



A comprehensive artificial neural network model for gasification process prediction

Simon Ascher, William Sloan, Ian Watson, Siming You*

James Watt School of Engineering, University of Glasgow, University Avenue, Glasgow G12 8QQ, UK

HIGHLIGHTS

- Artificial Neural Network model has been used to predict 10 key gasification outputs.
- Model suitable across a wide range of feedstock, reactor, and gasifying agent options.
- The importance of categorical predictor variables was demonstrated.
- Optimisation routine successfully identified the best network structure.

ARTICLE INFO

Keywords:

Gasification
Biomass
Waste
Model
Machine learning
Artificial neural network

ABSTRACT

The viability and the relative merits of competing biomass and waste gasification schemes depends on a complex mix of interacting factors. Conventional analytical methods that are used to aid decision making rely on a plethora of poorly defined parameters. Here we develop a method that eschews the uncertainty in process representation by using a machine learning, data driven, approach to predicting a set of 10 key measures of gasification technology's performance. We develop an artificial neural network that is novel in its use of both categorical and continuous data inputs, which makes it flexible and broadly applicable in assessing gasification process designs. It is the first model applicable to a wide range of feedstock types, gasifying agents, and reactor options. A strong predictive performance, quantified by a coefficient of determination (R^2) of 0.9310, was confirmed. The approach has the potential to generate accurate input data for e.g., cost-benefit analysis (CBA) and life cycle sustainability assessment (LCSA) and thus allow for more transparency in the decisions made by policy makers and investors.

1. Introduction

The provision of sustainable energy and the transition to more sustainable waste treatment methods are two major challenges of the 21st century. In 2018 The World Bank estimated an annual production of 2.01 billion tonnes of municipal solid waste (MSW), a third of which is not managed in environmentally safe ways [1]. Due to climate change concerns, non-fossil fuel alternatives for the generation of energy are of greater importance than ever. Biomass as an abundant resource has shown great potential to aid with this transition in both developed and developing countries [2].

Thermochemical processes, such as gasification, represent a suitable way to tackle both problems by recovering energy from biomass and waste feedstocks [3]. Gasification converts a carbonaceous feedstock

into a raw gas, which is often referred to as syngas, composed of N_2 , H_2 , CO , CO_2 , CH_4 , C_2H_n and a solid product often referred to as biochar.

Whilst the gasification of various biomass and waste feedstocks has been studied extensively, the technology's wider environmental, economic, and social impacts need to be better understood to allow for scale-up and efficient implementation. For this, accurate process models are essential. In the past thermodynamic equilibrium, kinetic, and computational fluid dynamics (CFD) models have widely and successfully been used for this purpose. A more recent alternative has been the use of machine learning (ML) models, such as artificial neural networks (ANNs), which can effectively map highly non-linear input-output relationships and thus have the potential to predict gasification outputs more accurately than conventional models [4].

The use of ML within the field of bioenergy has just started gaining

* Corresponding author.

E-mail address: siming.you@glasgow.ac.uk (S. You).

<https://doi.org/10.1016/j.apenergy.2022.119289>

Received 14 December 2021; Received in revised form 26 April 2022; Accepted 13 May 2022

Available online 25 May 2022

0306-2619/© 2022 The Author(s). Published by Elsevier Ltd. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

momentum. ML methods have been shown to be an effective tool to optimise thermochemical processes [5]. Prediction models can also ensure the generation of products with desired properties by identifying suitable operating conditions [6,7]. Other opportunities for innovation through ML are, among others, on-line process control and waste-to-energy network design. For instance, image data has successfully been used to quantify the major gas species concentrations in a gasifier in real-time [8]. Robust data-driven optimisation models for waste-to-energy network designs can aid with finding the most optimal waste-to-energy networks in the light of uncertainties such as varying feedstock prices [9].

A neural network can refer to many different models. In this field a simple multilayer perceptron structure has generally been employed, which is made up of neurons arranged in an input and output layer and one or more hidden layers [4,10–15]. These existing studies showed that ANNs can predict gasification outputs with a high prediction accuracy. For example, George *et al.* trained an ANN model on lab scale data from a bubbling fluidised bed gasifier [11]. Their model was shown to predict the syngas production from local biomass with a high accuracy with a regression coefficient of $R = 0.987$ and mean squared error of $MSE = 0.71$. Baruah *et al.* developed an ANN model to predict the H_2 , CO , CO_2 , and CH_4 yields from a downdraft gasifier [12]. The model predicted the outputs excellently, with coefficients of determination of $R^2 > 0.98$ and an average relative error of 2.65% when compared to experimental data.

However, existing studies were generally narrow in scope and used small data sets for model training. This means the developed models are often only applicable to a few select feedstocks [11]. Additionally, existing models are unable to predict the influences of different operational choices, such as the reactor design and gasifying agent, on the gasification process. Combined with the fact that many models do not predict all important outputs, this leads to limitations regarding the applicability of existing models towards optimal system and process design. For example, Serrano and Castelló modelled the tar production from the gasification of woody biomass with good prediction accuracy ($R^2 > 0.97$) using an ANN [10]. They achieved this by collating a data set of 120 samples from literature. They reasoned that only samples with the same meta-data (e.g. bubbling fluidised bed reactor, reactor size between 50 and 300 mm, and operation at atmospheric pressure) could be added to the model. Shenbagaraj *et al.* developed a model to predict a wider range of outputs (syngas composition and yield) with a high accuracy of $R^2 > 0.97$ [15]. However, the developed model is specific to supercritical water gasification for a single feedstock type (food waste). The narrow scope of their work meant that only a small data set of 40 samples could be collected from literature. Li *et al.* trained a gradient boosting model on a larger data set of 295 data points, but their model is only suitable for hydrothermal gasification of three different waste feedstocks [16]. Similarly, Kargbo *et al.* developed a model for a specific reactor design and feedstock combination. Their model was trained on a small data set of 40 samples obtained from a two-stage gasifier using waste wood [17]. It has previously been identified that none of the existing ML-based gasification models are applicable across a wide range of feedstocks, gasifying agents, and reactor designs [4]. Hence, this work aims to significantly improve on existing studies by training a feedforward ANN on a large data set covering a broad range of feedstocks and operational conditions. For this, a highly generalisable model was developed. To the best of our knowledge, it is the first model applicable to a wide range of feedstock types (woody biomass, herbaceous biomass, plastics, municipal solid waste, and sewage sludge), gasifying agents (air, steam, oxygen, and combinations of these), and reactor options (fixed bed and fluidised bed). The model's generalisation capability has been maximised by the novel use of categorical data so that the model is highly practical for general gasification process design. Categorical variables (e.g. gasifying agent, reactor type, and bed material) allow factors that are critical to gasification, but which have been omitted by previous models, to be captured. To further improve the model's practicality for general process design, it was formulated to predict a wide

array of outputs that are essential for further modelling such as life cycle sustainability assessment (LCSA), which studies a system's environmental, economic, and social impacts. This in turn allows policy makers and investors to make optimal decisions by knowing the wider impacts of the system. The choice of ANN input features has been optimised. Furthermore, a novel optimisation superstructure to systematically compare and identify the best performing network structures and hyperparameter choices was employed to improve on previous studies which employed trial and error methods.

2. Materials and methods

2.1. Data collection and description of the data set

A data set of 312 samples was collected from literature [18–46] to train the proposed ANN model. To ensure a model with high generalisation capability, a wide range of feedstocks, gasifier types, and operating conditions were considered. The procedures for collecting data from literature were: (i) parameters were converted to the basis used in the data set (e.g. ultimate composition data given as wet basis (wb) would be converted to dry ash-free basis (daf) to fit in with the rest of the data set); (ii) the feedstock's lower heating value (LHV) was calculated from the feedstock's higher heating value (HHV) where necessary [47]; (iii) the quoted particle size refers to the lowest dimension of the particle (e.g. a pellet of $2 \times 2 \times 10$ mm would be quoted as 2 mm); (iv) all higher order hydrocarbons (C_2H_n) in the syngas were treated as C_2H_4 , as C_2H_4 is generally the dominant species [39,48]; (v) cold gas efficiency (CGE) and carbon conversion efficiency (CCE) were calculated where necessary (as defined in Appendix A). Due to this, some efficiencies of $> 100\%$ were obtained. The complete data set may be found in Appendix B.

