Machine Learning Methods for Modelling the Gasification and Pyrolysis of Biomass and Waste

Simon Ascher [a], Ian Watson [a], Siming You [a, *]

[a] James Watt School of Engineering, University of Glasgow, Glasgow, G12 8QQ, UK

* Corresponding author. Email: siming.you@glasgow.ac.uk

Submitted to Renewable & Sustainable Energy Reviews

October 2021

Abstract

Over the past two decades, the use of machine learning (ML) methods to model biomass and waste gasification/pyrolysis has increased rapidly. Only 70 papers were published in the 2000s compared to a total of 549 publications in the 2010s. However, the approaches and findings have yet to be systematically reviewed. In this work, the machine learning methods most commonly employed for modelling gasification and pyrolysis processes are discussed with reference to their applications, merits, and limitations. Whilst coefficients of determination (R^2) can be difficult to compare directly, due to some studies having vastly different approaches and aims, it has generally been found that most studies consistently achieved a high prediction accuracy with $R^2 > 0.90$. Artificial neural networks have been most widely used due to their potential to learn highly non-linear input-output relationships. However, a variety of methods (e.g. regression methods, tree-based methods, and support vector machines) are appropriate depending on the application, data availability, model speed, etc. It is concluded that ML has great potential for the development of models with greater accuracy. Some advantages of machine learning models over existing models are their ability to incorporate relevant non-numerical parameters and the power to generate a multitude of solutions for a wide range of input parameters. More emphasis should be placed on model interpretability in order to better understand the processes being studied.

Keywords: Machine Learning; Gasification; Pyrolysis; Biomass; Waste

Word count: 10,065 (excluding references, figure captions, etc.)

List of Abbreviations

ANN	Artificial neural network
ANFIS	Adaptive neuro-fuzzy inference system
CCS	Carbon capture and storage
CFD	Computational fluid dynamics
CHP	Combined heat and power
CNN	Convolutional neural networks
DT	Decision tree
HHV	Higher heating value
LCA	Life cycle assessment
LHV	Lower heating value
ML	Machine learning
MLR	Multiple linear regression
MSE	Mean squared error
MSW	Municipal solid waste
NARX	Autoregressive exogeneous neural networks
PCA	Principal component analysis
PR	Polynomial regression
PSO	Particle swarm optimisation
R^2	Coefficient of determination
RF	Random forest
RMSE	Root mean squared error
RMSPE	Root mean squared percentage error
SVM	Support vector machine
XGB	Extreme gradient boosting

1. Introduction

To date over 80% of the world's energy demand is supplied by fossil fuels [1]. However, due to climate change and global warming concerns, tremendous efforts have been made to search for more sustainable energy options. Renewable energy sources such as solar and wind represent excellent options, but also suffer from strong fluctuations in their energy output, suggesting a high reliance on energy storage technologies. Bioenergy on the other hand has the potential to provide a relatively consistent energy production and is less affected by meteorological conditions. Furthermore, bioenergy recovery from waste facilitates the development of sustainable waste management towards a circular economy.

Thermochemical methods have been extensively investigated for their great potential to convert biomass and waste into bioenergy and serve as a valid waste treatment option. Generally, one may differentiate between the following five conversion pathways: combustion, gasification, pyrolysis, carbonisation/torrefaction, and liquefaction [2]. This work focuses on gasification and pyrolysis due to their widespread applications and significant environmental benefits. During gasification, a feedstock is converted to syngas at elevated temperatures (550 – 1000 °C) in an oxygen-deficient environment [3,4]. Pyrolysis on the other hand takes place at more moderate temperatures of 300-650 °C in the absence of oxygen [2]. Thus, biomass can be thermally degraded into three products: solid residues consisting of biochar and ash, gas, and pyrolysis liquid (pyrolysis oil or bio-oil) [5].

Gasification and pyrolysis represent cost-effective ways to recover energy from biomass and waste feedstocks [5]. Whilst combustion is well suited for the direct generation of electricity and heat, pyrolysis and gasification can produce a wider range of products such as liquid transport fuels, bulk chemicals, and biochar. These products may serve to effectively reduce the carbon footprint of relevant development and offer greater flexibility towards demand management. Transport fuels, such as biodiesel, may reduce the environmental impact of the transport sector when appropriate production strategies are employed [6–8]. Whereas biochar represents a solid form of carbon which, along with carbon capture and storage (CCS), has the potential to turn bioenergy into a carbon-negative industry (e.g. upon soil application) [9–11]

Dependent on factors such as capacity and feedstock, gasification-based systems have been shown to outperform combustion-based systems when considering their energy recovery, conversion efficiency, and potential to generate revenue [12]. For instance, one study found that municipal solid waste (MSW) gasification resulted in a net energy production to the grid

of 685 kWh/tonne MSW as compared to 544 kWh/tonne MSW for waste incineration [12]. However, this potential increase in efficiency can be associated with greater process complexity [13].

Importantly, gasification and pyrolysis also have the potential to environmentally outperform the technologies based on waste combustion processes [14,15]. One study comparing four commercial plants found that the environmental impact of gasification was lower than incineration, due to an improved energy efficiency and syngas cleaning technology [14]. Environmental performance is expected to improve even more as these technologies develop and further improvements to, for instance, syngas cleaning are made.

The great complexity of thermochemical processes calls for the development of mathematical models for process design, optimisation, and intensification. Models can be used to identify optimal operating conditions of a system and study the effects of different parameters on the outputs. Models can also be used to achieve real-time process control and optimisation and aid with scale-up of a system or initial system sizing [2].

Fig. 1 shows how the research interest in modelling techniques for the gasification and pyrolysis process have increased over the last two decades. A rapid increase in publications can especially be observed for the second half of the 2010s. Data was obtained from the "Web of Science" database using the key words model/modelling/simulation, waste/biomass, and the respective thermochemical process (i.e. gasification or pyrolysis) [16].

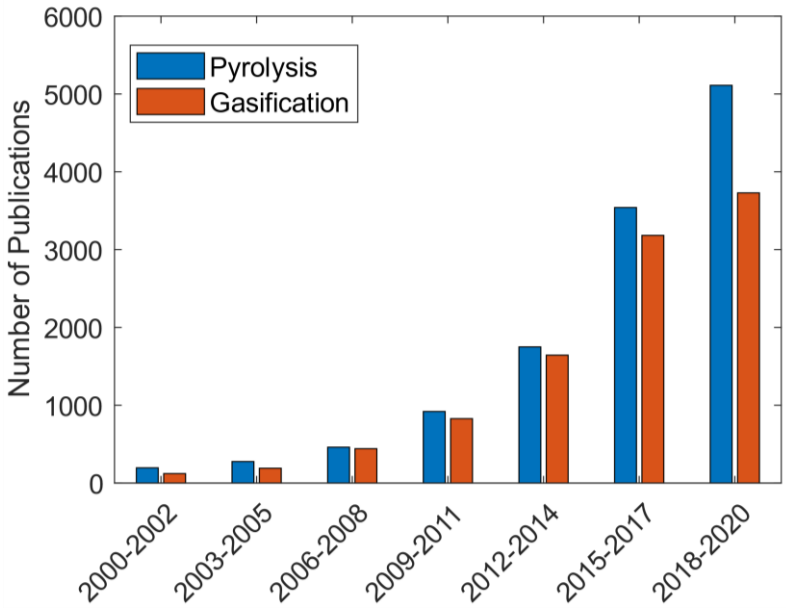

Fig. 1 Research interest in pyrolysis and gasification modelling throughout the years

A range of approaches have been used to model the complex chemical reactions, as well as mass and heat transfer inside the reactor, during gasification and pyrolysis. Conventionally, thermodynamic equilibrium, kinetic, and computational fluid dynamics (CFD) models have been the main focus. However, more recently machine learning (ML) techniques, such as artificial neural networks (ANNs), have received increasing attention due to their potential to efficiently model these highly complex processes.

Thermodynamic equilibrium models are effective at determining the theoretical limits of a process [17]. Main drawbacks of equilibrium models stem from the high number of assumptions made and include a lack of accuracy when it comes to modelling the effects of specific reactor designs. Furthermore, they are not well suited for conditions which are dissimilar to equilibrium conditions (e.g. low temperature gasification) [18]. Kinetic models can improve on the performance of equilibrium models by taking the reactor design into consideration. Meanwhile this also limits them to a specific reactor design [17,18]. Similar limitations are applicable to CFD models. Additionally, the high complexity of CFD models results in high computational demands.

ML models represent a different approach. Where conventional models try to accurately represent the physical and chemical processes occurring within a reactor, ML models are generally not concerned with the actual processes. Instead, an attempt is made at mapping inputs to the corresponding outputs. It is left up to the employed algorithm to understand the underlying mathematical relationships resulting from a vast number of highly complex processes. In this field it has the potential to more accurately model thermochemical processes and better predict their outputs than conventional models [19,20].

Most gasification review papers have a strong focus on thermodynamic equilibrium models and kinetic rate models [17,18,21–23], with Aspen Plus models [21] and CFD models [22,23] being reviewed to a lesser extent. ANN methodology was the only ML technique receiving limited attention in past review papers [17,21–23]. Moreover, ANN modelling was generally mentioned as a side note and not reviewed in greater detail. For example, Baruah and Baruah reviewed only two studies using ANNs to model gasification [22]. One of the reviewed studies showed that an ANN was able to better predict gasification products than a real-gas equilibrium model which incorporated tar, char, and permanent gas reactions [24]. It is important to highlight that other relevant ML options, such as support vector machine (SVM) and random forest (RF) approaches, have not been considered by any of the mentioned reviews [17,18,21–23].

A recent 2020 review by Kostetskyy and Broadbelt discusses the latest advances in biomass fast pyrolysis modelling [25]. A large proportion of this review focused on conventional approaches, such as atomistic and mechanistic techniques, to model the pyrolysis process. The authors included a section discussing recent advances in ML techniques and their potential applications within the wider field of chemical engineering. The successful implementation of other ML methods, such as SVM and RF approaches, as well as advanced neural networks, such as adaptive neuro-fuzzy inference systems (ANFIS) and convolutional neural networks (CNN), has been acknowledged [25].

On top of Kostetskyy and Broadbelt's review, several other closely related review papers have been published throughout the years. For example, Hameed *et al.* compared kinetic, mechanistic, and network/structural models [26]. Kaczor *et al.* paid special attention to how chemical reactions were implemented in different approaches and focused on solutions appropriate for subsequent CFD implementation [27]. In an earlier work Sharma *et al.* focused on kinetic, particle, and reactor-based models, as well as the effects of process parameters and catalyst use on the pyrolysis process [28]. Finally, Babu's 2008 review focused on mathematical techniques for the estimation of kinetic parameters to model biomass pyrolysis [29]. From this it becomes clear that ML-based techniques for pyrolysis modelling have received little attention, apart from Kostetskyy and Broadbelt's review.

In conclusion, a lack of a systematic review on the application of ML to model biomass and waste conversion via thermochemical methods, such as gasification or pyrolysis, is evident. Hence, this work aims to provide a clear picture of the current state of art of these methods within the field. To the author's knowledge, this work is the first review considering a wide range of ML methods for modelling gasification and pyrolysis.

