Fire Safety Journal 136 (2023) 103744

Contents lists available at ScienceDirect

Fire Safety Journal

journal homepage: www.elsevier.com/locate/firesaf

Inverse modelling of pyrolization kinetics with ensemble learning methods

Patrick Lauer [a], Lukas Arnold [b,c], Fabian Brännström [a,*]

[a] Chair of Fire Dynamics, University of Wuppertal, Gaußstraße 20, Wuppertal, 42119, Germany
[b] Institute for Advanced Simulation, Forschungszentrum Jülich, Wilhelm-Johnen-Straße, Jülich, 52428, Germany
[c] Chair of Computational Civil Engineering, University of Wuppertal, Pauluskirchstraße 7, Wuppertal, 42285, Germany

ARTICLE INFO

Dataset link: zenodo.org

Keywords:
Pyrolysis
Inverse modelling
Ensemble learning
Machine learning
Fire modelling
Extra trees
Decision trees

ABSTRACT

To simulate fire spread, especially the pyrolysis process – the thermal decomposition of a solid material – must be predicted. Yet, needed material dependent reaction kinetic parameters cannot be directly measured. Common methods infer them from small scale tests with inverse modelling approaches, which are computationally costly. Here, a novel machine learning based approach utilising extremely randomized trees (ERT) is presented and evaluated. It aims to derive these parameters almost instantly due to an inverse pre-trained surrogate model. The approach consists of an ERT classifier, a non-linear least squares optimiser and an ERT regressor. A thorough hyperparameter study was conducted. The model is evaluated with a synthetic thermogravimetric analysis (TGA) dataset. Calculated from an Arrhenius model, it contains data for more than $18 \cdot 10^6$ synthetic materials consisting of up to three components. The method is also applied on real experimental data, here polymethyl methacrylate (PMMA), gained from the Measurement and Computation of Fire Phenomena (MacFP) working group. Evaluation of the model demonstrated that it can instantly predict reaction kinetic parameters from TGA experiments for synthetic and real materials. Systematic analysis showed an overall R^2 score of 0.77 for the complete model predictions. The code and datasets are published as open access.

1. Introduction

Advanced fire modelling is based on computational fluid dynamics (CFD) and needs to consider fire spread to predict the progression of a fire. In contrast to design fires, fire spread modelling can be used to estimate the consequences of a fire and the effectiveness of counter measures on the fire performance of buildings [1,2], critical infrastructures [3] or vehicles [4,5]. In recent years, the prediction of wildfires emerged as another area where fire spread modelling is utilised [6].

To model the fire spread in all of these areas, the burning behaviour of materials has to be described. Burning of solid material can be separated into two mechanisms. One is the actual combustion in the gas phase and the other one is pyrolysis, i.e. the decomposition and transition from the condensed phase into the gas phase. To describe the condensed phase decomposition, the geometry, the heat transfer to the solid material, the heat conduction inside the material and the decomposition kinetics need to be characterised [7,8]. Adequate models exist to describe these processes [9–12]. Though, the material dependent input parameters for these models can rarely be directly measured. A common approach is to conduct small and bench scale experiments like thermogravimetrical analysis (TGA) and cone calorimeter tests, which both capture a temperature or time dependent mass loss rates (MLR). Then, the experimental results are used to adjust the involved model parameters with the aim to match the model and experiment data. This inverse modelling approach is mostly based on optimisation algorithms. This inverse modelling approach can be conducted with elaborated frameworks such as PROPTI, that was designed especially with pyrolysis modelling in focus [13,14]. Beside this approach, there is a plethora of other approaches [8,15–17].

Even though no scientific consensus exists on what is the best approach, the International Confederation for Thermal Analysis and Calorimetry (ICTAC) raised the concern that different methods of estimating kinetic parameters could be implicitly seen as in competition with each other. They rather suggest that different methods should be used complementary to each other to verify the estimated parameters [18]. While results from inverse modelling with optimisation algorithms can generally be considered to gain better results than other methods [16], computational costs for these approaches are high, spanning from hours to days on contemporary high performance computing clusters [19].

* Corresponding author.
E-mail addresses: lauer@uni-wuppertal.de (P. Lauer), l.arnold@fz-juelich.de, arnold@uni-wuppertal.de (L. Arnold), braennstroem@uni-wuppertal.de (F. Brännström).

https://doi.org/10.1016/j.firesaf.2023.103744
Received 16 September 2022; Received in revised form 24 November 2022; Accepted 10 January 2023
Available online 19 January 2023

To reduce the needed resources for the optimisation, this article proposes a machine learning approach in a supervised learning scenario. While application of machine learning methods in fire safety science and engineering in general had a substantial rise in the past years [20], they were rarely evaluated in the context of pyrolysis modelling for fire simulation. Recently, researchers started to evaluate machine learning methods for pyrolysis modelling in production of biochemical fuels [21–25]. Due to the emphasis on biochemical process optimisation in this application, the pyrolysis process happens in a controlled environment and can be modelled differently than in prediction of flame spread. Yet, flame spread during a fire happens in an uncontrolled environment, an aspect that needs to be taken into account. Hence, results cannot directly be adopted but give additional understanding for fire spread modelling applications or methodological insights.

In the proposed approach, an extremely randomized trees algorithm [26] is used to build a generalised inverse surrogate model to determine reaction kinetic parameters. These parameters are then used to solve the Arrhenius equation to compute the reaction rates. Synthetically generated TGA data is used for model training. Once trained and tested, this inverse surrogate model provides instantaneously the parameters needed to model pyrolysis. In the best case, the approach will lead to satisfactory parameter set, i.e. comparable to or better than those based on classical optimisation approaches. If not, it can be used as an accelerating preprocessing step for costlier, yet more precise, methods. In the following sections, it is shown how the training database is generated, how the model is assembled and how it is trained. Then, it is applied to a test dataset for validation and on real case experimental data. In the end, results are discussed.

2. Model

2.1. Forward model

The forward model is based on a TGA experiment. In this experiment, a specimen of only a few milligrams resting on a scale is heated in a furnace to a specified temperature. In fire science, the temperature of the experiment usually is determined by a constant heating rate β. The mass loss and its derivative, the mass loss rate, are captured with reference to temperature and time [7]. Thermal properties of the material itself are neglected, because usually the sample size is so small, that it is assumed to be in thermal equilibrium with the apparatus [27].

Reaction kinetics of pyrolysis of a material's component i undergoing a single reaction can be described with the Arrhenius approach displayed in Eq. (1). Here, $r_i(T_s)$ is the resulting reaction rate from the pre-exponential factor A_i, the activation energy E_i, the solid's temperature T_s, the components mass fraction $Y_{s,i}$ and the reaction order $n_{s,i}$. The index s refers to a solid material. [9]

$$r_i(T_s) = A_i Y_{s,i}^{n_{s,i}} \cdot \exp(-\frac{E_i}{RT_s}) \tag{1}$$

To model multiple reactions for a total number of components n and the interaction between the involved material components, individual instances of Eq. (1) can be summed up according to Eq. (2). Here, reaction rates r_i for all n components i are summed up to describe the mass change per time step, the mass loss rate. [9]

$$\frac{d\frac{m}{m_0}}{dt} = \sum_{i=1}^{n} \frac{dY_{s,i}}{dt} = -\sum_{i=1}^{n} r_i = -r(T) \tag{2}$$

2.2. Dataset generation

The forward model presented in the previous section is used to generate TGA experiment datasets synthetically. These sets are used to train the inverse surrogate model. Therefore, the pre-exponential factor A_i, the activation energy E_i and the initial material fraction $Y_{s,i,0}$ are randomly sampled. From these values the resulting reaction rate $r_i(T)$ can be calculated.

For simplification, only materials with up to three components are investigated in this contribution. The number of components a material consists of is denoted as n. Even if there were more reactions observed in a TGA experiment, most of the time, multi step-kinetics can be adequately described with two or three individual reactions [17]. Therefore, a maximum of three reactions was chosen. All components undergo an individual first order reaction. The reactions are considered as independent of each other and take place in parallel. The only material fractions that will be examined are the initial solid fractions, so $Y_{s,i,0}$ will be written as Y_i. All three reactions produce gaseous species or inert solid products. This does not need to be specified further, since only the reaction rate is investigated and none of the products will have any influence on it, when taken account for in the normalisation of the data. In this case, non-isothermal conditions are applied and the temperature in Eq. (1) is dependent on a constant heating rate β. All cases are investigated with four different heating rates. Using more than one heating rate helps to compensate for unwanted effects and errors in the experimental measurement. Since the modelled kinetic parameters generally are not temperature dependent and the forward model assumes a thermally thin sample, it would not have a significant effect on fitting when using synthetic data. For real data, it is recommended to find a fit for several different heating rates, as a small error in temperature measurement can lead to a significant error in the reaction kinetic parameters. So it is desirable to find kinetic parameters that fit experiments conducted at several heating rates, which combines occurring thermal effects. These effects can result for example from heat transfer in and to the sample, thermal inertia of the apparatus or from physical or chemical changes in the sample [28–30].

In general, values for the pre-exponential factor A_i and the activation energy E_i span a range of multiple orders of magnitude. Sampling uniformly will lead to an inadequate distribution and may result in reactions, which are either infinitely fast or will not start at all within the set boundaries. Thus, Eqs. (3) and (4) are used to limit the parameter space to a problem oriented parameter subspace. Here, the physical representation of the values $r_{i,p}$ and $T_{i,p}$ are the maximum reaction rate of a single reaction per component and the respective temperature at which it occurs. Y_i is the initial mass fraction of the material's component. [31–34]

$$E_i = \frac{RT_{i,p}^2 \exp(1) r_{i,p}}{\beta Y_i} \tag{3}$$

$$A_i = \frac{\exp(1) r_{i,p}}{Y_i} \exp\left(\frac{E_i}{RT_{i,p}}\right) \tag{4}$$

$r_{i,p}$, $T_{i,p}$ and Y_i are randomly sampled with a uniform distribution. With $r_{i,p}$ and $T_{i,p}$, the values for A_i and E_i are calculated using Eqs. (3) and (4) [31]. The boundaries for $r_{i,p}$ and $T_{i,p}$ given in Table 1 cover a broad range of values for commonly used real synthetic and natural polymers [7]. Sampling with the boundaries stated in Tables 1 and 2 leads to a parameter subspace bordered by the black line displayed in Fig. 1. By this, the parameter space is reduced to 16%, compared to uniform sampling of $\ln(A_i)$ and E_i, using their maximum values as upper boundaries. Any parameter combination that would be located outside of the marked area will react either completely in less than the chosen minimal time step or will not lead to any reaction at all within the chosen boundary conditions. The colours show the distribution of the samples in this subspace. Regions with higher sampling frequency are displayed in yellow, regions with lower frequencies are displayed in green, fading to blue.

Then, for the heating rate $\beta_1 \ldots \beta_4$, the temperature dependent mass loss rate $r(T)$ is calculated with the given parameters using Eqs. (1) and (2).

Each dataset $\mathbb{D}_n$ consists of N_u elements, where one element u contains $A_1^u, \ldots, A_n^u$, $E_1^u, \ldots, E_n^u$, $Y_1^u, \ldots, Y_n^u$ and $L_{\beta_1}^u, L_{\beta_2}^u, L_{\beta_3}^u, L_{\beta_4}^u$. Every $L_{\beta_l}^u =$

Table 1
Sampling range boundaries for peak reaction rate and temperature.

Parameter	Lower Boundary	Upper Boundary
$T_{i,p}$	100 °C	500 °C
$r_{i,p}$	$1 \cdot 10^{-4}$ s^{-1}	$1 \cdot 10^{-2}$ s^{-1}

Table 2
Sampling range boundaries and values for initial component fractions.

Parameter	n=1	n=2	n=3
Y_1	1	0...1	0...1
Y_2	–	$1 - Y_1$	0...Y_1
Y_3	–	–	$1 - Y_1 - Y_2$

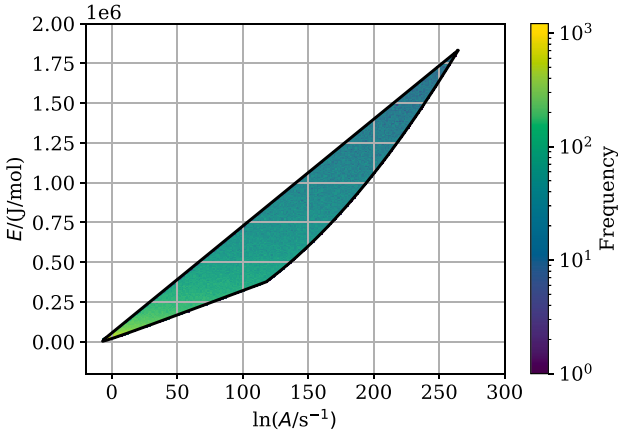


Fig. 1. Resulting parameter subspace from Eqs. (3) and (4) and the parameter ranges from Table 1. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 3
Generated datasets.

Dataset $\mathbb{D}_n$	Dataset size N_u	Components n	Heating Rates β_l / (K/min)
$\mathbb{D}_1$	$6.4 \cdot 10^6$	1	5, 10, 30, 40
$\mathbb{D}_2$	$6.4 \cdot 10^6$	2	5, 10, 30, 40
$\mathbb{D}_3$	$6.4 \cdot 10^6$	3	5, 10, 30, 40

$\{r(20\,°C), \dots, r(550\,°C)\}$ is a set of discrete values calculated from Eq. (2) for a range from 20 °C to 550 °C with a $\Delta T = 2$ K and a heating rate β_l given in Table 3. This leads to $m = 266$ discrete data points for each $L^u_{\beta_l}$. The properties of the three generated datasets are specified in Table 3. The Dataset size was limited by computational and memory capacities. Effects of this limitation are taken into consideration in Section 3.5.

As an example, a random dataset element for each number of components n is shown in Fig. 2. In the datasets, combinations of completely and partly overlapping peaks as well as completely separated curves for different components are present. For creating the data sets, standard python modules where used. The code is publicly available [35–37].