Tables 1 and 2 summarise the predictor and target variables collected for the database. In total 18 continuous and 8 categorical predictors were collected as part of the database. Three sets of predictors (i.e. ALL, OPT, and -CAT) were considered in this work to assess the effect of different parameter combinations on model training. Predictor set ALL uses all predictors for model training, except for predictors with a significant number of missing values (i.e. the feedstock's lignocellulosic composition (cellulose, hemicellulose, and lignin) and residence time). For predictor set OPT a range of additional parameters were excluded from the data set used for model training: (i) feedstock type and feedstock shape were excluded as it was assumed that the same information was already captured by the continuous feedstock composition data and feedstock particle size data, respectively; (ii) volatile matter content and fixed carbon content were excluded as they are considered dependent variables which are related to other compositional parameters [12]; (iii) feedstock LHV, feedstock oxygen content, and feedstock nitrogen content were excluded due to their strong correlations with other input parameters as identified in Section 2.3. Predictor set -CAT is used to assess the effect of removing categorical variables from the model. Hence set -CAT is a copy of set OPT but excludes all categorical variables and the remaining set only contains continuous variables. The three different predictor combinations considered in this work are summarised in Table 3.

The target variables used in the model are all variables listed in Table 2, except for the process' CGE and CCE. These were not considered as target variables during modelling, as they often could be estimated based on other predictor and target variables. In other words, they were considered functions of some of the other predicted parameters.

2.2. Data pretreatment

2.2.1. Encoding of categorical variables

As previously mentioned, the use of a range of categorical variables to improve the ANNs prediction accuracy is a major point of novelty of this work. Unlike other ML methods, such as random forest, categorical variables cannot be directly fed into an ANN [49]. Thus, it is necessary to

Table 1
Summary of predictor variables.

	Min	Max	Mean	SD*	No. of Samples
Feedstock Information					
Type	Woody biomass, Herbaceous biomass, Sewage sludge, Municipal solid waste, Plastics, Other				312
Shape	Chips, Dust, Fibres, Particles, Pellets, Other				312
Particle Size [mm]	0.25	70.00	4.82	6.71	270
LHV [MJ/kg wb]	11.50	42.90	18.44	6.55	312
Ultimate Analysis					
C [% daf]	40.07	86.03	51.42	9.23	312
H [% daf]	3.79	14.23	6.86	2.00	312
N [% daf]	0	7.32	1.07	1.87	312
S [% daf]	0	1.61	0.32	0.45	279
O [% daf]	0	53.40	39.89	10.88	312
Proximate Analysis					
Ash [% db]	0.27	44.00	7.43	11.90	312
Moisture [% db]	0	27.00	8.37	5.65	312
Volatile Matter [% db]	56.00	89.11	77.56	8.98	264
Fixed Carbon [% db]	9.07	23.82	15.60	3.18	222
Lignocellulosic Composition					
Cellulose [% db]	29.60	46.20	39.42	5.08	28
Hemicellulose [% db]	14.00	29.60	20.91	4.67	28
Lignin [% db]	14.00	33.00	24.99	5.39	28
Gasifier Conditions and Information					
Temperature [°C]	553.00	1050.00	799.13	80.88	312
Operation Mode	Batch, Continuous				312
Residence Time [min]	6.00	403.00	98.34	97.30	124
Steam/Biomass [wt/wt]	0	4.04	1.02	0.73	133
ER	0.09	0.87	0.30	0.11	235
Gasifying Agent	Air, Air + Steam, Air + Oxygen, Steam, Oxygen, Other				312
Reactor Type	Fixed-bed, Fluidised bed, Other				312
Bed Material	Silica, Alumina, Calcium Oxide, Olivine, Dolomite				234
Catalyst	Present, Not present				312
Scale	Lab, Pilot				312

* Standard deviation.

Table 2
Summary of target variables.

	Min	Max	Mean	SD*	No. of Samples
Syngas Composition					
N ₂ [vol.% db]	0	74.90	38.16	25.76	286
H ₂ [vol.% db]	3.10	73.85	22.49	17.28	312
CO [vol.% db]	2.20	50.00	19.43	9.59	309
CO ₂ [vol.% db]	0	38.25	16.10	6.45	309
CH ₄ [vol.% db]	0.25	16.00	5.17	3.30	309
C ₂ H _n [vol.% db]	0	9.50	2.06	1.45	229
Syngas LHV [MJ/Nm ³ db]	1.30	15.56	7.19	3.47	312
Tar content [g/Nm ³ db]	0	364.00	29.89	47.85	193
Yield and Process Efficiency					
Syngas yield [Nm ³ /kg wb]	0.49	6.19	1.76	0.96	278
Char yield [g/kg wb]	3.60	301.52	94.56	76.49	87
CGE [%]	16.70	122.87	61.27	20.76	278
CCE [%]	20.61	126.89	70.98	17.35	275

* Standard deviation.

Table 3
Predictor combinations used for model training.

	Included Predictor Variables
Set ALL	Carbon, Hydrogen, Sulphur, Nitrogen, Oxygen, Ash, Moisture, Volatile Matter, Fixed Carbon, Particle Size, LHV, Temperature, ER, Steam to Biomass Ratio, Operation Mode, Catalyst Use, System Scale, Feedstock Type, Feedstock Shape, Gasifying Agent, Reactor Type, Bed Material
Set OPT	Carbon, Hydrogen, Sulphur, Ash, Moisture, Particle Size, Temperature, ER, Steam to Biomass Ratio, Operation Mode, Catalyst Use, System Scale, Gasifying Agent, Reactor Type, Bed Material
Set -CAT	Carbon, Hydrogen, Sulphur, Ash, Moisture, Particle Size, Temperature, ER, Steam to Biomass Ratio

develop methods through which the categorical data can be represented by numerical values.

To the best of the authors' knowledge, only Serrano *et al.* previously made use of categorical variables to improve a gasification model [13]. In their study the bed material of a fluidised bed gasifier was considered a categorical variable. The authors encoded this as an ordinal variable. However, the values of the variable follow no intrinsic order or scale and, therefore, using an encoding method such as one-hot encoding would seem more appropriate. Thus, here we used one-hot encoding for our nominal variables. A categorical variable which can take on n distinct categorical values is transformed into n binary variables, where values of 1 and 0 indicate the presence and absence of the variable, respectively. This means a categorical input variable with five categories will require five input neurons. An example of this is shown in Appendix C for the variable gasifying agent. The following variables were encoded using this method: feedstock type, feedstock shape, gasifying agent, reactor type, and bed material.

All categorical variables that can only take on two values (e.g. the process' operation mode (batch/continuous), system scale (lab/pilot), and catalyst use (present/not present)) were encoded as binary values. This means the variable can take on a value of 0 or 1 (-1 and 1 after normalisation). In the current database all data came from lab or pilot scale systems and, therefore, the system scale was encoded using this method. One-hot encoding or ordinal encoding could be employed if data from e.g. industrial scale plants were to be added. Ordinal encoding simply encodes each label as an integer, retaining the feature's order. Many papers which were used to compile the presented database simply stated whether the studied system was of lab/bench or pilot scale. However, some papers quoted the system's power output instead. In this case, a system was considered of lab scale for power ratings of < 100 kW_e or < 150 kW_{th}. A pilot scale system was defined as < 1 MW_e or < 1.5 MW_{th}.

It is obvious that the addition of a catalyst can strongly affect the

gasification process. Various types of catalysts have been shown to improve factors such as the process' efficiency, tar removal, or hydrogen yield [20,41,50,51]. However, it can be hard to quantify the exact effect of a catalyst numerically. Hence the authors chose a simplified model which encodes the use of a catalyst as a binary as previously explained (i. e. 0 for no catalyst present and 1 for catalyst present). This way the network can draw useful information and learn from whether a catalyst is present during the process.