This review is structured as follows: In Section 2 thermochemical processes for the conversion of waste and biomass are discussed. Section 3 represents a brief introduction to numerous popular ML techniques which have been used in the field, their merits and relative advantages and disadvantages are compared. Section 4 takes a closer look at how researchers have used ML to model thermochemical processes. Section 5 outlines the current challenges in the successful deployment of ML and how they may be overcome, before Section 6 concludes the findings of this work.

2. Thermochemical Methods

A basic understanding of pyrolysis and gasification as well as the focus of the latest research is deemed essential in understanding how ML methods can aid with research in this field. As such, this section is intended as a brief introduction to these two thermochemical processes. The reader is referred to designated review papers, such as [30–32] and [22,33,34] for more detailed information on the pyrolysis and gasification processes, respectively.

2.1 Pyrolysis

During the pyrolysis process, the feedstock is converted into bio-oil, biochar, and gases by breaking down larger hydrocarbon molecules into smaller ones [2]. The process temperature, heating rate, and residence time are key parameters determining product yields. Generally, higher temperatures (>600 °C) favour gas production due to the dominance of secondary cracking reactions [35]. Bio-oil production generally peaks at temperatures in the range of 400-550 °C [30].

Another essential parameter is the process's heating rate. In combination with the residence time, it defines the type of pyrolysis, i.e. flash, fast, or slow pyrolysis. Gases and bio-oil are predominantly produced at high heating rates and low residence times. Slow pyrolysis on the other hand, characterised by an increase in residence time and decrease in heating rate leads to an increase in the yield of biochar. The influence of the reaction conditions has previously been reviewed in greater detail by Kan *et al.* [30].

Typical product yields for the pyrolysis and gasification of wood have previously been summarised by Bridgwater [32]. Product yield splits for liquids, solids, and gases of 75/12/13, 30/35/35, and 5/10/85 have been found for fast pyrolysis, slow pyrolysis, and gasification, respectively. Here, fast pyrolysis occurred at ~500 °C with a hot vapour residence time of ~ 1s. Slow pyrolysis on the other hand took place at ~400 °C with a residence time in the range of days. Finally, a gasification temperature of 700-900 °C was considered.

Recent research efforts in the field of biomass and waste pyrolysis focused on improved process and reactor designs as well as product upgrading and applications. For instance, work on the efficient and cost-effective upgrading of bio-oil is essential in making the process commercially viable [36]. Additionally, turning biochar into a high-value product may significantly help with economic feasibility. One such product may be activated biochar with

properties engineered for adsorption, catalysis, and wastewater treatment applications [37]. Furthermore, work on improved pyrolysis methods and reactor designs continues and novel approaches include microwave-assisted and solar pyrolysis [37–39]. Especially, microwave-assisted pyrolysis might be a valid option to resolve the drawbacks of conventional pyrolysis, making it an environmentally friendly method to produce biofuels from different municipal waste streams [39].

2.2 Gasification

In general, gasification aims to convert a carbonaceous feedstock into a gas product, often called syngas or producer gas, of usable heating value. Syngas can either be used for energy generation via an internal combustion engine, combined heat and power (CHP) plant or upgraded to value-added chemicals [2]. Gasification typically involves the following four processes: (i) drying, (ii) pyrolysis, (iii) partial combustion and tar cracking, and (iv) reduction.

Air, steam, oxygen, or a combination of the three have been used as the gasifying agent. Air is commonly employed in small or medium sized plants due to its cheapness. Due to the high nitrogen concentration in air, the produced syngas is diluted with nitrogen resulting in a relatively low heating value syngas of 4-7 MJ/Nm^3. The use of steam or oxygen as gasifying agent results in greater capital and operating costs but increases the heating value of the syngas to 10-18 MJ/Nm^3 and 12-28 MJ/Nm^3, respectively [2]. Steam improves the heating value by promoting the formation of H_2, resulting in a final gas with a higher H/C ratio. Using oxygen on the other hand, yields a gas high in CO and CO_2 with a much smaller proportion of H_2.

Significant research efforts in the field of gasification have been focused on the production of biofuels [40]. Hydrogen production by gasification has also been of interest to researchers [41]. To ensure industrial-scale deployment, efforts have been made to identify the optimal operating conditions which allow for efficient and economic production of hydrogen [42]. Other ongoing research is closely linked to aiding industrial-scale deployment, e.g. tar reduction and its removal are major technical issues as tar can cause significant problems to downstream equipment [43,44]. *In-situ* or post gasification tar removal are often expensive. Therefore, the use of catalysts, which are still potentially costly, to produce a cleaner syngas more suitable for industrial-scale deployment has been explored [43].

3. Machine Learning Techniques

Four ML methods have been commonly employed to model pyrolysis and gasification processes: regression analysis, ANNs, tree-based methods such as RF, and SVMs. Sections 3.1-3.4 provide a brief introduction to these methods.

Fig. 2 shows the rapid increase in publications applying ML methods to model gasification and pyrolysis processes. Web of Science search key words used to obtain the data illustrated by Fig. 2 may be found in Appendix A.

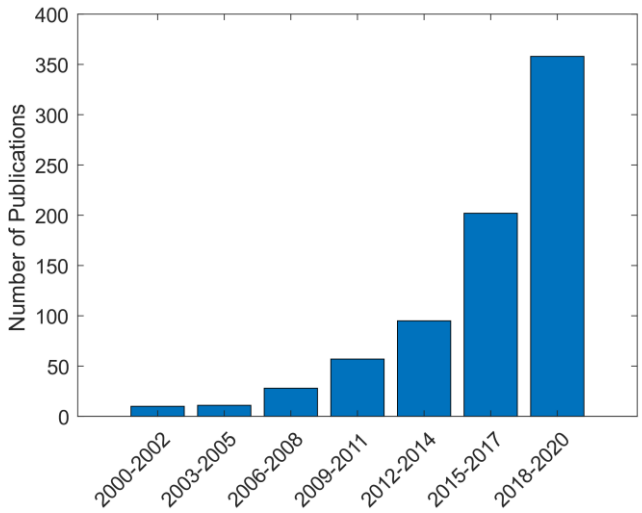

Fig. 2 Prevalence of machine learning use for gasification and pyrolysis modelling applications

3.1 Regression Methods

Regression analysis can be used to assess the relationship between dependent (target) variables and independent variables (inputs). The simplest form of regression analysis is linear regression for which a straight-line equation is fitted to describe the relationship between one dependent and one independent variable. This is generally done using the least squares method [45].

In addition to linear regression, many other types of regression exist with various applications. For example, logistic regression is well suited for discrete problems whereas shrinkage methods such as ridge regression can make sense of highly correlated independent variables. Polynomial regression may be used to fit a n^{th} degree polynomial to data that does not follow linear relationships. Kernel smoothing methods represent another option which

splits the domain into segments and locally fits several simple models. Combining these simpler models will in turn yield a fit to the overall data set [45].

Whilst other ML methods, such as ANNs, are often considered a black box, regression analysis generally allows for easy model interpretation. For instance, linear regression analysis clearly indicates how an independent variable affects a dependent variable. However, the method's simplicity is also its downfall as poor performance must be expected for data following non-linear relationships. More advanced regression techniques, such as polynomial regression, can better fit non-linear data while retaining the strong interpretability. However, in this case, model tuning becomes more challenging.

3.2 Artificial Neural Networks

ANNs are a programming paradigm inspired by the workings of the biological neural network within the human brain. The term neural network describes a larger number of different models and methods. Fig. 3 shows the simple and widely employed single layer perceptron, or single hidden layer back-propagation network.

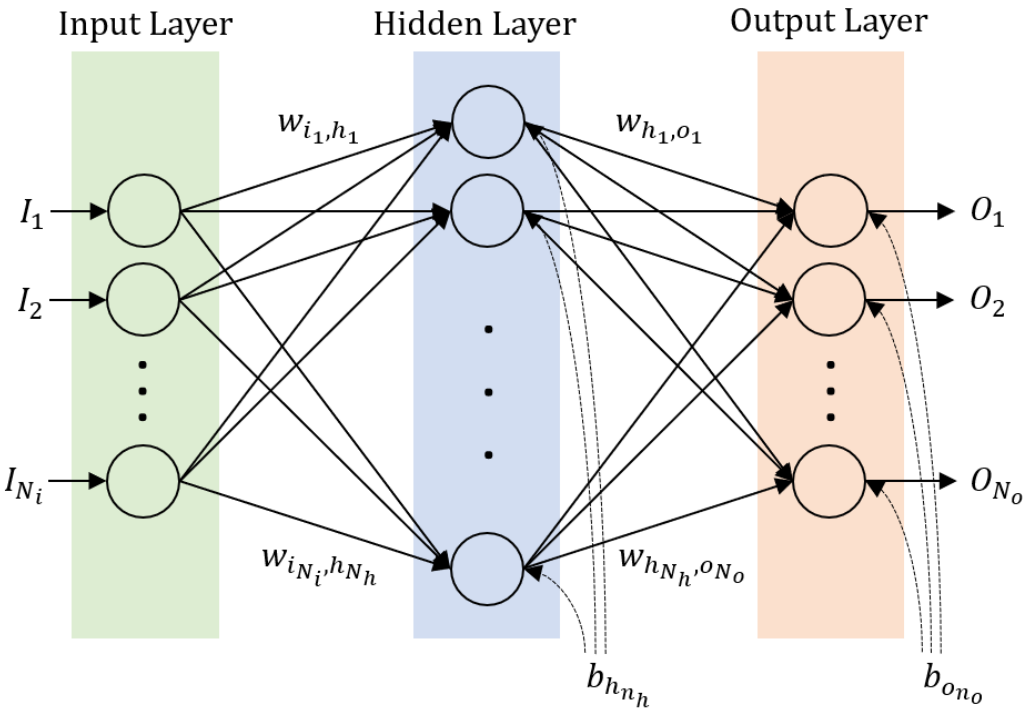

Fig. 3 Illustration of single layer perceptron (single hidden layer back-propagation network) architecture. I and O represent the perceptron's inputs and outputs, respectively. N_i, N_h, and N_o indicate the number of neurons in the ANN's input, hidden, and output layer, respectively. Weights between two neurons are denoted by w, where the two subscripts denote the neurons which are connected by the weight (e.g. w_{i_1,h_1} is the weight connecting the first input layer neuron with the first hidden layer neuron). Finally, b represents bias terms feeding into hidden and output layer neurons with the subscripts h_{n_h} and o_{n_o} representing each hidden and output layer neuron, respectively.

An ANN is made up of nodes called artificial neurons. Each neuron is connected to other neurons to which it can pass on received signals. Neurons are arranged into different layers, namely an input and output layer, as well as any number of hidden layers (one for a single layer perceptron). Each neuron has an associated weight, as well as a bias for hidden and output layer neurons. In short, the network learns by feeding back its predictions, comparing them to the corresponding inputs, and adjusting weights accordingly [45].

ANNs are not a new idea as the first functional networks were published in the 1960s and the idea itself existed for even longer [46]. However, more recently with the increase of processing power ANNs have experienced a renaissance. Their potential to model highly non-linear problems makes them appealing for modelling thermochemical processes. In fact, ANNs can be considered a universal function approximator. This means, in theory a neural network can be created to model any given function [47].

