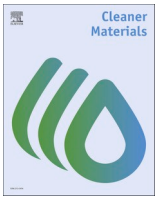
Cleaner Materials 9 (2023) 100206

Contents lists available at ScienceDirect

Cleaner Materials

journal homepage: www.journals.elsevier.com/cleaner-materials



A review on biomass thermal-oxidative decomposition data and machine learning prediction of thermal analysis

Yuying Chen [a,b], Zilong Wang [a,*], Shaorun Lin [c], Yunzhu Qin [a], Xinyan Huang [a,*]

[a] Research Center for Fire Safety Engineering, Department of Building Environment and Energy Engineering, The Hong Kong Polytechnic University, Kowloon, Hong Kong
[b] School of Civil and Environmental Engineering, University of Technology Sydney, Sydney, Australia
[c] Department of Mechanical Engineering, University of California, Berkeley, CA, USA

ARTICLE INFO

Keywords:
Biofuel database
Pyrolysis
Oxidative degradation
Thermogravimetric analysis
Artificial intelligence

ABSTRACT

Thermochemical conversion is the most economical approach to recovering energy and alternative fuels from biomass feedstock. This work first reviews the literature data on thermal-oxidative decomposition for common biomass types and forms a database of 18 parameters, including element, proximate, and thermogravimetric analysis (TGA). Then, an Artificial Neural Network (ANN) model is developed for the prediction of TGA data. Pearson correlation coefficient analysis reveals that the influence of environment heating rate on biomass thermal decomposition is larger than that of fuel properties. By inputting biomass elemental/proximate analysis and heating rate, the ANN model successfully predicts 8 key TGA parameters, namely, pyrolysis-onset temperature, peak pyrolysis temperature, oxidation-dominant temperature, peak oxidation temperature, oxidation-end temperature, peak pyrolysis rate, oxidation-dominant rate, and peak oxidation rate, with R^2 values greater than 0.98. A better performance can be achieved when all ten input features are considered. Final, an open-access online software, Intelligent Fuel Thermal Analysis (IFTA), is developed to predict thermal-oxidative decomposition across a wide range of heating rates and biomass types. This work provides a better understanding of biomass thermal-oxidative decomposition dynamics and a shortcut to obtain key parameters of biomass degradation without TGA tests.

1. Introduction

Biomass is an organic material, and the common biomass includes woody material (twigs, barks, wood processing residues, etc.), agricultural waste (wheat straw, corn stove, leaf litter, etc.), and forestry residues. Biomass is a vital resource for providing alternative clean and renewable energy in a sustainable world. Thermochemical conversion of biomass (including pyrolysis, torrefaction, gasification, and combustion) has been regarded as an economical and promising way to recover energy from biomass feedstock by producing heat, biofuel, biochar, and bulk chemicals (Madadian et al., 2020). Hence, understanding the thermal decomposition properties (or thermal stability) of biomass is crucial for the efficient and safe utilization of bioenergy (Vassilev et al., 2010).

Thermogravimetric analysis (TGA) is one of the most common and reliable analytical techniques utilized in biomass thermal analysis to identify the degradation properties of biomass rapidly and efficiently (Teh et al., 2021). During the thermal decomposition of biomass, the TG curve (i.e., the mass loss curve, see black dashed lines in Fig. 1) records the mass of the sample versus time or temperature while the sample is heated in a controlled atmosphere (inert or oxidative) and a programmed heating rate. Simultaneously, the derivative thermogravimetry (DTG) curve (see red lines in Fig. 1) that exhibits the mass loss rate (MLR) of the sample as a function of the time or temperature is also calculated and recorded. When heated in an inert atmosphere (Fig. 1 (a)), the thermal decomposition of biomass is generally called pyrolysis, which is an endothermic process. A typical DTG curve of pyrolysis shows two typical peaks corresponding to the loss of water or very light volatile compounds in a lower temperature range and pyrolysis in a higher temperature range, respectively.

Fundamentally, pyrolysis is a group of parallel reactions, and the peaks of these reactions can fully or partially overlap in the DTG curves, showing more than on peaks or some shoulder regions (Anca-Couce et al., 2014; Olatunji et al., 2018; Li et al., 2014; Niu and Liu, 2015). TGA of biomass in an inert atmosphere has been widely conducted and investigated to understand its pyrolysis process and products and its flammability as a fuel (Carrier et al., 2011; Chen et al., 2015; Saldarriaga

* Corresponding authors at: ZS832, 181 Chatham Road South, Kowloon, HK (X. Huang); ZN808, 181 Chatham Road South, Kowloon, HK (Z. Wang).
E-mail addresses: zilong.wang@connect.polyu.hk (Z. Wang), xy.huang@polyu.edu.hk (X. Huang).

https://doi.org/10.1016/j.clema.2023.100206
Received 25 June 2023; Received in revised form 1 September 2023; Accepted 3 September 2023
Available online 4 September 2023

Nomenclature

Symbols	
B	bias
cov	covariance
h	transfer (activation) function
N (n)	sample size
r	Pearson correlation coefficient
R^2	coefficient of determination
T	temperature (℃)
W	weight of the input
$\bar{x}$	mean value of variable x
X	input of the model
$\bar{y}$	mean value of variable y
Y	target of the model
Greeks	
σ	standard deviation
Subscripts	
ox,0	onset of oxidation domination
ox,p	oxidation peak
ox,e	end of oxidation
py,0	onset of pyrolysis
py,p	pyrolysis peak
Abbreviation	
FC	fixed carbon
HR	heating rate (K/min)
LOOCV	Leave-One-Out Cross Validation
MLR	mass loss rate (10^{-3}/K)
MSE	mean squared error
MAE	mean absolute error
RMSE	root mean squared error
VM	volatile matter

et al., 2015; Xiao et al., 2020; Yang et al., 2007). On the other hand, once heated in an oxidative atmosphere like air, the virgin biomass and biochar can be further oxidized. As shown in Fig. 1(b), there are three major peaks on a typical DTG curve in an oxidative atmosphere, that is, drying, pyrolysis and char oxidation with increasing temperature. The thermal-oxidative decomposition process under an oxidative atmosphere (e.g., air or pure oxygen) is of vital importance for understanding the heterogeneous kinetics of solid fuels (Di Blasi, 2009; Wang et al., 2017). Biomass and woody materials are important fuels in fires, especially wildfires. Thus, better understandings of the thermal stability of biomass under an oxidative atmosphere is critical to control biomass combustion, thermal processing, and wildfire hazards (Belcher, 2013). However, compared to the thermal biomass in an inert atmosphere, data and reviews of the thermal-oxidative decomposition are much fewer. Moreover, these extra oxidation reactions make the TGA curves more complex, so these data are more difficult to analyze and recognize patterns.

In recent years, machine learning (ML) has been widely applied in various research fields and industrial sectors (Manatura et al., 2023; khan et al., 2023). Researchers have employed different ML models for biomass utilization, such as artificial neural networks (ANN), random forest (RF), support vector machine (SVM), and decision tree (DT). These ML models help investigate the impact of various biomass features on the outputs, determine the optimal performance parameters of a system, and forecast the yields of by-products (Huang et al., 2022; Alabdrabalnabi et al., 2022; Alabdrabalnabi, 2021; Li et al., 2023; Zhu et al., 2019; Zhang et al., 2022; Kartal and Özveren, 2022; Ni et al., 2022; Ni et al., 2021). For example, Zhu *et al.* (Zhu et al., 2019) developed a RF model to predict the yield and carbon contents of biochar from pyrolysis of lignocellulosic biomass. A similar RF model was used by Zhang *et al.* (Zhang et al., 2022) to predict the characteristics of pyrolytic bio-oil. Alaba *et al.* (Adeniyi et al., 2020) examined the accuracy of an ANN model for the prediction of the thermal degradation behaviors of the rice husk in an inert atmosphere. TG data of coal slime and coffee waste co-pyrolysis were also predicted by ML (Ni et al., 2022; Ni et al., 2021).

This paper aims to establish a dataset of the physicochemical parameters of biomass as well as their thermal decomposition properties in an oxidative atmosphere from the literature. Based on the dataset, an ANN model is trained for predicting the TGA results by inputting the biomass characteristics and heating environment. This work provides a better understanding of biomass thermal-oxidative decomposition and a shortcut to obtaining key thermal-oxidative degradation parameters of a given biomass in the absence of TGA.

2. Methodology

2.1. Data collection

This work first reviews the TGA data of thermal-oxidative decomposition, as well as the physiochemical properties of the biomass, including the elemental analysis (EA) and proximate analysis (PA), and

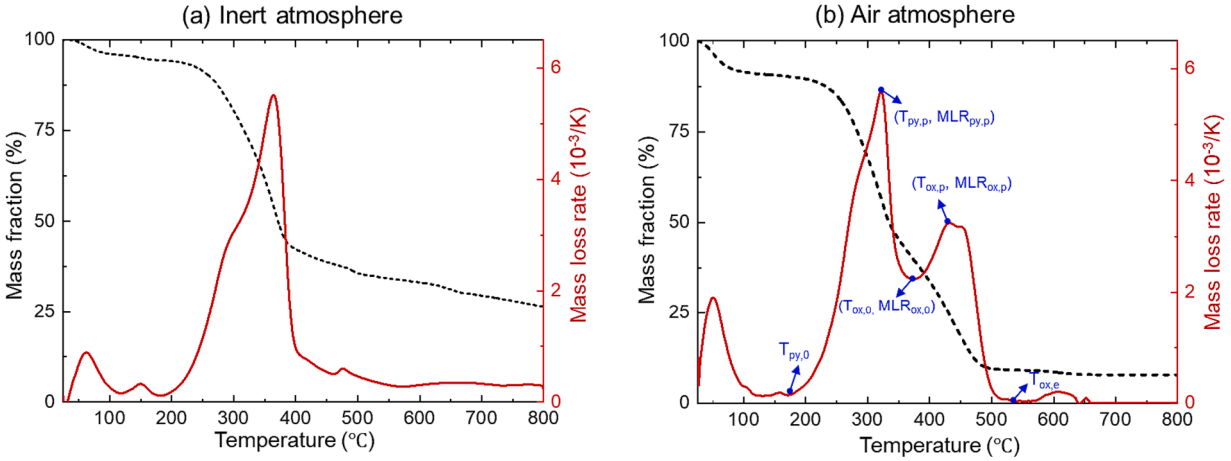


Fig. 1. Typical TG and DTG curves of wood in (a) inert and (b) air atmosphere at heating rate of 10 K/min (Chen et al., 2023).