2.2.2. Data normalisation

Data normalisation is an essential prerequisite for ANN modelling because it speeds up learning and ensures that all input features are within a common range. Eq. (1) has been used to normalise the data set between -1 and 1 :

$$x_{norm} = l + \frac{x - x_{min}}{x_{max} - x_{min}}(u - l) \quad (1)$$

where x is the observed value and x_{norm} is the normalised value. x_{max} and x_{min} represent the maximum and minimum values of the parameters, respectively. l and u are the lower and upper bound values between which the data is to be normalised. In this work, a lower and upper bound of -1 and 1 were selected.

2.2.3. Mean substitution of missing values

As the database used for model training was collected from many studies (29), it is unavoidable that there are inconsistencies in the format and types of data reported and some studies do not quote all parameters collected as part of the database. However, missing values can lead to issues during model training. For instance, the authors noticed a large loss of target-prediction pairs when a significant number of missing predictor data can result in the ANN being unable to make a prediction. In turn this led to fewer comparisons between the network's predicted value and the experimental target value which caused issues regarding the reproducibility of results. For this reason, the mean substitution method was employed for all predictors before model training, where a variable's missing values are replaced with the variable's mean value [52]. As the testing stage is the key stage of interest to evaluate the model's performance, this method was not employed for target variables and thus only affected model training.

2.3. Preliminary data analysis

Some initial analysis of the raw data was conducted by calculating the mean and standard deviation (SD) of variables as reported in Tables 1 and 2. Then Spearman's correlation coefficients (SCC) were calculated for the continuous predictor data to measure the monotonic relationship between data pairs. SCC was employed instead of the more commonly used Pearson correlation coefficient as the data was not known to follow a Gaussian distribution and SCC is less affected by outliers. Eq. (2) was used to calculate SCC:

$$SCC = \frac{\sum_{i=1}^N (R(x_i) - \bar{R}(x))(R(y_i) - \bar{R}(y))}{\sqrt{\sum_{i=1}^N (R(x_i) - \bar{R}(x))^2} \sqrt{\sum_{i=1}^N (R(y_i) - \bar{R}(y))^2}} \quad (2)$$

where N is the sample size, $R(x_i)$ and $R(y_i)$ are the ranks of individual samples of the two tested variables, and $\bar{R}(x)$ and $\bar{R}(y)$ indicate the mean ranks of the two variables.

A SCC of 0 indicates that two parameters are not correlated. The closer to ± 1 the stronger the monotonic relationship. In this work, a SCC of $\geq |0.6|$ was used to indicate a strong relationship which would lead to the removal of one of the parameters before model training. Fig. 1 shows the SCC matrix of all continuous predictors, except for parameters with a significant number of samples missing. The analysis of SCC led to the subsequent removal of the feedstock's nitrogen content

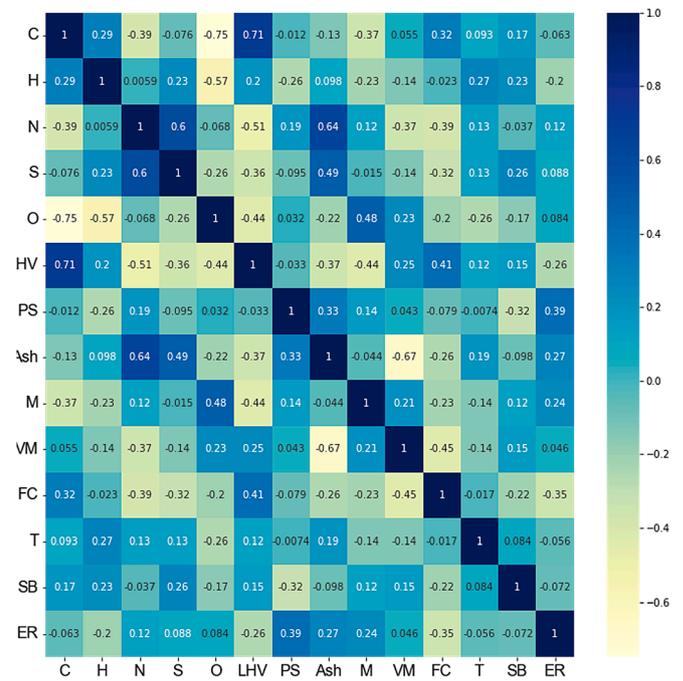


Fig. 1. Spearman correlation coefficient matrix of the following continuous input features: C – Carbon, H – Hydrogen, N – Nitrogen, S – Sulphur, O – Oxygen, LHV – LHV, PS – Particle Size, M – Moisture, VM – Volatile Matter, FC – Fixed Carbon, T – Temperature, SB – Steam to Biomass Ratio, and ER – Equivalence Ratio.

due to a significant correlation with the feedstock's sulphur and ash contents. Additionally, the feedstock's LHV and oxygen content were removed due to their correlations with the feedstock's carbon content. Highly correlated features are removed before training as they contain similar information and retaining both generally reduces model performance by adding additional noise to the model [53].

2.4. Artificial neural network development

2.4.1. Artificial neural network background

Conceptually, an ANN is a non-linear model which can fit any given function making it an universal function approximator [54]. It can be used to predict an arbitrary number of response variables given one or more explanatory variables. The type of ANN utilised in this work is often referred to as a feed-forward multilayer perceptron (MLP). A MLP is made up of an input and output layer and one or more hidden layers. Each layer contains nodes with weighted connections which allow for the transfer of information from one layer to the next.

Training is generally achieved through the backpropagation of errors [55]. This means, through an iterative process the optimum weights which minimise the prediction error for an independent test data set are

Table 4
Optimised and fixed hyperparameters used during ANN training.

Parameters	Type/Value
Optimised parameters	
No. neurons 1st hidden layer	2–24 (2)
No. neurons 2nd hidden layer	0–25 (5)
Split test set	10–20% (2.5%)
Training function	LM, BR
Hidden layer transfer function	TANSIG, LOGSIG
Fixed parameters	
Performance function	MSE
Learning rate	0.1
Output layer transfer function	Linear function
Maximum number of learning epochs	1000

found. Over the years various training algorithms have been developed. In this work, Levenberg-Marquardt backpropagation (LM) and Bayesian regularisation backpropagation (BR), which are implemented as part of the MATLAB deep learning toolbox, were considered [56–58]. The output of a fitted ANN is given by Eq. (3).

$$y_k = f_o \left(\sum_h w_{hk} f_h \left(\sum_i w_{ih} x_i \right) \right) \quad (3)$$

where y_k is the network's estimation of the response variable [59]. This is calculated from the sum of products of the weights w for i input variables x for the hidden nodes h . These are transferred by the activation functions f_o and f_h for the output and hidden nodes, respectively [60].

2.4.2. Artificial neural network optimisation and performance assessment

To find the best possible network configuration, five essential hyperparameters affecting the performance of the ANN were varied. The optimised parameters were: (i) number of neurons in 1st hidden layer, (ii) number of neurons in 2nd hidden layer, (iii) data split for test set, (iv) training function, (v) hidden layer transfer function. Initial tests resulted in the use of the parameter ranges and options indicated in Table 4. A test data split of 15% has frequently been employed in literature, hence a range from 10 to 20% was selected for optimisation purposes [4]. By initially considering various training functions, LM and BR were found to be the two preferred options which were further considered in the optimisation routine. Furthermore, initial tests allowed for the elimination of all but the hyperbolic tangent sigmoid (TANSIG) and logistic sigmoid (LOGSIG) function as hidden layer transfer function options. The step sizes (given in brackets in Table 4) for the numerical parameters (i)-(iii) were selected by weighing required accuracy against computational demands. Other parameters, such as the output layer's transfer function and the network's performance function were kept constant throughout and are also given in Table 4. Finally, 5-

fold cross validation was used to minimise the risk of overfitting and improve the confidence in the results.