ANNs have widely been accepted as a powerful tool for the modelling of highly non-linear processes. However, their black box nature and lack of interpretability has been a drawback highlighted by many authors [19,48–50]. Whilst this issue cannot be neglected, significant efforts have been made in creating robust methods which improve their interpretability [51]. Olden *et al.* found that the connection weights and partial derivatives methods predicted the true parameter importance better than for instance Garson's algorithm [51]. Model-agnostic methods are a flexible way to interpret the output of any ML model. Feature permutation or the use of global surrogate models are two examples of model-agnostic methods [52].

3.3 Decision Trees and Related Methods

Decision trees are a powerful methodology suitable for regression and classification problems. While growing a decision tree the algorithm splits the feature space into several regions. A split point is determined for each region and the splitting process is repeated for each resulting region. Initially a large tree structure is grown this way which likely results in overfitting. Here a strategy called pruning comes into play which removes excess branches. This is often done by introducing a cost-complexity criterion which captures the trade-off between the tree's generalisation capability (i.e. the model's ability to adapt properly to new, previously unseen data) and goodness of fit [46].

Many algorithms attempt to improve on individual decision trees by utilising methodologies such as bagging, boosting, or ensemble methods. One such method is called RF. RF

methodology creates and ensemble of decision trees. Importantly, each tree is trained on only a fraction of the data set. The data used to train individual trees is randomly chosen. Ultimately, for regression, the outputs of all decision trees are averaged to compute the RF model's final prediction. RF for classification uses majority voting instead for which each tree casts a vote indicating its predicted class. The class with the most votes wins and is returned as the RF model's predicted class [45,53]. A simplified illustration of decision tree and RF methodology is shown by Fig. 4.

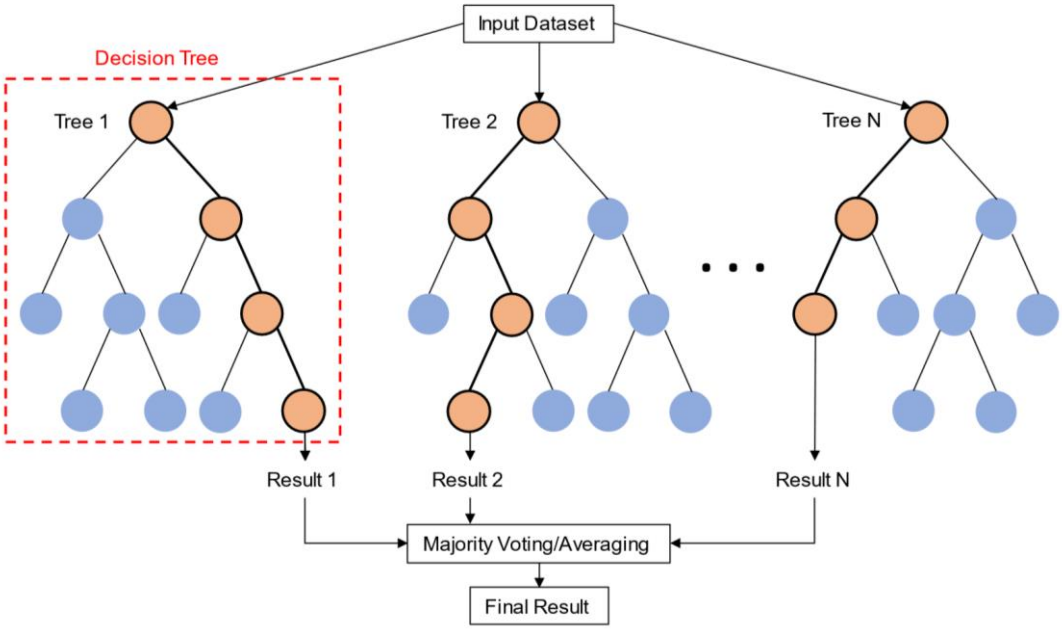

Fig. 4 Illustration of decision tree and RF methodology. The red box shows an individual decision tree, whereas the complete figure illustrates RF methodology. Branches highlighted in orange show the paths yielding a tree's prediction for a given input. Furthermore, the figure shows how RF averages (majority voting for classification) the results of all trees to calculate the model's final prediction.

Whilst simple decision trees can be trained rapidly, they are also prone to overfitting. Additionally, the method's hierarchical nature leads to high variance. This means small changes in the training data could result in a very different tree. Ensemble methods using a vast number of trees can help with overcoming this problem. Furthermore, these methods result in models robust to overfitting as many ensemble algorithms have in-built regularisation mechanisms. However, the large number of trees also significantly increases computational demands [45].

Two key ensemble methods are RF and gradient boosting. The fundamental difference between the two methods is that RF builds many large trees, whereas gradient boosting uses

many weak learners (i.e. shallow trees which are small trees with a low depth). A well-tuned gradient boosting model tends to outperform a RF model [45]. However, tuning can be challenging. RF on the other hand has fewer hyperparameters and generally performs well even for sub-optimal hyperparameter choices.

Other issues with basic decision trees are the lack of smoothness of their underlying functions and difficulties in capturing additive structures (e.g. noise in data set). These issues have been addressed by the development of e.g. multivariate adaptive regression splines [45].

In practice, tree-based methodologies are often simpler to implement than other ML methods. Data generally requires minimal pre-processing and especially missing predictors are easily handled due to the nature of the algorithms. Furthermore, the implementation of categorical predictors is straight forward, as these do not need to be turned into numerical values [45]. Finally, while ANNs are often considered a black box, tree-based methods allow the user to extract feature importance with relative ease. For example, one can determine how a feature effects splits in the tree structure to determine the relevance of the corresponding feature. This can help with troubleshooting the model, improve trust in the model, and increase understanding of the problem at hand which in turn may allow for model improvements.

3.4 Support Vector Machine

SVMs are a type of algorithm most commonly used for binary classification tasks, but the methodology is also well suited for other classification and regression problems. A number of applications are illustrated in Section 4.3.3. Fig. 5 exemplifies the functional principle of SVMs using a simple example of two linearly separable classes. The training algorithm finds the maximum-margin hyperplane which divides two classes in such a way that the distance between the hyperplane and the nearest data points from either class is maximised. Construction of the hyperplane is achieved by using support vectors – data points from either class closest to the hyperplane.

For applications where data are not linearly separable, soft-margin methodology can be employed by using for instance the hinge loss function. For the hinge loss function a high cost is associated with incorrectly classified instances. Correctly identified instances incur a low or no cost for instances with a small or large margin from the hyperplane, respectively [45]. To extend SVM methodology to non-linear problems, so called kernel functions can be used to map inputs into a higher dimensional feature space. Kernel functions are a set of

mathematical functions that transform the input into the required form. Common choices for kernel functions are n^{th} degree polynomial or radial basis functions [45]. Finally, the concepts illustrated for classification tasks can be extended to regression problems as first illustrated by Drucker *et al.* [54].

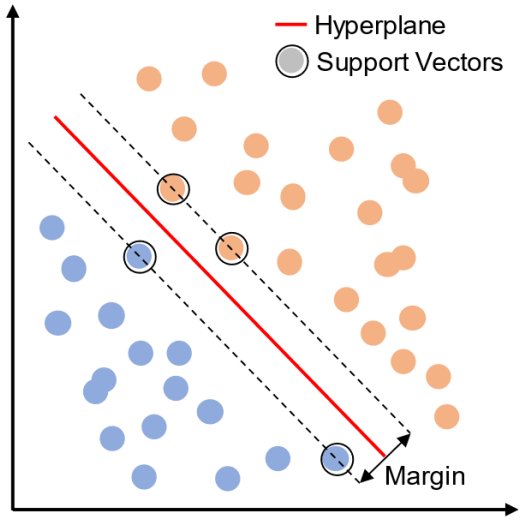

Fig. 5 SVM classification illustrated using a simple example of two linearly separable classes.

Empirical evidence suggests that SVMs perform well for many real-world applications [45]. They have been shown to yield good results in terms of prediction accuracy for data sets of a limited size. Especially high-dimensional data sets make SVMs a good option [55,56]. Another advantage of SVMs is their low sensitivity to outliers which results from the role the support vectors play in constructing the optimal hyperplane [45]. Additionally, it is often thought that SVMs may have an edge over the curse of dimensionality. However, Hastie *et al.* showed that this is not necessarily the case [45]. Finally, the selection of an appropriate kernel function is essential in building a well-functioning model [45].

4. Machine Learning Applications

Table 1 gives an overview of publications employing ML methods to model gasification and pyrolysis. Only publications published in 2016 or later were considered. For studies which developed a range of models, performance parameters are given as a range. If a study considered more than one ML method, performance parameters are quoted for all considered methods. Generally, models achieved an acceptable prediction performance. The high

predictive accuracy of ANNs stands out with R^2 values close to 1 and generally >0.9. Multi-output and single-output models have successfully been employed without either method outperforming the other.

Table 1. Summary of ML models used for the modelling of gasification and pyrolysis.

ML Model	Year	Feedstock Type	Predicted Parameters	Size Data set	Number of Inputs/Outputs	Prediction Performance		Comments	Ref.
						R^2	RMSE		
Gasification									
ANN	2016	MSW	Syngas tars, char LHV, and syngas yield	67	9/(1-3)	0.899-0.983	-	Double hidden layer better than single layer	[49]
ANN	2017	Woody and herbaceous biomass	CO, CO_2, H_2, and CH_4 yields	63	6/1	0.986-0.993	0.052-0.092	-	[57]
ANN	2018	Woody and herbaceous biomass	CO, CO_2, H_2, and CH_4 yields	70	7/4	0.914	-	-	[58]
SVM, RF	2018	Woody biomass	CO, CO_2, H_2, CH_4 yields, and syngas HHV	5237[a]	8/1	-	0.348-1.520, 1.796-5.144	Problem turned into classification problem	[59]
ANNs, NARX	2019	Woody biomass	CO, CO_2, H_2, CH_4 yields, and syngas HHV	3831[a]	8/1 or 9/1	0.930-0.999	0.020-0.853	NARX showed better accuracy than other ANNs	[60]
PR, SVM, DT, ANN	2020	Woody biomass	CO, CO_2, H_2, CH_4 yields, and syngas HHV	4826[a]	3/1	0.541-0.865, 0.713-0.915, 0.807-0.944, 0.823-0.944	0.101-0.582, 0.097-0.278, 0.058-0.180, 0.045-0.198	PCA used to reduce 16 inputs features to 3	[20]
NARX, PR	2020	Woody biomass	Reactor temperature, syngas composition, and syngas HHV	3500[a]	15/6	0.984-0.994, 0.832-0.874	0.016-0.034 ,0.054-0.094	PR suitable for real-time process control	[61]