2.3. Inverse model

2.3.1. General model introduction

The forward model described in Section 2.1 to calculate the reaction rate $r(T)$ with activation energy E_i, pre-exponential factor A_i, and material fractions Y_i as input parameters can be summarised in general as in Eq. (5).

$$f(A_1, \dots, A_n, E_1, \dots, E_n, Y_1, \dots, Y_n) = (L^u_{\beta_1}, L^u_{\beta_2}, L^u_{\beta_3}, L^u_{\beta_4}) \tag{5}$$

Since only $r(T)$ is easy to measure directly with experiments like TGA and A_i and E_i are unknown and needed for forward modelling,

Table 4
Subsets build from initial datasets.

Dataset	
$\mathbb{D}_{1,2,3}$	$= \mathbb{D}_1^{2\cdot10^6} \cup \mathbb{D}_2^{2\cdot10^6} \cup \mathbb{D}_3^{2\cdot10^6}$
$\mathbb{D}_{1,2,3,train}$	$\subseteq \mathbb{D}_{1,2,3},\ \lvert\mathbb{D}_{1,2,3,train}\rvert = 4.5 \cdot 10^6$
$\mathbb{D}_{1,2,3,test}$	$= \mathbb{D}_{1,2,3} \setminus \mathbb{D}_{1,2,3,train}$
$\mathbb{D}_{2,gauss}$	$\subseteq \mathbb{D}_2,\ \lvert\mathbb{D}_2\rvert = 0.5 \cdot 10^6$
$\mathbb{D}_{3,gauss}$	$\subseteq \mathbb{D}_3,\ \lvert\mathbb{D}_3\rvert = 0.5 \cdot 10^6$
$\mathbb{D}_{1,train}$	$\subseteq \mathbb{D}_1^{6\cdot10^6},\ \lvert\mathbb{D}_{1,train}\rvert = 4.5 \cdot 10^6$
$\mathbb{D}_{1,test}$	$= \mathbb{D}_1^{6\cdot10^6} \setminus \mathbb{D}_{1,train}$
$\mathbb{D}_{2,train}$	$\subseteq \mathbb{D}_2^{6\cdot10^6},\ \lvert\mathbb{D}_{2,train}\rvert = 4.5 \cdot 10^6$
$\mathbb{D}_{2,test}$	$= \mathbb{D}_2^{6\cdot10^6} \setminus \mathbb{D}_{2,train}$
$\mathbb{D}_{3,train}$	$\subseteq \mathbb{D}_3^{6\cdot10^6},\ \lvert\mathbb{D}_{3,train}\rvert = 4.5 \cdot 10^6$
$\mathbb{D}_{3,test}$	$= \mathbb{D}_3^{6\cdot10^6} \setminus \mathbb{D}_{3,train}$
$\mathbb{D}_{1,2,3,eval}$	$\subseteq (\mathbb{D}_1 \cup \mathbb{D}_2 \cup \mathbb{D}_3) \setminus \mathbb{D}_{1,2,3},\ \lvert\mathbb{D}_{1,2,3,eval}\rvert = 1.5 \cdot 10^5$

an inverse surrogate model to predict these parameters $\hat{A}_i$, $\hat{E}_i$ and $\hat{Y}_i$ from the reaction rate $r_i(T)$ can be formulated. The resulting inverse surrogate model is displayed in Eq. (6).

$$g(L^u_{\beta_1}, L^u_{\beta_2}, L^u_{\beta_3}, L^u_{\beta_4}) = (\hat{A}_1, \dots, \hat{A}_n, \hat{E}_1, \dots, \hat{E}_n, \hat{Y}_1, \dots, \hat{Y}_n) \tag{6}$$

To build such an inverse model, a machine learning method called extremely randomized trees (ERT) is used. This algorithm, adapted from the random forests algorithm [38], is a decision tree based method for supervised classification and regression problems [26]. It was chosen because it outperformed other ensemble learning algorithms in image classification [39] and gene network inference [40,41] and was shown to perform well on multi-target regression [42]. It is easy to use in a parallel computing environment like high performance computing clusters, gives a robust model, does not tend to overfit and is efficient on huge datasets [26,43,44]. ERT can pretty well interpolate in the parameter space that was used for training but has limitations on extrapolating outside the parameter space used for training [44]. The ERT was used as implemented in SciKit-Learn [45]. A detailed explanation on ERT is given in Appendix.

2.3.2. Dataset description

Eq. (7) shows the complete dataset, which generation was described in Section 2.2. Here, each x^u is the input value and y^u is the corresponding output value. Subsets of the initial datasets described in Table 3 are build for training and testing the individual submodels described in Sections 2.3.3 to 2.3.7. The composition of these subsets is given in Table 4. Here, the superscript defines a subset with elements from 1 to the value of the superscript for u from the original set. If only the cardinality is given, it describes the number of random elements from the original set.

$$\mathbb{D}_n = \{(x^u, y^u) : u = 1, \dots, N_u\} \tag{7}$$

2.3.3. Prediction concept

The training data is used to train the ERT model by providing the inputs and outputs from the samples. The ERT model adopts according to the provided training dataset during the training process and builds a forest of random decision trees. During the prediction process different combinations of the datasets are used as explained in the correspondent subsections.

There are three different types of predictions to be made. The number of components $\hat{n}$, the component's fractions $\hat{Y}$ and the reaction kinetic parameters $\hat{A}$ and $\hat{E}$ need to be predicted. Thus, the prediction process is split up in three submodels as shown in Fig. 3. First, an ERT classification model is trained, tested and applied to predict the number of independent reactions taking place, which equals to the number of components n. Then, a Gaussian curve fitting model is applied to predict the initial fractions Y_i of each component i. Individual model parts are built for materials consisting of two and three components.

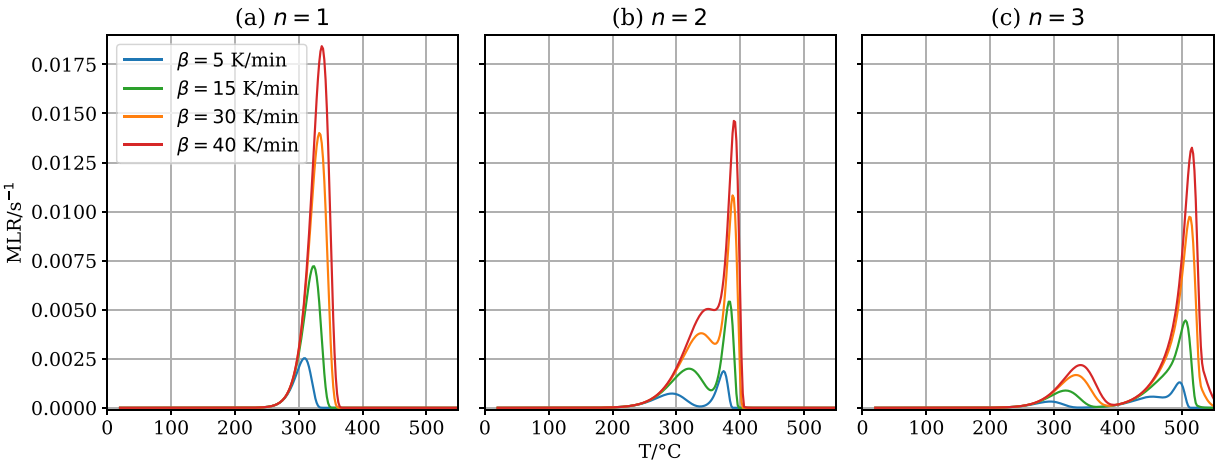


Fig. 2. Random examples for the generated reaction rates, with (**a**) 1 component, (**b**) 2 components and (**c**) 3 components.

Table 5
Generated submodels.

No.	Submodel name	Type	Input x	Output $\hat{y}$
1	Number of components	Classification	$L^u_{\beta_1}, L^u_{\beta_2}, L^u_{\beta_3}, L^u_{\beta_4}$	$\hat{n}$
2	Fraction of components	Gaussian fitting	$L^u_{\beta_1}, L^u_{\beta_2}, L^u_{\beta_3}, L^u_{\beta_4}$	$\hat{Y}_1, \dots, \hat{Y}_n$
3	Reaction kinetics	Regression	$L^u_{\beta_1}, L^u_{\beta_2}, L^u_{\beta_3}, L^u_{\beta_4}, \hat{Y}_1, \dots, \hat{Y}_n$	$\hat{A}_1, \dots, \hat{A}_n, \hat{E}_1, \dots, \hat{E}_n$

There is no need to predict the fraction of materials consisting of only one component, since it equals always to 1. Finally, an ERT regression model is utilised to predict the actual reaction kinetics, with an individual part for each number of components. These three submodels are sequentially applied. An overview of input and output quantities is given in Table 5.

On each of the submodels and on the complete model, several performance and error metrics are applied in a multi-fitness approach. They are described in the respective subsections [46]

2.3.4. Prediction of number of components submodel

In the first submodel, an ERT classifier is used to predict if the investigated process consists of one, two or three individual components. For this, there are subsets created from each $\mathbb{D}_n$, with N_u of $2 \cdot 10^6$ samples, leading to a new dataset $\mathbb{D}_{1,2,3}$ with a total of $6 \cdot 10^6$ samples. Here, each $x^u = \left(L^u_{\beta_1}, L^u_{\beta_2}, L^u_{\beta_3}, L^u_{\beta_4}\right)$ is the input value and $y^u = n$ is the expected output value for training, while $\hat{y}^u = \hat{n}$ is the predicted classification. $\mathbb{D}_{1,2,3}$ is split up into a training dataset $\mathbb{D}_{1,2,3,train}$, containing 75 % of $\mathbb{D}_{1,2,3}$ and a testing dataset $\mathbb{D}_{1,2,3,test}$, containing the other 25 %. $\mathbb{D}_{1,2,3,train}$ is used to train the classifier, while $\mathbb{D}_{1,2,3,test}$ is used to test the model by predicting $\hat{y}^u$ from x^u and comparing it to the given y^u. Evaluation metrics for this comparison are accuracy Ac (correct predictions in relation to all predictions, Eq. (8)), precision Pr (ratio of correct positive predicted values to all positive values, Eq. (9)) and recall Re (probability that an input is correctly assigned as belonging to a certain number of reactions, Eq. (10)). These metrics are applied to any individual number of components n as well as an average over all n. For cross validation of the training dataset, a k-fold strategy with $k = 5$ is used. Here, the data is split into k parts, where all combinations of $k - 1$ parts are used for training and then the remaining part is used for testing [47].

$$\text{Accuracy:} \quad Ac = \frac{\text{correct classifications}}{\text{all classifications}} \tag{8}$$

$$\text{Precision:} \quad Pr = \frac{\text{true positives}}{\text{true positives} + \text{false positives}} \tag{9}$$

$$\text{Recall:} \quad Re = \frac{\text{true positives}}{\text{true positives} + \text{false negatives}} \tag{10}$$

2.3.5. Prediction of component fractions submodel

The second submodel is used to predict the initial fractions of different components present in the material. For this, a Gaussian fitting model is used for materials consisting of two and three components. If the material consists only of one component, the fraction of the component is one. Gaussian curves in the form given by Eq. (11) with parameters μ_i, the position of the peak value, K_i, the peak value and σ_i^2, the variance, of each corresponding component i are fitted to the mass loss rate curves in the test data. The number of curves that are fitted to the data is equal to the number of components prescribed or predicted in the step before. For each component, three parameters need to be fitted. Generic, non-linear least squares routines with a trust region reflective algorithm [48] as implemented in [49] were applied to find the parameters of each of the Gaussian curves. In each iteration of these routines, a set of the parameters $\gamma = \{\mu_1, \dots, \mu_n, K_1, \dots, K_i, \sigma_1^2 \dots, \sigma_n^2\}$ are estimated, $p(z, \gamma)$ with $z = (20...550)$ is calculated and the squared error between y and $p(z, \gamma)$ is determined. Then, iteratively, this error is reduced by optimising γ with a trust region reflective algorithm.

$$p(z, \gamma) = \sum_{i=1}^{n} K_i \cdot \frac{1}{\sigma_i \sqrt{2\pi}} \exp\left(-\frac{1}{2} \frac{(x - \mu_i)^2}{\sigma_i^2}\right) \tag{11}$$

A fitting example is shown in Fig. 4. On the left, the mass loss rate from a TGA experiment at an arbitrary heating rate of a material consisting of three components is shown. The plot on the right side presents the resulting Gaussian curves fitted to the individual peaks and the surfaces S_1, S_2 and S_3 that are enclosed by each Gaussian curve.

