

## Full Length Article

## Predicting co-pyrolysis of coal and biomass using machine learning approaches

Hao Wei<sup>a</sup>, Kun Luo<sup>a,\*</sup>, Jiangkuan Xing<sup>a,b,\*</sup>, Jianren Fan<sup>a</sup><sup>a</sup> State Key Laboratory of Clean Energy Utilization, Zhejiang University, Hangzhou 310027, China<sup>b</sup> Department of Mechanical Engineering and Science, Kyoto University, Kyoto 615-8540, Japan

## ARTICLE INFO

## Keywords:

Biomass  
Coal  
Co-pyrolysis  
Random forest (RF)  
Extremely tree (ET)

## ABSTRACT

Coal and biomass co-thermochemical conversion has caught significant attentions, in which the co-pyrolysis is always the primary process. The traditional pyrolysis kinetic models are developed individually for coal and biomass, in which the synergistic effect wasn't comprehensively considered. In the present study, we innovatively explored a new method to accurately model this process using machine learning approaches, specifically the random forest algorithm based on *classification and regression trees* and *extremely trees*. First, a co-pyrolysis database is constructed from experimental data in published literatures, then divided into several sub-sets for training, application, and optimization, respectively. The machine learning models are trained on the training data-set, tested on the test data-set, and applied on the new data-set. The training and test results demonstrate both models are able to well predict the co-pyrolysis ( $R^2 > 0.999$ ), and the application results demonstrate models also perform well at outside data ( $R^2 > 0.873$ ), with model based on extremely trees performs better owing to its better accuracy, generalization and less overfitting. It also demonstrates the known of biomass pyrolysis will be better than known of coal pyrolysis. In addition, the suggestion of input feature groups is given through parametric study, and variable importance measurement are explored.

## 1. Introduction

Coal remains an important role in China's energy structure, and this situation will remain in the foreseeable future[1]. While excessive exploitation and consumption of coal will not only gradually exhaust its storage but also cause severe environmental problems[2,3]. In contrast, bioenergy can be massively produced[4] and the use of it is recognized as carbon neutral[5,6]. Co-thermochemical conversion, including co-pyrolysis, co-gasification and co-combustion, of biomass and coal could solve these problems and utilize coal in a cleaner method[7]. Among those co-thermochemical conversion processes, co-pyrolysis is one of the fundamental and initial steps of co-gasification and co-combustion[8,9], which has significant effects on the performance of the further process[10]. Besides, co-pyrolysis can also be regard as an independent co-thermochemical conversion process to produce a variety of fuels and chemicals[11,12]. Therefore, it is of great significance to study the co-pyrolysis behaviour of coal/biomass and then accurately model this process, which plays a vital role in designing and operating co-thermochemical utilization system.

During the whole co-pyrolysis process of coal and biomass, the

presence of biomass may also affect the pyrolysis of coal, whereas the existence of coal may affect the pyrolysis of biomass. The synergistic effect between coal and biomass is influenced by the above two aspects [10,13,14]. The thermal events during pyrolysis of coal like lignite can be divided into four stages(dewatering, decarboxylation, depolymerization, polycondensation), while most biomass can be divided into three stages(dewatering, volatile components removed, carbonization)[15], which makes the synergistic effect during coal and biomass co-pyrolysis quite complicated. And how the synergistic effect affects co-pyrolysis is also complicated. Sonobe et al.[14] found that synergistic effect exists in co-pyrolysis of lignite and corncobs, which has a positive effect on the liquid and gas production. Chen et al.[16] observed an interaction between the solid phases of *C. vulgaris* and coal blend, which has an inhibitive effect on thermal composition. In contrast, some researchers found that there are no obvious synergistic effects in co-pyrolysis process of coal and biomass. Sadhukhan et al.[17] found an absence of synergistic effect between lignite coal and waste-wood by comparisons between the experiment and calculation results. Ferrara et al. [18] found the synergistic effect between the considered fuels (South African coal and pine wood chips) are very slight.

\* Corresponding authors at: State Key Laboratory of Clean Energy Utilization, Zhejiang University, Hangzhou 310027, China.

E-mail addresses: [zjulk@zju.edu.cn](mailto:zjulk@zju.edu.cn) (K. Luo), [zjuxjk@zju.edu.cn](mailto:zjuxjk@zju.edu.cn) (J. Xing).

Obviously, the influence of synergistic effect is still controversial, not fully understood, and should be considered in the mathematical modelling of co-pyrolysis of coal and biomass.

Generally, thermal gravimetric analysis (TGA), is the used to analyse co-thermochemical behaviours of different kinds of fuels [19]. Then kinetic parameters of the pyrolysis are extracted from the TGA with pre-assumed reaction kinetics. For the individual pyrolysis of coal and biomass, there are many available methods for kinetic parameter calculation, such as the Kissinger–Akahira–Sunose method [11,16,20], the Flynn–Wall–Ozawa method [16,20], and the Friedman method [20,21]. However, an inherent deficiency of these methods is their inapplicability to confirm three kinetic parameters simultaneously [22]. In addition, most of them are established for modelling individual pyrolysis of coal and biomass without considering the complexity of synergistic effect of co-pyrolysis as mentioned before, which has handicapped the development of a unifying model. As for the blends, currently, there have been none direct models with high accuracy to well reproduce the co-pyrolysis process.

With the development of machine learning (ML), some ML systems, such as artificial neural network, random forest based on *classification and regression trees* [23] and *extremely trees* [24], are widely accepted as the technology that can be applied to handle non-linear problems, which have proven their potential in prediction of process parameters in energy-related processes [25–30]. Among those ML systems, random forest has the following advantages: 1) its structural parameters are less than those of other machine learning approaches like ANN, providing a higher applicability [31]; 2) it employs a number of decision trees and the final result is the average of each tree's prediction result, which effectively avoids over-fitting and the effect of singular samples [32]. It has been successfully applied in bio-oil yield prediction [33], pyrolytic gas yield and compositions prediction [34] and biomass's kinetic parameters prediction [28]. And there are also some researches focus on the co-pyrolysis process of coal and biomass [22,35,36]. Abbas et al. [35] proposed an approach based on neural networks to predict the volatiles released by the particles of coal and biomass. But the model's performance was not well for the blends. Yildiz et al. [22] studied the co-combustion of hazelnut husk-lignite coal blends of various composition by the artificial neural network theory. And it is only used for predicting co-pyrolysis of the selected coal and biomass, with the model's performance on wide ranges of fuel types and operation conditions remaining unclear and needing further exploration.

Based on the above backgrounds, the objective of the present work is to develop general models for accurately predicting the co-pyrolysis of coal and biomass with machine learning approaches, specifically the random forest algorithm based on *classification and regression trees* (named RF for brevity) and *extremely trees* (named ET for brevity). Firstly, the co-pyrolysis database is established from TGA experiment results in available published literatures, and then it has been divided into three sets named training set, test set and new set for training, testing and application, respectively. Secondly, the RF and ET models are trained on the training set with three different feature groups, and the hyper-parameters optimization is done by five-folds cross-validation methods. Thirdly, the models with optimal hyper-parameters are evaluated on three different new sets to validate their performances on outside samples. Finally, the variable importance measurement is explored to assess the relative importance of each factor on the co-pyrolysis behaviour. About the synergistic effect, our models concern it by training models with large number of co-pyrolysis data from experiment, and the co-pyrolysis is predicted by the trained models instead of calculated with combining the individual pyrolysis by the blending ratio through a linear superimpose. However, we admit that this kind of synergistic effect might not been fully modelled, which could be improved by updating the database in the future works.