ANN	2020	Woody biomass	Syngas tar content	120	8/1	>0.970	-	No significant differences between single and two hidden layer ANN	[48]
ANN	2020	Woody and herbaceous biomass, sewage sludge, MSW, plastics	Syngas composition (CO, CO_2, H_2, CH_4) and yield	203	9/1	>0.940	-	Considered effect of bed material	[62]
ANN	2020	Simulated (Woody, animal, and herbaceous biomass)	Power production of downdraft gasifier	1032	11/1	0.999	0.496	Training data obtained from thermodynamic equilibrium model	[63]
ANNs	2021	Woody biomass, Plastics	H_2 production	30	4/1	0.990	-	-	[64]
ANN	2021	Food waste	CO, CO_2, H_2, and CH_4 yields	40	3/1	>0.980	<0.300	-	[65]
Ensemble	2021	MSW	Syngas yield and LHV and LHV of other products	67	9/3	0.972-0.990	0.065-0.152	Ensemble method combines DT, XGB, RF, ANN, and SVM	[66]
		Pyrolysis							
ANN, SVM	2016	Animal biomass	Biochar mass yield	33	5/1	0.804, 0.963	0.835[c], 0.365[c]	SVM performed better than ANN	[67]
ANN	2017	Woody, animal, and herbaceous biomass	Kinetic constants	150	3/3	0.930-0.940	-	-	[68]
ANN, SVM	2018	Biomass (unspecified)	Char, gas, bio-oil yields, and bio-oil LHV	71-175[b]	9/1	0.800-0.932, 0.895-0.938	1.030-4.500, 0.980-3.340	SVM performed better than ANN	[69]
ANN	2018	Herbaceous biomass	Temperature dependent mass loss	1000	2/1	0.970-0.990	-	-	[70]

ANFIS	2019	Herbaceous biomass	Kinetic constants	192	4/3	>0.970	<5.006	-	[71]
RF	2019	Woody and herbaceous biomass	Biochar yield and C content	128-245[b]	(8-12)/2	0.755-0.864	3.403-6.939	Input features were varied	[72]
MLR, RF	2020	Woody and herbaceous biomass, Algae,	Bio-oil yield and H content	137-264[b]	(7-8)/2	0.016-0.352, 0.790-0.920	1.410-7.960, 0.540-3.050	MLR unable to predict outputs	[73]
RF	2020	Woody and herbaceous biomass, Algae,	Biochar yield and characteristics	276-538[b]	7/1	0.780–0.870	-	ML model integrated with LCA and economic analysis	[74]
ANN	2021	Herbaceous biomass, sewage sludge	Char, tar, and gas yield	-	13/3	0.984	2.427	-	[75]

DT – decision tree; HHV – higher heating value; LCA – life cycle assessment; LHV – lower heating value; MLR – multiple linear regression; NARX – autoregressive exogeneous neural networks; PR – polynomial regression; PCA – principal component analysis; R^2 – coefficient of determination; RMSE – root mean squared error; XGB – extreme gradient boosting

a – refers to individual data samples; b – different sized data sets were used for different model outputs; c – given in g instead of normalised form

4.1 Artificial Neural Networks for Process Modelling

ANNs have arguably been the most popular type of ML approach used for the process modelling of gasification and pyrolysis. These thermochemical processes are extremely complex which calls for advanced models [76]. ANNs lend themselves to the modelling of these processes due to their capability to learn highly non-linear relationships between inputs and outputs.

In the early 2010s, various researchers observed that ANNs had scarcely been employed to predict gasification product yields and properties [17,76]. Puig-Arnavat *et al.* explored the potential of ANNs to model the thermochemical conversion of biomass. Models for a circulating fluidised bed and bubbling fluidised bed gasifier were presented. For the training of the two models only wood gasification at atmospheric pressure was considered to ensure high data homogeneity. However, the collected data sets were small containing data from only 18 and 36 experimental runs. Despite the proposed models achieving a high predictive

accuracy ($R^2 > 0.97$), a major drawback is their limitation to a single type of feedstock and the specific range of gasifier conditions used in model training [76].

Baruah *et al.* (2017) addressed the issue of other studies (e.g. [76,77]) associated with the consideration of only one type of feedstock by using experimental data for different types of woody biomass for model development. Furthermore, the study identified that fixed bed downdraft gasifiers in particular had received little attention when it comes to ANN-based modelling despite being one of the most commonly employed types of gasifiers for small-scale applications. It is stated that expanding the used data set is desirable to improve the range of cases to which the model is applicable [57].

This lack of generalisation capability in earlier ANN studies in the field of biomass gasification was also identified by George *et al.* For this reason, the authors chose to train their model on five different feedstocks to ensure generalisation. The model considered the feedstock's carbon, hydrogen, oxygen, moisture, and ash content, as well as operating temperature and equivalence ratio as the key input parameters. However, it was acknowledged that a large number of parameters could be relevant and further exploration of factors such as the feedstock's polymeric composition or reactor geometry was essential. Additionally, it is stated that predicting other critical factors, which are essential to our understanding of the gasification process, such as tar and char yields might be an interesting direction for future research [58].

An attempt at including additional input parameters which are relevant to the process was recently made by including the reactors bed material as a parameter for model training. The effect of different bed materials has widely been studied but was never considered in an ANN before. Serrano *et al.* simply implemented this using ordinal encoding [62]. This technique assigns an integer value to each unique bed material option. This simplistic approach may result in acceptable results in certain cases, but for many cases more advanced methods to encode categorical variables are called for. Especially, parameters which do not have a clear order or ranking (e.g. bed material, gasifier type, or gasifying agent) need to be treated with care. For example, one-hot encoding is a commonly used method, which has been employed in other fields, to handle these types of variables [78]. One-hot encoding transforms a variable, which can take on *n* distinct categorical values, to *n* binary variables. A value of 1 indicates the presence of the variable and a value of 0 the absence.

Many early studies used trial and error to decide network structure and layout [24,67,71,79,80]. However, choices on network structure, training algorithms, threshold functions, etc. can all have a significant effect on the ANNs performance. Serrano *et al.* compared the effects of different activation functions and network topologies on the predictive power of ANNs. They compared two ANN types (multi-layer feed forward and cascade forward network) and found that no single combination resulted in the best performance for all parameters [62]. In another work the performance of a single hidden layer model and a two hidden layer model was compared [48]. It was found that the potential performance improvement obtained from a two hidden layer model was not justifiable considering the increase in computational power required [48]. Results such as this are case specific and it is important to consider and compare a range of different options.

Pandey *et al.* chose Levenberg-Marquardt back-propagation as the training algorithm for their ANN [49]. This training algorithm has been used in many comparable studies which also found that it performs well [19,24,48,49,60,79]. Nonetheless, a lack of comparison between training algorithms becomes apparent and future research should aim to improve this. However, the study attempted to improve previous works which often used trial and error or empirical equations (e.g. [58]) to decide on the optimum ANN configuration (e.g. number of hidden layers and nodes in hidden layers) by approaching network hyperparameter tuning in a more systematic fashion. For this, a procedure employing 100 Monte Carlo runs with different initial guesses, data segmentations, etc. was implemented to decide on the best network structure [49].

As previously mentioned, the prediction of other factors than syngas composition and yield is important to fully map the gasification process and extend the capacity of model-based process design. As tar generation can be one of the limiting factors hindering industrial scale deployment of gasification by affecting downstream applications, some authors attempted to predict tar generation using ANNs [48,49]. Serrano and Castelló identified the heterogenous nature of tar generation data, as well as its scarcity, as key reasons for the limited development of tar prediction models [48]. In their study three different models were trained for the different tar sampling methods, as well as a fourth trained on the entire data set. A data set of 120 data patterns was used for the model development. In general, the proposed ANN models were found to predict the tar generation well, with coefficients of determination >0.97 for test and validation and relative errors below 20% for most samples. The models

behaved well in most of the operating region, with exceptions in the boundary regions where insufficient data was available for the model to learn the data structure well [48].

A point frequently made when talking about ANNs is their "black box" nature which limits interpretability [48,49]. This means the process linking inputs to outputs is opaque. This makes it challenging to draw conclusions about the fundamental mechanisms of a considered process from the network's prediction. Many factors, such as heat and mass transfer effects, which are considered by e.g. kinetic models are not, and often cannot easily be, incorporated in ANNs [48]. There have been attempts to improve interpretability by evaluating the relative importance of impact factors. Garson's equation has frequently been used for this purpose [48,57,60,76,81]. It returns the percentage contribution of any given independent variable on the considered dependent variable. For this the absolute weights between the hidden-to-output layer nodes are partitioned into components linked to each input neuron [81]. The equation is further described in Appendix B.

The performance of Garson's equation for correctly identifying variable importance has previously been found to be lacking due to its use of absolute connection weight values [51,82]. It has been shown that different methods for quantifying variable importance may well result in different importance rankings [51]. For this reason, researchers should consider using a range of methods to be sure of their findings. One method that was found to perform well is the "Connection Weights" method introduced by Olden and Jackson [51,83]. Like Garson's equation, the input-hidden and hidden-output connection weights are used for the calculation. However, importantly, the raw connection weights are used instead of absolute values [51,83]. Another approach to improve interpretability of results is the hybridisation of theory-based models, such as thermodynamic equilibrium models, with ANNs. This approach is discussed in more detail in the following section (Section 4.1.1).

4.1.1 Hybridised use of ANNs with existing thermodynamic equilibrium models

To improve interpretability, some authors attempted to combine thermodynamic equilibrium models with ANN methodology. In the approach taken by Guo *et al.,* an ANN was used to fill in the gaps of a partial first principles model – i.e. the ANN was used as an estimator for parameters which could not be measured and proved challenging to model from first principles [84]. Whilst this approach is promising, this early study leaves open numerous questions. For example, inputs are matched with the corresponding outputs for most ANNs.

However, it was stated that the desired outputs were not measurable, making it unclear which parameters the inputs were matched with instead. Additionally, overfitting behaviour was identified but not further addressed. Finally, discrepancies between model and experimental results were identified but their size was not statistically quantified [84].

In a recent work [63], the role of ANN and thermodynamic equilibrium model was reversed as compared to the earlier work by Guo *et al.* [84]. The study identified that proper ANN model training required large amounts of data. To obtain the data, large-scale experiments are required which are both expensive and time-consuming. Hence, it was suggested that the data obtained from a thermodynamic equilibrium model should be used for ANN training [63,84].

The main benefits of this combined approach are simplicity and convenience. The approach allows for fast computation of many different input data sets once the model is sufficiently trained. This offers the user a convenient way to assess the effects of variations in for example feedstock composition due to different growing conditions or climatic differences on the systems net output power. Thus, intermediate calculation steps become obsolete. A very high R^2 >0.999 was achieved for the training and test stage of the ANN. This is not surprising as the training data comes from the equilibrium model which uses the same underlying equations throughout the data set. It is unclear why a validation stage was not included, as the data set appears to be sufficiently large to do so. Finally, it needs to be stressed that the underlying model is still a thermodynamic equilibrium model. As such the hybrid model is constrained by the shortcomings and assumptions made within the equilibrium model [63].

4.1.2 Time-series approaches

Many early ML-based models adopted a time independent approach [61]. Whilst this may predict the final outputs of the process well, a time dependent model is required for on-line process control and optimisation. Kinetic and CFD models are generally not suitable for on-line process control or implementation as model predictive controllers due to their computational intensiveness. Equilibrium models, on the other hand, are far less computationally intensive and represent a valid option for on-line control purposes. However, they only capture a snapshot of the process due to their time independent approach. Thus, they are better suited for the analysis and optimisation of a static process at set operating conditions [61,77].

Early studies used conventional ANNs for time dependent modelling. More recent studies utilised autoregressive exogeneous neural networks (NARX) which are more catered towards time-series data [60,61]. NARX models contain a feedback mechanism which allows them to learn a time-series by using input/output pairs of previous time steps as inputs to the current time step. In addition, this type of model uses an exogeneous time-series.