As input values, $x^u = \left(L^u_{\beta_1}, L^u_{\beta_2}, L^u_{\beta_3}, L^u_{\beta_4}\right)$ from the datasets $\mathbb{D}_{2,gauss}$ and $\mathbb{D}_{3,gauss}$ are used. The component fractions are calculated for each heating rate individually and then their mean is calculated. $y^u = Y^u_1, \dots, Y^u_n$ is the true output value for validation, while $\hat{y}^u = \hat{Y}^u_1, \dots, \hat{Y}^u_n$ is the predicted value. For evaluation, the root mean square error (RMSE) is calculated according to Eq. (12) between y^u and $\hat{y}^u$ [50].

$$\text{RMSE}(y^u, \hat{y}^u) = \frac{1}{\sqrt{n}} \sqrt{\sum_{i=1}^{n} (\hat{Y}^u_i - Y^u_i)^2} \tag{12}$$

2.3.6. Prediction of reaction kinetics submodel

The third submodel predicts the actual reaction kinetics. Individual ERT regression models are built for materials consisting of one, two and

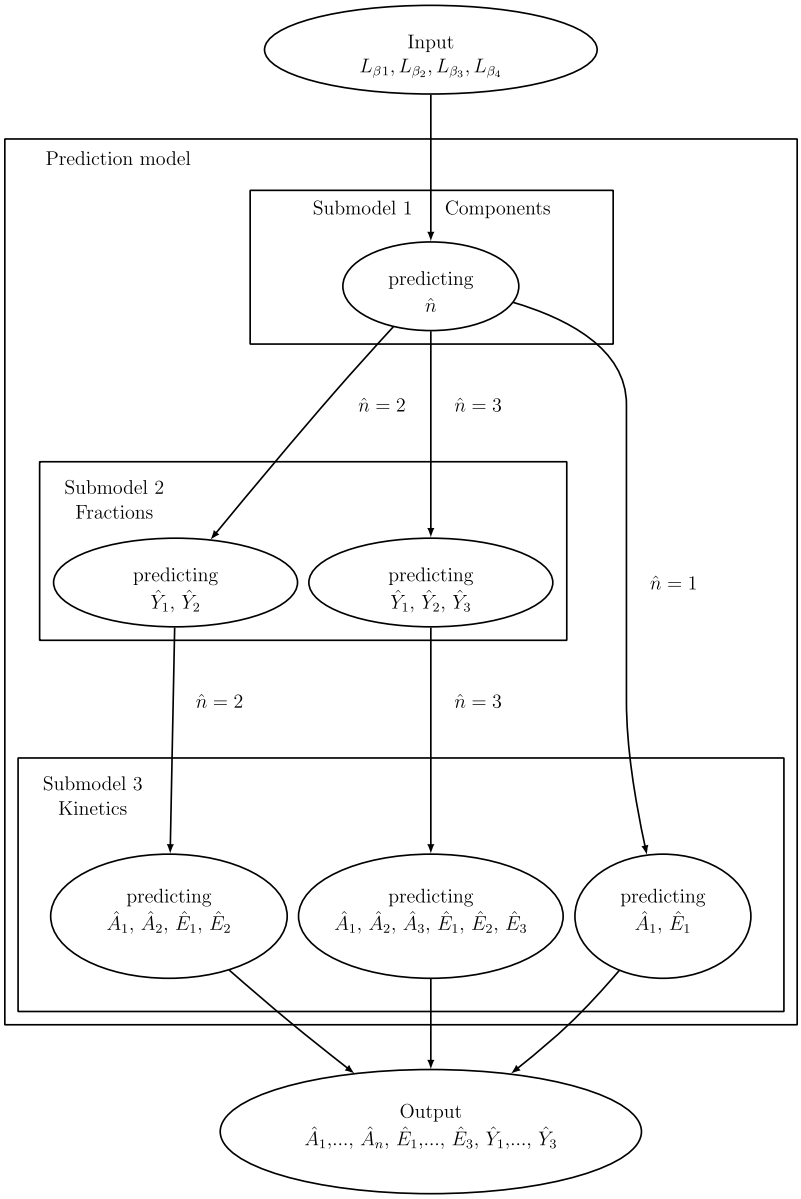
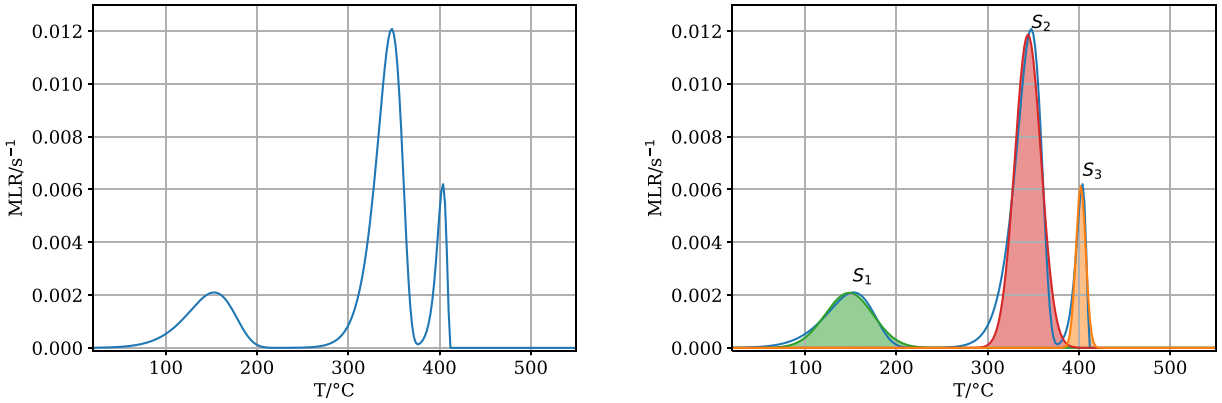


Fig. 3. Concept of the inverse surrogate model.



(a) Example dataset to be fitted with a set of three Gaussian curves.

(b) The resulting Gaussian curves fitted to the dataset.

Fig. 4. Example of the component fractions' prediction.

Table 6
Values of hyperparameter study.

No.	submodel	N_u 10^6	Number of estimators	Maximum depth
1	Number of components	1.5, 2.25, 4.5, 6	50, 100, 1000, 1500, 2000, 3000, 5000	1, 5, 10, 50, 100, 1064
3	Reaction kinetics	0.1, 0.75, 1.5, 3, 6	1, 50, 100, 200	1, 5, 10, 50, 1064

three components. For each number of components n with $n \in \{1,2,3\}$, a separate model is built. $\mathbb{D}_{n,train}$ is used for training and $\mathbb{D}_{n,test}$ is used for testing the model. The datasets are composed as described in Table 4. Here, each $x^u = \left(L^u_{\beta_1}, L^u_{\beta_2}, L^u_{\beta_3}, L^u_{\beta_4}, Y^u_1, \ldots, Y^u_n\right)$ is the input value and $y^u = A^u_1, \ldots, A^u_n, E^u_1, \ldots, E^u_n$ is the expected output. The model predicts $\hat{y}^u = \hat{A}^u_1, \ldots, \hat{A}^u_n, \hat{E}^u_1, \ldots, \hat{E}^u_n$.

The evaluation criterion is the R^2 score as shown in Eq. (13). The mean value and the distribution of values are used for evaluation. The test data $\mathbb{D}_{n,test}$ is used to determine the accuracy of the trained ERT model.

$$R^2(y^u, \hat{y}^u) = \left[\sum_{i=1}^{n}\left(1 - \frac{\sum_{u=1}^{N_u}(A^u_i - \hat{A}^u_i)^2}{\sum_{u=1}^{N_u}(A^u_i - \bar{A}^u)^2}\right) + \sum_{i=1}^{n}\left(1 - \frac{\sum_{u=1}^{N_u}(E^u_i - \hat{E}^u_i)^2}{\sum_{u=1}^{N_u}(E^u_i - \bar{E}^u)^2}\right)\right]/2n \tag{13}$$

$\hat{y}^u$, predicted with the inverse model from input x^u, is used to calculate $\hat{x}^u$ from Eqs. (1) and (2) for evaluation. Then, $\hat{x}^u$ and x^u are compared by calculating the normalised root mean square error (NRMSE) according to Eq. (14). Median value and distribution of values are used for evaluation. A 5-fold cross validation as in the first model is also conducted here.

$$\mathrm{NRMSE}(x^u, \hat{x}^u) = \frac{1}{(x^u_{max} - x^u_{min})\sqrt{4m}}\sqrt{\sum_{t=1}^{4m}(\hat{x}^u_t - x^u_t)^2} \tag{14}$$

2.3.7. Complete model

After the submodels are evaluated individually, a final evaluation, i.e. the combination of all three submodels as shown in Fig. 3, is conducted. The dataset $\mathbb{D}_{1,2,3,eval}$ is fed to the complete model with $x^u = \left(L^u_{\beta_1}, L^u_{\beta_2}, L^u_{\beta_3}, L^u_{\beta_4}\right)$ as the input and as the expected output $y^u = A_1, \ldots, A_n, E_1, \ldots, E_n, Y_1, \ldots, Y_n$. This dataset only contains data that was not used to build any of the submodels. Then, $\hat{y}^u = \hat{A}_1, \ldots, \hat{A}_n, \hat{E}_1, \ldots, \hat{E}_n, \hat{Y}_1, \ldots, \hat{Y}_n$ is predicted by the model and used to calculate $\hat{x}^u$ from Eqs. (1) and (2). $\hat{x}^u$ and x^u are used to calculate the NRMSE according to Eq. (14). To encounter possible misallocation, the best NRMSE for any order in $\hat{y}$ is used. Mean value and distribution of values are used for evaluation. In addition, the R^2 score for y^u and $\hat{y}^u$ is used to evaluate the model.

2.3.8. Hyperparameter tuning

The ERT algorithm has algorithmic parameters that are used to adjust the learning process; these parameters will be referred as hyperparameters. Studies for the best settings for these hyperparameters for the submodels were conducted.

For the number of components submodel, using an ERT classifier, the hyperparameters considered as most sensitive in literature are the training dataset size, the number of estimators (number of trees in the forest) and the maximum depth of a single tree [26,51]. The same is true for the reaction kinetics submodel. The fraction of components submodel has no hyperparameters to tune. Hyperparameters are evaluated one at a time, ceteris paribus. Their values are shown in Table 6. Results for the operating submodels in Section 3 are only presented for the parameters underlined in this table. Except for one value, where next to no change in outcome by increasing this value was observed, only the upper boundaries of the hyperparameter studies were chosen.

Table 7
Evaluation values for first submodel.

Components	Accuracy (%)	Precision (%)	Recall (%)
$n = 1$	99.97	98.73	99.97
$n = 2$	96.60	83.09	96.60
$n = 3$	80.27	97.29	80.27
total	92.28	93.04	92.28

2.3.9. Real case application

For real case application evaluation, TGA experimental data of black Polymethylmethacrylate (PMMA) is used. The data is taken from experiments conducted by the University of Lille published in context of the Measurement and Computation of Fire Phenomena Condensed Phase Working Group (MaCFP) [52]. The mass loss rates from experiments conducted with four different heating rates (5 K/min, 10 K/min, 20 K/min, 50 K/min) is used as input for the inverse model to predict the reaction kinetics. From the publicly available total mass over time, the relative mass loss rate (in s^{-1}) is calculated and smoothed by calculating the moving average over 10 data points. Then, the data is downsampled with linear interpolation to 266 evenly distributed mass loss rate input values for a temperature range from 20 °C to 550 °C per heating rate.

Since the previously presented model is designed for a different set of experimental heating rates, an adjusted model equivalent to the presented process was built and used. Except for the training dataset generated for heating rates of 5 K/min, 10 K/min, 20 K/min, 50 K/min, anything else is the same as explained before.

For comparison, NRMSE between the experimental and predicted mass loss rates is calculated. Another indicator is the peak mass loss rate and the corresponding temperature found by other laboratories contributing to the MaCFP group. They estimated the reaction kinetic parameters for the same black PMMA that was used to gain the experimental data utilised here with different methods and experiments. With the estimated reaction kinetic parameters, TGA experiments with heating rates of 10 K/min and 100 K/min are modelled and the peak mass loss rate and the corresponding temperature is compared [53,54].

3. Results

3.1. Prediction of number of components

A dataset $\mathbb{D}_{1,2,3,test}$ of $1.5 \cdot 10^6$ samples in total, $0.5 \cdot 10^6$ samples each for one, two and three components, was used to evaluate the performance of the classifier to predict the number of components from a given mass loss rate over temperature curve. Table 7 shows the accuracy, precision and recall in total and for each subset. With 92.28%, 93.04% and 92.28%, accuracy, precision and recall are each above 90%. Fig. 5 shows a confusion matrix. It gives the ratio between the number of predicted values $\hat{y}^u$ and true values y^u for each category and is normalised over the true values y^u. It can be seen that materials with one component are classified correctly with 100% probability, as well as materials with two components are classified correctly with 97% probability. Materials with three components are classified correctly with 80% probability, while 20% of reactions with three components are misclassified as reactions with two components. 5-fold cross validation leads to an average accuracy of 92.28% (standard deviation (SD): $1.91 \cdot 10^{-4}$).

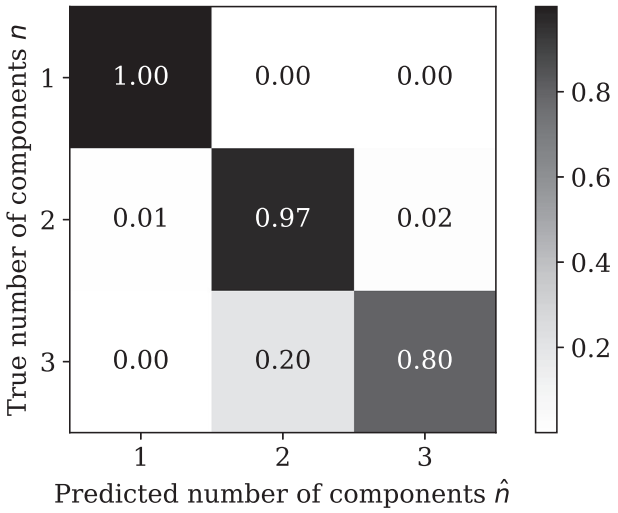


Fig. 5. Confusion matrix with relative fraction of prediction for first submodel.

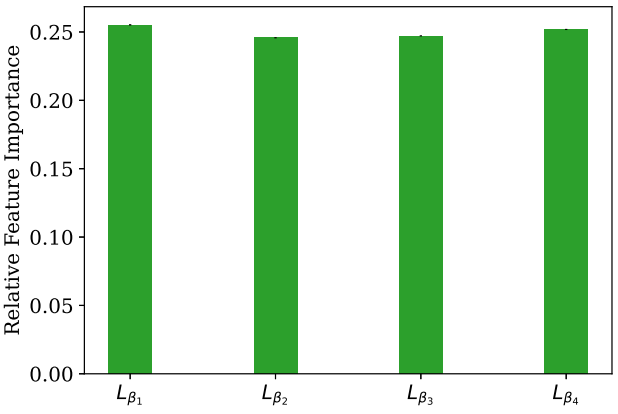


Fig. 6. Feature importance for input x of first submodel.

Fig. 6 shows the relative impurity-based input feature importance in the first submodel as a measure for how important a feature is to the prediction. The bars give the summed feature importance for $L_{\beta_1}, \ldots, L_{\beta_4}$ with error bars showing the standard deviation of the single features.

3.2. Prediction of fraction of components

Fig. 7 shows histograms for both of the two components fractions prediction submodels, i.e. for two and three components. On the x-axis is the NRMSE between y^u and $\hat{y}^u$ of the respective dataset $\mathbb{D}_{n,gauss}$, with n, number of components, separated in 100 bins, equally distributed on a log scale. The test sets $\mathbb{D}_{n,test}$ have each $0.5 \cdot 10^6$ elements. On the y-axis there is the relative frequency at which these values occur. The green bars display the relative frequency of each bin and the blue line shows the cumulative distribution.

