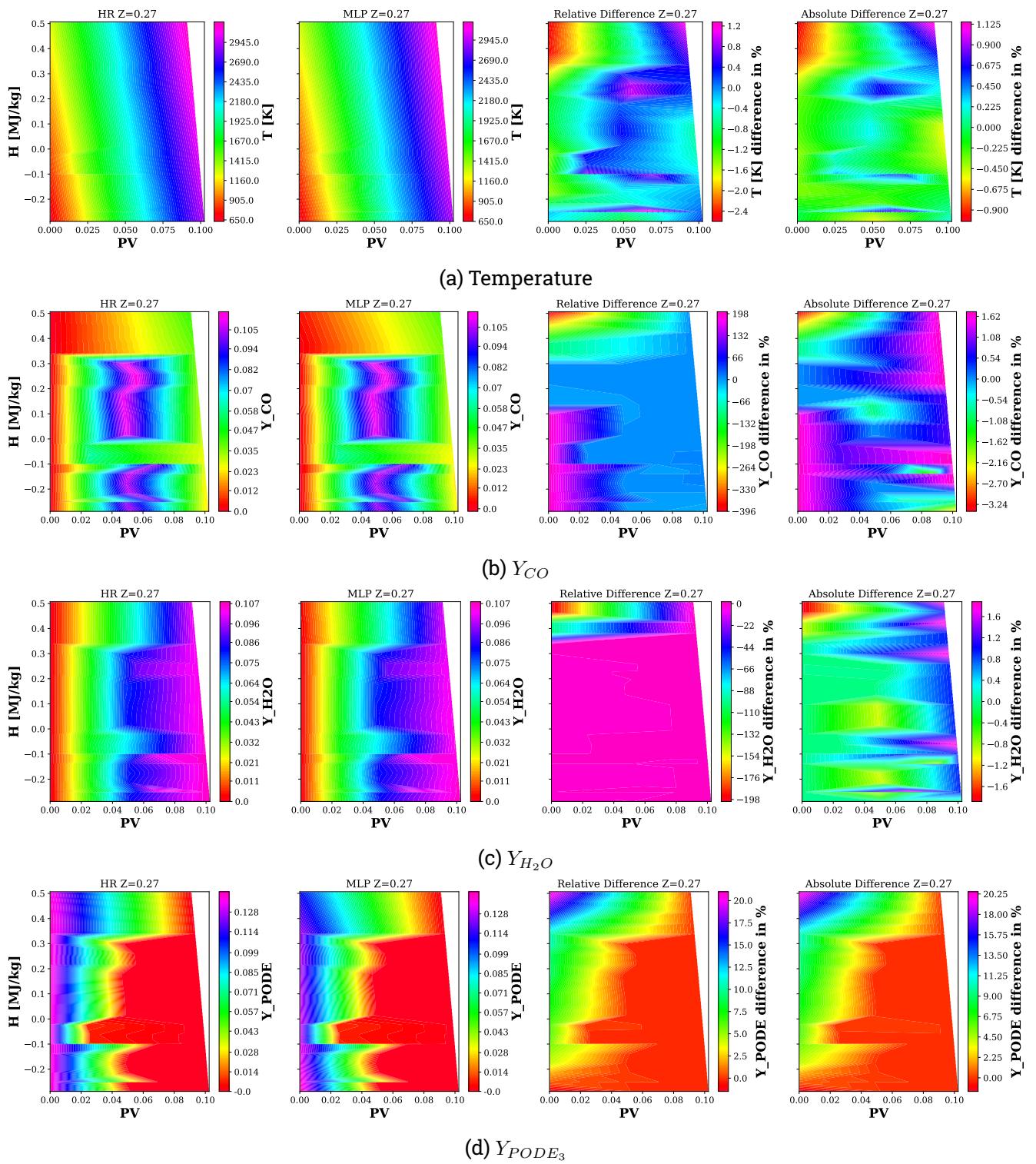


Figure A.10.: The output of the HR model and the corresponding output of the MLP using the models and Y_c for the comparison is shown. Evaluation is possible due to the relative and absolute deviation between the two output. Comparison is performed for $\phi = 1.0$ (corresponding $Z = 0.27$), a temperature range of 650 to 1250K and a pressure of 40bar.