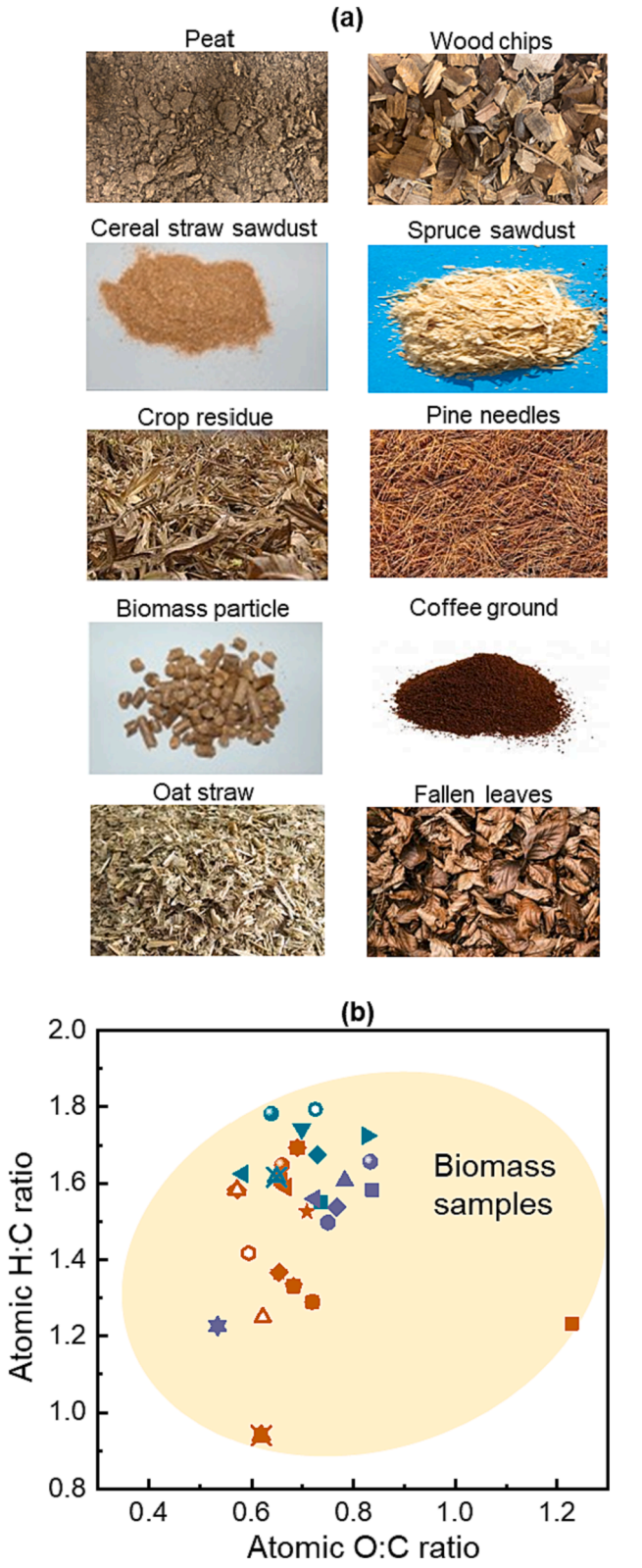


Fig. 2. (a) Photos of representative biomass in the database and (b) Van Krevelen diagram with H:C and O:C molar ratios of the biomass samples in the database, where the orange samples represent the woody biomass, the green samples are the agricultural residues, and the purple samples indicate the peat samples. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

the heating condition (i.e., the constant heating rate) from the literature (Chen et al., 2023; Niu and Liu, 2015; Hameed et al., 2020; Shen et al., 2009; Lin et al., 2019; Mlonka-Mędrala et al., 2019; Ekinci et al., 2021; Mureddu et al., 2018; Mil et al., 2022; Huang and Rein, 2016; Su et al., 2012; Chen et al., 2022; Zhao et al., 2014; Chen et al., 2011; Magdziarz and Wilk, 2013; Cancellieri et al., 2012). Specifically, two databases were selected for the literature review, Web of Science and Scopus. The keywords used for searching include "biomass + thermal decomposition", "biomass + thermogravimetric", "wood + thermal decomposition", "wood + thermogravimetric", "pine + thermal decomposition", "pine + thermogravimetric", "peat + thermal decomposition", "peat + thermogravimetric", "oxygen + thermal decomposition", and "oxygen + thermogravimetric".

If both the TG analysis under air and the elemental/proximate analysis are included in the researched papers, then they will be selected and included in this database. As shown in Fig. 2(a), typical biomass types in the database include different species of wood, peat, cotton, agricultural residues, etc. Before each test, the raw material was dried and ground into smaller size particles for all the biomass samples in the literature. The C-H-O-N compositions and the fixed carbon, volatile matter, and ash contents are all expressed on a dry mass basis (wt.%).

Fig. 2(b) further plots the O:C atomic ratios of the biomass samples, ranging from 0.54 to 1.23, and H:C atomic ratios, ranging from 0.94 to 1.79, within the literature range (Vassilev et al., 2010). The wide range of O:C and H:C ratios also reflect the diversity of the biomass in the established database. The TG tests were conducted under an air atmosphere and programmed with different heating rates. In total, 65 datasets (or a matrix of 65 rows and 18 columns) are acquired, as shown in Table A1, where 59 datasets include the elemental analysis, 52 datasets include proximate analysis, and 47 datasets consist of both analyses. Specifically, the C, H, O, N, H:C atomic ratio, O:C atomic ratio, fixed carbon, volatile matters, fixed carbon, and heating rate, or a combination of several of them will be used as the input features for the machine learning model.

To better characterize the TGA curves, we define eight parameters ($T_{py,0}$, $T_{py,p}$, $T_{ox,0}$, $T_{ox,p}$, $T_{ox,e}$, $MLR_{py,p}$, $MLR_{ox,0}$, and $MLR_{ox,p}$) by extracting five key data points from the DTG curve, as shown in Fig. 1 (b), and they serve as the output targets of the machine learning model.

1) **Pyrolysis-onset temperature ($T_{py,0}$)** refers to the temperature when the fuel starts to thermal decomposition (pyrolysis);
2) **Oxidation-dominant temperature ($T_{ox,0}$)** refers to the temperature when the fuel oxidation starts to dominate or the temperature when the pyrolysis-to-oxidation transition occurs;
3) **Oxidation-dominant rate ($MLR_{ox,0}$)** is the mass loss rate at the oxidation-dominant temperature;
4) **Peak pyrolysis rate ($MLR_{py,p}$)** is the maximum mass loss rate during the pyrolysis, usually the second peak after the first drying peak (not discussed here);
5) **Peak pyrolysis temperature ($T_{py,p}$)** is the temperature when the mass loss rate reaches $MLR_{py,p}$ on the DTG curve;
6) **Peak oxidation rate ($MLR_{ox,p}$)** is the maximum mass loss rate during the char oxidation, usually the third oxidation peak after the second pyrolysis peak;
7) **Peak oxidation temperature ($T_{ox,p}$)** indicates the temperature when the mass loss rate reaches $MLR_{ox,p}$; and
8) **Oxidation-end temperature ($T_{ox,e}$)** indicates the temperature when the oxidation or the whole degradation ends.

These eight parameters were directly acquired from the tables in the literature or extracted from figures with GetData Graph Digitizer.

2.2. Artificial neural network

An artificial neural network (ANN) is a deep sequencing data processing network, which can learn from prior events and establish a

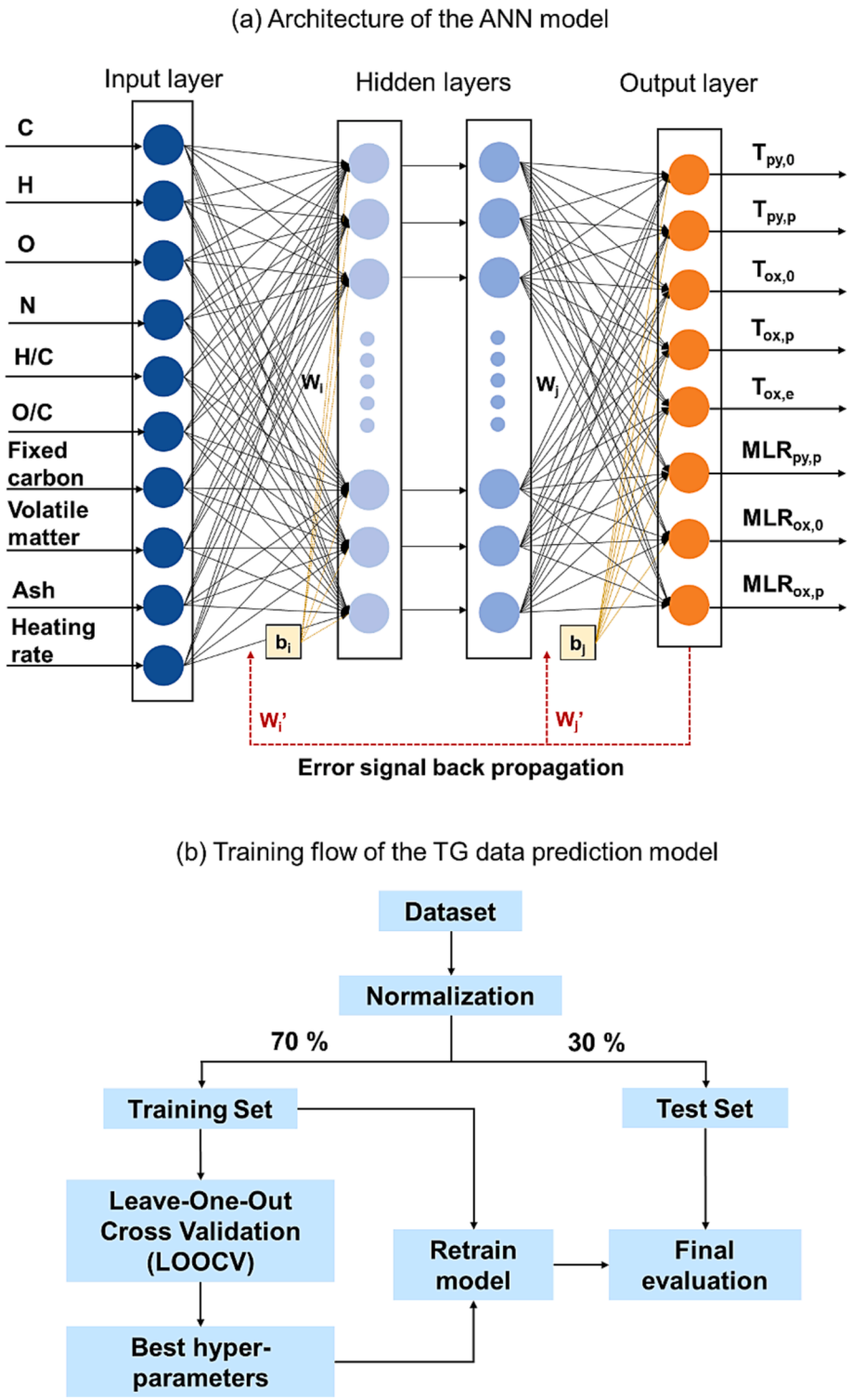


Fig. 3. (a) Architecture of the developed ANN model and (b) training flow of the TG data prediction model.

pattern to generate the desired result (Basu, 2010). This work is based on the back propagation neural network, which is a widely applied ANN model in thermal conversion research (Zhu et al., 2019; Ni et al., 2022; Ni et al., 2021; Adeniyi et al., 2020). The back propagation neural network is a multi-layer feedforward network trained according to an error back propagation algorithm (Li et al.,). As shown in Fig. 3(a), the ANN network typically has an input layer, an output layer, and one or multiple hidden layers with interconnected perceptron. Based on the established database, the elemental analysis (C, H, O, and N contents, and the H/C and O/C atomic ratios), the proximate analysis (fixed carbon, volatile matters, and ash contents), and the heating rate are chosen as features in the input layer. It should be noted that the prediction results with different groups of input features will also be discussed in this paper. The output layer contains eight parameters captured from the DTG curve, i.e., $T_{py,0}$, $T_{py,p}$, $T_{ox,0}$, $T_{ox,p}$, $T_{ox,e}$, $MLR_{py,p}$, $MLR_{ox,0}$, and $MLR_{ox,p}$.