The rest of this paper is organized as follows. The database construction, principle of random forest (including the principle of classification and regression trees and the extra-trees) and the evaluation

**Table 1**  
List of databases' samples.

No [Refs.]	Coal type	Biomass type
1[22]	Lignite	Hazelnut husk
2[37]	High volatile coal	Municipal solid waste
3[16]	Semi-anthracite	Chlorella vulgaris
4[38]	Lignite	Energy grass
5[39]	Sub-bituminous	White pine
6[15]	Lignite	Pine sawdust
7[40]	High volatile bituminous	Cherry pits
8[41]	Bituminous	corn stalks
9[42]	Sub-bituminous	Corn cob
10[43]	Hard coal	Sugarcane bagasse
11[17]	Lignite	Wood waste
12[18]	Sub-bituminous	Wood
13[14]	Bituminous	Corn cob
14[10]	Lignite	Platanus wood
15[44]	Bituminous	Potato starch
16[45]	Shenfu coal	Walnut seed
17[46]	Anthracite	Wood chips
		Wood chips 250
		Wood chips 300
18[47]	Bituminous	Corn stalks
19[48]	Sub-bituminous	Pinus sawdust
20[49]	Hard coal	Sugarcane bagasse from South Africa
		Corn cobs from South Africa
21[50]	Bituminous	Corn cob
		Hardwood
22[51]	Anthracite	Miscanthus Sacchariflorus
	Lignite*	
	Bituminous*	
23[13]	Bituminous	Walnut sawdust
		Rice husk
		Corn stalks*

indicators are introduced in [Section 2](#). Results and discussions, including the feature selection methods, hyper-parameter optimization, training and test performances, as well as the variable importance measurement, are presented in [Section 3](#). The final section provides some conclusion remarks.

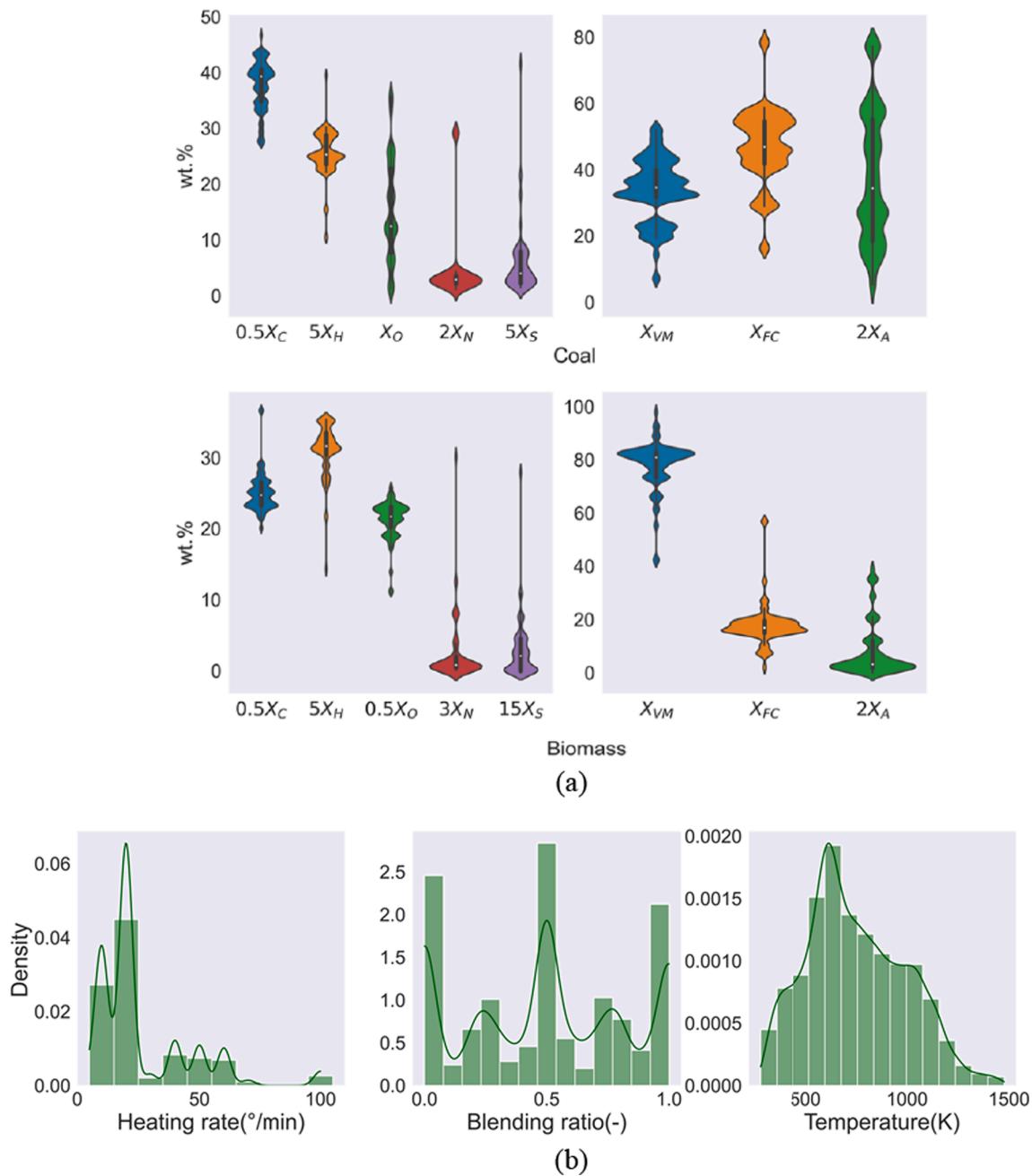
## 2. Materials and methods

### 2.1. Database construction

In the present study, co-pyrolysis database, including 30 biomass and 28 coal samples, is constructed from available literature listed in [Table 1](#), in which the types of coal and biomass are also given. The distribution of their proximate analysis (wt.%, dry basis) and ultimate analysis (wt.%, dry-ash-free basis) are shown in the [Fig. 1\(a\)](#), and it has been timed values for a clearer observation. The carbon, hydrogen, oxygen, nitrogen, sulfur, volatile matter, fixed carbon and ash fractions of coal are within the ranges of 55.38 ~ 93.46%, 2.9 ~ 7.9%, 1.21 ~ 35.72%, 0.53%~14.58%, 0.28 ~ 8.32%, 7.22 ~ 52.10%, 16.40 ~ 78.95% and 2.59 ~ 38.60%, respectively. While those of biomass are within the ranges of 40.18 ~ 73.30%, 2.87 ~ 7.08%, 22.35 ~ 51.58%, 0 ~ 10.04%, 0 ~ 1.86%, 42.53 ~ 97.94%, 1.98 ~ 56.89% and 0 ~ 19.51%, respectively. We have tried our best to collect data as much as possible, and we will follow up the appropriate literature and update the database in our future study.

where the sample with \* forms the new sets (more details are presented in the following paragraph).

The co-pyrolysis is determined not only by the samples properties but also by the experiment conditions such as the heating rate and the blending ratio. In this work, the blending ratio is calculated by Eq. (1).



**Fig. 1.** The distribution of database; (a) The distribution of coal and biomass's proximate analysis (wt.%, dry basis) and ultimate analysis (wt.%, dry-ash-free basis); (b) The distribution of heating rate( $^{\circ}\text{C}/\text{min}$ ), blending ratio and particle temperature (K).

$$\text{blendingratio} = \frac{m_{\text{coal}}}{m_{\text{coal}} + m_{\text{biomass}}} \quad (1)$$

where  $m_{\text{coal}}$  and  $m_{\text{biomass}}$  are the coal and biomass mass, respectively. The distribution of experiment heating rate and blending ratio from available literature are shown in Fig. 1(b). It can be seen that the range of heating rate, blending ratio and temperature are  $10 \sim 100(^{\circ}\text{C}/\text{min})$ ,  $0 \sim 1(-)$ ,  $293.15 \sim 1476.03(\text{K})$ , respectively.

Then the database is divided into the following three sub-sets: (1) training set; (2) test set; (3) new set. Data from No. 1 to No. 23 listed in Table 1(except the sample with \*, which consist of new set) consist of training set and test set. Training set is used for training model while the test set is used to test how the model perform on unseen data. The stratified sampling method based on blending ratio have been used to divide the training set and test set with the proportions of 70% and 30% randomly.