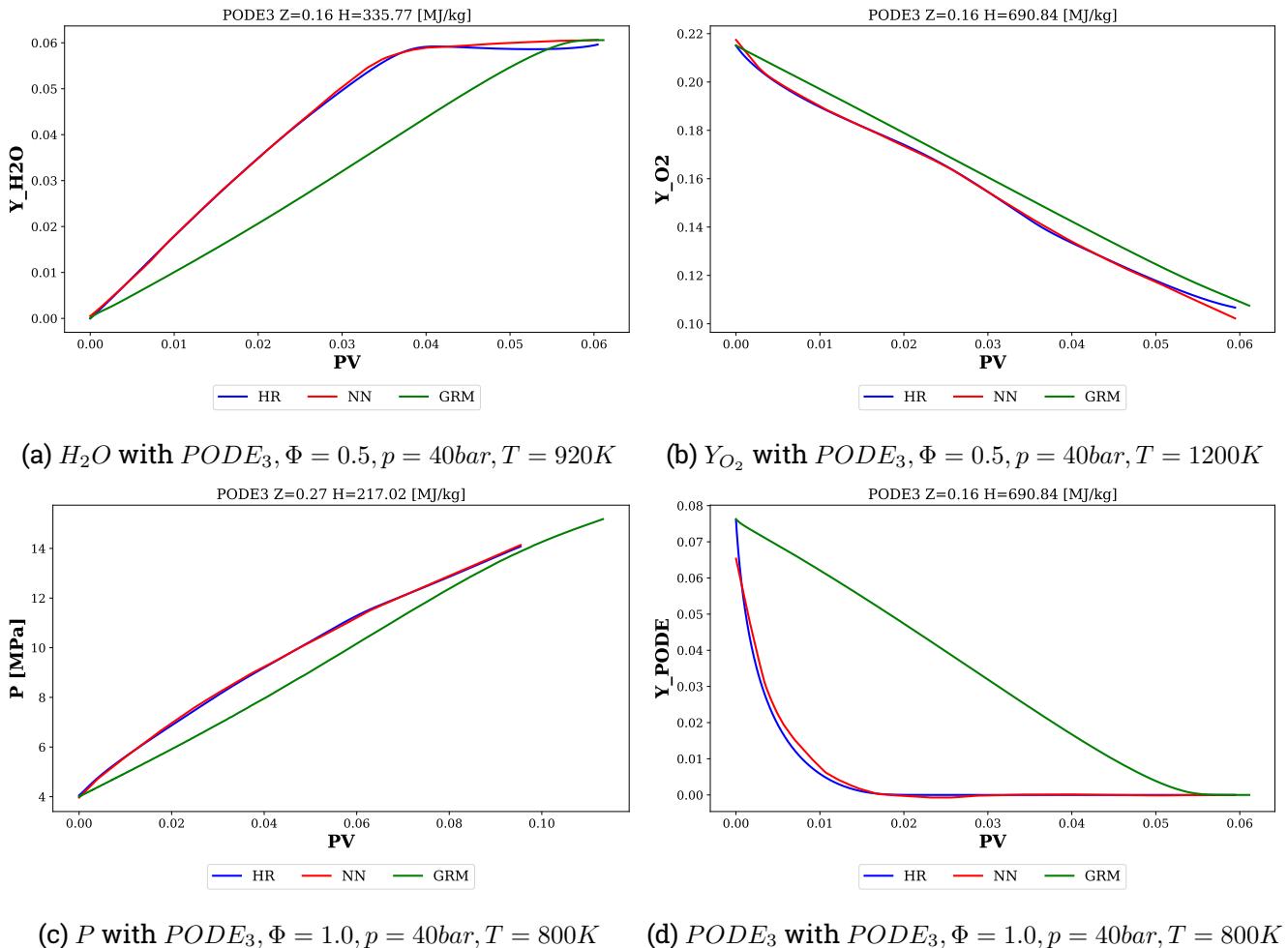


Figure A.11.: Comparison between the species progress in Y_c between the predicted results of the derived MLP model, the GRM model by Haspel [27] and the HR output