For each perceptron shown in Fig. 3(a), the output of it is calculated by Eq. (1):

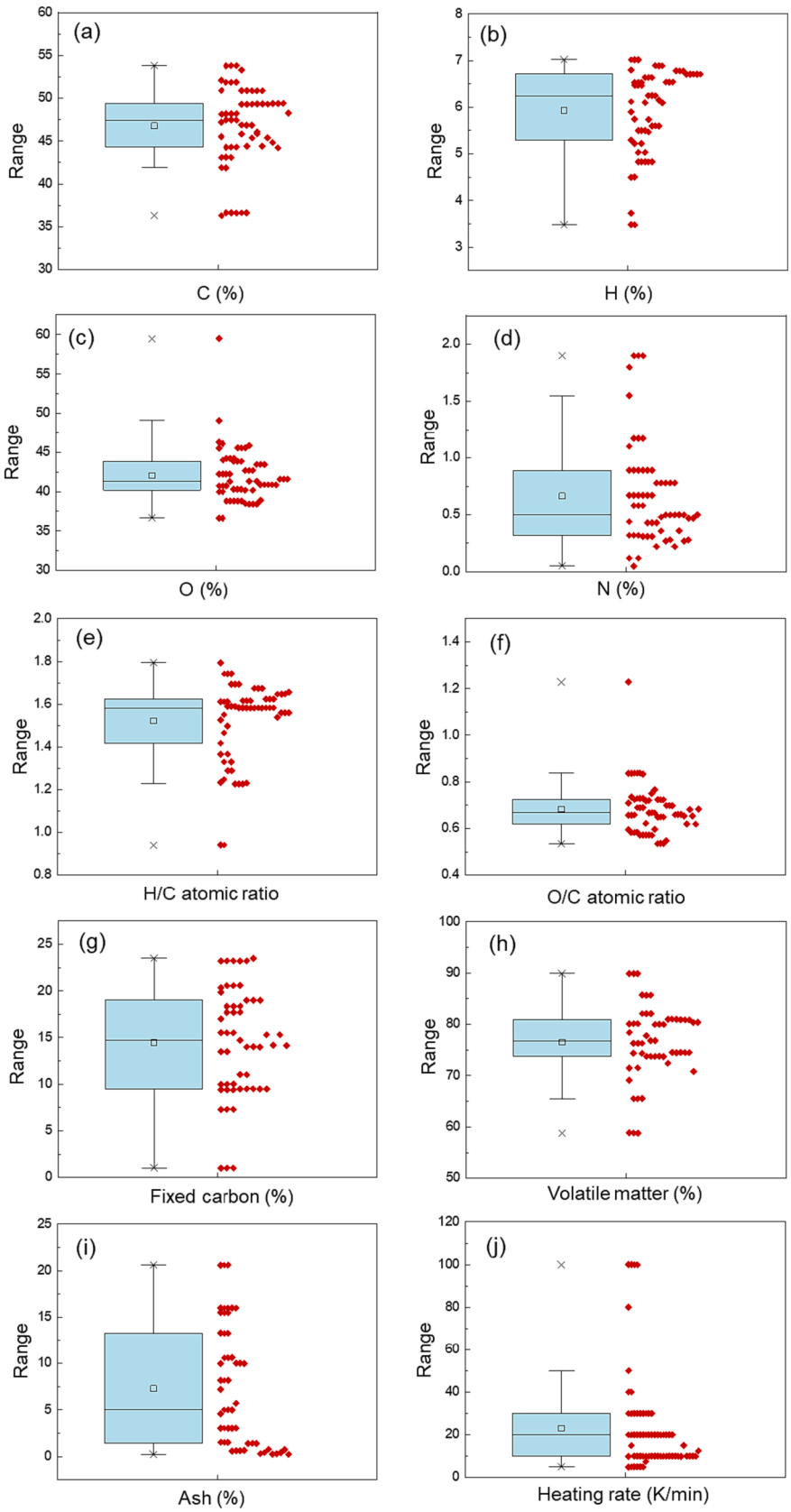


Fig. 4. Box plot of all input features in the dataset: (a) C, (b) H, (c) O, (d) N, (e) H:C atomic ratio, (f) O: C atomic ratio, (g) fixed carbon, (h) volatile matter, (i) ash, and (j) heating rate.

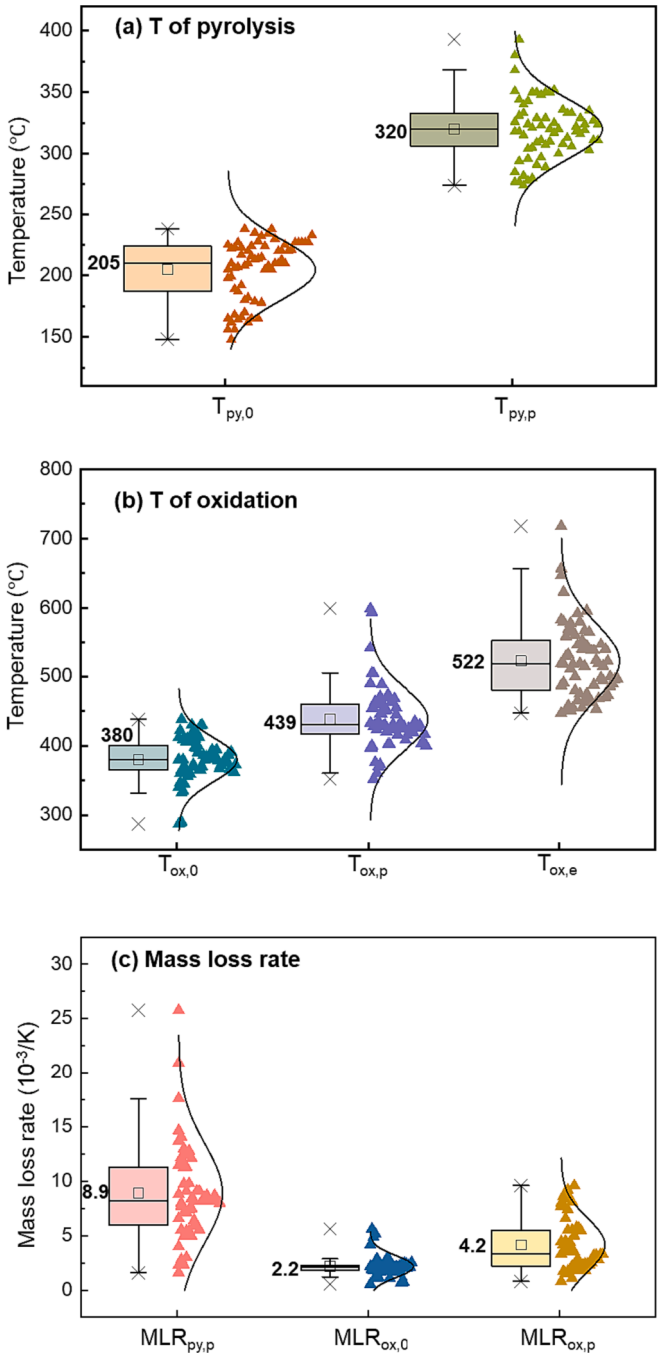


Fig. 5. Box plot of the output variables: (a) onset temperature of pyrolysis ($T_{py,0}$) and temperature of the pyrolysis peak ($T_{py,p}$), (b) pyrolysis-to-oxidation transition temperature ($T_{ox,0}$), temperature of the oxidation peak ($T_{ox,p}$), and end temperature of oxidation ($T_{ox,e}$), and (c) peak pyrolysis rate ($MLR_{py,p}$), oxidation-dominant rate ($MLR_{ox,0}$) and peak oxidation rate ($MLR_{ox,p}$), with the mean value marked.

$$Out_j = h\left(\sum_{i=1}^{N} W_{ji}X_i + B_j\right) \quad (1)$$

where X_i is the input from the preceding layer neuron, W_{ji} is the weight of X_i, B_j is the bias, and h is the transfer (activation) function. Many types of activation functions are available, such as *sigmoid*, *relu*, *tanh*, *tansig*, and so on. In this work, the sigmoid function was set as the transfer function between each layer.

The "Scikit-Learn" open-source library in Python was used to code the ANN model in this work. Fig. 3(b) shows the training flow of the TG data prediction model. Firstly, the whole dataset is normalized before it is randomly split as 70% training set and 30% test set. Then, the best hyper-parameters are determined by the trial-and-error method within the training set with Leave-One-Out Cross Validation (LOOCV), considering the limited size of the dataset. The cross-validation technique is commonly used to prevent the overfitting of the training data, and it is also a re-sampling procedure used to evaluate a model if we have limited data (Géron, 2022).

In LOOCV, each time, only one sample is used as a test set while the rest are used to train the model. This process will be repeated for n times (where n is the total number of observations in the training dataset) and a RMSE will be calculated each time. Then, the average RMSE for n times will be calculated. Finally, the hyper-parameters resulting in the smallest average RMSE will be selected. After that, the model will be

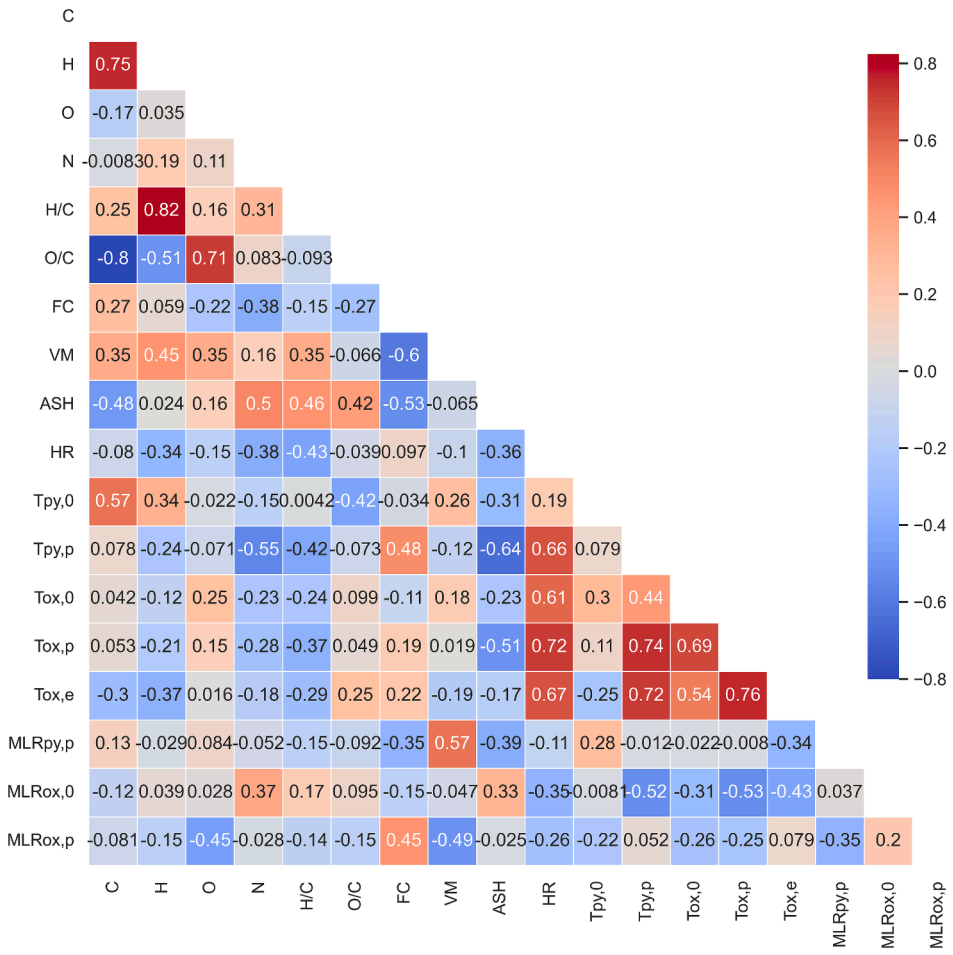


Fig. 6. Pearson correlation coefficient (*r*) among any two variables in the dataset.