The model's prediction performance was measured by considering the root mean square error (RMSE), R^2 , and adjusted coefficient of determination (R_{adj}^2). R_{adj}^2 is a modified version of the R^2 which considers the number of predictors in the model. Whilst R^2 generally increases with the addition of new predictors, despite potentially not adding any explanatory power, this is not the case for R_{adj}^2 . Thus, by using R_{adj}^2 an increase in R^2 due to the addition of unnecessary predictors is avoided, allowing for a fairer comparison between models trained with different input features. Eqs. (4)–(6) describe the three performance indicators:

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (y_i^o - y_i^p)^2} \quad (4)$$

$$R^2 = 1 - \frac{\sum_{i=1}^N (y_i^o - y_i^p)^2}{\sum_{i=1}^N (y_i^o - \bar{y}^o)^2} \quad (5)$$

$$R_{adj}^2 = 1 - \left[\frac{(1 - R^2)(N - 1)}{N - k - 1} \right] \quad (6)$$

where N is the number of samples and y_i^o and y_i^p are the observed and predicted values, respectively. \bar{y}^o and k represent the mean of all predicted values and the number of explanatory variables in the model (i.e. number of predictor variables used for model training), respectively.

2.4.3. Workflow summary

This study's workflow is illustrated by Fig. 2. The figure shows how the various steps and procedures introduced throughout Section 2 are linked to each other. It is shown how continuous and categorical variables have initially been pretreated separately before they were combined again after the normalisation step. At this point the data was ready

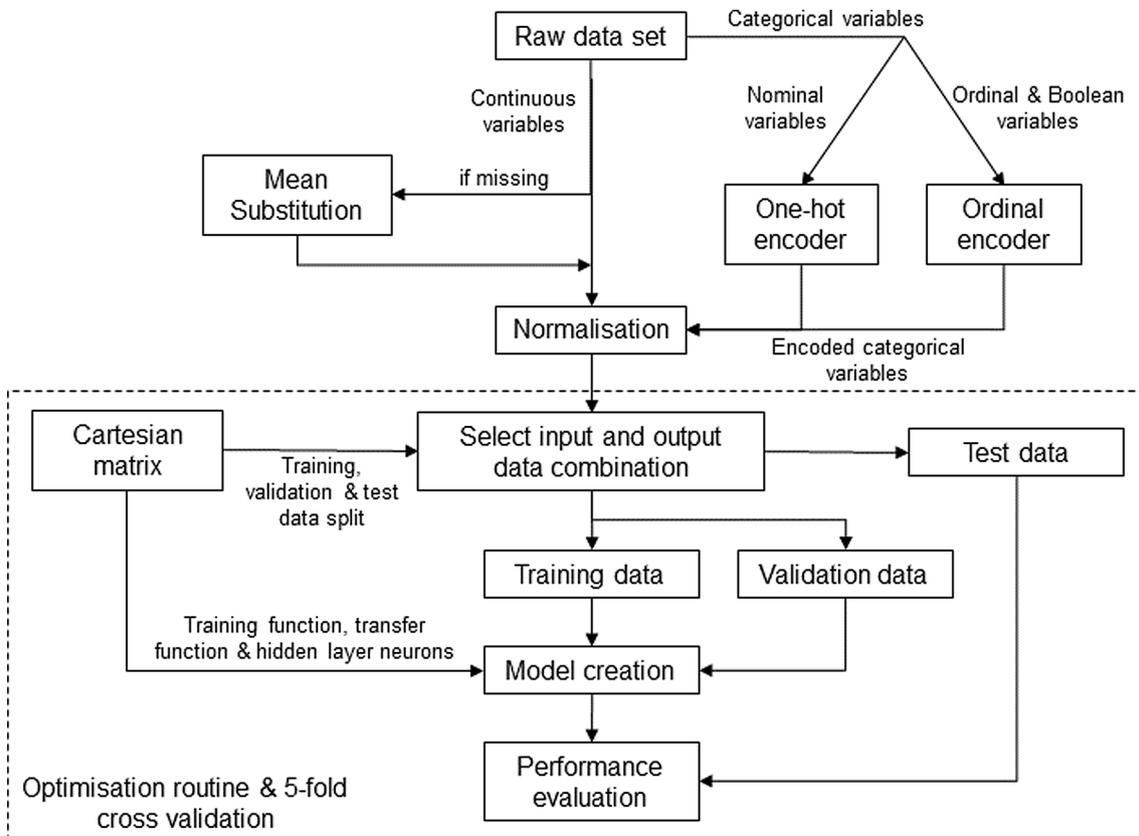


Fig. 2. Flowchart of the procedure employed for the model development.

for model training. Data was fed into the optimisation and cross validation stage of this work which is shown by the dashed box in the bottom half of the figure. Here the cartesian matrix contains all possible hyperparameter combinations introduced in Section 2.4.2. Models have been trained and cross validated for each of these combinations until the code had iterated through all combinations. Finally, the trained models' performance has been evaluated by considering their RMSE, R^2 , and R_{adj}^2 to find the best performing model. All data analysis and modelling work has been implemented in the MATLAB programming environment [61].

3. Results and discussion

3.1. Preliminary data analysis

The raw data's range, mean, and SD were calculated and are shown by Tables 1 and 2. Some of the key findings are highlighted here. When considering the syngas composition, it was found that the nitrogen content showed the highest variation of all variables with a SD of 25.76 vol% db. This can be explained by the use of various gasifying agents which result in syngas with vastly different nitrogen contents. The syngas LHV is similarly varied for the same reason. For instance, using air as a gasifying agent produces a syngas of a lower LHV, as compared to steam or oxygen, due to a strong dilution by nitrogen.

Studying the gasifiers operational conditions reveals a high variation in temperatures with a SD of 80.88 °C. This was expected as the temperature is one of the key factors widely varied during many experimental studies. The ER on the other hand showed very little variation with a SD of 0.11. Many studies considered an ER of 0.30 which was also found to be the parameter's mean value.

Looking at the feedstocks' compositional data the particle size showed some outliers (70 mm); however most the samples were > 10 mm which led to a small SD of 6.71 mm. When studying the feedstocks' elemental composition, the carbon contents were found to differ heavily with a SD of 9.23% daf. The high SD can be explained by the large range of different feedstocks being collected as part of the database. Whilst plastic waste was found to have carbon contents of up to 86.03% daf, most herbaceous biomass feedstocks have carbon contents < 50% daf.

The number of samples collected for a given variable is another important factor to consider. ML relies on the availability of large amounts of data and models tend to have a harder time learning the relationships effecting a variable when fewer samples are available. For instance, gasification is mainly considered a technology to produce syngas. Hence, many authors did not report on the process' char yields as shown by the relatively small number of samples (87 total) collected for this variable. This in turn was found to diminish the model's prediction accuracy for this variable as further discussed in Section 3.2.

3.2. Artificial neural network model optimisation

Fig. 3 shows the five-fold cross validated results of the best performing model structures trained on the three different data sets described in Table 3. Results are shown in normalised form as they are shown across all output variables. Considering the best performing model trained with all predictors collected as part of the database (set ALL) resulted in a well performing model with $RMSE = 0.1283$, $R^2 = 0.9329$, and $R_{adj}^2 = 0.9199$. This is clearly illustrated by the fact that the data points in Fig. 3 (a) are tightly distributed around the dotted perfect fit line with unity slope. Using an optimised selection of predictors (set OPT) yielded a model with a similarly high prediction accuracy and low error ($RMSE = 0.1307$, $R^2 = 0.9310$, and $R_{adj}^2 = 0.9254$). Generally, a model with more predictors will spuriously achieve a higher R^2 than a model with fewer predictors. However, when considering R_{adj}^2 instead, where excessive predictors are penalised, the model trained on set OPT was found to be the best performing model. The removal of all categorical variables (set -CAT), as shown by Fig. 3 (c), led to a significant