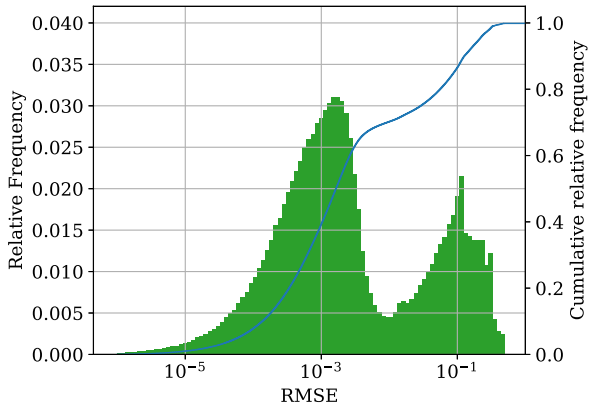
In Fig. 7(a), prediction of the fraction of two components, 70% of all predictions are in the bins with a RMSE below 0.0032 and 95% have a RMSE below 0.18. The R^2 score between y and $\hat{y}$ is 0.75.

In Fig. 7(b), prediction of the fraction of three components, 70% are in the bins with a RMSE below 0.034 and 95% have a RMSE below 0.18. Here, the R^2 score equals to 0.63.

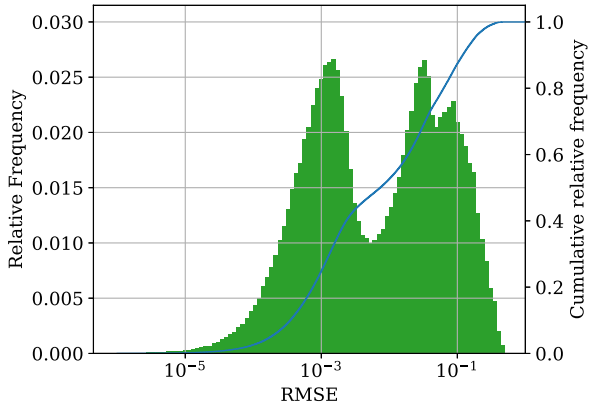
3.3. Prediction of reaction kinetics

Figs. 8 to 10 show histograms for each of the three reaction kinetics prediction submodels, i.e. for one, two and three components. On the log scale x-axis is the NRMSE between x^u and $\hat{x}^u$ of the respective dataset $\mathbb{D}_{n,test}$ with n, number of components, separated in 100 equally distributed bins. The test sets $\mathbb{D}_{n,test}$ have each $1.5 \cdot 10^6$ elements. On the y-axis there is the relative frequency at which these values occur. The green bars display the relative frequency of each bin and the blue line shows the cumulative distribution.

Table 8 gives the 70th and 95th percentiles for NRMSE$(x, \hat{x})$ for the three reaction kinetic submodels. Table 9 gives the R^2 scores and the results from 5-fold cross validation.



(a) Materials with two components.



(b) Materials with three components.

Fig. 7. Histograms of RMSE between y and $\hat{y}$ for materials with (**a**) two components and (**b**) three components. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 8
70th and 95th percentiles for NRMSE$(x, \hat{x})$ for the three reaction kinetic submodels.

Submodel	NRMSE 70th percentile	NRMSE 95th percentile
1 Component	$2.5 \cdot 10^{-6}$	$0.12 \cdot 10^{-3}$
2 Components	0.0965	0.163
3 Components	0.132	0.206

Table 9
$R^2(y, \hat{y})$ and cross validation results for the three reaction kinetic submodels.

Submodel	R^2 score	R^2 mean $\pm$ σ 5-fold Cross Validation
1 Component	1.0000	1.0000(0)
2 Components	0.9392	0.9408(2)
3 Components	0.5560	0.5588(3)

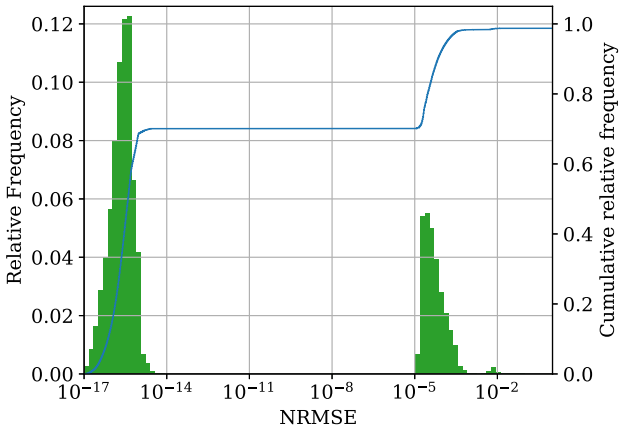


Fig. 8. Histogram of fitness between x and $\hat{x}$ for materials with one component, third submodel. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

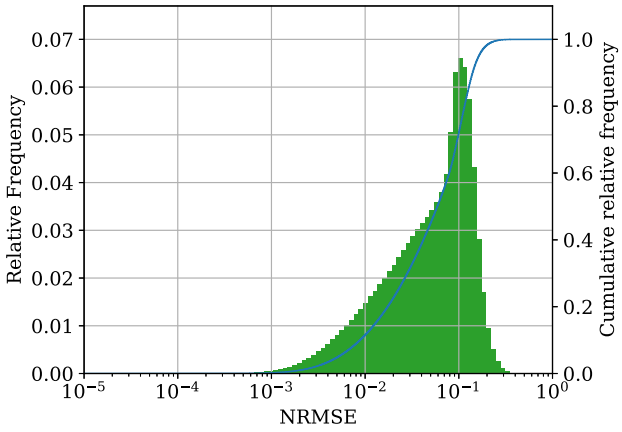


Fig. 9. Histogram of NRMSE fitness between x and $\hat{x}$ for materials with two components, third submodel. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

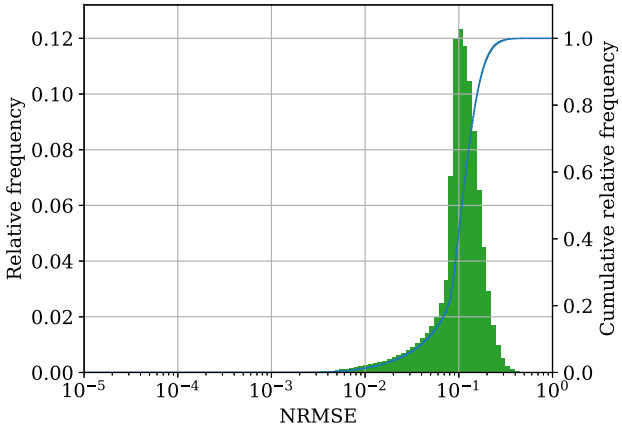
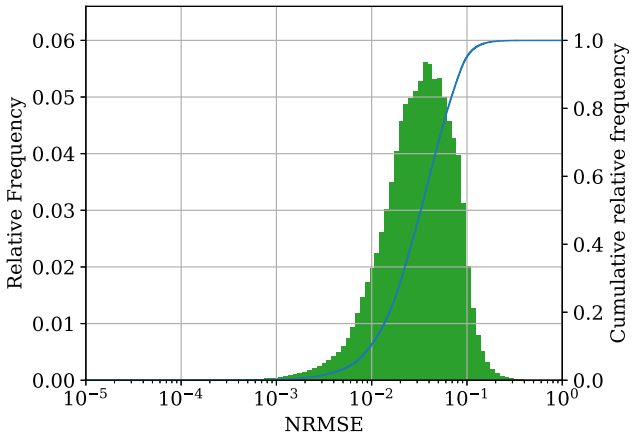
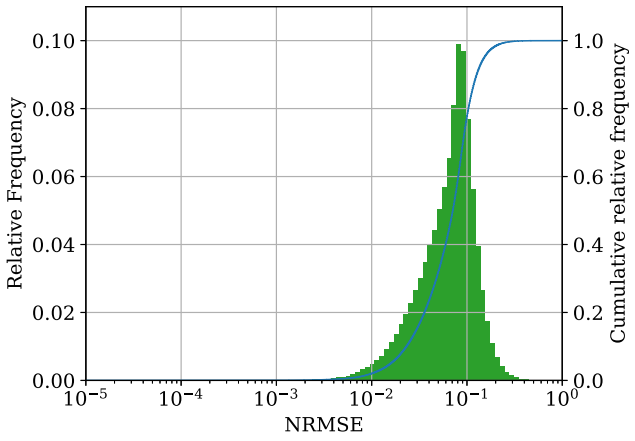


Fig. 10. Histogram of fitness between x and $\hat{x}$ for materials with three components, third submodel. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Results for the prediction of reaction kinetic parameters for two and three components came to two oppositional outcomes at different hyperparameter settings. A model consisting of only a single estimator, which reduces the model to a simple decision tree model, showed not the behaviour expected from literature studies. Here, NRMSE($x^u, \hat{x}^u$) was lower than in any other configuration, The R^2 score for this configuration was the poorest of any model, which was expected. Increasing the number of estimators in the models leads to improvement of the R^2 score, but slight disimprovement of the NRMSE($x^u, \hat{x}^u$). Consequently, both results are presented in Fig. 11 and discussed briefly in Section 4. The R^2 score for y^u and $\hat{y}^u$ in the submodel with only one estimator is 0.7682 with a 5-fold cross validation average of 0.7735 (SD: $5.24 \cdot 10^{-3}$) for two components and −0.0672 with a 5-fold cross validation average of −0.0624 (SD: $2.09 \cdot 10^{-3}$) for three components.



(a) Materials with two components, 1 estimator.



(b) Materials with three components, 1 estimators.

Fig. 11. Histogram of NRMSE fitness between x and $\hat{x}$ for materials with two and three components, third submodel.

Fig. 12 shows the relative input feature importance in all submodels. The bars give the summed feature importance for $L_{\beta_1}, \ldots, L_{\beta_4}$ with error bars showing the standard deviation of the single features and the values for $Y_1, \ldots, Y_3$.

3.4. Complete model prediction

The results from the prediction for the complete model as shown in Fig. 3 are presented in this section. A dataset $\mathbb{D}_{1,2,3,eval}$, consisting of $1.5 \cdot 10^5$ elements was used to evaluate the complete model. It is a combination of three times $0.5 \cdot 10^5$ samples, each for materials consisting of one, two and three components. The overall $R^2(y, \hat{y})$ for the test dataset is 0.77. A breakdown of this overall R^2 score is shown in Table 10. The table displays individual R^2 scores for each number of components n and each group of parameters A_i, E_i and Y_i.

Figs. 13 to 15 show a histogram of the true y vs. predicted $\hat{y}$ value for each of the nine parameters. Here, anything on the diagonal from

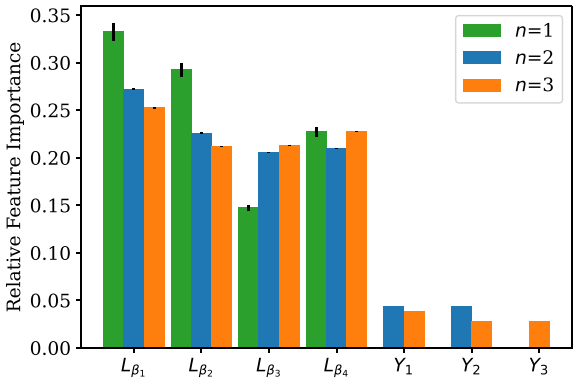


Fig. 12. Feature importance for input x of third submodel.

Table 10
R^2 scores for individual parameters and number of components.

Parameter	$n = 1$	$n = 2$	$n = 3$	total
A_i	1	0.90	0.24	
E_i	1	0.91	0.31	
Y_i	–	0.77	0.56	
total	1	0.86	0.37	0.77

the origin is an ideal prediction. For bins above the diagonal line, predictions are higher than the true values and for bins below the diagonal line, predictions are lower than the true values. Values of each dataset are sorted by descending material's component fraction. This means, that $Y_1 \geq Y_2 \geq Y_3$. Bins with a count of less than five are omitted for visual clarity. Fig. 13 shows a histogram of the values for A_i and $\hat{A}_i$. For first and second component, most of the predictions are on the diagonal, while a similar, but less distinct result can be observed for the third component. The same accounts for the true values E_i vs. the predicted values $\hat{E}_i$, displayed in Fig. 14.

Fig. 15 shows a histogram for the true Y_i vs. the predicted $\hat{Y}_i$. Here, for all three cases, there is a strong diagonal line visible. There also occurs scattering throughout the parameter space. Note, that here in one third of all cases Y_1 equals to 1, since one third of the cases have only one component. This might distort the optical impression of the distribution, since there is only one yellow dot in the first plot at $(1, 1)$.

Fig. 16 shows a histogram of the NRMSE between the true x and the predicted $\hat{x}$ calculated from the predicted $\hat{y}$ for the complete prediction model. On the x-axis is the NRMSE between x^u and $\hat{x}^u$ of the dataset $\mathbb{D}_{1,2,3,eval}$, separated in 100 equally distributed bins. The test set $\mathbb{D}_{1,2,3,eval}$ has 150000 elements. On the y-axis there is the relative frequency at that these fitness values occur. The orange, green and red bars display the relative frequency of each bin, broken up into the number of components in the true values. The blue line shows the cumulative distribution. 50% of the results have a NRMSE below 0.0429, while 95% have a NRMSE of below 0.147.

For a better visualisation of the implications of the NRMSE, Fig. 17 shows the mass loss rates for the 5th, 50th and 95th percentile of the distribution shown in Fig. 16. The true value x for all four heating rates is shown as a solid line while the predicted value $\hat{x}$ is shown as a dashed line. The rows are broken down to the number of components. For $n = 1$, there is no difference between the true and predicted values. For $n = 2$, differences between the true and the predicted values can be observed for the 50th and 95th percentile. For $n = 3$, differences for all percentiles can be observed.