Although the test set is different from the training set, the data in it may be from the same curves with some data in the training set, which means it is inside of training set. In order to value the model at outside data, the rest of the data (the sample with \* of No.22 and No. 23) forms the new set. It has been divided into three different parts, specifically new set A, new set B and new set C. New set A is composed of the co-pyrolysis data of lignite and Miscanthus Sacchariflorus from No 22, while the individual pyrolysis data of No 22's coal and biomass are in training set, which aims to test how the model perform when the individual pyrolysis of coal and biomass are available (the "available" here refers to the availability of data in the training set). New set B is composed of the pyrolysis data of bituminous as well as the co-pyrolysis data of coal and Miscanthus Sacchariflorus from No 22, in which only the individual pyrolysis of biomass is available during training, and it aims to test how the model perform only the pyrolysis of biomass are

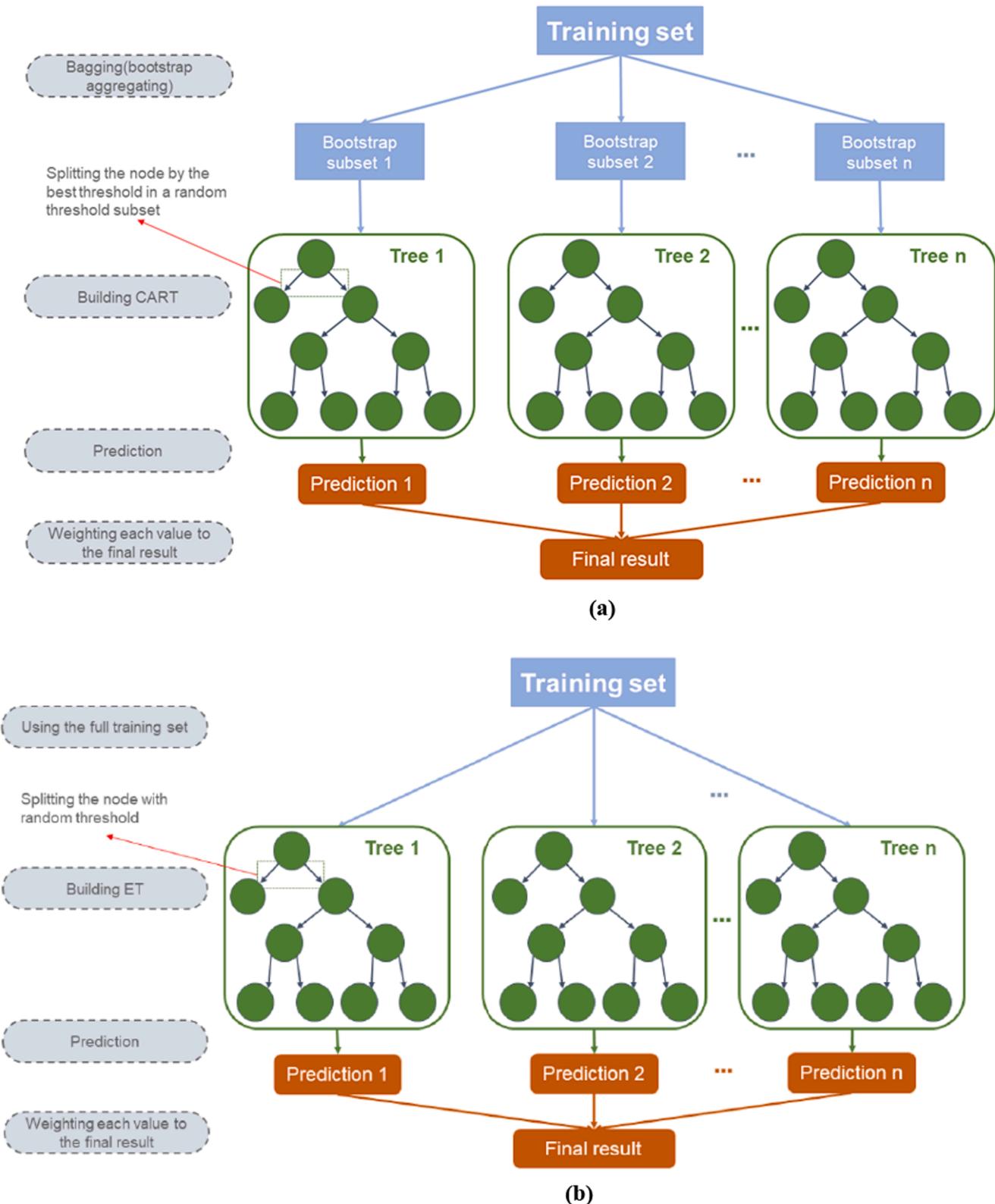


Fig. 2. The RF(a) and ET(b) algorithmic diagram.

available. New set C is composed of the pyrolysis data of cornstalks as well as the co-pyrolysis data of bituminous and cornstalks from No 23, in which only the individual pyrolysis of bituminous is available during training, and it aims to test how the model perform only the pyrolysis of coal are available. The reason for dividing the new sets in this way is that the individual pyrolysis of coal (or biomass) is known for the coal-fired

(or biomass-fired) boilers. And the good performance of co-pyrolysis model will be meaningful for the development of coal-fired (or biomass-fired) boilers into coal-biomass boilers.

## 2.2. Machine learning methods

### 2.2.1. Random forest

Random Forest is one of the most famous machine learning approach, which was originally developed by Breiman [23]. The schematic topological of Random Forest approach can be seen in Fig. 2 (a). As an ensemble methodology, it employs a number of decision trees and is generally trained via the bagging method that randomly generates the subsets from training set. Then the generated subsets are used for the learning process of Classification And Regression Tree(CART) algorithm [52] respectively (the details about learning process of CART will be discussed in next part). For regression problem, the final result is the average of each tree's prediction result.

The CART is working in the following way. Firstly, the algorithm splits the set (the bootstrap subset in Fig. 2(a)) into two subsets by the best threshold (in a random threshold subset) which minimizes the mean square error (MSE, calculated by Eqs. (2), (3) and (4)). Secondly, if it has successfully split the training set into two, it splits the subsets using the same logic, then the sub-subsets and so on, recursively. It stops recursing when it cannot find a split that will reduce MSE, or if it reaches the maximum depth (defined by the *max\_depth* hyperparameter in *Scikit-Learn*, discussed in the following part). Each leaf node predicts a value which is simply the average target value of the instances associated to this leaf node. Generally, the algorithm splits each region in a way that makes most training instances as close as possible to that predicted value.

$$J(k, t_k) = \frac{n_{\text{left}}}{n} \text{MSE}_{\text{left}} + \frac{n_{\text{right}}}{n} \text{MSE}_{\text{right}}, \quad (2)$$

$$\text{MSE}_{\text{node}} = \sum_{i \in \text{node}} \left( \hat{y}_{\text{node}} - y^{(i)} \right)^2 \quad (3)$$

$$\hat{y}_{\text{node}} = \frac{1}{n} \sum_{i \in \text{node}} y^{(i)} \quad (4)$$

where the  $J(k, t_k)$  is the cost function of CART,  $\text{MSE}_{\text{left/right}}$  measures the MSE of the left/right subset,  $n$  is the number of total instances in subset,  $n_{\text{left/right}}$  is the number of instances in the left/right subset,  $y^{(i)}$  is the value of  $i$  instance in subset and  $\hat{y}_{\text{node}}$  is average value of them, in this work, it represent the percent of weight loss.