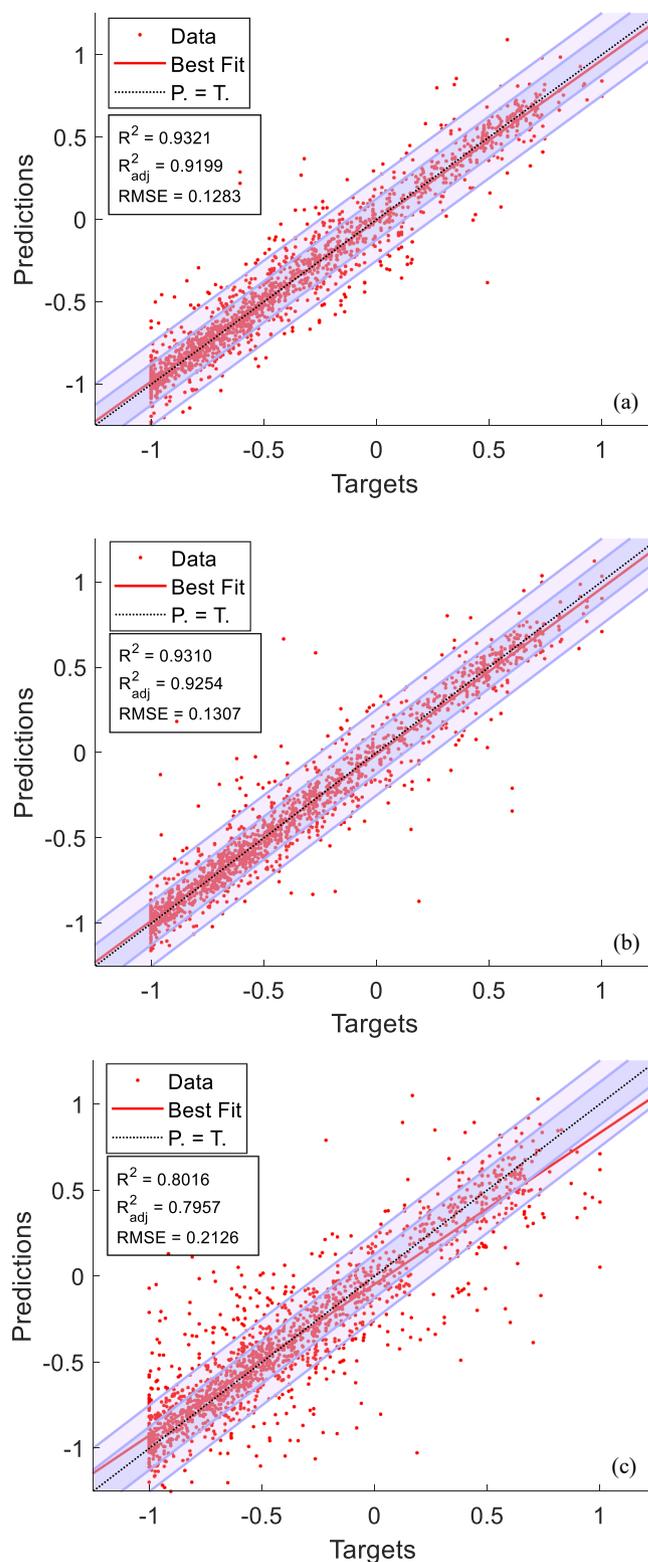


Fig. 3. Distribution of normalised model predictions against target values for all 5 cross validation runs of best performing model. (a), (b), and (c) show the results for predictor combinations set ALL, set OPT, and set -CAT respectively. The black dotted line represents the perfect fit line for the case where all predictions equal their target values, and the solid red line indicates the best fit to the data points. The inner blue region shows the 10% error range, whereas the lighter outer region shows the 20% error range. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 5
Summary of performance indicators and hyperparameter choices for best performing models.

	Performance Indicators ^a				Hyperparameters			
	RMSE	R ²	R _{adj} ²	Training time [s]	Hidden layer neurons	Training function	Hidden layer transfer function	Test split [%]
Set ALL	0.1283 (0.0145)	0.9321 (0.0153)	0.9199 (0.0175)	29.53 (9.64)	8, 10	BR	TANSIG	12.5
Set OPT	0.1307 (0.0194)	0.9310 (0.0289)	0.9254 (0.0305)	43.40 (16.15)	18	BR	TANSIG	12.5
Set -CAT	0.2126 (0.0280)	0.8016 (0.0540)	0.7957 (0.0557)	4.35 (1.44)	20,25	LM	LOGSIG	12.5

a – values in brackets indicate the SD over all cross validation runs.

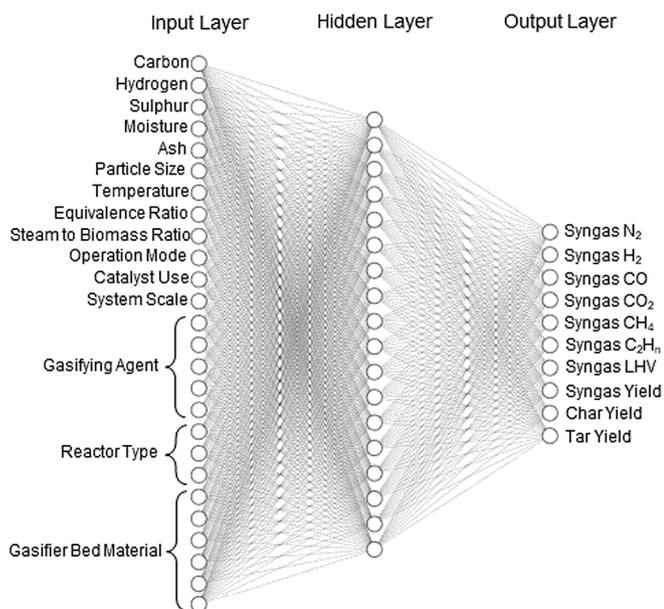


Fig. 4. Network structure of best performing ANN trained on set OPT.

drop in model accuracy as illustrated by the wider dispersion of data points from the perfect fit line and increased RMSE = 0.2126 and decreased R² = 0.8016 and R_{adj}² = 0.7959. This highlights the importance of incorporating key categorical variables to improve the model's accuracy. In conclusion, these findings indicate that the use of a well optimised combination of predictors, including meaningful categorical variables, can significantly increase the prediction accuracy of an ANN model. A summary of these findings is presented in Table 5.

Based on these findings, set OPT was considered as the most suitable combination of predictors for model training and the model's optimised structure is illustrated by Fig. 4. Hence models trained on this data set were further evaluated. When considering the next best performing network structures, it becomes clear that various network structures and hyperparameter choices can result in a well performing model. For example, a model trained with a double hidden layer structure of 24 and 15 neurons utilising the LM training function and a model trained with the BR training function and a single hidden layer with 16 neurons resulted in a similar predictive performance to the best performing model with R² > 0.92 and RMSE < 0.14. This agrees with existing literature where both BR and LM were found to be suitable training functions for the ANN modelling of various thermochemical processes [13,62–64]. Furthermore, various data splits for testing and validation ranging from 12.5 to 17.5% were found to produce acceptable results. Finally, the TANSIG hidden layer transfer function was found to generally be the best choice for the trained models.

A limitation of this work is the use of lab and pilot scale data only for model development. At the time of development, no industrial scale data

has been available to train the model on. It remains uncertain how well the model performs on unseen industrial scale data. However, through the novel use of categorical variables, the presented approach could easily be adapted were industrial scale data to become available as highlighted in Section 2.2.1.

In general, adding a 2nd hidden layer was not shown to improve the model's prediction accuracy. This agrees with theory and the universal approximator theorem which states that a single hidden layer ANN is sufficient to represent any given function arbitrarily well [54]. Thus, the increased complexity from adding a 2nd hidden layer cannot be justified. When it comes to the choice of training function, it was found that the BR function generally resulted in a better performing model than LM. This was in particular the case for single hidden layer network structures, albeit at the cost of an increased computational demand.

Fig. 5 shows the prediction accuracy for all ten individual target outputs for the best performing model structure trained with set OPT. The model was found to predict some outputs with greater accuracy than others, as highlighted by the wider dispersion of data points for some parameters. The performance indicators for the individual model outputs are summarised in Table 6. It is clear that some factors (e.g. the process' syngas yield and H₂ content) can be predicted with higher accuracy than others (e.g. char yield and syngas C₂H_n content). The highest R² of 0.9791 and lowest RMSE of 0.0731 were obtained for the syngas N₂ content and tar yield, respectively. Conversely, the lowest R² of 0.5202 and highest RMSE of 0.3020 were obtained for the syngas C₂H_n content and char yield, respectively. Understanding and accounting for the uncertainty in certain predictions is essential for further system analysis such as LCSA.