3.5. Hyperparameter tuning

Hyperparameter tuning was conducted for the first and the third submodel. The second submodel does not have any hyperparameters. For the number of components prediction submodel, results are presented in Fig. 18 and results for the reaction kinetics prediction submodel in Fig. 19.

Fig. 18 shows the influence of the hyperparameters dataset size, maximum tree depth and number of trees on the accuracy of the number of components prediction submodel determined with the test dataset. It shows that the dataset size and the number of trees in the model have more pronounced influence in the considered intervals. A four fold increase of dataset size from 1.5 million sets to 6 million sets increases the accuracy by 0.026 from 0.897 to 0.923. The increase of the maximum tree depth continuously increases the accuracy value up to a maximum tree depth of 100, where the maximum tree depth converges to the maximum accuracy. A hundredfold increase in the number of trees increases accuracy by 0.008.

Fig. 19 shows the influence of the hyperparameters dataset size, maximum tree depth and number of trees on the R^2 score of the kinetic parameter prediction submodels for two and three components, determined with the test dataset. A hyperparameter study for the kinetic parameter prediction of one component was not conducted, because a single decision tree already reached a R^2 score of 0.9999999. The R^2 score increases from 0.70 to 0.94 for the two components prediction and from 0.36 to 0.56 for the three components prediction with increasing dataset size from 0.1 million sets to 6 million sets. For a maximum tree depth, R^2 rises from 0.07 to 0.94 and from 0.04 to 0.56 respectively, with maximum tree depth values from 1 to 50. The R^2 for changing the number of trees from 1 to 200 changes from 0.07 to 0.94 and from −0.07 to 0.56 for the respective model.

Table 11
Values from model prediction for PMMA.

Parameter	Component 1	Component 2
$\ln(A_i)/\mathrm{s}^{-1}$	57.958	29.478
$E_i/(\mathrm{J\,mol}^{-1})$	250.527×10^3	176.876×10^3
Y_i	0.032	0.968

3.6. Real case application

Based on the experimental data from MaCFP, the model predicted that the PMMA decomposition can be described with two components, each undergoing a first order reaction. The related components fractions and reaction kinetic parameters are displayed in Table 11. These parameters are used to model TGA experiments and compare the modelled mass loss rates to the original experimental data. The mass loss rates for experimental and predicted values for the four different heating rates is shown in Fig. 20. A NRMSE between experimental and modelled mass loss rate data of 0.0468 is calculated.

In another comparison, peak mass loss rate values and the temperatures at that they occur from predictions of the presented inverse model are compared to the mean of predictions made by contributors of the MaCFP working group. The mean values are calculated from 16 contributions. In Table 12, the peak mass loss rate values and the temperatures at that they occur are displayed for modelled TGA experiments with heating rates of 10 K/min and 100 K/min. [53,54]

4. Discussion

4.1. Submodels

It was shown in Fig. 5 that the prediction of the number of components works exceptionally well for materials consisting of one and two components. These were classified correctly with only minor noise. While only 80% of the materials consisting of three materials where classified correctly as materials with three components, it is likely that a majority of the 20% of the materials misclassified as materials with two components can reasonably well be modelled as materials with two components due to exactly overlapping peaks in reaction rates [30].

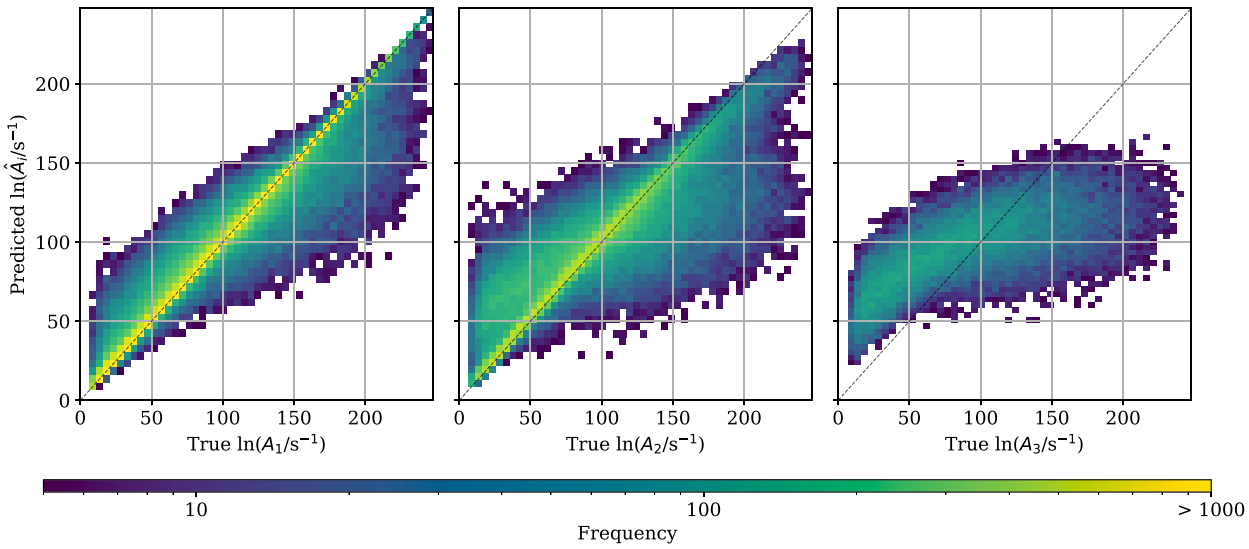


Fig. 13. Model results: predicted vs. true values for A_i.

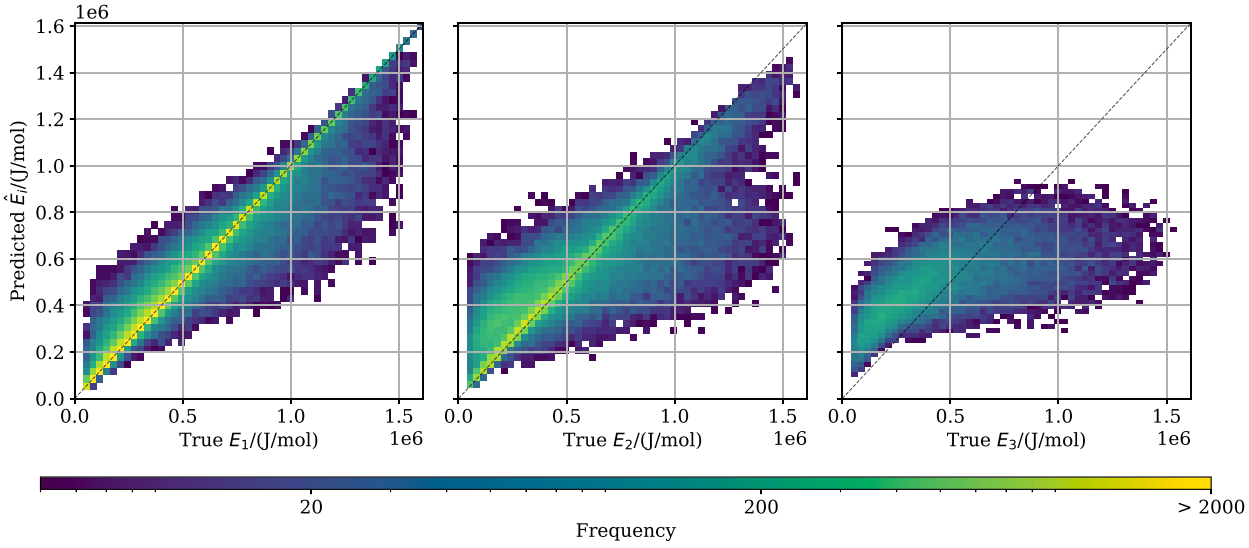


Fig. 14. Model results: predicted vs. true values for E_i.

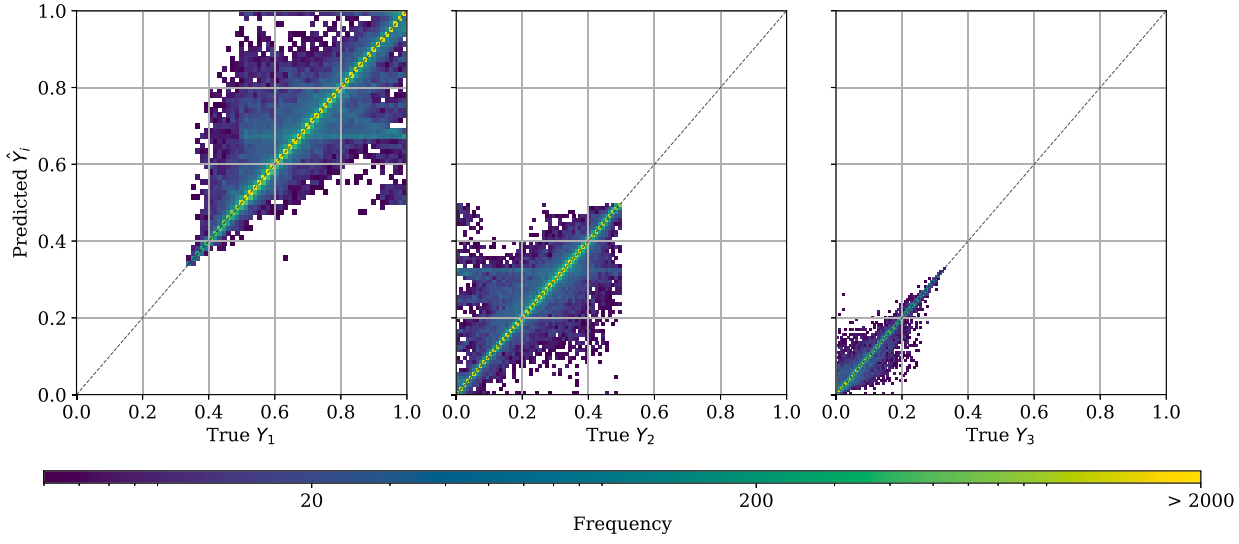


Fig. 15. Model results: predicted vs. true values for Y_i. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 12
Peak value comparison for mass loss rates from TGA experiment and from predicted reaction kinetic parameters.

Parameter	Heating rate	Model prediction	MaCFP Mean $\pm 2\sigma$
$r_{peak}/(\mathrm{mg\,s^{-1}})$	10 K/min	0.0172	0.0152±0.0030
T_{peak}/K	10 K/min	622.6	635.6±15.8
$r_{peak}/(\mathrm{mg\,s^{-1}})$	100 K/min	0.152	0.130±0.038
T_{peak}/K	100 K/min	664.8	673.5±40.8

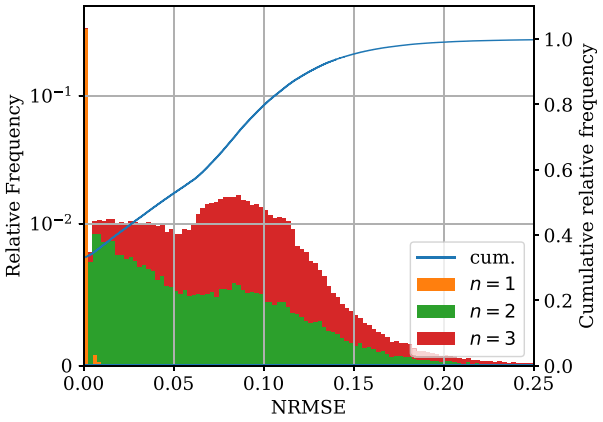


Fig. 16. Histogram of fitness between x and $\hat{x}$, complete model. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

The importance of the input features, here $L_{\beta_1}, \ldots, L_{\beta_4}$ is uniformly distributed, as shown in Fig. 6. It can be summarised that this submodel works well for the intended application.

The prediction of the fractions of the components with a Gaussian fitting model can be evaluated with the histograms in Figs. 7(a) and 7(b). In both figures, a distribution with two peaks can be observed. In both models are groups with a fit at around 10^{-3} covering about 70% (two components) and 50% (three components) of the results. A second peak at around 10^{-1} covers about 30% (two components) and 50% (three components) of the results. Resulting R^2 scores of 0.75 and 0.63 for two and three components lead to a moderate but still usable result as can be seen in the overall model.

The submodel predicting reaction kinetic parameters for materials consisting of only one component works ideal. In Fig. 8, two peaks for the NRMSE between x and $\hat{x}$ can be observed. One peak covers roughly 70% of the results lie below 10^{-16}, which implies a perfect fit with noise introduced through the precision of the used datatype (double-precision floating-point). The other peak, covering roughly 30% of the results, shows actual divergence, but this divergence is small that is still negligible. For two components, the prediction is not as good as for one component, but is still reaching a R^2 score of 0.94 for the 200 estimator model comparing y and $\hat{y}$. Fig. 9 confirms this outcome for a comparison between x and $\hat{x}$. The model for three components only reaches a R^2 score of 0.56 for y and $\hat{y}$ in the 200 estimator model. Fig. 10 illustrates that the fit between x and $\hat{x}$ is not as good as in the models for one and two components. The feature importance for all three models as given in Fig. 12 shows that the importance of $L_{\beta_1}, \ldots, L_{\beta_4}$ for predicting reaction kinetics of materials consisting of two and three components is close to uniform distribution with slight accentuation on L_{β_1}. The submodel for only one component shows significant non-uniform distribution, that is a result of only a single decision tree used in this model. This also explains the higher standard deviation shown. The summed relative feature importance of the estimated fractions Y_n is 0.087 for two components and 0.095 for three components.

The contradictory results for the two and three components submodels shown in Fig. 11, where a single decision tree only seems to reach better results comparing x and $\hat{x}$ but way worse results comparing y and $\hat{y}$ could be explained by the kinetic compensation effect where different values for kinetic parameters lead to similar reaction rate profile [55,56].

Almost all submodels yielded a standard deviation below 10^{-3} in the 5-fold cross validation for R^2 scores between y and $\hat{y}$. Exceptions are the models predicting reaction kinetics for two and three components with only one estimator, where the standard deviation was below $2.3 \cdot 10^{-3}$. This shows that the partition into training and testing datasets did not have any significant influence on the model performance.