### 2.2.2. Extra-Trees

The schematic topological of *Extremely Randomized Trees* (or *Extra-Trees*) approach [24] can be seen in Fig. 2 (b). It looks similar with RF's schematic topological, and in fact, there are only two difference between the two. One is how the tree generated, and the other is which set used for training. For the trees of RF, at each node, a random subset of the threshold is considered for splitting and the best will be selected. In order to make trees even more random, random threshold are used instead of searching for the best possible threshold in a random subset mentioned before. This kind of trees is called *Extra-Trees*, and a forest of such trees is regarded as an *Extremely Randomized Trees* ensemble (ET). The other difference between RF and ET is which set used for training. ET methods use the full original training set rather than a bootstrap replica generated from training set to grow trees.

Compared with CART, ET trades more bias for a lower variance as

well as it's much faster to train owing to finding the best possible threshold in a random subset at every node is one of the most time-consuming tasks of growing a tree. But it is difficult to tell in advance whether a RT will perform better or worse than an ET without comparing the prediction results. So, it will be discussed in the following part.

### 2.2.3. Variable importance measurement (VIM)

To quantitatively explore the effect of each factor on the co-pyrolysis behaviour, here we use the Mean Decrease Impurity importance (MDI) method[53] to measure the variable importance, which is executed in the following way. In the context of forest, the importance of a variable  $X_j$  can be evaluated by adding up the weighted impurity decreases  $p(t)\Delta i(s_t, t)$  for all trees  $\varphi_m$  (form = 1, ..., M) in the forest:

$$\text{Imp}(X_j) = \frac{1}{M} \sum_{m=1}^M \sum_{t \in \varphi_m} 1(j_t = j) [p(t)\Delta i(s_t, t)] \quad (6)$$

where  $p(t) = \frac{N_t}{N}$ , is the proportion of samples,  $j_t$  denotes the identifier of the variable used for splitting node  $t$ , and  $\Delta i(s_t, t)$  is the decrease impurity (MSE for regression).

## 2.3. Evaluation indicators

In order to quantitatively evaluate the performances of RF and ET, several evaluation indicators have been introduced, including the determination coefficients ( $R^2$ ) and the root mean square error (RMSE), as defined as below.

$$R^2 = 1 - \frac{\sum_{i=1}^N \left( y_{\text{pred}}^{(i)} - y_{\text{exp}}^{(i)} \right)^2}{\sum_{i=1}^N \left( y_{\text{exp}}^{(i)} - \hat{y}_{\text{exp}} \right)^2} \quad (7)$$

$$\text{RMSE} = \left( \frac{1}{N} \sum_{i=1}^N \left( y_{\text{exp}}^{(i)} - y_{\text{pred}}^{(i)} \right)^2 \right)^{\frac{1}{2}}, \quad (8)$$

$$\hat{y}_{\text{exp}} = \frac{1}{N} \sum_{i=1}^N y_{\text{exp}}^{(i)} \quad (9)$$

where  $y_{\text{exp}}^{(i)}$  is the experimental value,  $y_{\text{pred}}^{(i)}$  is the predicted value and  $\hat{y}_{\text{exp}}$  is the average experimental value.

## 3. Results and discussions

In this section, we will first introduce the process of training, including the feature selection methods and hyper-parameter optimization, and then the performances of RF and ET models on training, test and new sets. Finally, the variable importance measurement is conducted to explore the relative importance of each input on the co-pyrolysis behavior.

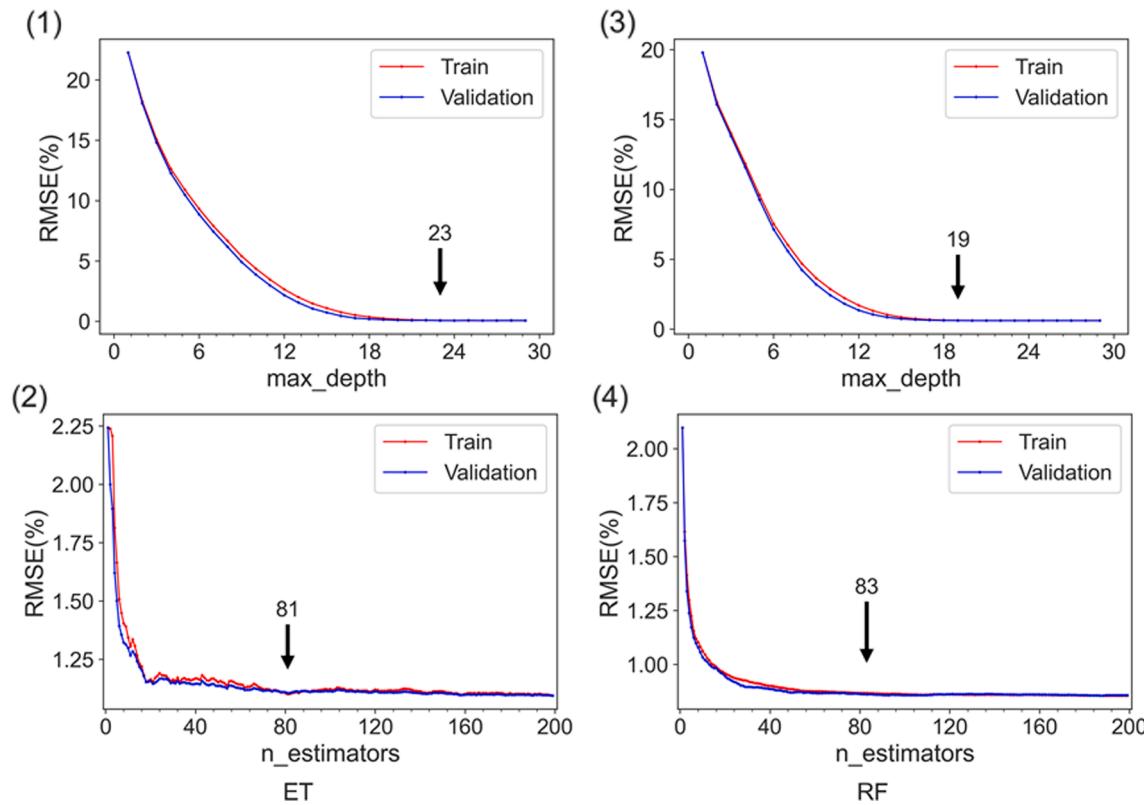
### 3.1. Training

#### 3.1.1. Feature selection

In machine learning, feature selection is a key pre-processing step.

**Table 2**  
Features selection methods.

Feature	Experiment conditions			Temperature	Sample features					Proximate analysis		
	Heating rate	Blending ratio			C	H	O	N	S	VM	FC	A
F01		✓						✓			✓	
F02		✓						✓			✗	
F03		✓						✗			✓	



**Fig. 3.** Hyper-parameters optimization results of ET and RF models with F01 by five-folds cross-validation.

The statistical features are given as inputs for a machine learning algorithm. As mentioned before, for co-pyrolysis, there are two kinds of features: the experiment conditions and the sample features. The experiment conditions include the heating rate and the blending ratio, while the sample features include the proximate analysis and ultimate analysis. Both proximate analysis and ultimate analysis are used to feature the overall composition and the elemental conservation of the solid fuel, respectively. When one of the ultimate and proximate analysis data is unavailable due to experimental devices limitations, whether we could get acceptable models to well describe the devolatilization process as both of them are known or not. This feature division method is used to answer the above concern. It can prove which feature group is best for different new sets. And there is also some similar work. Tang et al. [34] study the effect of pyrolysis conditions and fuel types on the gas yield with machine learning, in which the fuel type is characterized with both proximate and ultimate analysis data. Xing et al. [54] predicted the heat value of biomass from proximate or ultimate analysis data with machine learning approaches. The feature selection in our work are done by employed three types of feature groups as introduced below and listed in Table 2, in which F01 includes all experiment conditions and sample features, F02 includes all experiment conditions and the ultimate analysis and F03 includes all experiment conditions and the proximate analysis.