When considering R_{adj}², only a minor decrease from R² was found for the outputs which the model predicted well. However, the two outputs which the model predicted poorly, the syngas C₂H_n content and char yield, showed a significant decrease. The former decreased by 13% and the latter by 21%. This decrease and poor performance can be explained by the significantly smaller sample size available for the training of these two features compared to other features. For instance, only 87 samples could be collected for the process' char yield, whereas the data set contained 312 samples for the excellently predicted features syngas LHV and H₂ content. Adding additional samples for these outputs would likely allow the model to better learn the relationships affecting the outputs, leading to an increase in prediction accuracy.

As previously mentioned, one of the key aims of this study is to create a highly generalisable model which is suitable across a wide range of feedstocks and gasifier design choices. For this reason, a like-for-like comparison between existing studies and this work is not always possible or sensible due to vastly different scopes and aims. Existing studies often developed gasifier specific models or models specific to a select few feedstocks [65–67]. For instance, Kargbo *et al.* optimised a two-stage gasification process using an ANN [17]. Their model was found to generally predict most gasification outputs excellently with R² = 0.99. However, the tar production was predicted less well with R² of 0.74–0.82. Whilst their model generally performs better than the model presented in this study, their model is only applicable for the specific

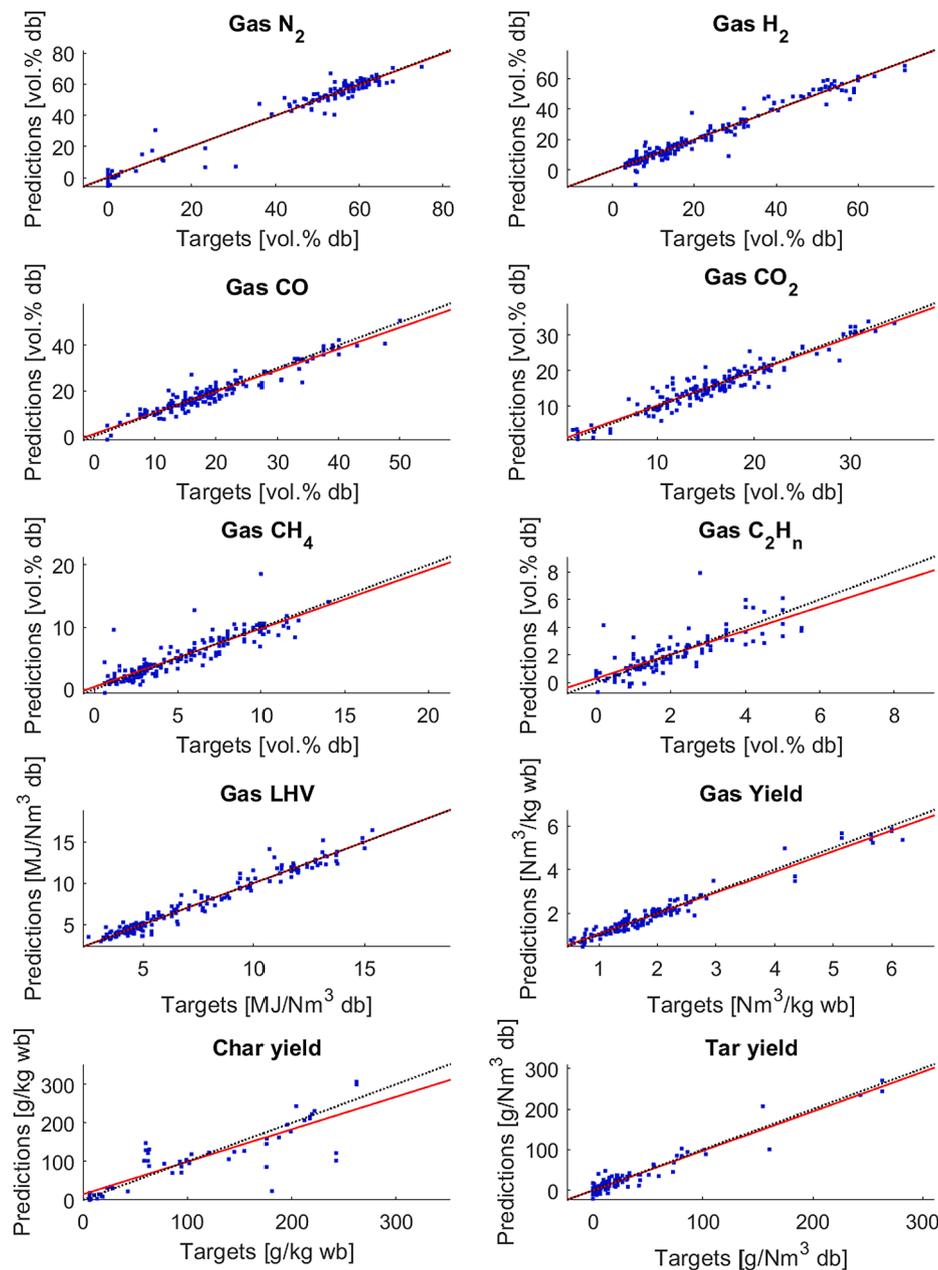


Fig. 5. Distribution of model predictions against target values by output for all 5 cross validation runs for model trained using set OPT.

Table 6

Model performance indicators by output.

	N ₂	H ₂	CO	CO ₂	CH ₄	C ₂ H _n	Gas LHV	Gas yield	Char yield	Tar yield
RMSE	0.1047	0.1094	0.1084	0.1078	0.1796	0.1883	0.0989	0.0821	0.3020	0.0731
R2	0.9791	0.9515	0.9261	0.8954	0.799	0.5202	0.9571	0.9558	0.7106	0.9190
R2adj	0.9770	0.9471	0.9193	0.8858	0.7805	0.4584	0.9532	0.9512	0.5852	0.9063

two-stage gasifier design used in the study. Furthermore, their model was trained on a small data set of 40 samples containing only a single feedstock (woody biomass). This limits the model's use to these specific conditions. Elmaz *et al.* applied four regression techniques to data obtained from a 10 kW downdraft fixed-bed gasifier [68]. Decision tree regression and MLP were found to predict the gasification outputs best with $R^2 > 0.9$. Serrano *et al.* trained ANN models for different gasifier bed material options with a prediction accuracy of $R^2 = 0.90$ – 0.97 [13]. Whilst the mean absolute percentage error (MAPE) for a model

considering only silica sand as bed material was acceptable (MAPE = 2.25–24.89%), larger errors (MAPE = 8.18–38.91%) were obtained for a model designed for a range of bed material options. In comparison, a MAPE of 11.14% was obtained in this work when averaged across all outputs for a model suitable across a range of bed material options.

The developed model could be used for a variety of applications. Its power to make predictions over a wide range of feedstock types, gasifying agents, and reactor options makes it particularly suitable for modelling and designing optimal gasification systems and process

conditions. Since the model can predict more gasification outputs than existing models, it has the potential to simplify further modelling and analysis relying on these outputs. For instance, techno-economic analysis or LCSA could be streamlined by using the predictions made by the model, which can be further combined with multi-objective optimisation to decide on optimal gasification process conditions.

4. Conclusions

A highly generalisable artificial neural network (ANN) model of excellent prediction accuracy ($R^2 = 0.9310$ and $RMSE = 0.1307$) has been developed to model the gasification of biomass and waste. The developed model is the first model able to predict gasification outputs over a wide range of feedstock, reactor, and gasifying agent options. Thus, investors and policy makers can quickly and conveniently compare a multitude of system options and their performance. This in turn reduces the need for costly and time-consuming experimental studies. Since the model predicts many key gasification outputs, the suggested approach has the potential to facilitate the development of integrated gasification design models by combining the developed model with e.g. life cycle sustainability assessment (LCSA), cost-benefit analysis, and multi-objective optimisation.

CRediT authorship contribution statement

Simon Ascher: Conceptualization, Methodology, Software, Writing – original draft. **William Sloan:** Writing – review & editing, Funding acquisition. **Ian Watson:** Writing – review & editing. **Siming You:** Conceptualization, Supervision, Writing – review & editing.