4.2. Complete model

The overall model reaches a R^2 score of 0.77. Score breakdown in Table 10 shows that the prediction for materials consisting of one component works almost ideal. The prediction of the actual kinetic parameters for materials consisting of two components works also well, but less than in the evaluation of the single submodel due to inaccurate prediction of the yields of components. The same pattern can be seen for predicting kinetic parameters for materials consisting of three components. Here, it shows an even higher impact compared to the particular submodel due to the influence of misclassification of the number of components and inaccurate prediction of the yields of components.

These results are also visualised in Figs. 13 to 15, where the true vs. predicted values are plotted as a 2d histogram. In all plots except for the third components A_3 and E_3, the highest density is on the diagonal line, which represents an ideal prediction.

In the histogram shown in Fig. 16, comparing the fit between x and $\hat{x}$, it can be seen that more than one third of the calculations have a fit that counts in the lowest bin. This mostly is an effect of the materials consisting of only one component. The results for materials with two and three components are represented by the rest of this distribution. Evaluating the distribution of results further as shown in Fig. 17 confirms this. Predictions for materials consisting of one component almost always fit perfectly to the true values. Predictions for materials consisting of two components are distributed from an almost perfect match to results that reproduce the general mass loss profile but are slightly shifted in peak temperatures or peak mass loss rates. This is even more true for materials consisting of three components. Here, the results are also distributed from an almost perfect fit to deviations in position and height of peaks.

In this model, several submodels are linked in sequence and in parallel. Since different error metrics are applied, the error propagation is not calculated directly. To assess the error propagation, the single submodel performance and error metrics can be compared to the performance and error metrics of the complete model. Also, the error propagation can be traced semi-quantitatively. The wrong classifications from the first submodel propagate to the second submodel directly. This can be quantified by the precision shown in Table 10, calculating the true positives of one class to all classifications into this class. Then, the error propagates further from the second submodel to the third submodel. The estimations from the second submodel represent only 2 of 1066 input features (two components) respectively 3 of 1067 input features (three components) of the third submodel. To quantify the influence on the prediction of the third submodel, the relative feature importance can be evaluated, which is 0.087 for two components and 0.095 for three components.

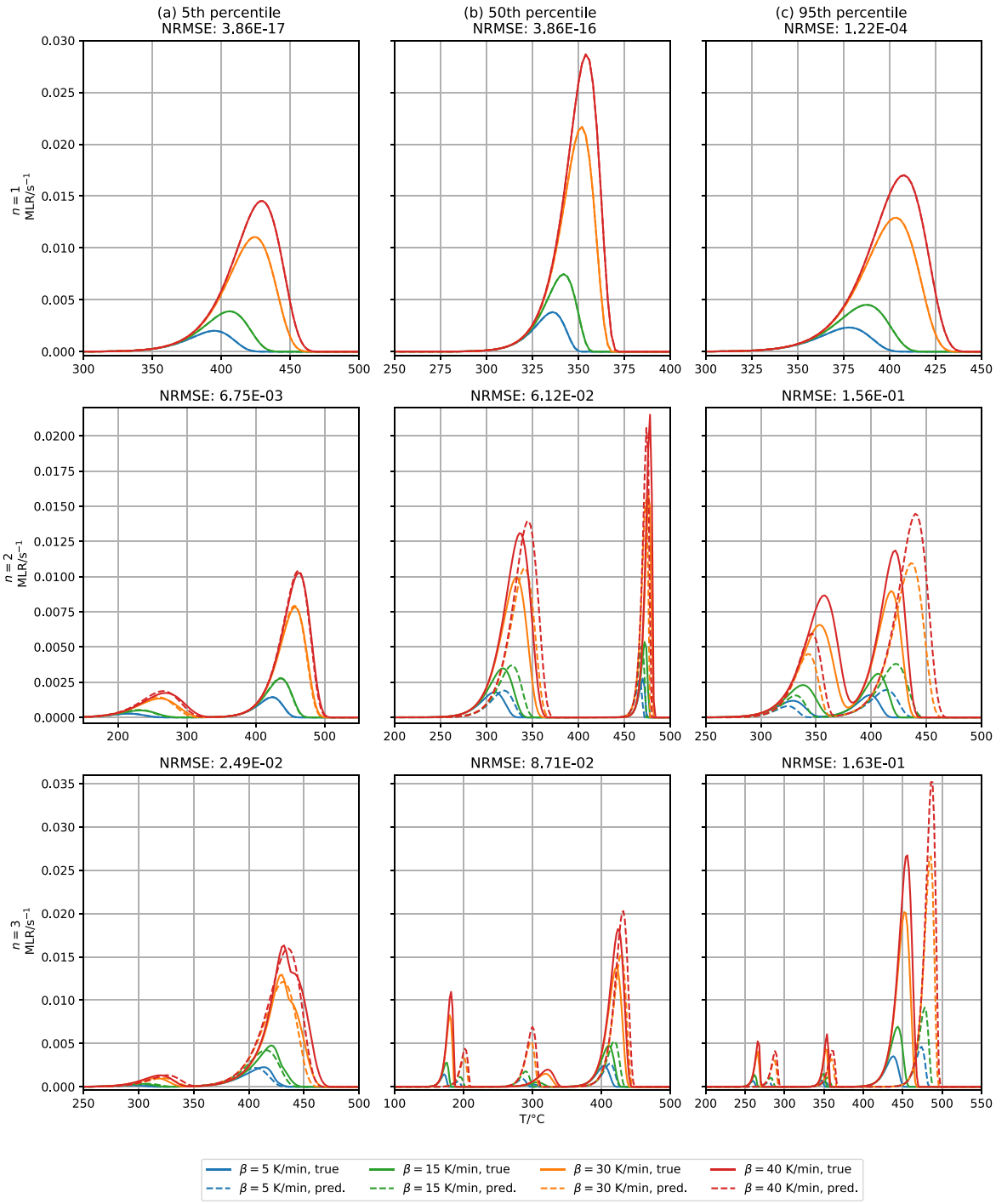


Fig. 17. Mass loss rates of (a) 5th, (b) 50th and (c) 95th NRMSE percentile for $n = 1$, $n = 2$ and $n = 3$.

Since the presented approach utilises a huge dataset of synthetically experiments and other approaches mainly focus on a limited set of real or synthetically produced experiments, a general quantitative comparison to other approaches for the complete inverse surrogate model is not feasible. Though, what can be made is an evaluation of individual results, which is done in Section 4.4 and qualitative comparison to other approaches. The presented approach is capable of qualitatively reaching at least about the same level of accuracy comparing the experimental and predicted TGA mass loss rate curves to selected other methods in several cases [57–59]. Although, in certain cases, especially if more components are involved, the model fails to make precise predictions.

The main advantage of this approach over other methods is the combination of getting results instantly for TGA experiments with multiple heating rates on materials that can be described as consisting of up to three components. Other methods are missing at least one of these three properties. They are either computational expensive, work only on well separated pathways or can only make use of TGA experiments with a single heating rate at a time. A disadvantage of the presented approach is, in some cases the results are either slightly deviated in position and height of peaks or do not fit at all. But due to using an inverse pretrained surrogate model, the computational costs are way less expensive than any other method that combines the advantages of this approach and it can be implemented as a first step in estimating

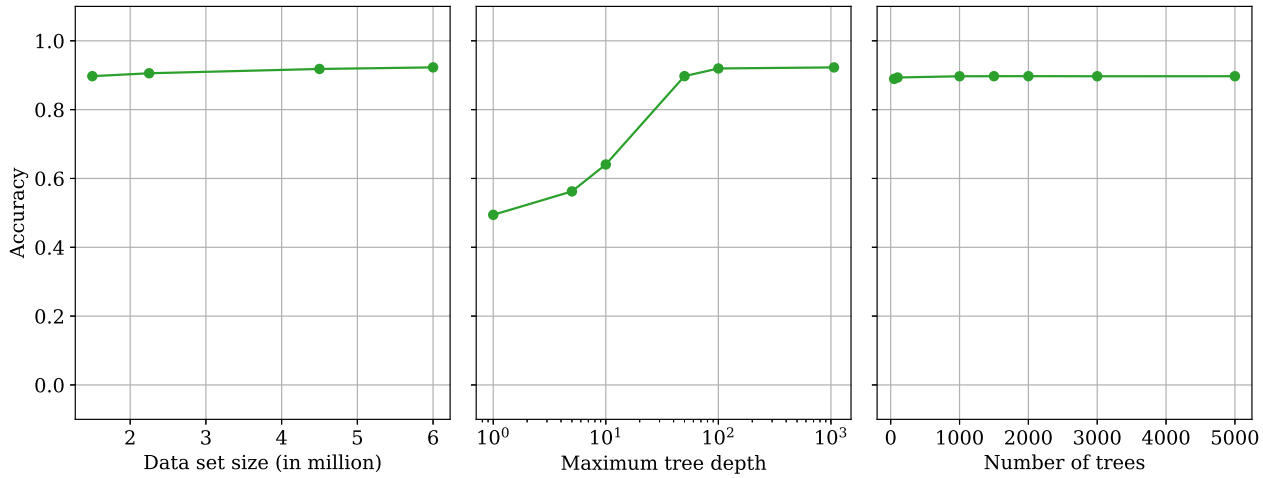


Fig. 18. Hyperparameters for number of components prediction submodel.

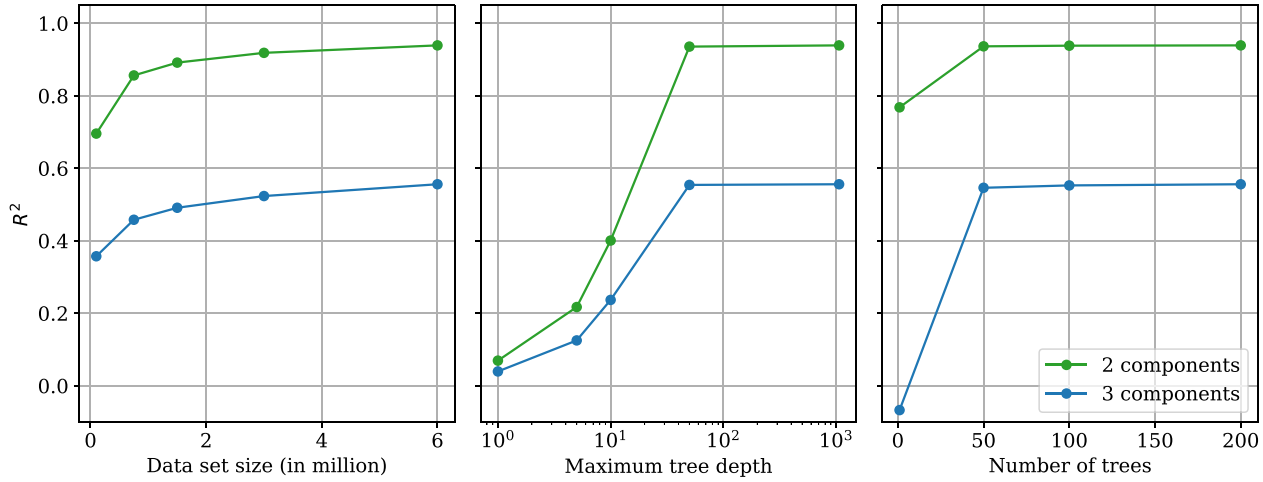


Fig. 19. Hyperparameters for kinetic parameter prediction submodel.

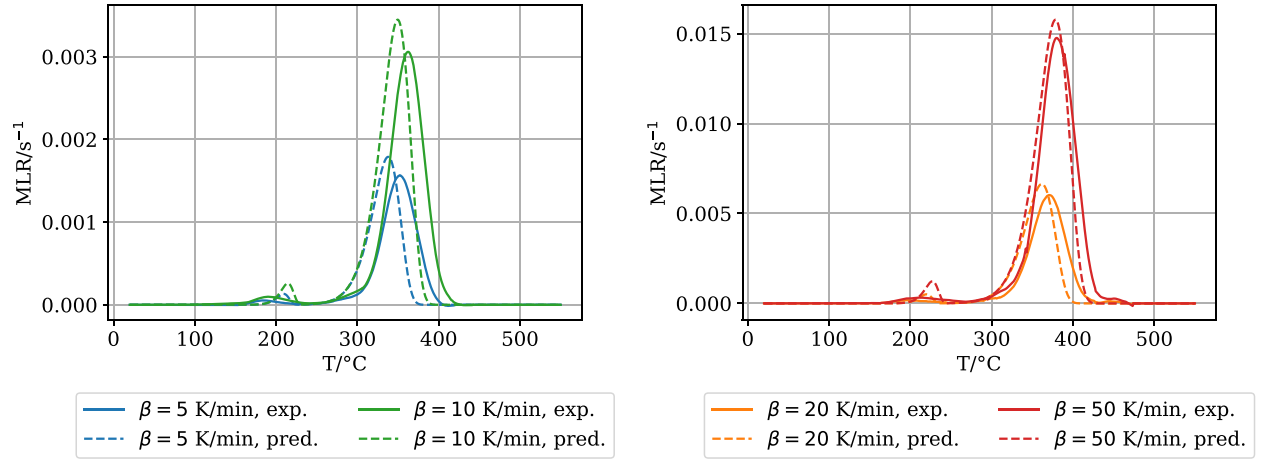


Fig. 20. Mass loss rates from TGA experiment and from predicted reaction kinetic parameters.

reaction kinetic parameters. If the initial estimation of this approach is already good, further investigation with e.g. an inverse metaheuristic optimisation approach can be waived. Else, if the results slightly deviate from the experiment, these reaction kinetic parameters can be used as an initial guess for the inverse metaheuristic optimisation approach and could reduce computational costs compared to a random initial guess.

A common counter argument against using inverse modelling methods in pyrolysis modelling in general and especially applying such machine learning methods, is that the inverse surrogate model cannot be verified as it is acting as a black box and the model results cannot be validated. Using this ERT approach, once a model is trained, it acts deterministic, i.e. a given input value will always produce the same

output. Even the pathways along the decision trees can be traced to comprehend why a given input leads to a certain output. This is a step towards explainable machine learning models [60]. To validate the output of the surrogate model, i.e. the reaction kinetic parameters $\hat{y}$, they can be used to simulate the exact same TGA experiments that were used as the input for the inverse surrogate model with a direct pyrolysis model of choice. Calculating the mass loss rate in the simulation, the direct model output $\hat{x}$ and the experimental values x can be compared. Furthermore, with additional TGA experiments for different heating rates other than the ones used in training the surrogate model, the extrapolation capacities of the reaction kinetic parameters can be evaluated. Though, this means, a set of reaction kinetic parameters are always effective model parameters only validated to the specific combination of experimental setup and the direct pyrolysis model used for validation. But this is true for any other approach as well.