### 3.1.2. Hyper-parameters optimization

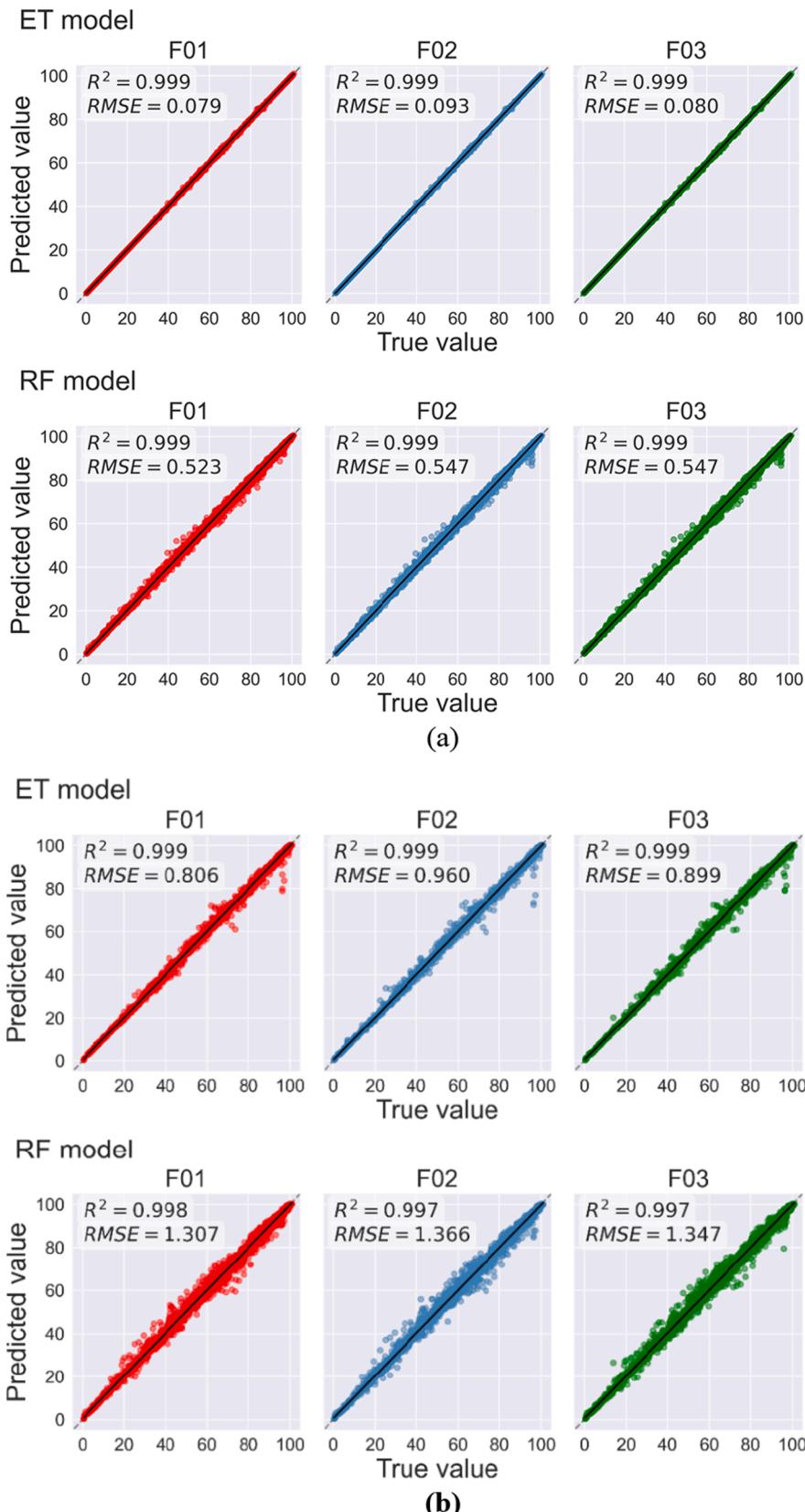
Hyper-parameters optimization has a great influence on the model's performance, robustness and generalization capability. This sub-section details the selection of optimal hyper-parameters of RF and ET. One of the most important hyper-parameters is the *n\_estimator*, which stands for the tree number, directly related to the computational cost. To decrease the possibility of over-fitting, after the *n\_estimator* selected, the *max\_depth*, which is the maximum depth of each trees, would be discussed as another hyper-parameter. A stepwise searching method is used to find optimal values of model's hyper-parameters, and the step is 1 for

*n\_estimator* changing from 1 to 200 as well as *max\_depth* changing from 1 to 30. At the same time, the K-fold cross-validation is used for a better evaluation. It divides the training data set into *k* subsets of equal size named folds. Each fold is used as a validation dataset to test the model, whereas the left *k*-1 datasets are used for training. To balance the training time and the evaluation result, in our study, the five-folds is used.

The results of hyper-parameters optimization of ET and RF models are shown in Fig. 3, in which (1), (2), (3) and (4) stand for optimization of *max\_depth*, *n\_estimators* of ET and RF models, respectively, and the blue and red lines stand for RMSE of train and validation sets, respectively. Owing to the results of hyper-parameters optimization models with F02 and F03 presenting the similar trend with F01, here only the results of models with F01 are shown for brevity. At the beginning, RMSE decreases sharply with the increasing of parameters, and then stabilizes, showing a threshold of *n\_estimator* and *max\_depth* of RF and ET models. When the parameter value is larger than the threshold, the RMSEs of both train and validation sets don't significantly decrease, which means the model's accuracy won't be improved obviously with the increase of parameters, but the computational time will. Therefore, the optimal parameters are chosen for a good compromise between accuracy and computational time, which are 81/23, 83/19, 72/22, 107/18, 123/22, 117/20 for *n\_estimator* and *max\_depth* of the ET and RF models with F01, F02 and F03 respectively.

### 3.2. Training and testing performance

After determining the optimal hyper-parameters, models have been retained and then evaluated in training and test set with inputting each selected feature selection as listed in Table 2. The performances of both models in training set are shown in Fig. 4 (a) and the results of RMSE and  $R^2$  are also shown in Fig. 4 (a), and those in test set are shown in Fig. 4 (b). In which black line stands for the regression line, the grey dotted line stands for the best fitting line ( $R^2 = 1$ , perfect models, which is not

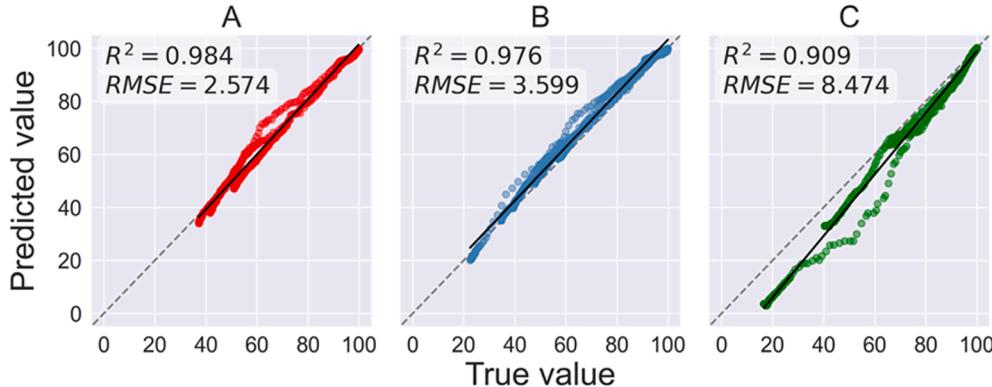
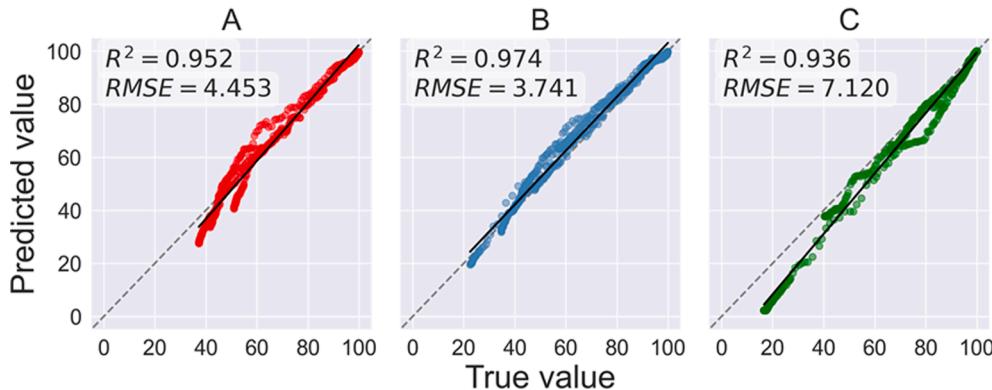


**Fig. 4.** Comparisons of training (a) and test (b) performance of RF and ET models with optimal hyper-parameters.

**Table 3**

RMSE [%] and  $R^2$  [-] results of ET and RF models with different feature groups on new set A, B and C.