Table 1
Hyper-parameters of the developed ANN models with three different groups of input features.

Hyper-parameters	Group (1) (All features)	Group (2) (PA and HR)	Group (3) (EA and HR)
hidden_layer_size	(16, 16)	(9, 9)	(9, 15)
activation function	logistic	logistic	logistic
solver	adam	adam	adam
learning_rate_init	0.0005	0.0005	0.0005
max_iter	20,000	25,000	25,000
tol	2×10^{-10}	2×10^{-10}	2×10^{-10}

retrained with the best hyper-parameters and tested within the 30% test data.

2.3. Performance evaluation

The performance of the developed model is measured by the difference between the predicted and actual values. There are three common criteria to evaluate the model: MSE (mean squared error) in Eq. (2) and R^2 (coefficient of determination) in Eq. (3). The larger R^2 value and smaller MSE indicate the better ANN model.

$$MSE = \frac{1}{N}\sum_{i=1}^{N}\left(Y_{predicted} - Y_{actual}\right)^2 \quad (2)$$

$$R^2 = 1 - \frac{\sum_{i=1}^{N}\left(Y_{predicted} - Y_{actual}\right)^2}{\sum_{i=1}^{N}\left(Y_{predicted} - Y_{mean}\right)^2} \quad (3)$$

where $Y_{predicted}$ is the predicted value by the developed model, Y_{actual} is the target output, Y_{mean} is the average value of target outputs, and N is the number of instances.

3. Results and discussion

3.1. Dataset description

A database of biomass properties and heating environment (input) as well as the TGA in the air atmosphere (output) is established. Specifically, the input features are C, H, O, N, H/C atomic ratio, O/C atomic ratio, fixed carbon (FC), volatile matters (VM), ash, and heating rate (HR). The output layer consists of eight targets, $T_{py,0}$, $T_{py,p}$, $T_{ox,0}$, $T_{ox,p}$, $T_{ox,e}$, $MLR_{py,p}$, $MLR_{ox,0}$, and $MLR_{ox,p}$. Fig. 4 & 5 show the description of all the input features and output targets, respectively, where the box chart on the left includes the maximum, upper bound, 75th percentile, median, mean, 25th percentile, lower bound, and a minimum of the data and the right side shows the specific data points. The bounds (whiskers) of the box chart are determined by the 1.5IQR method, and any data point less than the lower bound or more than the upper bound is considered an outlier (Seo, 2006).

As shown in Fig. 4(a-d), the C, H, O, and N contents of the biomass samples are distributed from 36.30 to 53.81 %, 3.48 to 7.02 %, 36.59 to 59.50 %, and 0.05 to 1.91 %, respectively. And the median C, H, O, and N contents are located at 47.49 %, 6.25 %, 41.29 %, and 0.50 %,

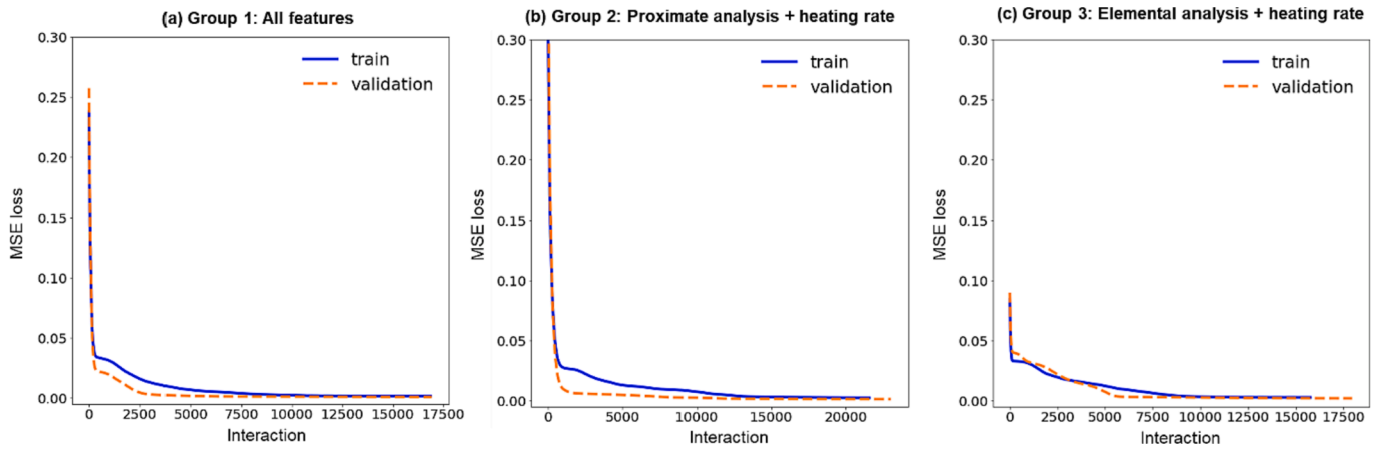


Fig. 7. MSE losses of training and validation of the ANN models during the training process with the inputs of (a) Group (1) (all features), (b) Group (2) (proximate analysis and heating rate), and (c) Group (3) (elemental analysis and heating rate).

respectively. As demonstrated, the carbon content of biomass is generally slightly higher than the oxygen content. Furthermore, Fig. 4(e, f) show the H:C ratio is in the range of 0.94–1.79 with an average value of 1.52, while the O:C ratio ranges from 0.54 to 1.23 with an average value located at 0.68. These ratios show good consistency with the previous research (Vassilev et al., 2010).

In terms of the proximate analysis, Fig. 4(g-i) show the mass fraction of fixed carbon (FC), volatile matter (VM), and ash are in the range of 1.00–23.50 %, 58.82–89.90 %, and 0.24–20.60 %, respectively, which are all dispersed over a wide range. The heating rate is the most important condition controlled in the TGA test and an essential parameter influencing the thermal-oxidative decomposition of fuel. Fig. 4(j) indicates that a vast range of heating rates from 5 to 100 K/min are included in the dataset. As exhibited, the most widely used heating rates in the literature are 10 and 20 K/min. Overall, the input features show a wide range, demonstrating the diversity of the dataset.

The box plots of the output variables are shown in Fig. 5. Fig. 5(a) summarizes the onset temperature of pyrolysis ($T_{py,0}$) and the temperature of the pyrolysis peak ($T_{py,p}$). Generally, the biomass will start pyrolysis at a temperature ($T_{py,0}$) from 148 to 238 ℃ with the median value concentrated at 210 ℃. In comparison, $T_{py,p}$ is in the range of 274–393 ℃ with a median temperature of 320 ℃, meaning that the DTG curve usually reaches its pyrolysis peak at around 320 ℃. At this pyrolysis peak, the mass loss rate ($MLR_{py,p}$) can reach a value ranging from 1.6 to 25.77 10^{-3}/K with an average of 8.9 10^{-3}/K, as shown in Fig. 5(c).

Regarding the oxidation-dominated stage, Fig. 5(b) shows that $T_{ox,0}$, $T_{ox,p}$, and $T_{ox,e}$ is in the range of 287–438 ℃, 352–599 ℃, and 447–718 ℃, respectively. Moreover, $T_{ox,0}$ and $T_{ox,p}$ are almost in accord with normal distribution, which is concentrated at 380 and 430 ℃, respectively. Fig. 5(c) shows the mass loss rate of the oxidation peak ($MLR_{ox,p}$) ranges from 0.8 to 9.6 10^{-3}/K averaging at 4.2 10^{-3}/K. Besides, the mass loss rate of the oxidation peak ($MLR_{ox,p}$) is generally larger than that of the $MLR_{py,p,}$ except for several outlier points. Overall, we can see that most variables in Fig. 5 are almost in accord with normal distribution and scattered over a wide range, which also reflects the diversity of the biomass property and heating environment in the dataset.

3.2. Analysis of Pearson correlation coefficient

Pearson correlation coefficient (r) is a common way of measuring a linear correlation between any two variables. Historically, it is the first formal measure of correlation and still the most widely used method (Kader and Franklin, 2008). The correlation coefficient is described as follows:

$$r = \frac{cov_{x,y}}{\sigma_x \bullet \sigma_y} = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum (x_i - \bar{x})^2}\sqrt{\sum (y_i - \bar{y})^2}} \quad (4)$$

where $cov_{x,y}$ is the covariance between variable x and y; σ_x and σ_y are the standard deviation of variables x and y, respectively; and $\bar{x}$ and $\bar{y}$ are the mean values of x and y. A large positive coefficient value (close to 1.0) indicates a strong positive correlation, while a small negative value (near −1.0) indicates a strong negative correlation.

Furthermore, to determine if a correlation coefficient is statistically significant, the t-score (Eq. (5)) and p-value are calculated.

$$t = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}} \quad (5)$$

where n is the sample size, where the degrees of freedom are n-2, and r is the Pearson correlation coefficient calculated by Eq. (5). After that, the p-value is calculated as the corresponding two-sided p-value for the t-distribution with n-2 degrees of freedom. Generally, most scholars refer to statistically significant as $p < 0.05$ and statistically highly significant as $p < 0.001$ (Thiese et al., 2016).

Fig. 6 shows the half Pearson correlation coefficient (r) matrix, where each row represents a variable in the dataset, and all the columns represent the same variables as rows. Each cell represents the value of the correlation coefficient between two variables. The reliability of the linear correlation (p-value) is summarized in Table A2. As reported, there is a high correlation ($p < 0.01$) between the following input features and output targets: C-$T_{py,0}$, N-$T_{py,p}$, H/C-$T_{py,p}$, FC-$T_{py,p}$, Ash-$T_{py,p}$, HR-$T_{py,p}$, HR-$T_{ox,0}$, Ash-$T_{ox,p}$, HR-$T_{ox,p}$, HR-$T_{ox,e}$, VM-$MLR_{py,p}$, O-$MLR_{ox,p}$, FC-$MLR_{ox,p}$, and VM-$MLR_{ox,p}$.

Fig. 6 shows the carbon content has a positive correlation with $T_{py,0}$, with r of 0.57. This means that the biomass with a higher carbon content often has higher thermal stability. $T_{py,p}$ is negatively correlated with nitrogen content, H/C ratio, and ash, with the Pearson correlation coefficient of −0.55, −0.42, and −0.64, respectively, whereas its correlation with fixed carbon is positive ($r = 0.48$). This means the effects of carbon content on $T_{py,p}$ is consistent with that on $T_{py,0}$. The heating rate is closely positively related to many output variables, and its Pearson correlation coefficients with $T_{py,p}$, $T_{ox,0}$, $T_{ox,p}$, and $T_{ox,e}$ are all larger than 0.6. It can be inferred that the heating environment has significant effects on the thermal-oxidative decomposition behaviour of biomass, and a higher heating rate may cause the DTG curves to shift to the higher temperature. This is possible because the heat transfer in the biomass particles is better when the heating rate is smaller (Cai et al., 2018).