4.3. Hyperparameter tuning

The hyperparameter study showed for all evaluated set ups and all investigated hyperparameters a positive correlation between the investigated parameter and the resulting R^2 score as shown in Figs. 18 and 19. For all cases with parameter values greater than 1, the R^2 score showed an improvement with increasing parameter values but the improvement slowed down with increasing parameter values.

The fraction of components prediction model reacted most sensitive on maximum tree depth in the considered boundaries. Only from a maximum tree depth of 100 it started to converge and only slightly increased by 0.003 in R^2 score at a maximum depth level of 200. Dataset size and number of trees did not had a huge influence and only slightly improved by increasing size from values of 50 and above.

The kinetic parameter prediction model also showed to be most sensitive to the maximum tree depth number, increasing from 1 to 50 resulting in a R^2 score rise from 0.07 to 0.94 for the two component model and from 0.04 to 0.56 for the three component model with no significant rise afterwards. From a dataset size of 1.5 million, a slow rise for both submodels was observed. While values for just a single estimator were low, only slight improvement was reached raising the number of estimators from 50 to 200.

4.4. Real case application

In a real case application, the presented inverse surrogate model was evaluated with TGA experiment data for black PMMA. The mass loss rates calculated from the predicted parameters show good agreement with the experimental data with a NRMSE of 0.0468. The peak mass loss rates of the prediction are slightly higher and occur at slightly lower temperatures than in the real TGA experiments (Fig. 20). Compared to predictions other groups found with different methods on the same material, the predictions of the presented model for peak mass loss rates and temperatures, at that they occur, lie well within the mean $\pm 2\sigma$ range of results other groups found (Table 12).

5. Conclusions

In this contribution, a novel method to predict reaction kinetic parameters from TGA experiments was presented and evaluated. An inverse surrogate model utilising ensemble learning techniques as a novel method is used for prediction. The model is based on an extremely randomized trees algorithm and a Gaussian curve fitting algorithm. It consists of three submodels. The first submodel predicts the number of components the material consists of, the second submodel predicts the fraction of each component and the third submodel predicts the actual kinetic parameters for each component.

Evaluation of the model demonstrated that it can predict reaction kinetic parameters from TGA experiments with an overall R^2 score of 0.77. While it performed best on materials consisting of one component, it still performed good on materials with two components and least good on materials with three components. The submodel predicting the number of components in the examined material shows strongest results. The prediction of the fractions of the single components has appeared as the least accurate of the model, compared to the other parameters that the model predicted.

An example real case application of TGA experiment data of PMMA showed that predictions both fit to the experimental data and are in good agreement with predictions other groups made for the same material.

The inverse surrogate model is able to predict reaction kinetic parameters that can directly be used for further modelling in most of the cases. In cases where the prediction is not sufficient for direct further use, the values can be used as a starting point for different conventional strategies like metaheuristic optimisation.

Leading to almost instant results, even if they might not be exact, is one of the biggest benefits of this inverse surrogate model. The combination of this feature with the ability to use TGA experiments with four different heating rates and the capability of predicting reaction kinetic parameters of up to three components makes this approach unique.

While this presented inverse surrogate model has certain limitations in model complexity, the general concept is capable of considering higher model complexity. The presented approach is limited to a maximum of three parallel first-order reactions. This was chosen to not exceed common computational resources. The training data can be extended for more complex reaction networks with any number of nth order reactions and residue resulting from charring and then can be used to build an inverse surrogate model capable of more complex reactions. The presented inverse surrogate model already can handle materials that are producing residual mass fractions, when the experimental input data is preprocessed with a normalisation step that takes the residual mass fraction into account.

Since the code for this approach is open source and modular, parts or the whole method can easily integrated into other approaches.

With the ability to validate the results by using the estimated parameters to predict the experimental values and the limitations highlighted before, the model in the current state best would be used complementary to other methods. Further refinement and reinforcement of the model is needed, before it can be used as a main method to predict reaction kinetic parameters. Most efficiently, the second submodel should be revised, since it has the weakest performance compared to the other submodels. The whole model will improve if the error of the second submodel decreases and subsequently reduces the error of the total model, especially through error propagation. Automated deconvolution analysis would be a good starting point [17]. Increasing the training data set size in combination with limiting the boundaries closer to common values seems also a promising way to achieve better performance in all of the three submodels. The next step would be a statistical analysis to evaluate if there are specific nodes in the random forests that are a common reason for failed predictions.

Future work will include studies to better justify and quantify the selection of the used distribution and boundary values for dataset generation and to understand the influence of the kinetic compensation effect on the model. Another point will be the improvement of the submodel that predicts the fractions of the components. Different machine learning algorithms, e.g. neural networks, for the first and third submodel, will be evaluated as well. Due to the modular conception of the proposed method, this is possible with a reasonable amount of time. The influence of using synthetic and real experimental data gained through microscale combustion calorimetry (MCC) [61] needs to be investigated. While TGA data gives temperature resolved mass information, MCC data gives temperature resolved heat of combustion information. The results of the inverse surrogate model would benefit from using training data gained through MCC, either as sole input or combined with TGA data. The general structure of the surrogate model allows such investigation.

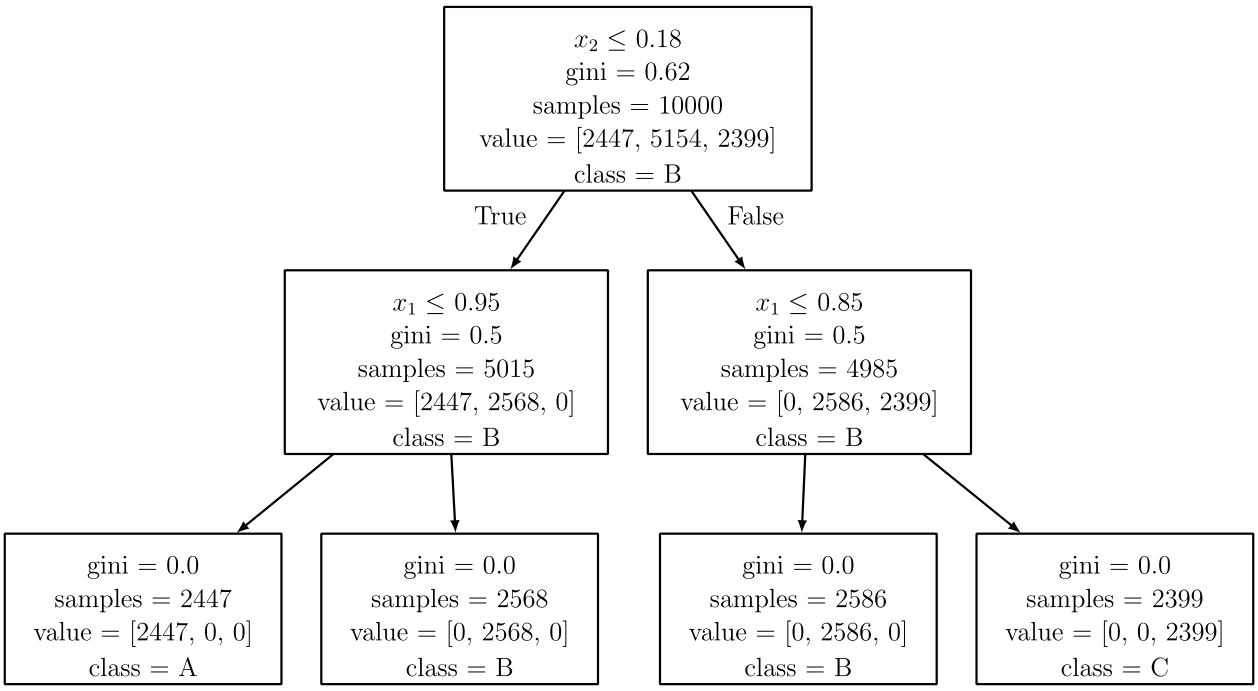


Fig. 21. Random example for a decision tree used in ERT algorithm.

CRediT authorship contribution statement

Patrick Lauer: Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Data curation, Writing – original draft, Writing – review & editing, Visualization. **Lukas Arnold:** Resources, Writing – review & editing. **Fabian Brännström:** Writing – review & editing, Supervision.

Declaration of competing interest

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

Data availability

Code [35] and dataset [36] used in this publication as well as a pre built reference model [37] are available on zenodo.org.

Acknowledgements

The authors gratefully acknowledge the computing time granted (project: cjjsc27; application number: 19595) by the JARA-HPC Vergabegremium and VSR commission on the supercomputer JURECA at Forschungszentrum Jülich [62,63]. Part of the computation was conducted on the extension of the computing cluster Pleiades, funded by the German Federal Ministry of Education and Research, Germany within the CoBra project (funding project number 13N15497). We acknowledge support from Frederick Sabath, who did helpful preparative programming work in his bachelors thesis.

Appendix. Extremely randomized trees

Extremely randomized trees [26] (ERT) is an ensemble learning method derived from the random forests method [38] and follows a supervised learning approach. It can be used for classification and regression tasks. Ensemble learning methods rely on the idea of building a prediction model as a combination of a collection of more basic models [64].

Random forests act on this principle by building a plethora of decision trees by a specified set of rules from a training data set which contains the input as well as the expected output. With the trained prediction model, each of the decision trees transforms an input to an output prediction. The final result is determined by a majority vote out of the aggregated individual decision tree results. In regression tasks, the arithmetic average is used [65]. With averaging the results, a random forest can combine the advantages of a single decision tree (universal approximation, robustness to outliers, linear computational complexity, robustness to irrelevant and redundant inputs and good interpretability) with a reduction of its main disadvantage, high variance [51]. A prediction model built with ERT works the same in terms of how a final result from several decision trees is achieved, but differs on how these single decision trees are constructed. This leads to a significant improvement in precision and to a reduced computational complexity [51]. To comprehend these techniques that are used by ERT to build a decision tree, lets imagine a trivial example. A set of two input values (features) $\{x_1, x_2\}$ shall be classified in one of the three classes A, B and C. x_1 and x_2 are randomly assigned with values of either 0 or 1. The output classes represent the sum of x_1 and x_2. A is for $x_1 + x_2 = 0$, B is for $x_1 + x_2 = 1$ and C for $x_1 + x_2 = 2$.

For training an ERT model, 10000 random sets of $\{x_1, x_2\}$ were generated and the output class was calculated. Fig. 21 shows an exemplary random decision tree for this classification task. This whole exemplary ERT model consists of 100 of these decision trees. The decision making is started from the top and yields a result at the bottom. Both of the two upper levels represents a decision taken upon one of the values of the features. Starting in the first node with the decision, if x_2 of the given input is smaller or equal to 0.18. If this is true, the left path is taken, else the right path. The "gini" impurity is a criterion for the quality of the split and ideally should be minimal. It is the probability of an incorrect classification of a random sample. "Samples" gives the number of samples of the training set were considered in this node and "value" gives the number of how many of the samples are either in class A, B or C. If the decision in this node is true, the decision in the next node is based on $x_1 \leq 0.95$. The bottom level gives the prediction of the

model, where each letter represents one class. The final output of the complete model is determined by a majority vote out of the aggregated individual decision tree results.

In context of prediction of the number of components in this contribution, the input x is the mass loss rate and the output classes are the number of components (either 1, 2 or 3) predicted for an input.

The technique to build this decision tree that is particular to ERT is the random split generation. For each node a set of randomly chosen features with random splitting points are evaluated and the one with the best result, i.e. at this classification task the lowest gini impurity, is used for this node. The size of the set can be adjusted as anything from one (completely random tree) to the number of features. Suggestion by the authors of the original paper is for classification models the square root of the number of features and for regression models the number of features [26].

Two further important parameters are the number of decision trees a forest consists of and the number of minimum samples in a node. The latter correlates to the maximum depth of a tree, which was considered as hyperparameter to investigate in this contribution. The optimal values for these two parameters are problem specific and can be determined by hyperparameter studies.

References

[1] F. Vermina Lundström, P. van Hees, É. Guillaume, A review on prediction models for full-scale fire behaviour of building products, Fire Mater. 41 (3) (2017) 225–244.

[2] C. Wade, D. Hopkin, J. Su, M. Spearpoint, C. Fleischmann, Enclosure fire model for mass timber construction—Benchmarking with a kinetic wood pyrolysis sub model, in: Interflam: 15th International Conference on Fire Safety Engineering, 2019, pp. 1–3.

[3] T. Hehnen, L. Arnold, P. van Hees, S. La Mendola, Simulation of fire propagation in cable tray installations for particle accelerator facility tunnels, in: Proceedings from the 8th International Symposium on Tunnel Safety and Security, Borås (Sweden), 14 Mar 2018 - 16 Mar 2018, in: Eighth International Symposium on Tunnel Safety and Security, RISE Research Institutes of Sweden AB, Stockholm, 2018, pp. 503–514, URL https://juser.fz-juelich.de/record/844686.

[4] F. Brännström, Application of fire simulation within the railway industry, in: Proceedings from the 14th International Conference and Exhibition on Fire Science and Engineering Interflam, Windsor, UK, 14 Mar 2018 - 16 Mar 2018, in: 14th International Conference and Exhibition on Fire Science and Engineering Interflam, Interscience Communications Ltd, 2016, pp. 119–130.

[5] E. Guillaume, A. Camillo, T. Rogaume, Application and limitations of a method based on pyrolysis models to simulate railway rolling stock fire scenarios, Fire Technol. 50 (2) (2014) 317–348.

[6] A. Bakhshaii, E.A. Johnson, A review of a new generation of wildfire–atmosphere modeling, Can. J. Forest Res. 49 (6) (2019) 565–574.