	ET			RF		
	F01 $R^2$	RMSE	F02 $R^2$	RMSE	F03 $R^2$	RMSE
A	0.98	2.57	0.98	2.57	0.97	3.39
B	0.98	3.60	0.97	4.22	0.98	3.66
C	0.91	8.47	0.98	3.74	0.94	6.64
			F01 $R^2$	RMSE	F02 $R^2$	RMSE

**ET model****RF model****Fig. 5.** ET and RF models' performance on new set A, B and C with feather group F01.

easy to see in Fig. 4 owing to both two models showing a great performance on training and test set and the grey dotted line being covered by the colored scatters) and the colored scatters stand for the prediction result of models. It can be seen that both RF and ET models show a great prediction performance on the training set. The RMSEs of both models are quite small especially for ET (all less than 0.1), and  $R^2$ s are very close to 1(all equal to 0.999) for two models with all feature groups. As for the test set, the performances of both models are also quite good, which  $R^2$ s are approximately equal to 1 (all above 0.997) and RMSEs are quite small, which are less than 1 and 1.4 of ET and RF respectively. And in both sets, the results of ET model are better than those of RF model with a smaller RMSE and higher  $R^2$ . The maximum RMSE of ET is smaller than the minimum RMSE of RF while the minimum  $R^2$  of ET is larger than the maximum  $R^2$  of RF. It shows both the ET model and RF model predict well for co-pyrolysis TG curves, while ET is slightly better than RF in RMSE. More important is that the difference of both models' performance between training and test set is very slight, which prove that there is no significant overfitting of both models.

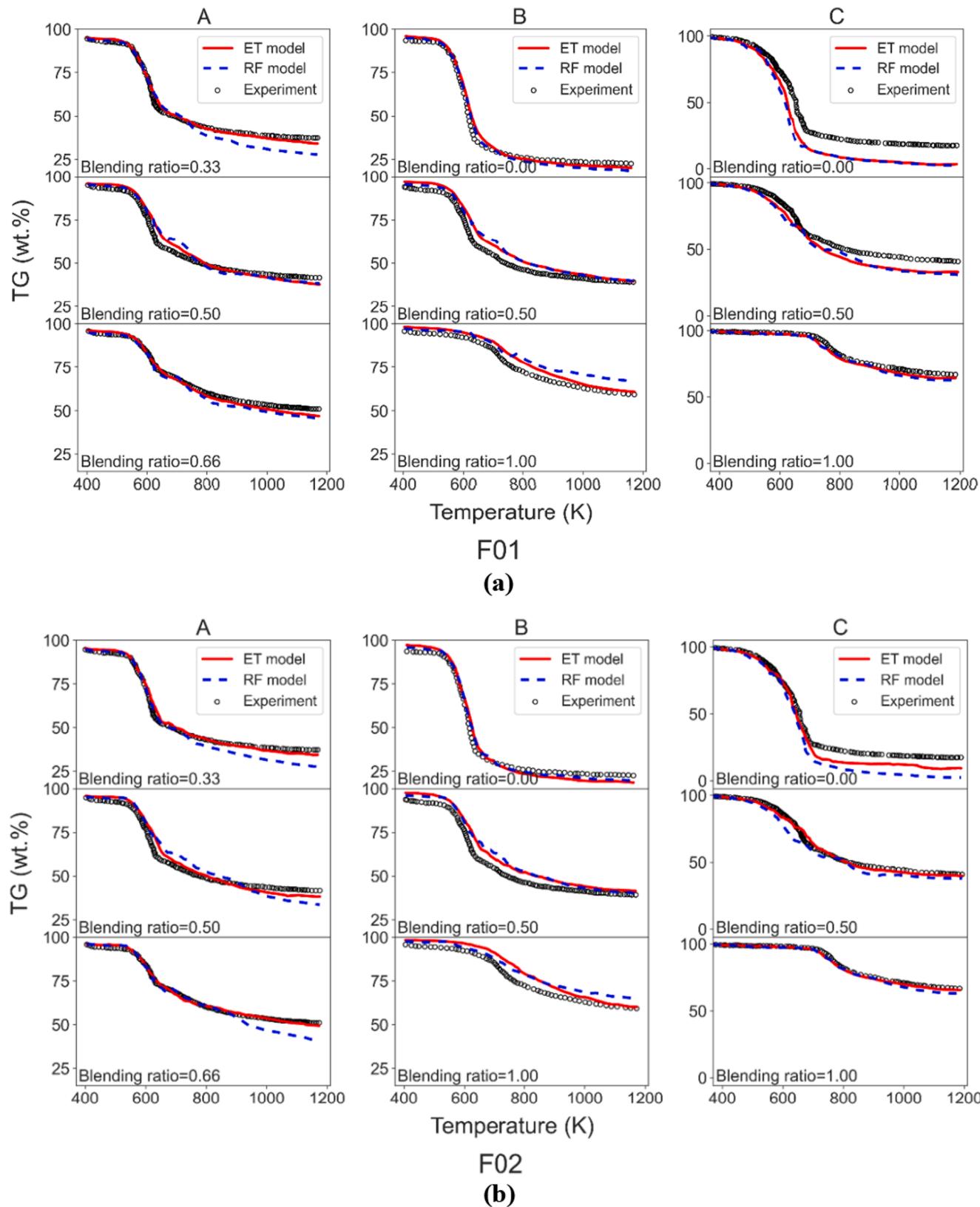
From the above discussions, it can be concluded that both the RF and ET models are capable for predicting the co-pyrolysis of coal and biomass at inside data. While how it performs on outside data remains unclear, and will be discussed in the next part.

### 3.3. Application performances

To comprehensively evaluate the performance of RF and ET models, both RF and ET models have been used to predict the co-pyrolysis processes of new sets (defined in section 2.1), which are outside of the training set.

Both ET and RF models' corresponding evaluation indicators on new set A, B and C with F01, F02 and F03 are listed in Table 3. For brevity, Fig. 5 only shows the comparison of performance of models on new set A, B and C with F01, in which the symbol is the same as Fig. 4. Owing to most data used for training, the scatters of Fig. 5 are less than that of Fig. 4, and it's enough to evaluate the models. Obviously, the models' performance on new sets are pretty good with  $R^2$  generally large than 0.9, which shows the models' generalization is also acceptable. And ET model performs better than RF model on most cases of new sets just like it behaves on training and test sets.