Declaration of Competing Interest

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

Acknowledgements

The authors would like to acknowledge the financial support from the Engineering and Physical Sciences Research Council (EPSRC) Studentship and Programme Grant (EP/V030515/1). Siming You would like to thank the financial support from Supergen Bioenergy Hub Rapid Response Funding (RR 2022_10) and Royal Society Research Grant (RGS \R1\211358). All data supporting this study are provided in full in the 'Methodology' and 'Results and Discussion' sections and Appendices of this paper.

Appendix A. - C Supplementary material

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.apenergy.2022.119289>.

Reference

- [1] Kaza S, Yao LC, Bhada-Tata P, Van Woerden F. What a Waste 2.0. Washington DC: Washington, DC: World Bank; 2018. <https://doi.org/10.1596/978-1-4648-1329-0>.
- [2] IEA Bioenergy. The Availability of Biomass Resources for Energy: Summary and Conclusions from the IEA Bioenergy ExCo58 Workshop; 2008.
- [3] Bridgewater AV, 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(3):181–246.
- [4] Ascher S, Watson I, You S. Machine learning methods for modelling the gasification and pyrolysis of biomass and waste. *Renew Sustain Energy Rev* 2022;155:111902.
- [5] Wang K, Zhang J, Shang C, Huang D. Operation optimization of Shell coal gasification process based on convolutional neural network models. *Appl Energy* 2021;292:116847. <https://doi.org/10.1016/j.apenergy.2021.116847>.
- [6] Wang S, Wen Y, Shi Z, Nuran Zaini I, Göran Jönsson P, Yang W. Novel carbon-negative methane production via integrating anaerobic digestion and pyrolysis of organic fraction of municipal solid waste. *Energy Convers Manag* 2022;252:115042. <https://doi.org/10.1016/j.enconman.2021.115042>.
- [7] Tang Q, Chen Y, Yang H, Liu M, Xiao H. Machine learning prediction of pyrolytic gas yield and compositions with feature reduction methods: Effects of pyrolysis conditions and biomass characteristics. *Bioresour Technol* 2021;339:125581. <https://doi.org/10.1016/j.biortech.2021.125581>.
- [8] Ögren Y, Tóth P, Garami A, Sepman A, Wiinikka H. Development of a vision-based soft sensor for estimating equivalence ratio and major species concentration in entrained flow biomass gasification reactors. *Appl Energy* 2018;226:450–60. <https://doi.org/10.1016/j.apenergy.2018.06.007>.
- [9] Ning C, You F. Data-driven Wasserstein distributionally robust optimization for biomass with agricultural waste-to-energy network design under uncertainty. *Appl Energy* 2019;255:113857. <https://doi.org/10.1016/j.apenergy.2019.113857>.
- [10] 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>.
- [11] 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(20):9558–68. <https://doi.org/10.1016/j.ijhydene.2018.04.007>.
- [12] Baruah D, Baruah DC, Hazarika MK. Artificial neural network based modeling of biomass gasification in fixed bed downdraft gasifiers. *Biomass Bioenergy* 2017;98:264–71. <https://doi.org/10.1016/j.biombioe.2017.01.029>.
- [13] 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>.
- [14] 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(6):9580–94. <https://doi.org/10.1002/er.6483>.
- [15] Shenbagaraj S, Sharma PK, Sharma AK, Raghav G, Kota KB, Ashokkumar V. 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(24):12739–57. <https://doi.org/10.1016/j.ijhydene.2021.01.122>.
- [16] Li J, Suvarna M, Pan L, Zhao Y, Wang X. A hybrid data-driven and mechanistic modelling approach for hydrothermal gasification. *Appl Energy* 2021;304:117674. <https://doi.org/10.1016/j.apenergy.2021.117674>.
- [17] Kargbo HO, Zhang J, Phan AN. Optimisation of two-stage biomass gasification for hydrogen production via artificial neural network. *Appl Energy* 2021;302:117567. <https://doi.org/10.1016/j.apenergy.2021.117567>.
- [18] Arena U, Di Gregorio F. Gasification of a solid recovered fuel in a pilot scale fluidized bed reactor. *Fuel* 2014;117:528–36. <https://doi.org/10.1016/j.fuel.2013.09.044>.
- [19] Arena U, Zaccariello L, Mastellone ML. Fluidized bed gasification of waste-derived fuels. *Waste Manag* 2010;30(7):1212–9. <https://doi.org/10.1016/j.wasman.2010.01.038>.
- [20] Baratieri M, Pieratti E, Nordgreen T, Grigiante M. Biomass gasification with dolomite as catalyst in a small fluidized bed experimental and modelling analysis. *Waste Biomass Valorization* 2010;1(3):283–91. <https://doi.org/10.1007/s12649-010-9034-6>.
- [21] Luo S, Zhou Y, Yi C. Syngas production by catalytic steam gasification of municipal solid waste in fixed-bed reactor. *Energy* 2012;44(1):391–5. <https://doi.org/10.1016/j.energy.2012.06.016>.
- [22] Khan Z, Yusup S, Ahmad MM, Chin BLF. Hydrogen production from palm kernel shell via integrated catalytic adsorption (ICA) steam gasification. *Energy Convers Manag* 2014;87:1224–30. <https://doi.org/10.1016/j.enconman.2014.03.024>.
- [23] Wang J, Cheng G, You Y, Xiao B, Liu S, He P, et al. Hydrogen-rich gas production by steam gasification of municipal solid waste (MSW) using NiO supported on modified dolomite. *Int J Hydrogen Energy* 2012;37(8):6503–10. <https://doi.org/10.1016/j.ijhydene.2012.01.070>.
- [24] Karatas H, Akgun F. Experimental results of gasification of walnut shell and pistachio shell in a bubbling fluidized bed gasifier under air and steam atmospheres. *Fuel* 2018;214:285–92. <https://doi.org/10.1016/j.fuel.2017.10.061>.
- [25] Van HC, Kong SC. Performance characteristics of a pilot-scale biomass gasifier using oxygen-enriched air and steam. *Fuel* 2013;103:987–96. <https://doi.org/10.1016/j.fuel.2012.09.033>.
- [26] Erkiaga A, Lopez G, Amutio M, Bilbao J, Olazar M. Influence of operating conditions on the steam gasification of biomass in a conical spouted bed reactor. *Chem Eng J* 2014;237:259–67. <https://doi.org/10.1016/j.cej.2013.10.018>.
- [27] Ruoppolo G, Ammendola P, Chirone R, Miccio F. H₂-rich syngas production by fluidized bed gasification of biomass and plastic fuel. *Waste Manag* 2012;32(4):724–32. <https://doi.org/10.1016/j.wasman.2011.12.004>.
- [28] Liu H, Hu J, Wang H, Wang C, Li J. Experimental studies of biomass gasification with air. *J Nat Gas Chem* 2012;21(4):374–80. [https://doi.org/10.1016/S1003-9953\(11\)60379-4](https://doi.org/10.1016/S1003-9953(11)60379-4).
- [29] Gai C, Dong Y. Experimental study on non-woody biomass gasification in a downdraft gasifier. *Int J Hydrogen Energy* 2012;37(6):4935–44. <https://doi.org/10.1016/j.ijhydene.2011.12.031>.
- [30] Lahijani P, Zainal ZA. Gasification of palm empty fruit bunch in a bubbling fluidized bed: A performance and agglomeration study. *Bioresour Technol* 2011;102(2):2068–76. <https://doi.org/10.1016/j.biortech.2010.09.101>.
- [31] Narváez I, Orío A, Aznar MP, Corella J. Biomass Gasification with Air in an Atmospheric Bubbling Fluidized Bed. Effect of Six Operational Variables on the Quality of the Produced Raw Gas. *Ind Eng Chem Res* 1996;35(7):2110–20.