[7] A. Witkowski, A.A. Stec, T.R. Hull, Thermal decomposition of polymeric materials, in: SFPE Handbook of Fire Protection Engineering, Springer, 2016, pp. 167–254.

[8] M.M. Khan, A. Tewarson, M. Chaos, Combustion characteristics of materials and generation of fire products, in: M.J. Hurley (Ed.), SFPE Handbook of Fire Protection Engineering, Springer New York, New York, NY, 2016, pp. 1143–1232, http://dx.doi.org/10.1007/978-1-4939-2565-0_36.

[9] K. McGrattan, S. Hostikka, J. Floyd, R. McDermott, M. Vanella, Fire dynamics simulator user's guide, version 6.7.4, 2020, http://dx.doi.org/10.6028/NIST.SP.1019.

[10] C. Lautenberger, C. Fernandez-Pello, Generalized pyrolysis model for combustible solids, Fire Saf. J. 44 (6) (2009) 819–839.

[11] J. Lachaud, N.N. Mansour, Porous-material analysis toolbox based on OpenFOAM and applications, J. Thermophys. Heat Transfer 28 (2) (2014) 191–202.

[12] S.I. Stoliarov, R.E. Lyon, Thermo-kinetic model of burning for pyrolyzing materials, in: Proceedings of the Ninth International Symposium on Fire Safety Science, Vol. 9, 2008, pp. 1141–1152.

[13] L. Arnold, T. Hehnen, P. Lauer, C. Trettin, A. Vinayak, PROPTI–a generalised inverse modelling framework, J. Phys. Conf. Ser. 1107 (2018) 032016.

[14] L. Arnold, T. Hehnen, P. Lauer, C. Trettin, A. Vinayak, Application cases of inverse modelling with the PROPTI framework, Fire Saf. J. (2019) 102835, http://dx.doi.org/10.1016/j.firesaf.2019.102835, URL www.sciencedirect.com/science/article/pii/S0379711219300438.

[15] A. Brown, M. Bruns, M. Gollner, J. Hewson, G. Maragkos, A. Marshall, R. McDermott, B. Merci, T. Rogaume, S. Stoliarov, et al., Proceedings of the first workshop organized by the IAFSS working group on measurement and computation of fire phenomena (MaCFP), Fire Saf. J. 101 (2018) 1–17.

[16] T. Nyazika, M. Jimenez, F. Samyn, S. Bourbigot, Pyrolysis modeling, sensitivity analysis, and optimization techniques for combustible materials: A review, J. Fire Sci. 37 (4–6) (2019) 377–433.

[17] S. Vyazovkin, A.K. Burnham, L. Favergeon, N. Koga, E. Moukhina, L.A. Pérez-Maqueda, N. Sbirrazzuoli, ICTAC kinetics committee recommendations for analysis of multi-step kinetics, Thermochim. Acta 689 (2020) 178597.

[18] M. Brown, M. Maciejewski, S. Vyazovkin, R. Nomen, J. Sempere, A.a. Burnham, J. Opfermann, R. Strey, H. Anderson, A. Kemmler, et al., Computational aspects of kinetic analysis: Part A: The ICTAC kinetics project-data, methods and results, Thermochim. Acta 355 (1–2) (2000) 125–143.

[19] P. Lauer, C. Trettin, F.-W. Wittbecker, L. Arnold, Performance of optimization algorithms for deriving material data from bench scale tests, in: Fire and Evacuation Modelling Technical Conference 2016, 2016.

[20] M. Naser, Mechanistically informed machine learning and artificial intelligence in fire engineering and sciences, Fire Technol. (2021) 1–44.

[21] S. Ascher, I. Watson, S. You, Machine learning methods for modelling the gasification and pyrolysis of biomass and waste, Renew. Sustain. Energy Rev. (2021) 111902.

[22] B.R. Hough, D.T. Schwartz, J. Pfaendtner, Detailed kinetic modeling of lignin pyrolysis for process optimization, Ind. Eng. Chem. Res. 55 (34) (2016) 9147–9153.

[23] C. Tsekos, S. Tandurella, W. de Jong, Estimation of lignocellulosic biomass pyrolysis product yields using artificial neural networks, J. Anal. Appl. Pyroly. 157 (2021) 105180, http://dx.doi.org/10.1016/j.jaap.2021.105180, URL https://www.sciencedirect.com/science/article/pii/S0165237021001662.

[24] H. Wei, K. Luo, J. Xing, J. Fan, Predicting co-pyrolysis of coal and biomass using machine learning approaches, Fuel 310 (2022) 122248, http://dx.doi.org/10.1016/j.fuel.2021.122248, URL https://www.sciencedirect.com/science/article/pii/S0016236121021220.

[25] W. Ji, F. Richter, M.J. Gollner, S. Deng, Autonomous kinetic modeling of biomass pyrolysis using chemical reaction neural networks, Combust. Flame 240 (2022) 111992.

[26] P. Geurts, D. Ernst, L. Wehenkel, Extremely randomized trees, Mach. Learn. 63 (1) (2006) 3–42.

[27] F. Richter, G. Rein, The role of heat transfer limitations in polymer pyrolysis at the microscale, Front. Mech. Eng. 4 (2018) 18.

[28] S. Vyazovkin, Comments on multiple publications reporting single heating rate kinetics, Appl. Organomet. Chem. (2022) e6929.

[29] S. Vyazovkin, A.K. Burnham, J.M. Criado, L.A. Pérez-Maqueda, C. Popescu, N. Sbirrazzuoli, ICTAC kinetics committee recommendations for performing kinetic computations on thermal analysis data, Thermochim. Acta 520 (1–2) (2011) 1–19.

[30] S. Vyazovkin, K. Chrissafis, M.L. Di Lorenzo, N. Koga, M. Pijolat, B. Roduit, N. Sbirrazzuoli, J.J. Suñol, ICTAC kinetics committee recommendations for collecting experimental thermal analysis data for kinetic computations, Thermochim. Acta 590 (2014) 1–23.

[31] K. McGrattan, R. McDermott, W. Mell, G. Forney, J. Floyd, S. Hostikka, Modeling the burning of complicated objects using lagrangian particles, in: Proceedings of the 2010 Interflam Conference, Nottingham, 2010, pp. –1, URL https://tsapps.nist.gov/publication/get_pdf.cfm?pub_id=905798.

[32] R.E. Lyon, Heat release kinetics, Fire Mater. 24 (4) (2000) 179–186.

[33] R.E. Lyon, N. Safronava, A comparison of direct methods to determine n-th order kinetic parameters of solid thermal decomposition for use in fire models, J. Therm. Anal. Calorim. 114 (1) (2013) 213–227.

[34] K. McGrattan, S. Hostikka, J. Floyd, R. McDermott, M. Vanella, Fire dynamics simulator technical reference guide volume 2: Verification, version 6.7.4, 2020, http://dx.doi.org/10.6028/NIST.SP.1018.

[35] P. Lauer, p-lauer/inversepyrolysis-ert: inversepyrolysis-ert v.0.1, Zenodo, 2022, http://dx.doi.org/10.5281/zenodo.6606248.

[36] P. Lauer, Training data set for inverse pyrolysis machine learning, Zenodo, 2022, http://dx.doi.org/10.5281/zenodo.6337389.

[37] P. Lauer, Inverse pyrolysis surrogate model, Zenodo, 2022, http://dx.doi.org/10.5281/zenodo.6346476.

[38] L. Breiman, Random forests, Mach. Learn. 45 (1) (2001) 5–32.

[39] R. Marée, P. Geurts, J. Piater, L. Wehenkel, Random subwindows for robust image classification, in: 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Vol. 1, CVPR'05, IEEE, 2005, pp. 34–40.

[40] V.A. Huynh-Thu, A. Irrthum, L. Wehenkel, P. Geurts, Inferring regulatory networks from expression data using tree-based methods, PLoS One 5 (9) (2010) e12776.

[41] J. Ruyssinck, V.A. Huynh-Thu, P. Geurts, T. Dhaene, P. Demeester, Y. Saeys, NIMEFI: Gene regulatory network inference using multiple ensemble feature importance algorithms, PLoS One 9 (3) (2014) e92709.

[42] D. Kocev, M. Ceci, Ensembles of extremely randomized trees for multi-target regression, in: N. Japkowicz, S. Matwin (Eds.), Discovery Science, Springer International Publishing, Cham, 2015, pp. 86–100.

[43] P. Geurts, G. Louppe, Learning to rank with extremely randomized trees, in: Proceedings of the Learning to Rank Challenge, PMLR, 2011, pp. 49–61.

[44] S. Galelli, A. Castelletti, Assessing the predictive capability of randomized tree-based ensembles in streamflow modelling, Hydrol. Earth Syst. Sci. 17 (7) (2013) 2669–2684, http://dx.doi.org/10.5194/hess-17-2669-2013, URL https://hess.copernicus.org/articles/17/2669/2013/.

[45] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, E. Duchesnay, Scikit-learn: Machine learning in Python, J. Mach. Learn. Res. 12 (2011) 2825–2830.

[46] M. Naser, A.H. Alavi, Error metrics and performance fitness indicators for artificial intelligence and machine learning in engineering and sciences, Archit. Struct. Construct. (2021) 1–19.

[47] T. Hastie, R. Tibshirani, J. Friedman, Model assessment and selection, in: The Elements of Statistical Learning: Data Mining, Inference, and Prediction, Springer New York, New York, NY, 2009, pp. 219–259, http://dx.doi.org/10.1007/978-0-387-84858-7_7.

[48] M.A. Branch, T.F. Coleman, Y. Li, A subspace, interior, and conjugate gradient method for large-scale bound-constrained minimization problems, SIAM J. Sci. Comput. 21 (1) (1999) 1–23.

[49] P. Virtanen, R. Gommers, T.E. Oliphant, M. Haberland, T. Reddy, D. Cournapeau, E. Burovski, P. Peterson, W. Weckesser, J. Bright, S.J. van der Walt, M. Brett, J. Wilson, K.J. Millman, N. Mayorov, A.R.J. Nelson, E. Jones, R. Kern, E. Larson, C.J. Carey, İ. Polat, Y. Feng, E.W. Moore, J. VanderPlas, D. Laxalde, J. Perktold, R. Cimrman, I. Henriksen, E.A. Quintero, C.R. Harris, A.M. Archibald, A.H. Ribeiro, F. Pedregosa, P. van Mulbregt, SciPy 1.0 Contributors, SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python, Nature Methods 17 (2020) 261–272, http://dx.doi.org/10.1038/s41592-019-0686-2.

[50] P. Lauer, C. Trettin, L. Arnold, F. Brännström, T. Hehnen, Role of the cost function for material parameter estimation, in: Fire and Evacuation Modelling Technical Conference 2020, 2020.

[51] L. Wehenkel, D. Ernst, P. Geurts, Ensembles of extremely randomized trees and some generic applications, in: Proceedings of Robust Methods for Power System State Estimation and Load Forecasting, 2006.

[52] R. McDermott, I. Leveton (Eds.), Measurement and Computation of Fire Phenomena (MaCFP): Condensed Phase Material Database, MaCFP Working Group, 2019, Commit: 7335c4470ec2cb0327ac0fe8386ecfeee6b3c6b1URL https://github.com/MaCFP/matl-db.

[53] M. Bruns, Macfp 2021 condensed phase workshop, 2021, https://github.com/MaCFP/matl-db/releases/tag/v1.1.0.

[54] I. Leventon, B. Batiot, M. Bruns, S. Hostikka, Y. Nakamura, P. Reszka, T. Rogaume, S. Stoliarov, et al., The MaCFP condensed phase working group: A structured, global effort towards pyrolysis model development, in: ASTM Selected Technical Papers (STP), Atlanta, GA, US, 2021.

[55] N. Koga, A review of the mutual dependence of Arrhenius parameters evaluated by the thermoanalytical study of solid-state reactions: The kinetic compensation effect, Thermochim. Acta 244 (1994) 1–20.

[56] M. Bruns, J. Koo, O. Ezekoye, Population-based models of thermoplastic degradation: Using optimization to determine model parameters, Polym. Degrad. Stab. 94 (6) (2009) 1013–1022.

[57] M.C. Bruns, I.T. Leventon, Automated fitting of thermogravimetric analysis data, Fire Mater. 45 (3) (2021) 406–414.

[58] A. Matala, et al., Methods and Applications of Pyrolysis Modelling for Polymeric Materials, VTT Technical Research Centre of Finland, 2013.

[59] L. Hasalová, J. Ira, M. Jahoda, Practical observations on the use of shuffled complex evolution (SCE) algorithm for kinetic parameters estimation in pyrolysis modeling, Fire Saf. J. 80 (2016) 71–82, http://dx.doi.org/10.1016/j.firesaf.2016.01.007, URL https://www.sciencedirect.com/science/article/pii/S0379711216300042.

[60] M. Naser, An engineer's guide to explainable artificial intelligence and interpretable machine learning: Navigating causality, forced goodness, and the false perception of inference, Autom. Constr. 129 (2021) 103821.

[61] ASTM D 7309-21b Test Method for Determining Flammability Characteristics of Plastics and Other Solid Materials Using Microscale Combustion Calorimetry, ASTM International, 2021, http://dx.doi.org/10.1520/d7309-21b.

[62] J.S. Centre, JURECA: General-purpose supercomputer at Jülich supercomputing centre, J. Large-Scale Res. Facil. 2 (2016) http://dx.doi.org/10.17815/jlsrf-2-121.

[63] D. Krause, P. Thörnig, JURECA: Modular supercomputer at Jülich supercomputing centre, J. Large-Scale Res. Facil. 4 (2018) A132, http://dx.doi.org/10.17815/jlsrf-4-121-1.

[64] T. Hastie, R. Tibshirani, J. Friedman, Ensemble learning, in: The Elements of Statistical Learning: Data Mining, Inference, and Prediction, Springer New York, New York, NY, 2009, pp. 605–624, http://dx.doi.org/10.1007/978-0-387-84858-7_16.

[65] T. Hastie, R. Tibshirani, J. Friedman, Random forests, in: The Elements of Statistical Learning: Data Mining, Inference, and Prediction, Springer New York, New York, NY, 2009, pp. 587–604, http://dx.doi.org/10.1007/978-0-387-84858-7_15.