In general, researchers found that time-series approaches are better suited for on-line control and model predictive control than equilibrium models which have conventionally been used. This is especially true for NARX which were found to learn a system's dynamics well [60,61,77]. However, to train these types of models, all training data is obtained from experiments based on one specific reactor. The high homogeneity of the training data allows for greater accuracy, but also means that the model is only applicable to the specific reactor from which the training data has been obtained.

One strength of the NARX approach is its ability to easily implement real-time readings from sensors (e.g. thermocouples) installed throughout the reactor. For example, Yucel *et al.* [60] explored the effects of the temperature distribution throughout a fixed-bed downdraft gasifier on the syngas composition and HHV. It is generally accepted that the gasification temperature is one of the key parameters dictating the process. It was found that the temperature distribution within the gasifier accounted for 34.53% of the variation in the outputs and especially the temperatures of the reduction, combustion, and drying zones were found to be most influential [60].

Elmaz and Yücel compared a NARX with a simpler polynomial regression model [61]. It was found that the conventional validation methods used for ML applications, such as k-fold validation, might be suboptimal for a time-series approach. Other methods, such as rolling window analysis, may be better suited for time-series approaches to evaluate the stability and performance of the model [61].

In addition to choosing appropriate validation techniques future research considering time-series approaches needs to improve on previous researchers' works by: (i) using formal statistical descriptors (e.g. R^2 and RMSE) to evaluate model accuracy, (ii) providing sufficient description of training data and how it has been obtained, and (iii) describing the network/model structure to an appropriate level of detail. Improving on these points is essential for allowing other researchers to understand the methods employed and reproduce findings if necessary.

4.2 Hybrid Machine Learning Models

The hybridisation of ML models with other ML techniques or advanced statistical methods has shown great potential in the wider energy sector [85]. More recently, researchers have also looked towards hybrid models to model thermochemical processes [74,86,87]. Compared to ordinary models, hybrid models generally have higher accuracy and generalisation ability at the cost of an increase in computational load. For instance, ensemble methods, such as gradient boosting, can turn many weak and poorly performing models into an effective one [88]. A recent study by Kardani *et al.* [66] nicely illustrated how an ensemble model could improve model robustness despite only limited data being available for model training. Finally, incorporating statistical methods, such as Monte Carlo simulation, can greatly improve the credibility of a model's results by accounting for the related uncertainties [74]. Two frequently considered methods – particle swarm optimisation (PSO) for ANN training and dimensionality reduction algorithms – are reviewed in greater detail in Sections 4.2.1 and 4.2.2.

4.2.1 Use of particle swarm optimisation for network learning

As identified in Section 4.1, existing research often paid little attention to identifying the optimal network structure. As part of this, the effect of choosing the most suitable training algorithm has often been neglected. More recently, PSO has been considered as an advanced algorithm for model training [89]. PSO is a type of metaheuristic inspired by the movement of a fish school or flock of birds. As a metaheuristic, few or no assumptions are made about the problem which is to be optimised [90].

Chiñas-Palacios [86] explored the potential of a hybrid model combining PSO and ANN methodology to model a biomass gasification plant. The hybrid PSO neural network was compared to two more conventional networks: cascade forward propagation and feed forward back propagation. The optimal ANN configurations, for the various considered networks, were found after performing 215 simulation runs. The number of hidden neurons were varied between 3, 10, and 100 and the particle population for the PSO were varied between 10, 100, 600, and 1000. Intermediate values were not considered which might result in significant performance penalties. It was found that using PSO resulted in poor performance when using suboptimal particle populations. However, when using the optimal configuration, a mean squared error (MSE) reduction ranging from 11-37% was achieved for a range of parameters

when compared to the two more conventional ANN configurations. This highlights the potential performance improvements with PSO, as well as the importance of appropriate parameter choices [86].

In an earlier work, Aghbashlo *et al.* used a PSO algorithm for the tuning of a different type of neural network [71]. ANFIS methodology was used which combines the learning ability of neural networks with fuzzy logic principles. Upon using previously unseen biomass samples, their model was shown to be robust and adequately accurate with a R^2 value of >0.91 [71]. Obafemi *et al.* identified that there was a lack of comparison between different training algorithms [80]. In the past, many researchers simply chose the Levenberg-Marquardt algorithm for model training [49,58,62,91–93]. Future research needs to consider and compare different training options (e.g. Levenberg-Marquard algorithm, Bayesian regularisation, PSO, etc.) to find the most suitable one. Finally, PSO may also have the potential to automatically design an ANN by determining the most suitable hyperparameters [93]. This type of methodology may make choosing the optimal network structure and hyperparameters by trial and error obsolete.

4.2.2 Dimensionality reduction methods

Generally, the purpose of dimensionality reduction is the transformation of training data from a high- to low-dimensional space whilst retaining the data's variation. This is often desirable for predictive models with a large number of observations. In practice, principal component analysis (PCA) is the most commonly employed method, but other methods such as kernel-based methods or autoencoder methodology do exist [94]. For a more in-depth, general review on the topic, the reader is referred to Van Der Maaten *et al.* (2009) [94].

Recently, Shahbeig and Nosrati used PCA to identify the principal reactions occurring during the pyrolysis of biological waste. The authors showed that PCA allowed for the use of high dimensional experimental data, such as differential thermogravimetric data, for ML model training [87]. Elmaz *et al.* trained four different ML models to predict biomass gasification products. The training data set consisted of an initial 16 features which were reduced to three using principal component analysis. High R^2 values of up to 0.944 were achieved for the best performing ML methods, ANN and RF [20].

Hosseinpour *et al.* successfully combined an ANN-based model with PCA to predict biomass HHV based on the feedstock's proximate composition (i.e. fixed carbon, volatile matter, and

ash). Clearly, the initial data set already has low dimensionality. Thus, in this case, PCA was used to eliminate linear correlation between predictors before the data was used for ANN training [95].

It was found that linear techniques such as PCA do not always perform well with complex non-linear data. Some researchers showed that non-linear techniques outperformed linear techniques on selected artificial data sets. However, these performance improvements of non-linear techniques could not always be replicated on real-world data sets [94]. Nonetheless, non-linear techniques for dimensionality reduction may be a worthwhile direction to explore in the future for thermochemical process data sets.

4.3 Use of Other Machine Learning Methods

ML algorithms are commonly split into supervised and unsupervised learning. For this review, a further split by learning task (i.e. regression, classification, and clustering) is desirable. Identifying the learning task at hand allows the ML practitioner to select an appropriate ML algorithm. For gasification and pyrolysis modelling, most learning tasks are regression problems. A lot of the early research within the field considered solely ANNs (as reviewed in Section 4.1). However, more recently, other ML techniques have received increasing attention. The three most frequently employed methods beyond ANNs are discussed below in Sections 4.3.1-4.3.3.

4.3.1 Regression methods

A range of different types of regression analysis exist. Linear regression has been used for many years in statistics and is one of the simplest forms of analysis to assess the relationship between dependent and independent variables [45]. Like many other real-world phenomena, thermochemical processes are known to follow complex highly non-linear relationships. It is no surprise, therefore, that linear regression has been found to be inadequate for the modelling of thermochemical processes such as the pyrolysis of biomass [73].

However, regression methods are not limited to linear regression and other methods (e.g., polynomial regression, stepwise regression, and lasso regression) exist that are better suited for higher dimensionality data and the modelling of non-linear relationships [73]. One of the key benefits of regression-based methods compared to other ML methods is their lower

computational demands. Elmaz and Yücel compared a NARX model to a simpler polynomial regression model. The polynomial regression model did not perform as well achieving R^2>0.80 for all outputs as compared to R^2>0.98 for the NARX model outputs [61]. In a separate study, even larger performance differences were identified upon comparing quadratic and cubic polynomial regression with support vector regression, decision tree regression and an ANN [20]. However, a less accurate model may still have its merits especially when it displays a much lower complexity and computational cost. Importantly, the authors noted that the polynomial regression model was still able to predict trends in the data. This makes this type of model well suited for real-time model predictive control applications implemented with a conventional microcontroller. Ultimately, required model accuracy and computational demands need to be weighed against each other to find the most appropriate approach in the face of trade-offs.

4.3.2 Decision tree-based methods

In practice, tree-based methods, such as regression trees or more advanced ensemble methods, have seen less frequent use than ANNs for the prediction of gasification or pyrolysis product yields and properties. However, in a recent comparison between the ANN and regression tree methodology for the prediction of syngas composition and heating value, both methods were found to achieve similar prediction accuracy [20]. This highlights the importance of comparing different ML algorithms on a case-by-case basis. When comparing various ML techniques, it is key to initially select techniques which are suitable for the problem at hand. For example, a recent study by Tang *et al.* [73] compared linear regression to RF methodology for the prediction of the bio-oil yield and bio-oil hydrogen content from biomass pyrolysis. The well-known non-linear nature of biomass pyrolysis made linear regression a poorly suited approach to the problem and a comparison with ANN or SVM methodology seemed more worthwhile.

Whilst a simple decision tree is easy to follow and understand, ensemble methods suffer from the same black box nature as ANNs [45]. However, interpretability of tree-based methods can still exceed that of ANNs, as the feature importance is generally more easily extracted. For example, Tang *et al.* [73] found that the biomass composition generally has a larger effect than the pyrolysis conditions on the process outputs. On the other hand, studies by Cheng *et al.* [74] and Zhu *et al.* [72] found that pyrolysis temperature had a strong effect on the

production of biochar from a range of biomass feedstocks. Interestingly, the findings by Zhu *et al.* illustrated that the feature importance obtained from RF methodology could significantly differ from the rankings obtained using Pearson correlation coefficients [72].

Cheng *et al.* [74] proposed the novel integration of a RF-based predictive model for pyrolysis with life cycle assessment and economic analysis. The aim of their study was to achieve a comprehensive evaluation of a range of different pyrolysis feedstocks on a more holistic basis by considering environmental and economic factors. This type of approach shows the power of ML techniques which allow for the rapid estimation of a range of different scenarios without the need for expensive experiments. This type of approach is not limited to RF, and other ML techniques and thermochemical processes may also be considered for similar implementations in the future.

Some other advantages of tree-based methods are their capability to handle missing predictor values, as well as the easy implementation of categorical values. Many authors used existing experimental studies to collect a process database for model training [62,68,69,72]. However, due to low homogeneity in published literature it can be challenging to amass a sufficiently large database without any missing values. Many ML methods struggle with handling missing predictor values. Tree based methods, on the other hand, are well suited for handling missing predictor values without having to discard observations which could reduce the size of the training set [45]. Furthermore, categorical variables which may have a significant effect on the model outputs are easily implemented in the tree structure of such models.