The $MLR_{py,p}$ claims a positive correlation with the volatile matter (VM) ($r = 0.57$), which means the mass loss rate of the pyrolysis peak of the biomass with a larger VM content might be larger. However, the

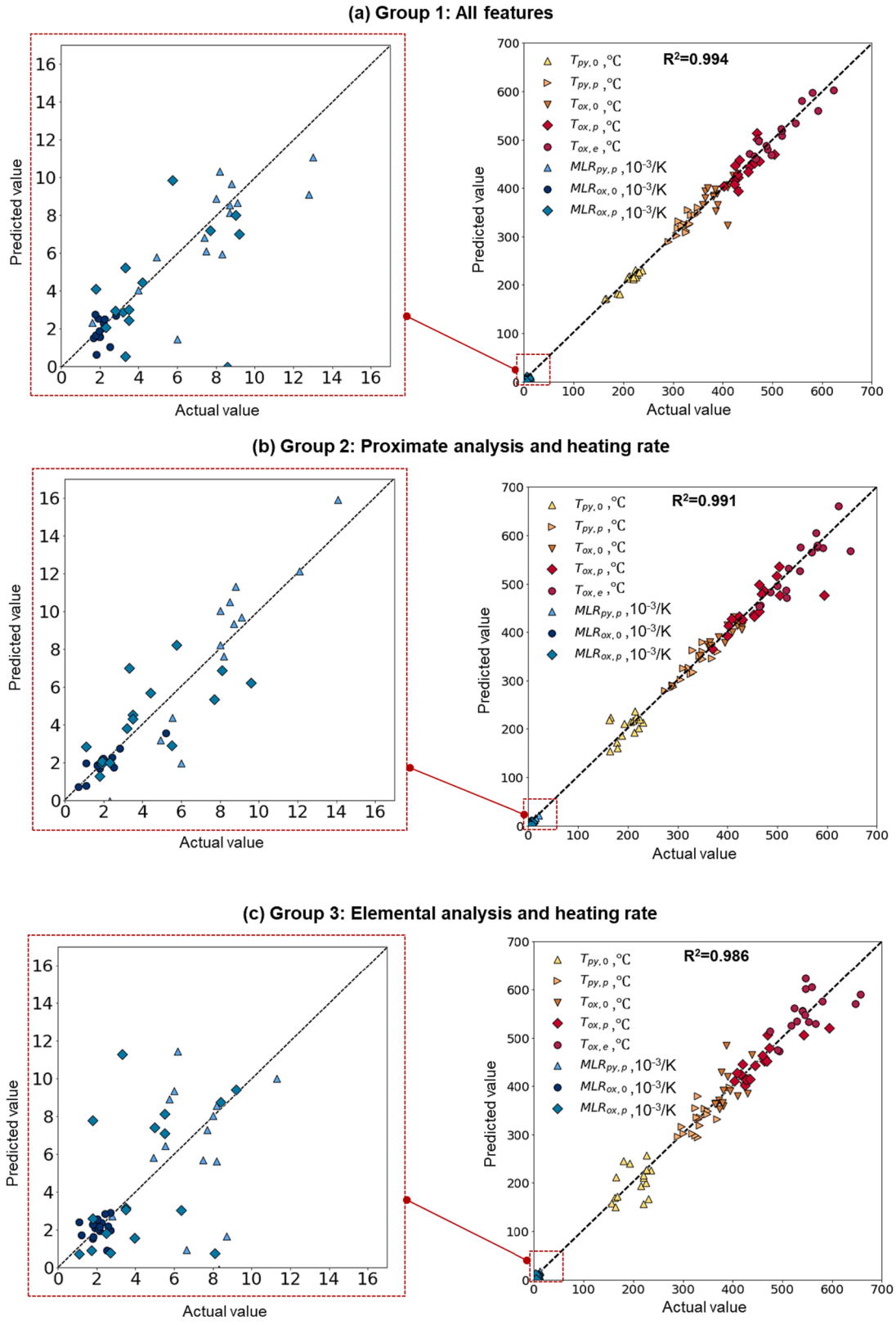


Fig. 8. Predicted value versus true value in the test set using the developed ANN model with the input of (a) Group (1) (all features), (b) Group (2) (proximate analysis and heating rate), and (c) Group (3) (elemental analysis and heating rate).

$MLR_{ox,p}$ shows a negative correlation with the volatile matter (r = -0.49), while is positively correlated with the fixed carbon (FC) (r = 0.45). The model with different groups of input features will be discussed in this paper. Also, the strong correlation between any two output variables indicates the rationality to output them as a group. For example, the r-values between $T_{py,p}$ and $T_{ox,p}$, and between $T_{ox,p}$ and $T_{ox,e}$ are 0.76 and 0.73, respectively. All in all, from the Pearson correlation coefficient analysis, it can be seen that the heating rate (HR), C content,

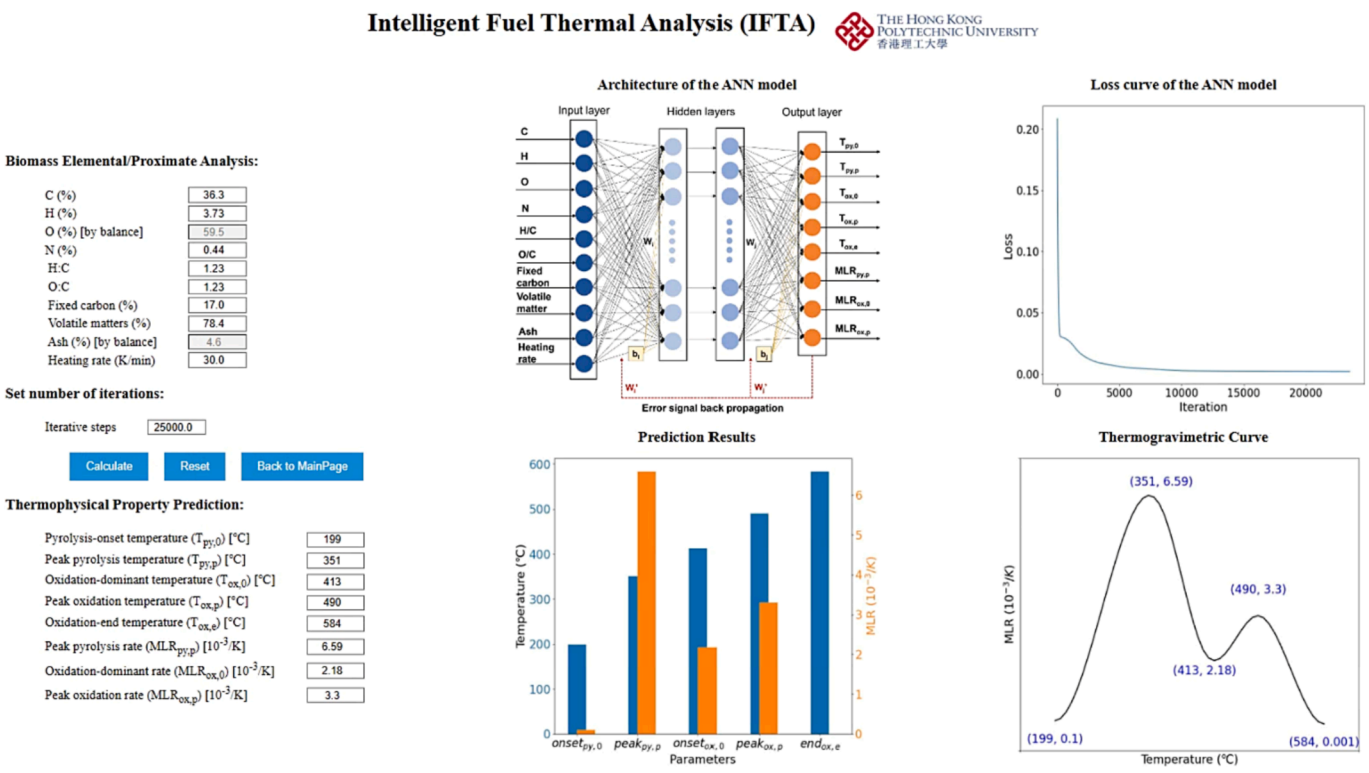


Fig. 9. User Interface of Intelligent Fuel Thermal Analysis (IFTA) (open-access at http://ai-fuel.firelabxy.com/).

fixed carbon (FC), volatile matter (VM), and ash content are the most related parameters with the DTG data of the biomass.

3.3. Evaluation of the ANN model performance

The hyper-parameters of the best training models are determined with Leave-One-Out Cross-Validation (LOOCV) and used to retrain the prediction model. The ANN models with different groups of input features are developed and evaluated, where Group (1) includes all the 10 features introduced in *Section 2.1*, Group (2) selects the proximate analysis (PA) and heating rate (HR) as inputs (4 features), and Group (3) chooses the elemental analysis (EA) and heating rate (HR) as inputs (5 features).

After optimization by the trial-and-error method within the training set, the hyper-parameters of the developed ANN models with different inputs are summarized in Table 1. Specifically, it could achieve cross-validation RMSE scores of 0.22, 0.19, and 0.26, respectively. As shown in Table 1, the developed ANN models with different groups of input all have the same architecture of two hidden layers, while having different nodes in each layer. Between each layer, the logistic sigmoid function is set as the transfer function, and a stochastic gradient-based optimizer is chosen for weight optimization.

The loss functions of training and validation during the training process are presented in Fig. 7. As shown, the developed models with different groups of inputs all reach convergence with a minimum MSE loss close to zero. Thus, the ANN model can accurately predict TGA data in the training dataset.

Moreover, the performances of the developed models evaluated on the test set are shown in Fig. 8. The evaluation of the ANN models on the test set all reach an acceptable level with an overall R^2 score larger than 0.98, which is significantly better than the predictions of key thermal decomposition temperatures in previous literature (Beste and Barnes, 2020). As shown in Fig. 8(a), when all the decomposition features are introduced to the model, the R^2 scores of the model reached a maximum value of 0.994. Fig. 8(b-c) show that when the input features are reduced, the accuracy of the ANN models will decrease but still keep at an acceptable level with the R^2 scores larger than 0.980. The reason for the reduced overall accuracy is that the model cannot precisely predict the MLR when only a single type of chemical analysis is input.

(b-c) exhibit that the R^2 scores for the ANN models with the input of proximate analysis and heating rate, and the input of elemental analysis and heating rate are 0.991 and 0.986, respectively. By comparison, we found the model with input features of proximate analysis in Fig. 8(b) has a better performance than that of input elemental analysis in Fig. 8 (c). Therefore, it can be concluded that with any of the proximate analysis, elemental analysis, and both, the developed ANN model can accurately predict the key points on the DTG curves of biomass. Overall, the results demonstrate that the developed ANN model can accurately predict key parameters of biomass thermal-oxidative decomposition.

With the established ANN model, an open-access online software named Intelligent Fuel Thermal Analysis (IFTA) (https://ai-fuel.firelabxy.com/) has been further developed for readers to get the biomass DTG curves under air quickly. Fig. 9 shows the operation panel of the IFTA. After inputting the biomass elemental/proximate analysis and iterative steps manually, this AI software can calculate the key TGA parameters, as displayed in the illustrative DTG curve. This code will be further updated by add more biomass TGA data into the database for the training of AI.

4. Conclusions

This work reports a comprehensive review of the thermal decomposition data of a wide range of biomass in the oxidative atmosphere. Based on the database collected from the literature, an ANN model is developed for the first time to predict the key points on the derivative thermogravimetry (DTG) curve using the fuel properties and the heating environment. Pearson correlation coefficient analysis reveals that the heating rate is more correlated to the biomass thermal-oxidative decomposition behavior than the biomass properties.

The developed ANN model can successfully predict the TG data

Table A1
Dataset of biomass properties and TG data.