- [32] Yin R, Liu R, Wu J, Wu X, Sun C, Wu C. Influence of particle size on performance of a pilot-scale fixed-bed gasification system. *Bioresour Technol* 2012;119:15–21. <https://doi.org/10.1016/j.biortech.2012.05.085>.
- [33] Song T, Wu J, Shen L, Xiao J. Experimental investigation on hydrogen production from biomass gasification in interconnected fluidized beds. *Biomass Bioenergy* 2012;36:258–67. <https://doi.org/10.1016/j.biombioe.2011.10.021>.
- [34] Biagini E, Barontini F, Tognotti L. Development of a bi-equilibrium model for biomass gasification in a downdraft bed reactor. *Bioresour Technol* 2016;201:156–65. <https://doi.org/10.1016/j.biortech.2015.11.057>.
- [35] Dutta PP, Pandey V, Das AR, Sen S, Baruah DC. Down Draft Gasification Modelling and Experimentation of Some Indigenous Biomass for Thermal Applications. *Energy Proc* 2014;54:21–34.
- [36] Herguido J, Corella J, Gonzalez-Saiz J. Steam Gasification of Lignocellulosic Residues in a Fluidized Bed at a Small Pilot Scale. Effect of the Type of Feedstock. *Ind Eng Chem Res* 1992;31:2295–310.
- [37] Premaux S, Beheshti SM, Ghassemi H, Shahsavan-Markadeh R. An experimental study on hydrogen-rich gas production via steam gasification of biomass in a research-scale fluidized bed. *Energy Convers Manag* 2015;91:427–32. <https://doi.org/10.1016/j.enconman.2014.12.048>.
- [38] Zhang Z, Pang S. Experimental investigation of biomass devolatilization in steam gasification in a dual fluidised bed gasifier. *Fuel* 2017;188:628–35. <https://doi.org/10.1016/j.fuel.2016.10.074>.
- [39] Iovane P, Donatelli A, Molino A. Influence of feeding ratio on steam gasification of palm shells in a rotary kiln pilot plant. Experimental and numerical investigations. *Biomass Bioenergy* 2013;56:423–31. <https://doi.org/10.1016/j.biombioe.2013.05.025>.
- [40] Lv PM, Xiong ZH, Chang J, Wu CZ, Chen Y, Zhu JX. An experimental study on biomass air-steam gasification in a fluidized bed. *Bioresour Technol* 2004;95(1):95–101. <https://doi.org/10.1016/j.biortech.2004.02.003>.
- [41] de Andrés JM, Narros A, Rodríguez ME. Air-steam gasification of sewage sludge in a bubbling bed reactor: Effect of alumina as a primary catalyst. *Fuel Process Technol* 2011;92(3):433–40. <https://doi.org/10.1016/j.fuproc.2010.10.006>.
- [42] Kaewluan S, Pipatmanomai S. Gasification of high moisture rubber woodchip with rubber waste in a bubbling fluidized bed. *Fuel Process Technol* 2011;92(3):671–7. <https://doi.org/10.1016/j.fuproc.2010.11.026>.
- [43] Sarker S, Nielsen HK. Assessing the gasification potential of five woodchips species by employing a lab-scale fixed-bed downdraft reactor. *Energy Convers Manag* 2015;103:801–13. <https://doi.org/10.1016/j.enconman.2015.07.022>.
- [44] Sarker S, Bimbela F, Sánchez JL, Nielsen HK. Characterization and pilot scale fluidized bed gasification of herbaceous biomass: A case study on alfalfa pellets. *Energy Convers Manag* 2015;91:451–8. <https://doi.org/10.1016/j.enconman.2014.12.034>.
- [45] Mansaray KG, Ghaly AE, Al-Taweel AM, Hamdullahpur F, Ugursal VI. Air gasification of rice husk in a dual distributor type fluidized bed gasifier. *Biomass Bioenergy* 1999;17(4):315–32.
- [46] Loha C, Chattopadhyay H, Chatterjee PK. Energy generation from fluidized bed gasification of rice husk. *J Renew Sustain Energy* 2013;5(4):043111. <https://doi.org/10.1063/1.4816496>.
- [47] Antonopoulos I-S, Karagiannidis A, Gkouletsos A, Perkoulidis G. Modelling of a downdraft gasifier fed by agricultural residues. *Waste Manag* 2012;32(4):710–8. <https://doi.org/10.1016/j.wasman.2011.12.015>.
- [48] Franco C, Pinto F, Gulyurtlu I, Cabrita I. The study of reactions influencing the biomass steam gasification process. *Fuel* 2003;82(7):835–42. [https://doi.org/10.1016/S0016-2361\(02\)00313-7](https://doi.org/10.1016/S0016-2361(02)00313-7).
- [49] Hastie T, Tibshirani R, Friedman J. In: *The Elements of Statistical Learning*. 2nd ed. New York, NY: Springer New York; 2009. <https://doi.org/10.1007/978-0-387-84858-7>.
- [50] Manuel De Andrés J, Roche E, Narros A, Encarnación RM. Characterisation of tar from sewage sludge gasification. Influence of gasifying conditions: Temperature, throughput, steam and use of primary catalysts. *Fuel* 2016;180:116–26. <https://doi.org/10.1016/j.fuel.2016.04.012>.
- [51] Tomishige K, Asadullah M, Kunimori K. Syngas production by biomass gasification using Rh/CeO₂/SiO₂ catalysts and fluidized bed reactor. *Catal Today* 2004;89(4):389–403. <https://doi.org/10.1016/j.cattod.2004.01.002>.
- [52] van Buuren S. *Flexible Imputation of Missing Data*. 2nd ed. CRC Press; 2018.
- [53] Li J, Pan L, Suvarna M, Wang X. Machine learning aided supercritical water gasification for H₂-rich syngas production with process optimization and catalyst screening. *Chem Eng J* 2021;426:131285. <https://doi.org/10.1016/j.cej.2021.131285>.
- [54] Hornik K, Stinchcombe M, White H. Multilayer feedforward networks are universal approximators. *Neural Networks* 1989;2(5):359–66. [https://doi.org/10.1016/0893-6080\(89\)90020-8](https://doi.org/10.1016/0893-6080(89)90020-8).
- [55] Rumelhart DE, Hinton GE, Williams RJ. Learning representations by back-propagating errors. *Nature* 1986;323(6088):533–6. <https://doi.org/10.1038/323533a0>.
- [56] Marquardt DW. An Algorithm for Least-Squares Estimation of Nonlinear Parameters. *J Soc Ind Appl Math* 1963;11(2):431–41.
- [57] Mt H, Mb M. Training feedforward networks with the Marquardt algorithm. *IEEE Trans Neural Networks* 1994;5:989–93. <https://doi.org/10.1109/72.329697>.
- [58] Foresee FD, Hagan MT. Gauss-Newton approximation to bayesian learning. *Proc Int Jt Conf Neural Networks* 1997.
- [59] Bishop CM. *Neural Networks for Pattern Recognition*. Oxford University Press, Inc.198 Madison Ave. New York, NY, United States; 1995.
- [60] Beck MW. *NeuralNetTools: Visualization and analysis tools for neural networks*. *J Stat Softw* 2018;85:1–20. <https://doi.org/10.18637/jss.v085.i11>.
- [61] MathWorks. *MathWorks - MATLAB* 2022. <https://uk.mathworks.com/products/matlab.html> [accessed April 18, 2022].
- [62] Nasrudin NA, Jewaratnam J, Hossain MA, Ganeson PB. Performance comparison of feedforward neural network training algorithms in modelling microwave pyrolysis of oil palm fibre for hydrogen and biochar production. *Asia-Pacific J Chem Eng* 2020;15:1–16. <https://doi.org/10.1002/apj.2388>.
- [63] Puig-Arnavat M, Bruno JC, Coronas A. Review and analysis of biomass gasification models. *Renew Sustain Energy Rev* 2010;14(9):2841–51. <https://doi.org/10.1016/j.rser.2010.07.030>.
- [64] Sreejith CC, Muraleedharan C, Arun P. Performance prediction of fluidised bed gasification of biomass using experimental data-based simulation models. *Biomass Bioenergy* 2013;3(4):283–304. <https://doi.org/10.1007/s13399-013-0083-5>.
- [65] Puig-Arnavat M, Hernández JA, Bruno JC, Coronas A. Artificial neural network models for biomass gasification in fluidized bed gasifiers. *Biomass Bioenergy* 2013;49:279–89. <https://doi.org/10.1016/j.biombioe.2012.12.012>.
- [66] 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>.
- [67] Özveren U, Kartal F, Sezer S, Özdoğan ZS. Investigation of steam gasification in thermogravimetric analysis by means of evolved gas analysis and machine learning. *Energy* 2022;239:122232.
- [68] 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>.