4.3.3 Support vector machines

SVMs have been commonly used for classification tasks. However, recently a number of researchers have looked towards SVM as an alternative to ANN for the modelling of thermochemical processes [20,67,69,87]. Upon the comparison between SVM and ANN, both Cao *et al.* [67] and Chen *et al.* [69] found that SVM showed better performance, in terms of RMSE and R^2, at predicting pyrolysis product distributions and properties. However, it is important to note that these results cannot necessarily be generalised. The two studies used small data sets of 33 and 71/166/175 (for the prediction of char and gas yield, bio-oil yield, and bio-oil HHV, respectively) data samples, respectively. SVM are known to generally perform well in a high dimensional space with small to medium sized data sets [55,56]. Correspondingly, the larger data sets used in Chen *et al.*'s [69] study reduced the

performance difference between SVM and ANN. In contrast, Elmaz *et al.* [20] gave the edge to ANN. The authors found that a multilayer perceptron performed better than a SVM at predicting the syngas composition and HHV from biomass gasification. The key difference between this study and [67,69] is the significantly larger data set that consists of 4,826 time-dependent data samples.

For the case of small data sets, SVM can show better robustness to overfitting than ANN as suggested by Cao *et al.* [67] and Chen *et al.*'s [69] studies where the ANN models had higher coefficients of determination and lower RMSE during model training than SVM. However, the subsequent decrease in performance during testing indicated overfitting.

Selecting an appropriate kernel function is essential in creating a well performing model. Choosing a linear function results in similar results to linear regression, making more advanced non-linear functions generally more attractive. The radial basis function has widely been accepted as a good option for the modelling of non-linear processes [96]. All studies reviewed here selected the radial basis function as the kernel function [20,67,69,87]. In conclusion, SVMs represents an interesting alternative to ANNs or other ML techniques, especially for small, high-dimensional data sets.

4.4 Use of Machine Learning in Bioenergy-related Fields

Researchers in a vast number of different fields have been looking towards ML for simulation and modelling tasks. Bioenergy systems tend to be complex systems spanning a range of fields. Hence, developing comprehensive ML models for e.g. sustainable waste management schemes will likely require knowledge across different disciplines.

A thorough review on how ML has been employed in related fields is beyond the scope of this review. However, several authors published dedicated reviews on these topics. For example, Guo *et al.* published a review on the use of ML methods in the field of recycling and organic solid waste treatment [97]. Various treatment methods were discussed, among them incineration, gasification, and pyrolysis are highlighted. The authors noticed that ANN were the most employed method accounting for more than half of all considered studies. Furthermore, it was noted that combining different ML techniques in integrated models or ML techniques with other innovative methods is seen as promising for improved performance. Finally, the authors stressed the importance of demystifying the "black box" of many ML techniques by, for example, quantifying the contributions of input parameters [97].

Sewsynker-Sukai *et al.* reviewed the use of ANNs for the modelling and optimisation of biofuels production through biological methods [98]. Coupling genetic algorithms with ANNs was seen as very attractive due to their power to solve a vast array of optimisation problems. Additionally, the authors discussed several papers comparing ANNs to response surface methodology for the optimisation of biofuel production processes. ANNs were found to considerably outperform models based on response surface methodology due to their capability to represent highly non-linear processes without being limited to quadratic equations. Finally, ANN applications in real-time process monitoring and control were suggested as a potential direction for further research [98].

Mosavi *et al.* (2019) shifted the focus from process modelling/optimisation applications to an energy systems approach. In their thorough review, a vast range of ML techniques for numerous research areas was considered. Reviewed ML applications ranged from wind speed forecasting and building energy demand estimations to district heating system optimisation. However, thermochemical biomass and waste conversion technologies were not discussed by the review. Two methods that stand out are ensemble methods and hybrid ML models due to their improved accuracy and generalisation ability. As a closing remark, the study pointed out that the case-specific development of ML models for a particular application resulted in the best performance despite requiring more time and effort [99].

One of the earliest uses of ANN has been for the prediction of feedstocks' heating values. Dong *et al.* predicted the LHV of MSW from its physical composition (i.e. weight fraction of different waste components). Upon comparing the proposed model to various conventional models based on the waste's physical, ultimate, and proximate compositions, the ANN model was found to predict the LHV best. In this early work, the distinction between the training, validation, and test stages is not clear [100].

A similar approach was used for the prediction of the gross calorific value of Indian coal. Proximate and ultimate data was used, and the model was found to have an improved performance when compared to conventional linear models. This is likely due to the fact that some parameters exhibited a linear relationship with the gross calorific value while others did not [101].

Hosseinpour *et al.* considered various soft computing techniques for the estimation of biomass' HHV using its proximate composition information [95]. All models performed well and had lower mean absolute percentage error scores than various empirical equations in the

literature. Upon comparing the various proposed models, the authors concluded that the advanced model combining ANN and partial least squares methodology performed best on the test data set (R^2 = 0.9732).

5. Challenges and Prospects

5.1 Challenges

Data availability for model development remains a significant challenge. Running experiments can be expensive and time consuming while usually a large amount of training data is required for developing ML-based models. Many studies collected experimental data from literature, but inconsistencies in how data is quoted can make this process challenging and subsequent comparison difficult. Furthermore, there is a trade-off between a model's prediction accuracy and its generalisation capability related to the type of training data used. Using highly homogeneous experimental data from e.g. only one type of gasifier generally results in a model with high predictive accuracy for modelling the process of the particular type of gasifier. However, this means the model is only applicable for the narrow type of training data used. It will be valuable to build more diverse models with a lower, but still satisfactory prediction accuracy. Again, this would increase training data requirements as the data needs to span a wider range of parameters.

A problem previously pointed out is the black box nature of some ML algorithms which limits data and model interpretability. Whilst simple methods, such as linear regression, do not suffer from this problem they have also been shown to have non-satisfactory prediction accuracy [73]. ANNs which have been the most employed algorithm can make it challenging to draw valuable mechanistic conclusions from their findings. Efforts have been made at illuminating ANNs and other ML methods [51,83,102]. Section 4.1 described the popularity of Garson's equation to determine the variable importance of input features. As this method has been found to be lacking, more recent approaches need to be considered to improve interpretability [51,82]. One such option is the use of a global surrogate model for which a simpler, more interpretable model is trained to approximate the behaviour of a more complex black box model [52]. Another option is feature permutation to assess the importance of different input parameters. This method shuffles the values of one feature at a time and computes the resulting increase in the prediction error.

A lack of clarity in published literature is another factor holding back the successful deployment of more ML models and the improvement of existing ones. For instance, a lack of clarity in how data is split for model training, validation, and testing has been identified. Additionally, some studies have been unclear on which stage their results refer to. The use of various statistical parameters to assess a model's performance can make comparison between models challenging. Generally, R^2 and RMSE have been used, but some studies used, for instance R [58] and MSE [49,58]. Not all studies quoted error values in normalised form [67]. Future studies should consider using R^2 (or adjusted R^2) and RMSE in normalised form to facilitate model comparison. Quoting error values as percentage errors (e.g. root mean squared percentage errors (RMSPE)) is another option which is generalisable and allows for an intuitive assessment of the error's size. Finally, vast differences in training data make model comparison challenging.

5.2 Prospects

Further work on illuminating the black box of ML algorithms to enhance explainability is considered essential. This way new interconnections between a system's inputs and outputs might be learned. The development of more comprehensive prediction models, capable of predicting a larger range of relevant factors, is essential for the creation of more holistic models. Understanding how the gasification/pyrolysis process ties into a larger scheme or system analysis will encourage deployment of these technologies by minimising uncertainties for stakeholders. For example, Cheng *et al.* used the results from a pyrolysis prediction model in a larger framework incorporating LCA and economic analysis [74]. The successful applications of ML-based methods for facilitating the design and optimisation of pyrolysis and gasification whole system will shape the direction of future research and development.

Existing models may be improved by considering a wider range of training algorithms, approaching hyperparameter tuning in a more systematic fashion, and incorporating additional relevant input parameters such as categorical data into the model structure. Another promising avenue of research has been the combination of various ML methods or the combination of ML methods with other advanced methods. One such example is the use of PSO for model training and optimisation (Section 4.2.1). Since data sets for model training could often be limited in size, ensemble methods might be another good option to improve model robustness as illustrated by Kardani *et al.* [66].

For control purposes models with a low computational demand (e.g. regression models) and acceptable accuracy may allow for better real-time process control as compared to conventional models. Conversely, more demanding and accurate models such as NARX can also be used for process optimisation. Combining these could allow for system scale up and incentivise a more widespread deployment of the studied technologies. Importantly, future models need to find a balance between computational demands and accuracy to allow for their implementation through conventional microcontrollers.

6. Conclusions

ML methods have shown great potential towards modelling the gasification and pyrolysis processes and predicting the processes' product yields and properties. Their predictive performance is already on a similar or even superior level to that of conventional models. Thus, it should not come as a surprise that the number of publications increased drastically over the past two decades. In the period from 2000-2009 a total of 70 papers have been published. On the other hand, 549 publications were found between 2010 and 2019 which represents an increase of 684%.

Whilst R^2 values can be difficult to compare, due to some studies having vastly different approaches and aims, it has generally been found that most studies achieved a high prediction accuracy with $R^2 > 0.90$. Despite their high accuracy, ML models do not lend themselves to interpretation, as for instance thermodynamic equilibrium models do, and it has thus been challenging to interpret modelling results at times. Developing more interpretable models is essential in improving the underlying understanding and design capacity of the processes.

Utilising the strong predictive performance of ML models by incorporating them as part of more comprehensive system models has the potential to encourage industrial scale deployment by minimising uncertainties for investors and policymakers. For this, the availability of sufficient training data is essential and considering the nature and availability of training data is key in deciding on the most appropriate ML method for any given use case. Finally, as data availability and computational power increase, new avenues for research become available. For instance, more complex models may be used for real-time process control and optimisation applications.

Acknowledgements

The authors would like to acknowledge the financial support from the Engineering and Physical Sciences Research Council (EPSRC) Studentship and Programme Grant (EP/V030515/1).

References

[1] International Energy Agency. IEA - World Energy Balances 2021. https://www.iea.org/data-and-statistics?country=WORLD&fuel=Energy supply&indicator=TPESbySource (accessed March 17, 2021).

[2] Basu P. Biomass gasification, pyrolysis and torrefaction: Practical design and theory. 2nd ed. Academic Press; 2018. https://doi.org/10.1016/C2016-0-04056-1.

[3] Patterson T, Esteves S, Dinsdale R, Guwy A. Life cycle assessment of biogas infrastructure options on a regional scale. Bioresour Technol 2011;102:7313–7323. https://doi.org/10.1016/j.biortech.2011.04.063.

[4] Moharir R V., Gautam P, Kumar S. Waste Treatment Processes/Technologies for Energy Recovery. In: Kumar S, Kuma R, Pandey A, editors. Curr. Dev. Biotechnol. Bioeng. Waste Treat. Process. Energy Gener., 2019, p. 53–77.

[5] Bridgwater A V, Toft AJ, Brammer JG. A techno-economic comparison of power production by biomass fast pyrolysis with gasification and combustion. Renew Sustain Energy Rev 2002;6:181–248.

[6] Searchinger TD. Biofuels and the need for additional carbon. Environ Res Lett 2010;5:24007. https://doi.org/10.1088/1748-9326/5/2/024007.

[7] Mathews JA. Carbon-negative biofuels. Energy Policy 2008;36:940–5. https://doi.org/10.1016/j.enpol.2007.11.029.

[8] Fargione J, Hill J, Tilman D, Polasky S, Hawthorne P. Land clearing and the biofuel carbon debt. Science (80-) 2008;319:1235–8. https://doi.org/10.1126/science.1152747.