References	Biomass type	Input										Output							
		C (%)	H (%)	O (%)	N (%)	H/C*	O/C*	FC (%)	VM (%)	Ash (%)	Rate (K/min)	$T_{py,0}$ (℃)	$T_{py,p}$ (℃)	$T_{ox,0}$ (℃)	$T_{ox,p}$ (℃)	$T_{ox,e}$ (℃)	$MLR_{py,p}$ (10^{-3}/K)	$MLR_{ox,0}$ (10^{-3}/K)	$MLR_{ox,p}$ (10^{-3}/K)
Su *et al.*_2012 (Su et al., 2012)	Pine wood	48.14	6.12	45.48	0.12	1.23	1.23	19.89	69.08	10	20	205	344	380	454	487	2.3	0.55	0.8
Niu *et al.*_2015 (Niu and Liu, 2015)	Pine branch	48.19	6.47	42.24	0.58	1.53	0.71	18.36	80.12	1.52	10	165	326	360	454	519	8.8	2.0	3.5
		48.19	6.47	42.24	0.58	1.61	0.66	18.36	80.12	1.52	15	162	331	365	464	518	8.0	1.9	3.5
		48.19	6.47	42.24	0.58	1.61	0.66	18.36	80.12	1.52	20	168	333	380	460	529	7.7	1.8	3.5
Wang *et al.*_2021 (Wang et al., 2021)	Wood sawdust	44.29	6.25	40.72	0.43	1.61	0.66	15.52	76.32	8.16	10	225	321	356	397	450	11.5	2.25	3.9
		44.29	6.25	40.72	0.43	1.69	0.69	15.52	76.32	8.16	20	222	318	359	397	458	11.9	2.2	2.2
		44.29	6.25	40.72	0.43	1.69	0.69	15.52	76.32	8.16	30	210	315	359	400	447	9.8	4.17	3.5
Chen *et al.*_2022 (Chen et al., 2023)	Wood chips	36.30	3.73	59.47	0.44	1.69	0.69	17	78.4	4.6	30	198	351	412	490	583	6.6	2.17	3.33
Ekinci *et al.*_2020 (Ekinci et al., 2021)	Wood (Pine)	50.90	5.30	42.22	1.55	1.25	0.62	14.7	77.8	0.6	10	223	329	369	471	511	13.7	2.88	4.46
		49.30	6.53	43.85	0.32	1.59	0.67	9.4	89.9	0.63	10	207	340	398	460	518	14.7	1.20	1.74
		49.30	6.53	43.85	0.32	1.59	0.67	9.4	89.9	0.63	40	217	380	422	599	580	14.1	1.10	1.55
		49.30	6.53	43.85	0.32	1.59	0.67	9.4	89.9	0.63	80	227	393	438	542	657	12.2	1.20	1.74
Medrala *et al.*_2019 (Mlonka-Mędrala et al., 2019)	Mixed cereal straw (MCS1)	49.32	6.64	42.70	0.89	1.62	0.65	7.3	82.1	10.6	5	208	277	340	428	451	9.1	2.08	2.7
		49.32	6.64	42.70	0.89	1.62	0.65	7.3	82.1	10.6	10	223	285	409	430	455	9.8	1.67	2.7
		49.32	6.64	42.70	0.89	1.62	0.65	7.3	82.1	10.6	20	238	294	409	423	480	7.1	2.39	2.2
	Mixed cereal straw (MCS2)	47.46	6.89	44.21	0.89	1.74	0.70	10	80	10	5	210	282	396	435	465	8.4	2.17	2.3
		47.46	6.89	44.21	0.89	1.74	0.70	10	80	10	10	216	296	405	446	475	8.2	2.11	2.4
		47.46	6.89	44.21	0.89	1.74	0.70	10	80	10	20	220	304	410	446	500	7.1	2.35	2.4
	Corn straw and MxG[1] and SV[2] (CS)	46.86	6.54	45.54	0.67	1.67	0.73	1	85.7	13.3	5	213	280	408	428	455	9.1	2.84	1.9
		46.86	6.54	45.54	0.67	1.67	0.73	1	85.7	13.3	10	224	291	410	430	466	9.1	2.84	2.3
		46.86	6.54	45.54	0.67	1.67	0.73	1	85.7	13.3	20	235	297	390	420	475	8.2	2.70	2.5
	Miscanthus and giganteus (MxG)	51.86	7.02	40.29	0.67	1.62	0.58	14	81	5	5	214	305	365	402	463	8.5	1.98	3.2
		51.86	7.02	40.29	0.67	1.62	0.58	14	81	5	10	228	309	365	402	463	8.7	1.98	3.2
		51.86	7.02	40.29	0.67	1.62	0.58	14	81	5	20	232	311	370	423	488	7.4	1.97	2.8
	Salix viminalis (SV)	49.39	6.78	43.46	0.31	1.65	0.66	17.7	80.9	1.4	5	224	307	386	431	453	13.0	2.21	4.2
		49.39	6.78	43.46	0.31	1.65	0.66	17.7	80.9	1.4	10	234	310	380	425	468	11.3	2.09	4.5
		49.39	6.78	43.46	0.31	1.65	0.66	17.7	80.9	1.4	20	238	317	384	451	498	8.2	1.75	3.3
Shen *et al.*_2009 (Shen et al., 2009)	Pine	41.89	4.50	40.19	0.22	1.29	0.72	15.31	71.49	0.30	10	215	329	381	445	494	11.3	2.10	5.5
	Aspens	45.84	5.22	39.97	0.36	1.37	0.65	11.03	80.37	0.41	10	205	320	382	416	470	12.4	2.14	8.0
	Birch	44.41	3.48	36.65	0.27	0.94	0.62	13.49	74.36	0.76	10	208	325	389	431	473	12.8	1.98	8.6
	Oak	45.37	5.03	41.29	0.28	1.33	0.68	14.16	76.82	0.24	10	205	320	394	452	486	12.1	2.06	4.4
	Pine	41.89	4.50	40.19	0.22	1.29	0.72	15.31	71.49	0.30	100	230	368	430	594	647	8.2	1.1	1.4
	Aspens	45.84	5.22	39.97	0.36	1.37	0.65	11.03	80.37	0.41	100	220	343	421	474	546	8.7	1.8	1.8
	Birch	44.41	3.48	36.65	0.27	0.94	0.62	13.49	74.36	0.76	100	220	350	410	470	559	8.3	1.83	1.83
	Oak	45.37	5.03	41.29	0.28	1.33	0.68	14.16	76.82	0.24	100	210	350	420	505	569	8.0	1.68	1.68
Mureddu *et al.*_2018 (Mureddu et al., 2018)	Stone pine	50.88	6.71	38.80	0.50	1.58	0.57	23.23	73.74	3.03	20	210	336	388	433	519	4.0	2.18	9.0
		50.88	6.71	38.80	0.50	1.58	0.57	23.23	73.74	3.03	10	225	333	381	425	490	7.5	2.26	9.2
		50.88	6.71	38.80	0.50	1.58	0.57	23.23	73.74	3.03	30	220	348	394	463	540	2.8	2.09	8.4
		50.88	6.71	38.80	0.50	1.58	0.57	23.23	73.74	3.03	40	221	350	400	463	578	2.3	1.83	8.1
		50.88	6.71	38.80	0.50	1.58	0.57	23.23	73.74	3.03	50	213	350	412	504	623	1.6	1.67	7.7
	Eucalyptus	48.30	5.90	38.47	0.12	1.47	0.60	20.38	3.03	7.21	20	210	333	379	427	515	3.0	2.46	7.9
Yang *et al.*_2016 (Yang et al., 2016)	Raw sphagnum moss peat	36.62	4.83	40.87	0.78	1.58	0.84	9.5	74.5	16	5	164	289	367	409	545	5.53	2.42	5.51
		36.62	4.83	40.87	0.78	1.58	0.84	9.5	74.5	16	10	156	300	372	416	540	4.94	2.04	5
		36.62	4.83	40.87	0.78	1.58	0.84	9.5	74.5	16	15	166	308	375	425	565	5.42	2.26	6.58
		36.62	4.83	40.87	0.78	1.58	0.84	9.5	74.5	16	20	170	313	377	420	561	5.41	2.11	7.39
		36.62	4.83	40.87	0.78	1.58	0.84	9.5	74.5	16	30	187	322	382	424	592	4.94	1.90	5.76
Chen *et al.*_2022 (Chen et al., 2022)	Peat	46.09	5.75	46.09	0.47	1.50	0.75	23.5	70.8	5.7	30	192	329	429	468	580	5.99	2.53	3.3

(*continued on next page*)

Table A1 (*continued*)

References	Biomass type	Input										Output							
		C (%)	H (%)	O (%)	N (%)	H/C*	O/C*	FC (%)	VM (%)	Ash (%)	Rate (K/min)	$T_{py,0}$ (℃)	$T_{py,p}$ (℃)	$T_{ox,0}$ (℃)	$T_{ox,p}$ (℃)	$T_{ox,e}$ (℃)	$MLR_{py,p}$ (10^{-3}/K)	$MLR_{ox,0}$ (10^{-3}/K)	$MLR_{ox,p}$ (10^{-3}/K)
Magdziarz *et al.*_2013 (Magdziarz and Wilk, 2013)	Wood biomas	47.20	6.10	46.30	0.05	1.55	0.74				10	183	325	378	458	529	12.0	1.84	6.0
	Oat	45.50	6.80	44.00	1.80	1.79	0.73				10	165	296	376	497	590	14.4	1.08	2.1
Lachman *et al.*_2022 (Mil et al., 2022)	Spruce	52.08	6.15	41.29	0.48	1.42	0.59				10	206	323	373	489	718	11.45	1.12	1.22
Huang *et al.*_2016 (Huang and Rein, 2016)	Irish moss peat	53.80	5.50	38.40	1.90	1.23	0.54				10	162	306	347	419	538	6.11	2.62	6.63
		53.80	5.50	38.40	1.90	1.23	0.54				20	165	318	377	435	553	5.76	2.58	6.35
		53.80	5.50	38.40	1.90	1.23	0.54				30	180	324	387	427	567	6.18	2.7	8.11
Cancellieri *et al.*_2012 (Cancellieri et al., 2012)	Scotland peat 1	53.32	5.47	38.90	1.10	1.23	0.55				20	227	316	366	426	596	4.45	1.94	3.3
	Siberian peat 2	44.81	5.74	45.84	0.47	1.54	0.77				20	227	335	391	430	538	6.55	1.88	13.7
	Siberian peat 3	43.09	5.60	41.56	1.17	1.56	0.72				20	227	327	364	403	524	6.65	2.16	3.95
		43.09	5.60	41.56	1.17	1.56	0.72				10	227	319	362	406	519	4.31	2.5	3.22
		43.09	5.60	41.56	1.17	1.56	0.72				30	227	324	373	417	547	3.85	2.15	2.71
Lin *et al.*_2019 (Lin and Huang, 2021)	Estonian peat	44.20	6.10	49.10	0.50	1.66	0.83				30	233	352	430	455	565	6.76	2.19	9.0
Zhao *et al.*_2014 (Zhao et al., 2014)	Changbai Mountain forest peat							20.58	58.82	20.6	10	148	303	332	352	515	6.41	5.62	5.70
								20.58	58.82	20.6	20	156	311	335	376	540	5.91	4.86	5.46
								20.58	58.82	20.6	30	179	311	345	400	522	5.56	5.19	9.60
Chen *et al.*_2011 (Chen et al., 2011)	Changbai Mountain forest peat							19	65.5	15.5	7.5	165	277	287	361	496	25.77	0.68	1.97
								19	65.5	15.5	10	178	274	288	370	500	20.93	0.70	1.92
								19	65.5	15.5	12.5	182	278	290	373	497	17.6	0.76	1.82

*: H/C and O/C indicate the atomic ratios which are calculated from C, H, and O mass fractions.

Table A2
The *p*-value of the *t*-test between any two variables (The red value indicates $p < 0.01$).