In order to lucubrate both models' behavior, Fig. 6 shows the RF and ET models' performance on new sets A, B and C with F01, F02 and F03, in which the black scatters are the experiment TG results from literature, and the red solid lines and the blue dashed lines stand for the devolatilization processes predicted by ET and RF models, respectively. It can be seen that ET performs better than RF in most cases like on new set A



**Fig. 6.** Comparisons of TG curves from the experiments and predicted by RF and ET models on new set A, B, and C with feather group F01 (a), F02 (b) and F03 (c)s.

with F01 at blending ratio = 0.33 and on new set B with F02 at blending ratio = 1, and this also prove ET shows a better generalization than RF. Compared with RF's prediction curves, the prediction curves of ET are

much smoother, especially on new set A with F01 at blending ratio = 0.50 and with F02 at blending ratio = 0.66, which indicates less overfitting of ET than RF. It also shows both RF and ET models' prediction

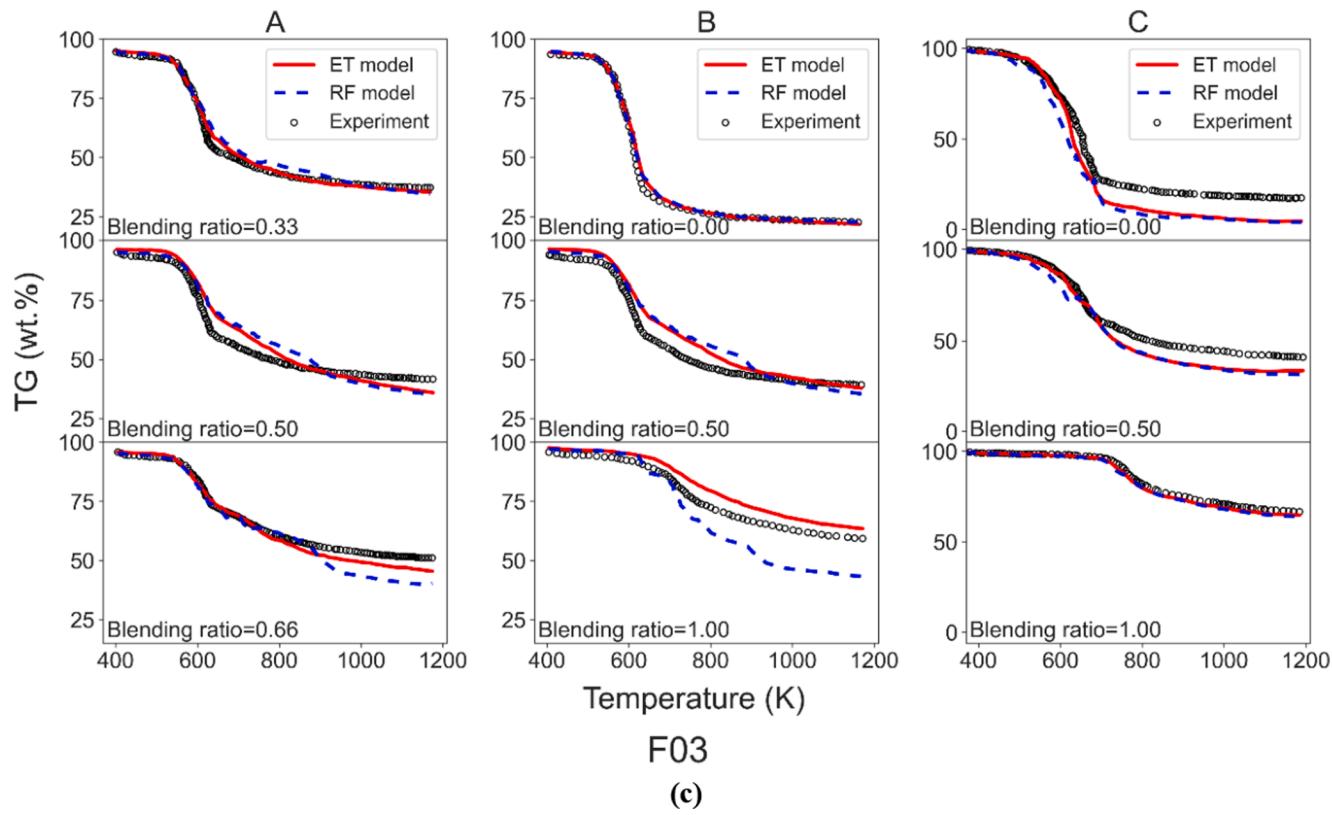


Fig. 6. (continued).

performance are best on new set A, followed by B, and then C in most cases. The best RMSE of models on new set A is 2.572, which is quite small. It proves that when the individual pyrolysis of coal and biomass are available at training (once again, the “available” here refers to the availability of data in the training set), the model can predict the co-pyrolysis at a very high accuracy. For new set B (only the individual pyrolysis of biomass is available during training), the model can also predict the co-pyrolysis at a high accuracy, which the best RMSE is 3.599, quite close to new set A. However, for new set C, the prediction performance isn't as good as the former two. The best RMSE is 3.744 for ET model with feature group F02, but under other conditions, the RMSEs are much larger which are changing from 6 to 8. The reasons for the declining performance of new set C are as follows. New set B is consisting of bituminous and Miscanthus Sacchariflorus, and the individual pyrolysis of Miscanthus Sacchariflorus is available during training. While for new set C, it is consisting of bituminous and cornstalks, in which the individual pyrolysis of bituminous is available during training. So, the declining performance of models on new set C lies on the different effect of the absence of biomass or the absence of coal during training. As Table 1 shows, there are 30 biomass and 28 coal samples, but the types of biomass are more than coal. It is a term for all organic material that stems from plants including lots types, such as land- and water-based vegetation, as well as all organic wastes[55]. And what's more, there are 11 different kinds of bituminous (including the bituminous of new set B) but only 2 kinds of cornstalks (including the cornstalks of new set C) showing in Table 1, which means only 1 kind of cornstalks, but 10 different kinds bituminous are used for training. So, the models perform better on new set B than C owing to more simples used during training.

It's interesting to find that RMSE varies not only on the different sets but also on different input features of same set from Table 3. It can be concluded that there is no model with specific inputting features groups that is best in all new sets. And since the difference of models on different sets has been discussed before, here we discusses the effect of

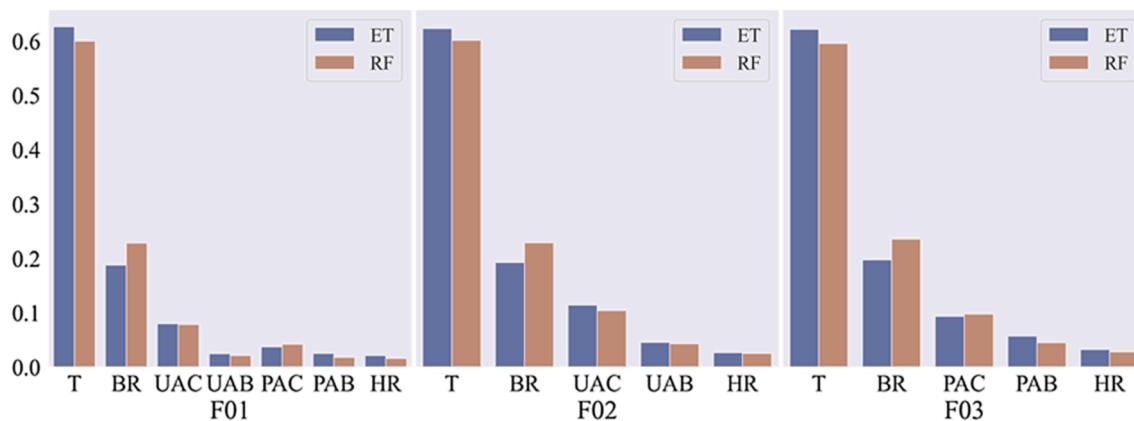
**Table 4**  
Suggestion of input feature group and model at different situation.

Known		Suggested feature group	Suggested model
Biomass pyrolysis	Coal pyrolysis		
✓	✓	F01 or F02	ET
✓	✗	F02	RF
✗	✓	F02	ET

different inputting feature groups on models' performance for the same set, which suggests the more suitable feature group for specific input data. Obviously, when the individual pyrolysis of coal and biomass are known, inputting either F01 or F02 as feature group into ET model is better. And inputting F01 into ET model is better when only the pyrolysis of biomass is known. For only the pyrolysis of coal is known, inputting F02 into ET model is much better than others. All of the conclusion discussed before are summarized and listed in Table 4.