[9] Lehmann J. A handful of carbon. Nature 2007;447:143–4. https://doi.org/10.1038/447143a.

[10] Roberts KG, Gloy BA, Joseph S, Scott NR, Lehmann J. Life cycle assessment of biochar systems: Estimating the energetic, economic, and climate change potential. Environ Sci Technol 2010;44:827–33. https://doi.org/10.1021/es902266r.

[11] Matovic D. Biochar as a viable carbon sequestration option: Global and Canadian perspective. Energy 2011;36:2011–6. https://doi.org/10.1016/j.energy.2010.09.031.

[12] Young GC. Introduction to Gasification/Pyrolysis and Combustion Technology(s).

Munic. Solid Waste to Energy Convers. Process. Econ. Tech. Renew. Comp., Wiley; 2010.

[13] Fiori L, Florio L. Gasification and combustion of grape marc: Comparison among different scenarios. Waste and Biomass Valorization 2010;1:191–200. https://doi.org/10.1007/s12649-010-9025-7.

[14] Dong J, Tang Y, Nzihou A, Chi Y, Weiss-Hortala E, Ni M. Life cycle assessment of pyrolysis, gasification and incineration waste-to-energy technologies: Theoretical analysis and case study of commercial plants. Sci Total Environ 2018;626:744–53. https://doi.org/10.1016/j.scitotenv.2018.01.151.

[15] Arena U, Ardolino F, Di Gregorio F. A life cycle assessment of environmental performances of two combustion- and gasification-based waste-to-energy technologies. Waste Manag 2015;41:60–74. https://doi.org/10.1016/j.wasman.2015.03.041.

[16] Analytics C. Web of Science 2021. https://apps.webofknowledge.com/UA_GeneralSearch_input.do?product=UA&search_mode=GeneralSearch&SID=F5PfPgYdxUvvmDBoMwA&preferencesSaved= (accessed March 17, 2021).

[17] Patra TK, Sheth PN. Biomass gasification models for downdraft gasifier: A state-of-the-art review. Renew Sustain Energy Rev 2015;50:583–93. https://doi.org/10.1016/j.rser.2015.05.012.

[18] Safarian S, Unnþórsson R, Richter C. A review of biomass gasification modelling. Renew Sustain Energy Rev 2019;110:378–91. https://doi.org/10.1016/j.rser.2019.05.003.

[19] Puig-Arnavat M, Bruno JC. Artificial Neural Networks for Thermochemical Conversion of Biomass. Recent Adv. Thermochem. Convers. Biomass, 2015, p. 133–56. https://doi.org/10.1016/B978-0-444-63289-0.00005-3.

[20] Elmaz F, Yücel Ö, Mutlu AY. Predictive modeling of biomass gasification with machine learning-based regression methods. Energy 2020;191:116541. https://doi.org/10.1016/j.energy.2019.116541.

[21] Puig-Arnavat M, Bruno JC, Coronas A. Review and analysis of biomass gasification models. Renew Sustain Energy Rev 2010;14:2841–51.

https://doi.org/10.1016/j.rser.2010.07.030.

[22] Baruah D, Baruah DC. Modeling of biomass gasification: A review. Renew Sustain Energy Rev 2014;39:806–15. https://doi.org/10.1016/j.rser.2014.07.129.

[23] Ramos A, Monteiro E, Rouboa A. Numerical approaches and comprehensive models for gasification process: A review. Renew Sustain Energy Rev 2019;110:188–206. https://doi.org/10.1016/j.rser.2019.04.048.

[24] Sreejith CC, Muraleedharan C, Arun P. Performance prediction of fluidised bed gasification of biomass using experimental data-based simulation models. Biomass Convers Biorefinery 2013;3:283–304. https://doi.org/10.1007/s13399-013-0083-5.

[25] Kostetskyy P, Broadbelt LJ. Progress in Modeling of Biomass Fast Pyrolysis: A Review. Energy & Fuels 2020;34:15195–216. https://doi.org/10.1021/acs.energyfuels.0c02295.

[26] Hameed S, Sharma A, Pareek V, Wu H, Yu Y. A review on biomass pyrolysis models: Kinetic, network and mechanistic models. Biomass and Bioenergy 2019;123:104–22. https://doi.org/10.1016/j.biombioe.2019.02.008.

[27] Kaczor Z, Buliński Z, Werle S. Modelling approaches to waste biomass pyrolysis: a review. Renew Energy 2020;159:427–43. https://doi.org/10.1016/j.renene.2020.05.110.

[28] Sharma A, Pareek V, Zhang D. Biomass pyrolysis - A review of modelling, process parameters and catalytic studies. Renew Sustain Energy Rev 2015;50:1081–96. https://doi.org/10.1016/j.rser.2015.04.193.

[29] Babu B V. Biomass pyrolysis: A state-of-the-art review. Biofuels, Bioprod Biorefining 2008;2:393–414. https://doi.org/10.1002/bbb.92.

[30] Kan T, Strezov V, Evans TJ. Lignocellulosic biomass pyrolysis: A review of product properties and effects of pyrolysis parameters. Renew Sustain Energy Rev 2016;57:1126–40. https://doi.org/10.1016/j.rser.2015.12.185.

[31] Dhyani V, Bhaskar T. A comprehensive review on the pyrolysis of lignocellulosic biomass. Renew Energy 2018;129:695–716. https://doi.org/10.1016/j.renene.2017.04.035.

[32] Bridgwater A V. Review of fast pyrolysis of biomass and product upgrading. Biomass

and Bioenergy 2012;38:68–94. https://doi.org/10.1016/j.biombioe.2011.01.048.

[33] Kumar A, Jones DD, Hanna MA. Thermochemical biomass gasification: A review of the current status of the technology. Energies 2009;2:556–81. https://doi.org/10.3390/en20300556.

[34] Arena U. Process and technological aspects of municipal solid waste gasification. A review. Waste Manag 2012;32:625–639. https://doi.org/10.1016/j.wasman.2011.09.025.

[35] Li J, Yan R, Xiao B, Wang X, Yang H. Influence of temperature on the formation of oil from pyrolyzing palm oil wastes in a fixed bed reactor. Energy and Fuels 2007;21:2398–407. https://doi.org/10.1021/ef060548c.

[36] Uddin MN, Techato K, Taweekun J, Rahman MM, Rasul MG, Mahlia TMI, et al. An overview of recent developments in biomass pyrolysis technologies. Energies 2018;11:3115. https://doi.org/10.3390/en11113115.

[37] Foong SY, Liew RK, Yang Y, Cheng YW, Yek PNY, Wan Mahari WA, et al. Valorization of biomass waste to engineered activated biochar by microwave pyrolysis: Progress, challenges, and future directions. Chem Eng J 2020;389:124401. https://doi.org/10.1016/j.cej.2020.124401.

[38] Sobek S, Werle S. Solar pyrolysis of waste biomass: Part 1 reactor design. Renew Energy 2019;143:1939–48. https://doi.org/10.1016/j.renene.2019.06.011.

[39] Wan Mahari WA, Azwar E, Foong SY, Ahmed A, Peng W, Tabatabaei M, et al. Valorization of municipal wastes using co-pyrolysis for green energy production, energy security, and environmental sustainability: A review. Chem Eng J 2021;421:129749. https://doi.org/10.1016/J.CEJ.2021.129749.

[40] Molino A, Larocca V, Chianese S, Musmarra D. Biofuels Production by Biomass Gasification: A Review. Energies 2018;11:811. https://doi.org/10.3390/en11040811.

[41] Cao L, Yu IKM, Xiong X, Tsang DCW, Zhang S, Clark JH, et al. Biorenewable hydrogen production through biomass gasification: A review and future prospects. Environ Res 2020;186:109547. https://doi.org/10.1016/j.envres.2020.109547.

[42] Shayan E, Zare V, Mirzaee I. Hydrogen production from biomass gasification; a theoretical comparison of using different gasification agents. Energy Convers Manag 2018;159:30–41. https://doi.org/10.1016/j.enconman.2017.12.096.

[43] Ren J, Cao JP, Zhao XY, Yang FL, Wei XY. Recent advances in syngas production from biomass catalytic gasification: A critical review on reactors, catalysts, catalytic mechanisms and mathematical models. Renew Sustain Energy Rev 2019;116:109426. https://doi.org/10.1016/j.rser.2019.109426.

[44] Valderrama Rios ML, González AM, Lora EES, Almazán del Olmo OA. Reduction of tar generated during biomass gasification: A review. Biomass and Bioenergy 2018;108:345–70. https://doi.org/10.1016/j.biombioe.2017.12.002.

[45] Hastie T, Tibshirani R, Friedman J. The Elements of Statistical Learning. Second. New York, NY: Springer New York; 2009. https://doi.org/10.1007/978-0-387-84858-7.

[46] Schmidhuber J. Deep Learning in neural networks: An overview. Neural Networks 2015;61:85–117. https://doi.org/10.1016/j.neunet.2014.09.003.

[47] Nielsen MA. Neural Networks and Deep Learning. Determination Press; 2015.

[48] Serrano D, Castelló D. Tar prediction in bubbling fluidized bed gasification through artificial neural networks. Chem Eng J 2020;402:126229. https://doi.org/10.1016/j.cej.2020.126229.

[49] Pandey DS, Das S, Pan I, Leahy JJ, Kwapinski W. Artificial neural network based modelling approach for municipal solid waste gasification in a fluidized bed reactor. Waste Manag 2016;58:202–13. https://doi.org/10.1016/j.wasman.2016.08.023.

[50] Li J, Pan L, Suvarna M, Tong YW, Wang X. Fuel properties of hydrochar and pyrochar: Prediction and exploration with machine learning. Appl Energy 2020;269:115116. https://doi.org/10.1016/j.apenergy.2020.115166.

[51] Olden JD, Joy MK, Death RG. An accurate comparison of methods for quantifying variable importance in artificial neural networks using simulated data. Ecol Modell 2004;178:389–97. https://doi.org/10.1016/j.ecolmodel.2004.03.013.

[52] Molnar C. Interpretable Machine Learning - A Guide for Making Black Box Models Explainable. 2019.

[53] Breiman L. Random Forests. Mach Learn 2001;45:5–32.

[54] Drucker H, Burges CJC, Kaufman L, Smola A, Vapoik V. Support Vector Regression Machines. Adv Neural Inf Process Syst 1996;9:155–61.

[55] Vapnik VN. The Nature of Statistical Learning Theory. Springer New York; 2000.

https://doi.org/10.1007/978-1-4757-3264-1.

[56] Noori R, Abdoli MA, Ameri Ghasrodashti A, Jalili Ghazizade M. Prediction of municipal solid waste generation with combination of support vector machine and principal component analysis: A case study of mashhad. Environ Prog Sustain Energy 2009;28:249–58. https://doi.org/10.1002/ep.10317.

[57] Baruah D, Baruah DC, Hazarika MK. Artificial neural network based modeling of biomass gasification in fixed bed downdraft gasifiers. Biomass and Bioenergy 2017;98:264–71. https://doi.org/10.1016/j.biombioe.2017.01.029.