		C	H	O	N	H/C	O/C	FC	VM	Ash	HR	$T_{py,0}$	$T_{py,p}$	$T_{ox,0}$	$T_{ox,p}$	$T_{ox,e}$	$MLR_{py,p}$	$MLR_{ox,0}$	$MLR_{ox,p}$
C	r	1	0.750545	−0.16935	−0.00828	0.246916	−0.80039	0.269172	0.352771	−0.48218	−0.08022	0.568292	0.07796	0.042223	0.078063	−0.30028	0.132203	−0.11853	−0.08072
	p	0	1.24E-09	0.255114	0.955961	0.094286	1.46E-11	0.067312	0.015011	0.000599	0.591933	3.09E-05	0.602459	0.778097	0.601979	0.040287	0.375708	0.427474	0.58963
H	r	0.750545	1	0.035024	0.190503	0.824145	−0.51311	0.059365	0.451093	0.023678	−0.34015	0.343	−0.23542	−0.12424	−0.15596	−0.36968	−0.02869	0.038616	−0.15287
	p	1.24E-09	0	0.815204	0.199619	1.1E-12	0.000226	0.691829	0.001461	0.874475	0.019314	0.018262	0.111166	0.405403	0.295158	0.010545	0.848153	0.796633	0.303961
O	r	−0.16935	0.035024	1	0.105664	0.160577	0.714081	−0.22373	0.34901	0.160082	−0.1518	−0.0222	−0.07097	0.249383	0.159311	0.016298	0.08436	0.027902	−0.45205
	p	0.255114	0.815204	0	0.479647	0.280931	1.75E-08	0.130602	0.016199	0.282434	0.308405	0.882276	0.635479	0.090938	0.284791	0.913413	0.572897	0.852308	0.001423
N	r	−0.00828	0.190503	0.105664	1	0.307499	0.082593	−0.38091	0.163309	0.496365	−0.37653	−0.14911	−0.55241	−0.23289	−0.28727	−0.1758	−0.05183	0.374491	−0.02783
	p	0.955961	0.199619	0.479647	0	0.035502	0.581	0.008254	0.272717	0.000387	0.00909	0.317144	5.68E-05	0.115169	0.050241	0.237196	0.729357	0.009503	0.852673
H/C	r	0.246916	0.824145	0.160577	0.307499	1	−0.09261	−0.14728	0.350195	0.457548	−0.43339	0.004245	−0.4168	−0.23548	−0.30668	−0.28722	−0.15168	0.169724	−0.14277
	p	0.094286	1.1E-12	0.280931	0.035502	0	0.535824	0.323196	0.015816	0.001222	0.002342	0.977406	0.003564	0.111078	0.036022	0.050288	0.308774	0.254063	0.338375
O/C	r	−0.80039	−0.51311	0.714081	0.082593	−0.09261	1	−0.27172	−0.06617	0.422604	−0.0393	−0.42024	−0.07303	0.09893	0.033955	0.250439	−0.09215	0.094630	−0.14985
	p	1.46E-11	0.000226	1.75E-08	0.581	0.535824	0	0.064664	0.658569	0.003084	0.793099	0.003272	0.625665	0.508232	0.820752	0.089533	0.537844	0.526925	0.314735
FC	r	0.269172	0.059365	−0.22373	−0.38091	−0.14728	−0.27172	1	−0.59932	−0.53035	0.097031	−0.03443	0.483746	−0.1079	0.127536	0.219297	−0.34775	−0.15126	0.451447
	p	0.067312	0.691829	0.130602	0.008254	0.323196	0.064664	0	8.54E-06	0.000126	0.516448	0.818307	0.000571	0.470359	0.392943	0.138597	0.016613	0.310159	0.001447
VM	r	0.352771	0.451093	0.34901	0.163309	0.350195	−0.06617	−0.59932	1	−0.06511	−0.10014	0.260455	−0.11821	0.177569	0.130431	−0.19181	0.56902	−0.04668	−0.4943
	p	0.015011	0.001461	0.016199	0.272717	0.015816	0.658569	8.54E-06	0	0.663684	0.50303	0.077033	0.428712	0.232443	0.382199	0.196505	3.00E-05	0.755339	0.000413
Ash	r	−0.48218	0.023678	0.160082	0.496365	0.457548	0.422604	−0.53035	−0.06511	1	−0.36402	−0.31097	−0.64212	−0.2278	−0.49517	−0.17086	−0.38759	0.328689	−0.0249
	p	0.000599	0.874475	0.282434	0.000387	0.001222	0.003084	0.000126	0.663684	0	0.01189	0.033372	1.14E-06	0.123565	0.000402	0.250865	0.007108	0.024085	0.867924
HR	r	−0.08022	−0.34015	−0.1518	−0.37653	−0.43339	−0.0393	0.097031	−0.10014	−0.36402	1	0.193622	0.662692	0.6076	0.656408	0.668899	−0.11087	−0.34584	−0.26173
	p	0.591933	0.019314	0.308405	0.00909	0.002342	0.793099	0.516448	0.50303	0.01189	0	0.192219	3.88E-07	5.92E-06	5.45E-07	2.76E-07	0.458123	0.017262	0.07553
$T_{py,0}$	r	0.568292	0.343	−0.0222	−0.14911	0.004245	−0.42024	−0.03443	0.260455	−0.31097	0.193622	1	0.078783	0.297242	0.109831	−0.24821	0.27559	−0.00810	−0.21847
	p	3.09E-05	0.018262	0.882276	0.317144	0.977406	0.003272	0.818307	0.077033	0.033372	0.192219	0	0.598622	0.042455	0.46239	0.092516	0.060793	0.956892	0.14012
$T_{py,p}$	r	0.07796	−0.23542	−0.07097	−0.55241	−0.4168	−0.07303	0.483746	−0.11821	−0.64212	0.662692	0.078783	1	0.444956	0.758975	0.721357	−0.01216	−0.52065	0.05193
	p	0.602459	0.111166	0.635479	5.68E-05	0.003564	0.625665	0.000571	0.428712	1.14E-06	3.88E-07	0.598622	0	0.001726	6.32E-10	1.06E-08	0.935345	0.000175	0.72885
$T_{ox,0}$	r	0.042223	−0.12424	0.249383	−0.23289	−0.23548	0.09893	−0.1079	0.177569	−0.2278	0.6076	0.297242	0.444956	1	0.677517	0.542974	−0.02164	−0.30987	−0.25673
	p	0.778097	0.405403	0.090938	0.115169	0.111078	0.508232	0.470359	0.232443	0.123565	5.92E-06	0.042455	0.001726	0	1.69E-07	8.04E-05	0.885195	0.034036	0.081505
$T_{ox,p}$	r	0.078063	−0.15596	0.159311	−0.28727	−0.30668	0.033955	0.127536	0.130431	−0.49517	0.656408	0.109831	0.758975	0.677517	1	0.72564	−0.00080	−0.52999	−0.25366
	p	0.601979	0.295158	0.284791	0.050241	0.036022	0.820752	0.392943	0.382199	0.000402	5.45E-07	0.46239	6.32E-10	1.69E-07	0	7.9E-09	0.957407	0.000127	0.085358
$T_{ox,e}$	r	−0.30028	−0.36968	0.016298	−0.1758	−0.28722	0.250439	0.219297	−0.19181	−0.17086	0.668899	−0.24821	0.721357	0.542974	0.72564	1	−0.34016	−0.43222	−0.07865
	p	0.040287	0.010545	0.913413	0.237196	0.050288	0.089533	0.138597	0.196505	0.250865	2.76E-07	0.092516	1.06E-08	8.04E-05	7.9E-09	0	0.01931	0.002414	0.599255
$MLR_{py,p}$	r	0.132203	−0.02869	0.084362	−0.05183	−0.15168	−0.09215	−0.34775	0.569015	−0.38759	−0.11087	0.275594	−0.01216	−0.02164	−0.00801	−0.34016	1	0.036687	−0.35090
	p	0.375708	0.848153	0.572897	0.729357	0.308774	0.537844	0.016613	0.000030	0.007108	0.458123	0.060794	0.935345	0.885195	0.957407	0.019312	0	0.806596	0.01559
$MLR_{ox,0}$	r	−0.11853	0.03862	0.027902	0.374491	0.169724	0.094631	−0.15126	−0.04669	0.328689	−0.34584	−0.00810	−0.52065	−0.30987	−0.52999	−0.43222	0.036687	1	0.20048
	p	0.427474	0.796633	0.852308	0.009504	0.254063	0.526925	0.310159	0.755339	0.024084	0.017262	0.956892	0.000176	0.034036	0.000127	0.002414	0.806596	0	0.176639
$MLR_{ox,p}$	r	−0.08072	−0.15287	−0.45205	−0.02783	−0.14277	−0.14984	0.45144	−0.49431	−0.02492	−0.26173	−0.21847	0.05192	−0.25673	−0.25365	0.078647	−0.35091	0.200480	1
	p	0.589633	0.304961	0.001423	0.852673	0.338375	0.314735	0.001447	0.000413	0.867924	0.075538	0.140120	0.728854	0.081505	0.085358	0.599255	0.015590	0.176639	0

based on either the proximate analysis and heating rate or elemental analysis and heating rate, with R^2 scores larger than 0.98. The best performance is found in the model will all the features input, and the model with proximate analysis performs better than that with elemental analysis. By implementing this ANN model to an open-access online software, Intelligent Fuel Thermal Analysis (IFTA), users can quickly get an accurate prediction of key TGA data for a wide range of heating rates and various biomass types. This work provides a better understanding of biomass thermal-oxidative decomposition and a shortcut to obtaining critical biomass information without the complex process of thermogravimetric analysis.

CRediT authorship contribution statement

Yuying Chen: Investigation, Methodology, Writing – original draft, Formal analysis. **Zilong Wang:** Investigation, Methodology, Software. **Shaorun Lin:** Resources, Formal analysis. **Yunzhu Qin:** Resources, Formal analysis. **Xinyan Huang:** Conceptualization, Methodology, Supervision, Writing – review & editing, Funding acquisition.

Declaration of Competing Interest

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

Data availability

Data will be made available on request.

Acknowledgements

This work is funded by the National Natural Science Foundation of China (No. 52322610, 51876183) and General Research Fund of Hong Kong Research Grant Council (No. 15221523).

Appendix

References

Adeniyi, P., Segun, A., Faisal, I.P., Ching, A., Lee, S., Ohunakin, O.S., et al., 2020. Thermal decomposition of rice husk: a comprehensive artificial intelligence predictive model. J. Therm. Anal. Calorim. 140, 1811–1823. https://doi.org/10.1007/s10973-019-08915-0.

Alabdrabalnabi, A., Gautam, R., Mani, S.S., 2022. Machine learning to predict biochar and bio-oil yields from co-pyrolysis of biomass and plastics. Fuel 328, 125303. https://doi.org/10.1016/j.fuel.2022.125303.

Alabdrabalnabi A. Utilization of Machine Learning to Predict Bio-Oil and Biochar Yields from CoPyrolysis of Biomass with Waste Polymers 2021.

Anca-Couce, A., Berger, A., Zobel, N., 2014. How to determine consistent biomass pyrolysis kinetics in a parallel reaction scheme. Fuel 123, 230–240. https://doi.org/10.1016/j.fuel.2014.01.014.

Basu P. Biomass Gasification and Pyrolysis: practical design and theory. 2010. https://doi.org/https://doi.org/10.1016/B978-0-12-374988-8.00001-5.

Belcher, C.M., 2013. Fire Phenomena and the Earth System. Wiley.

Beste A, Barnes BC. Prediction of thermal decomposition temperatures using statistical methods Prediction of Thermal Decomposition Temperatures Using Statistical Methods. AIP Conference Proceedings 2020;050004.