[58] George J, Arun P, Muraleedharan C. Assessment of producer gas composition in air gasification of biomass using artificial neural network model. Int J Hydrogen Energy 2018;43:9558–68. https://doi.org/10.1016/j.ijhydene.2018.04.007.

[59] Mutlu AY, Yucel O. An artificial intelligence based approach to predicting syngas composition for downdraft biomass gasification. Energy 2018;165:895–901. https://doi.org/10.1016/j.energy.2018.09.131.

[60] Yucel O, Aydin ES, Sadikoglu H. Comparison of the different artificial neural networks in prediction of biomass gasification products. Int J Energy Res 2019;43:5992–6003. https://doi.org/10.1002/er.4682.

[61] Elmaz F, Yücel Ö. Data-driven identification and model predictive control of biomass gasification process for maximum energy production. Energy 2020;195:117037. https://doi.org/10.1016/j.energy.2020.117037.

[62] Serrano D, Golpour I, Sánchez-Delgado S. Predicting the effect of bed materials in bubbling fluidized bed gasification using artificial neural networks (ANNs) modeling approach. Fuel 2020;266:117021. https://doi.org/10.1016/j.fuel.2020.117021.

[63] Safarian S, Mohammad S, Saryazdi E, Unnthorsson R, Richter C. Artificial neural network integrated with thermodynamic equilibrium modeling of downdraft biomass gasification-power production plant. Energy 2020;213:118800. https://doi.org/10.1016/j.energy.2020.118800.

[64] Ayodele BV, Mustapa SI, Kanthasamy R, Zwawi M, Cheng CK. Modeling the prediction of hydrogen production by co-gasification of plastic and rubber wastes using machine learning algorithms. Int J Energy Res 2021;45:9580–94. https://doi.org/10.1002/er.6483.

[65] Shenbagaraj S, Sharma PK, Sharma AK, Raghav G, Kota KB, Veeramuthu A. Gasification of food waste in supercritical water: An innovative synthesis gas composition prediction model based on Artificial Neural Networks. Int J Hydrogen Energy 2021;46:12739–57. https://doi.org/10.1016/j.ijhydene.2021.01.122.

[66] Kardani N, Zhou A, Nazem M, Lin X. Modelling of municipal solid waste gasification using an optimised ensemble soft computing model. Fuel 2021;289:119903. https://doi.org/10.1016/j.fuel.2020.119903.

[67] Cao H, Xin Y, Yuan Q. Prediction of biochar yield from cattle manure pyrolysis via least squares support vector machine intelligent approach. Bioresour Technol 2016;202:158–64. https://doi.org/10.1016/j.biortech.2015.12.024.

[68] Sunphorka S, Chalermsinsuwan B, Piumsomboon P. Artificial neural network model for the prediction of kinetic parameters of biomass pyrolysis from its constituents. Fuel 2017;193:142–58. https://doi.org/10.1016/j.fuel.2016.12.046.

[69] Chen X, Zhang H, Song Y, Xiao R. Prediction of product distribution and bio-oil heating value of biomass fast pyrolysis. Chem Eng Process - Process Intensif 2018;130:36–42. https://doi.org/10.1016/j.cep.2018.05.018.

[70] Çepelioğullar Ö, Mutlu İ, Yaman S, Haykiri-Acma H. Activation energy prediction of biomass wastes based on different neural network topologies. Fuel 2018;220:535–45. https://doi.org/10.1016/j.fuel.2018.02.045.

[71] Aghbashlo M, Tabatabaei M, Nadian MH, Davoodnia V, Soltanian S. Prognostication of lignocellulosic biomass pyrolysis behavior using ANFIS model tuned by PSO algorithm. Fuel 2019;253:189–98. https://doi.org/10.1016/j.fuel.2019.04.169.

[72] Zhu X, Li Y, Wang X. Machine learning prediction of biochar yield and carbon contents in biochar based on biomass characteristics and pyrolysis conditions. Bioresour Technol 2019;288:121527. https://doi.org/10.1016/j.biortech.2019.121527.

[73] Tang Q, Chen Y, Yang H, Liu M, Xiao H, Wu Z, et al. Prediction of Bio-oil Yield and Hydrogen Contents Based on Machine Learning Method: Effect of Biomass Compositions and Pyrolysis Conditions. Energy & Fuels 2020;34:11050–60. https://doi.org/10.1021/acs.energyfuels.0c01893.

[74] Cheng F, Luo H, Colosi LM. Slow pyrolysis as a platform for negative emissions technology: An integration of machine learning models, life cycle assessment, and

economic analysis. Energy Convers Manag 2020;223:113258. https://doi.org/10.1016/j.enconman.2020.113258.

[75] Gu C, Wang X, Song Q, Li H, Qiao Y. Prediction of gas-liquid-solid product distribution after solid waste pyrolysis process based on artificial neural network model. Int J Energy Res 2021:1–15. https://doi.org/10.1002/er.6707.

[76] Puig-Arnavat M, Hernández JA, Bruno JC, Coronas A. Artificial neural network models for biomass gasification in fluidized bed gasifiers. Biomass and Bioenergy 2013;49:279–89. https://doi.org/10.1016/j.biombioe.2012.12.012.

[77] Mikulandrić R, Lončar D, Böhning D, Böhme R, Beckmann M. Artificial neural network modelling approach for a biomass gasification process in fixed bed gasifiers. Energy Convers Manag 2014;87:1210–23. https://doi.org/10.1016/j.enconman.2014.03.036.

[78] Pai C, Potdar K. A Comparative Study of Categorical Variable Encoding Techniques for Neural Network Classifiers. Artic Int J Comput Appl 2017;175:975–8887. https://doi.org/10.5120/ijca2017915495.

[79] Buyukada M. Co-combustion of peanut hull and coal blends: Artificial neural networks modeling, particle swarm optimization and Monte Carlo simulation. Bioresour Technol 2016;216:280–6. https://doi.org/10.1016/j.biortech.2016.05.091.

[80] Obafemi O, Stephen A, Ajayi O, Nkosinathi M. A survey of artificial neural network-based prediction models for thermal properties of biomass. Procedia Manuf., vol. 33, Elsevier B.V.; 2019, p. 184–91. https://doi.org/10.1016/j.promfg.2019.04.103.

[81] Garson GD. A Comparison of Neural Network and Expert Systems Algorithms with Common Multivariate Procedures for Analysis of Social Science Data. Sociel Sci Comput Rev 1991;9:399–434.

[82] Gevrey M, Dimopoulos I, Lek S. Review and comparison of methods to study the contribution of variables in artificial neural network models. Ecol Modell 2003;160:249–64. https://doi.org/10.1016/S0304-3800(02)00257-0.

[83] Olden JD, Jackson DA. Illuminating the “‘black box’”: a randomization approach for understanding variable contributions in artificial neural networks. Ecol Modell 2002;154:135–50.

[84] Guo B, Li D, Cheng C, Lü ZA, Shen Y. Simulation of biomass gasification with a

hybrid neural network model. Bioresour Technol 2001;76:77–83. https://doi.org/10.1016/S0960-8524(00)00106-1.

[85] Scheidt F vom, Medinová H, Ludwig N, Richter B, Staudt P, Weinhardt C. Data analytics in the electricity sector – A quantitative and qualitative literature review. Energy AI 2020;1:100009. https://doi.org/10.1016/j.egyai.2020.100009.

[86] Chiñas-Palacios C, Vargas-Salgado C, Aguila-Leon J, Hurtado-Pérez E. A cascade hybrid PSO feed-forward neural network model of a biomass gasification plant for covering the energy demand in an AC microgrid. Energy Convers Manag 2021;232:113896. https://doi.org/10.1016/j.enconman.2021.113896.

[87] Shahbeig H, Nosrati M. Pyrolysis of biological wastes for bioenergy production: Thermo-kinetic studies with machine-learning method and Py-GC/MS analysis. Fuel 2020;269:117238. https://doi.org/10.1016/j.fuel.2020.117238.

[88] Friedman JH. Stochastic gradient boosting. Comput Stat Data Anal 2002;38:367–78. https://doi.org/10.1016/S0167-9473(01)00065-2.

[89] Grimaldi EA, Grimaccia F, Mussetta M, Zich RE. PSO as an effective learning algorithm for neural network applications. ICCEA 2004 - 2004 3rd Int. Conf. Comput. Electromagn. its Appl. Proc., 2004, p. 557–60. https://doi.org/10.1109/iccea.2004.1459416.

[90] Clerc M. Particle swarm optimization. Wiley & Sons; 2010.

[91] Ozveren U. An artificial intelligence approach to predict gross heating value of lignocellulosic fuels. J Energy Inst 2017;90:397–407. https://doi.org/10.1016/j.joei.2016.04.003.

[92] Gong S, Sasanipour J, Shayesteh MR, Eslami M, Baghban A. Radial basis function artificial neural network model to estimate higher heating value of solid wastes. Energy Sources, Part A Recover Util Environ Eff 2017;39:1778–84. https://doi.org/10.1080/15567036.2017.1370513.

[93] Garro BA, Vázquez RA. Designing Artificial Neural Networks Using Particle Swarm Optimization Algorithms. Comput Intell Neurosci 2015;2015. https://doi.org/10.1155/2015/369298.

[94] Van Der Maaten L, Postma E, Van Den Herik J. Dimensionality Reduction: A Comparative Review. J Mach Learn Res 2009;10:66–71.

[95] Hosseinpour S, Aghbashlo M, Tabatabaei M, Mehrpooya M. Estimation of biomass higher heating value (HHV) based on the proximate analysis by using iterative neural network-adapted partial least squares (INNPLS). Energy 2017;138:473–9. https://doi.org/10.1016/j.energy.2017.07.075.

[96] Suykens JAK, Vandewalle J. Least squares support vector machine classifiers. Neural Process Lett 1999;9:293–300. https://doi.org/10.1023/A:1018628609742.

[97] Guo H, Wu S, Tian Y, Zhang J, Liu H. Application of machine learning methods for the prediction of organic solid waste treatment and recycling processes: A review. Bioresour Technol 2021;319:124114. https://doi.org/10.1016/j.biortech.2020.124114.

[98] Sewsynker-Sukai Y, Faloye F, Kana EBG. Artificial neural networks: an efficient tool for modelling and optimization of biofuel production (a mini review). Biotechnol Biotechnol Equip 2017;31:221–35. https://doi.org/10.1080/13102818.2016.1269616.

[99] Mosavi A, Salimi M, Ardabili SF, Rabczuk T, Shamshirband S, Varkonyi-Koczy AR. State of the art of machine learning models in energy systems, a systematic review. Energies 2019;12:1301. https://doi.org/10.3390/en12071301.

[100] Dong C, Jin B, Li D. Predicting the heating value of MSW with a feed forward neural network. Waste Manag 2003;23:103–6. https://doi.org/10.1016/S0956-053X(02)00162-9.

[101] Patel SU, Jeevan Kumar B, Badhe YP, Sharma BK, Saha S, Biswas S, et al. Estimation of gross calorific value of coals using artificial neural networks. Fuel 2007;86:334–44. https://doi.org/10.1016/j.fuel.2006.07.036.

[102] Postma GJ, Krooshof PWT, Buydens LMC. Opening the kernel of kernel partial least squares and support vector machines. Anal Chim Acta 2011;705:123–34. https://doi.org/10.1016/j.aca.2011.04.025.