Cai, H., Zou, H., Liu, J., Xie, W., Kuo, J., Buyukada, M., et al., 2018. Thermal degradations and processes of waste tea and tea leaves via TG-FTIR: Combustion performances, kinetics, thermodynamics, products and optimization. Bioresour. Technol. 268, 715–725. https://doi.org/10.1016/j.biortech.2018.08.068.

Cancellieri, D., Leroy-Cancellieri, V., Leoni, E., Simeoni, A., Kuzin, A.Y., Filkov, A.I., Rein, G., 2012. Kinetic investigation on the smouldering combustion of boreal peat. Fuel 93, 479–485.

Carrier, M., Loppinet-Serani, A., Denux, D., Lasnier, J.-M., Ham-Pichavant, F., Cansell, F., Aymonier, C., 2011. Thermogravimetric analysis as a new method to determine the lignocellulosic composition of biomass. Biomass Bioenergy 35 (1), 298–307.

Chen, Z., Hu, M., Zhu, X., Guo, D., Liu, S., Hu, Z., Xiao, B.o., Wang, J., Laghari, M., 2015. Characteristics and kinetic study on pyrolysis of five lignocellulosic biomass via thermogravimetric analysis. Bioresour. Technol. 192, 441–450.

Chen, Y., Lin, S., Liang, Z., Surawski, N.C., Huang, X., 2022. Smouldering organic waste removal technology with smoke emissions cleaned by self-sustained flame. J. Clean. Prod. 362. https://doi.org/10.1016/j.jclepro.2022.132363, 132363.

Chen, Y., Liang, Z., Lin, S., Huang, X., 2023. Limits of sustaining a flame above smoldering woody biomass. Combust. Sci. Technol. 195, 2801–2819. https://doi.org/10.1080/00102202.2022.2041000.

Chen, H., Zhao, W., Liu, N., 2011. Thermal Analysis and Decomposition Kinetics of Chinese Forest Peat under Nitrogen and Air Atmospheres. Energy Fuel 25, 797–803. https://doi.org/10.1021/ef101155n.

Di Blasi, C., 2009. Combustion and gasification rates of lignocellulosic chars. Prog. Energy Combust. Sci. 35, 121–140. https://doi.org/10.1016/j.pecs.2008.08.001.

Ekinci, Z., Civan, M., Yurdakul, S., 2021. Effects of particle size on oxidative thermal decomposition kinetics and mechanisms of selected waste wood samples. Chem. Eng. Commun. 208, 1775–1788. https://doi.org/10.1080/00986445.2020.1817743.

Géron, A., 2022. Hands-on machine learning with Scikit-Learn, Keras, and TensorFlow. O'Reilly Media, Inc.

Hameed, Z., Naqvi, S.R., Naqvi, M., Ali, I., Taqvi, S.A.A., Gao, N., Hussain, S.A., Hussain, S., 2020. A comprehensive review on thermal coconversion of biomass, sludge, coal, and their blends using thermogravimetric analysis. J. Chem. 2020, 1–23.

Huang, Z., Manzo, M., Xia, C., Cai, L., Zhang, Y., Liu, Z., Nadda, A.K., Le, Q.V., Sonne, C., Lam, S.S., 2022. Effects of waste-based pyrolysis as heating source: Meta-analyze of char yield and machine learning analysis. Fuel 318. https://doi.org/10.1016/j.fuel.2022.123578.

Huang, X., Rein, G., 2016. Thermochemical conversion of biomass in smouldering combustion across scales: The roles of heterogeneous kinetics, oxygen and transport phenomena. Bioresour. Technol. 207, 409–421. https://doi.org/10.1016/j.biortech.2016.01.027.

Kader, G., Franklin, C., 2008. The evolution of Pearson's correlation coefficient. The Mathematics Teacher 102 (4), 292–299.

Kartal, F., Özveren, U., 2022. Investigation of the chemical exergy of torrefied biomass from raw biomass by means of machine learning. Biomass Bioenergy 159. https://doi.org/10.1016/j.biombioe.2022.106383.

khan, M., Raza Naqvi, S., Ullah, Z., Ali Ammar Taqvi, S., Nouman Aslam Khan, M., Farooq, W., Taqi Mehran, M., Juchelková, D., Štěpanec, L., 2023. Applications of machine learning in thermochemical conversion of biomass-A review. Fuel 332, 126055.

Li J, Cheng J, Shi J, Huang F. Brief introduction of back propagation (BP) neural network algorithm and its improvement. Advances in Computer Science and Information Engineering: Volume 2, Springer; 2012, p. 553–8.

Li, K., Huang, X., Fleischmann, C., Rein, G., Ji, J., 2014. Pyrolysis of medium density fibreboard: optimized search for kinetic scheme and parameters via genetic algorithm driven by Kissinger's method. Energy Fuel 28 (9), 6130–6139. https://doi.org/10.1021/ef501380c.

Li, J., Li, L., Tong, Y.W., Wang, X., 2023. Understanding and optimizing the gasification of biomass waste with machine learning. Green Chem. Eng. 4 (1), 123–133.

Lin, S., Huang, X., 2021. Quenching of smoldering: Effect of wall cooling on extinction. Proc. Combust. Inst. 38 (3), 5015–5022.

Lin, S., Sun, P., Huang, X., 2019. Can peat soil support a flaming wildfire? Int. J. Wildland Fire 28, 601–613. https://doi.org/10.1071/WF19018.

Madadian, E., Haelssig, J.B., Pegg, M., 2020. A Comparison of Thermal Processing Strategies for Landfill Reclamation: Methods, Products, and a Promising Path Forward. Resour. Conserv. Recycl. 160, 104876 https://doi.org/10.1016/j.resconrec.2020.104876.

Magdziarz, A., Wilk, M., 2013. Thermal characteristics of the combustion process of biomass and sewage sludge. J. Therm. Anal. Calorim. 114 (2), 519–529.

Manatura, K., Chalermsinsuwan, B., Kaewtrakulchai, N., Kwon, E.E., Chen, W.H., 2023. Machine learning and statistical analysis for biomass torrefaction: A review. Bioresour. Technol. 369, 128504 https://doi.org/10.1016/j.biortech.2022.128504.

Mil, P., Lachman, J., Lisý, M., Bal, M., Lis, H., 2022. Spent coffee grounds and wood co-firing: fuel preparation, properties, thermal decomposition, and emissions. Renew. Energy 193. https://doi.org/10.1016/j.renene.2022.05.003.

Mlonka-Mędrala, A., Magdziarz, A., Dziok, T., Sieradzka, M., Nowak, W., 2019. Laboratory studies on the in fl uence of biomass particle size on pyrolysis and combustion using TG GC / MS. Fuel 252, 635–645.

Mureddu, M., Dessì, F., Orsini, A., Ferrara, F., Pettinau, A., 2018. Air- and oxygen-blown characterization of coal and biomass by thermogravimetric analysis. Fuel 212, 626–637. https://doi.org/10.1016/j.fuel.2017.10.005.

Ni, Z., Bi, H., Jiang, C., Wang, C., Tian, J., Zhou, W., et al., 2021. Investigation of the co-pyrolysis of coal slime and coffee industry residue based on machine learning methods and TG-FTIR : Synergistic effect, kinetics and thermodynamic. Fuel 305, 121527. https://doi.org/10.1016/j.fuel.2021.121527.

Ni, Z., Bi, H., Jiang, C., Sun, H., Zhou, W., Tian, J., Lin, Q., 2022. Investigation of co-combustion of sewage sludge and coffee industry residue by TG-FTIR and machine learning methods. Fuel 309. https://doi.org/10.1016/j.fuel.2021.122082.

Niu, H., Liu, N., 2015. Thermal decomposition of pine branch: Unified kinetic model on pyrolytic reactions in pyrolysis and combustion. Fuel 160, 339–345. https://doi.org/10.1016/j.fuel.2015.07.108.

Olatunji, O.O., Akinlabi, S.A., Mashinini, M.P., Fatoba, S.O., Ajayi, O.O., 2018. Thermogravimetric characterization of biomass properties: A review. IOP Conf. Ser.: Mater. Sci. Eng. 423 https://doi.org/10.1088/1757-899X/423/1/012175.

Saldarriaga, J.F., Aguado, R., Pablos, A., Amutio, M., Olazar, M., Bilbao, J., 2015. Fast characterization of biomass fuels by thermogravimetric analysis (TGA). Fuel 140, 744–751.

Seo S. A review and comparison of methods for detecting outliers in univariate data sets 2006.

Shen, D.K., Gu, S., Luo, K.H., Bridgwater, A.V., Fang, M.X., 2009. Kinetic study on thermal decomposition of woods in oxidative environment. Fuel 88, 1024–1030. https://doi.org/10.1016/j.fuel.2008.10.034.

Su, Y., Luo, Y., Wu, W., Zhang, Y., Zhao, S., 2012. Characteristics of pine wood oxidative pyrolysis: degradation behavior, carbon oxide production and heat properties. J. Anal. Appl. Pyrol. 98, 137–143. https://doi.org/10.1016/j.jaap.2012.07.005.

Teh, J.S., Teoh, Y.H., How, H.G., Sher, F., 2021. Thermal analysis technologies for biomass feedstocks: a state-of-the-art review. Processes 9. https://doi.org/10.3390/pr9091610.

Thiese, M.S., Ronna, B., Ott, U., 2016. P value interpretations and considerations. J. Thorac. Dis. 8 (9), E928–E931.

Vassilev, S.V., Baxter, D., Andersen, L.K., Vassileva, C.G., 2010. An overview of the chemical composition of biomass. Fuel 89, 913–933. https://doi.org/10.1016/j.fuel.2009.10.022.

Wang, S., Dai, G., Yang, H., Luo, Z., 2017. Lignocellulosic biomass pyrolysis mechanism: A state-of-the-art review. Prog. Energy Combust. Sci. 62, 33–86. https://doi.org/10.1016/j.pecs.2017.05.004.

Wang, T., Rong, H., Chen, S.i., Zhou, Y.i., Li, J., Xiao, Y., Xue, Y., 2021. TG- MS study on in-situ sulfur retention during the co-combustion of reclaimed asphalt binder and wood sawdust. J. Hazard. Mater. 403 https://doi.org/10.1016/j.jhazmat.2020.123911.

Xiao, R., Yang, W., Cong, X., Dong, K., Xu, J., Wang, D., Yang, X., 2020. Thermogravimetric analysis and reaction kinetics of lignocellulosic biomass pyrolysis. Energy 201.

Yang, J., Chen, H., Zhao, W., Zhou, J., 2016. Combustion kinetics and emission characteristics of peat by using TG-FTIR technique. J. Therm. Anal. Calorim. 124, 519–528. https://doi.org/10.1007/s10973-015-5168-x.

Yang, H., Yan, R., Chen, H., Lee, D.H., Zheng, C., 2007. Characteristics of hemicellulose, cellulose and lignin pyrolysis. Fuel 86, 1781–1788. https://doi.org/10.1016/j.fuel.2006.12.013.

Zhang, T., Cao, D., Feng, X., Zhu, J., Lu, X., Mu, L., Qian, H., 2022. Machine learning prediction of bio-oil characteristics quantitatively relating to biomass compositions and pyrolysis conditions. Fuel 312. https://doi.org/10.1016/j.fuel.2021.122812.

Zhao, W., Chen, H., Liu, N., Zhou, J., 2014. Thermogravimetric analysis of peat decomposition under different oxygen concentrations. J. Therm. Anal. Calorim. 117, 489–497. https://doi.org/10.1007/s10973-014-3696-4.

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