### 3.4. Variable importance measurement

As mentioned before, not only experiment conditions but also sample features have a significant effect on co-pyrolysis, but the detail roles of them are still unclear. Owing to the advantage of RF and ET models, the VIM analysis is carried to assess the relative importance of each feature using the MDI method introduced in section 2.2.4. The results are shown in Fig. 7. The blue bars show the features importance of ET model, while the brown bars show that of RF model in three different feature groups. In the present study, the fuel type (coal and biomass) effect is considered with those ultimate or proximate analysis data. Therefore, those parameters are summed together to reflect the effect of fuel type on the devolatilization behavior here in terms of ultimate and proximate analysis aspects. Specifically, the mass fractions of carbon, hydrogen, oxygen, nitrogen and sulfur are summed to feature the fuel type effect in terms of ultimate analysis. Volatile matter, fixed carbon and ash are



**Fig. 7.** The VIM results of RF and ET models with three different input feature groups where UAC, UAB, PAC and PAB stand for ultimate and proximate analysis of coal and biomass, respectively.

summed to feature the fuel type in terms of proximate analysis. And for brevity, the ultimate analysis of coal, ultimate analysis of biomass, proximate analysis of coal and proximate analysis of biomass are named as UAC, UAB, PAC and PAB, respectively.

It shows the particle temperature (T) plays the most important role in co-pyrolysis, followed by blending ratio (BR). While the effect of proximate and ultimate analysis of coal and biomass are much smaller, commonly in the order of UAC, PAC, UAB, PAB. It proves that although the composition of coal and biomass have an influence on co-pyrolysis but it's weaker than the blending ratio. As for heating rate, its influence on co-pyrolysis is relatively small.

#### 4. Conclusion

Models based on RF and ET algorithm trained from available TG experiment results have been proposed for accurately predicting the co-pyrolysis of biomass and coal. Firstly, a co-pyrolysis database is constructed from experimental data in available published literatures, and divided into several sub-sets for model training, optimization, testing and application, respectively. Then the machine learning models are trained on the training set with different feature groups, optimized on training set with five-fold methods, tested on the test set and applied on three different new sets which are outside of training set. Both ET and RF models show good performances on predicting co-pyrolysis of coal and biomass, and ET model is better in most cases owing to its better accuracy, generalization and less overfitting. The suggestions of inputting feature groups and models for different input data are also discussed, which finds that known of biomass pyrolysis will be better for model to predict than known of coal pyrolysis. The VIM results indicate that temperature plays the most important role in co-pyrolysis, followed by blending ratio, then ultimate and proximate analysis of samples, and finally heating rate.

#### CRediT authorship contribution statement

**Hao Wei:** Conceptualization, Methodology, Software, Data curation, Writing – original draft, Writing – review & editing. **Kun Luo:** Methodology, Supervision, Project administration, Funding acquisition. **Jiangkuan Xing:** Conceptualization, Methodology, Writing – original draft, Writing – review & editing, Funding acquisition. **Jianren Fan:** Supervision, Project administration.

#### 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 are grateful for the support from National Natural and Science Foundation of China (Grant No: 51925603). JX especially thanks the financial support from Japan Society for the Promotion of Science (Grant No: 21P20351).

#### References

- [1] Normile D. China's bold climate pledge earns praise—but is it feasible? *Science* 2020;370(6512):17–8. <https://doi.org/10.1126/science.370.6512.17>.
- [2] Davis SJ, Caldeira K, Matthews HD. Future CO<sub>2</sub> Emissions and Climate Change from Existing Energy Infrastructure. *Science* 2010;329(5997):1330–3. <https://doi.org/10.1126/science.1188566>.
- [3] Scott V, Haszeldine RS, Tett SFB, Oschlies A. Fossil fuels in a trillion tonne world. *Nat Clim Change* 2015;5(5):419–23. <https://doi.org/10.1038/nclimate2578>.
- [4] Leturcq P. GHG displacement factors of harvested wood products: the myth of substitution. *Sci Rep* 2020;10:20752. <https://doi.org/10.1038/s41598-020-77527-8>.
- [5] Davis SJ, Lewis NS, Shaner M, Aggarwal S, Arent D, Azevedo IL, et al. Net-zero emissions energy systems. *Science* 2018;360(6396):eaas9793. <https://doi.org/10.1126/science.aas9793>.
- [6] Ragauskas AJ, Williams CJ, Davison BH, Britovsek G, Cairney J, Eckert CA, et al. The Path Forward for Biofuels and Biomaterials. *Science* 2006;311:484–9. <https://doi.org/10.1126/science.1114736>.
- [7] Masnadi MS, Habibi R, Kopyscinski J, Hill JM, Bi X, Lim CJ, et al. Fuel characterization and co-pyrolysis kinetics of biomass and fossil fuels. *Fuel* 2014;117:1204–14. <https://doi.org/10.1016/j.fuel.2013.02.006>.
- [8] Xing J, Luo K, Wang H, Gao Z, Jin T, Fan J. Comparative Study on Different Treatments of Coal Devolatilization for Pulverized Coal Combustion Simulation. *Energy Fuels* 2020;34(3):3816–27. <https://doi.org/10.1021/acs.energyfuels.9b04434>.
- [9] Xing J, Bai Y, Zhao C, Gao Z, Wang H. Numerical Studies of Coal Devolatilization Characteristics with Gas Temperature Fluctuation. *Energy Fuels* 2018;32(8):8760–7. <https://doi.org/10.1021/acs.energyfuels.8b01361>.
- [10] Meng H, Wang S, Chen L, Wu Z, Zhao J. Thermal behavior and the evolution of char structure during co-pyrolysis of platanus wood blends with different rank coals from northern China. *Fuel* 2015;158:602–11. <https://doi.org/10.1016/j.fuel.2015.06.023>.
- [11] Wu Z, Wang S, Zhao J, Chen L, Meng H. Thermal Behavior and Char Structure Evolution of Bituminous Coal Blends with Edible Fungi Residue during Co-Pyrolysis. *Energy Fuels* 2014;28(3):1792–801. <https://doi.org/10.1021/ef500261q>.
- [12] Yaman S. Pyrolysis of biomass to produce fuels and chemical feedstocks. *Energy Convers Manag* 2004;45(5):651–71. [https://doi.org/10.1016/S0196-8904\(03\)00177-8](https://doi.org/10.1016/S0196-8904(03)00177-8).
- [13] Chen X, Liu Li, Zhang L, Zhao Y, Qiu P. Pyrolysis Characteristics and Kinetics of Coal-Biomass Blends during Co-Pyrolysis. *Energy Fuels* 2019;33(2):1267–78. <https://doi.org/10.1021/acs.energyfuels.8b03987>.
- [14] Sonobe T, Worasuwannarak N, Pipatmanomai S. Synergies in co-pyrolysis of Thai lignite and corn cob. *Fuel Process Technol* 2008;89(12):1371–8. <https://doi.org/10.1016/j.fuproc.2008.06.006>.
- [15] Song Y, Tahmasebi A, Yu J. Co-pyrolysis of pine sawdust and lignite in a thermogravimetric analyzer and a fixed-bed reactor. *Bioresour Technol* 2014;174:204–11. <https://doi.org/10.1016/j.biortech.2014.10.027>.
- [16] Chen C, Ma X, He Y. Co-pyrolysis characteristics of microalgae Chlorella vulgaris and coal through TGA. *Bioresour Technol* 2012;117:264–73. <https://doi.org/10.1016/j.biortech.2012.04.077>.

- [17] Sadhukhan AK, Gupta P, Goyal T, Saha RK. Modelling of pyrolysis of coal–biomass blends using thermogravimetric analysis. *Bioresour Technol* 2008;99(17):8022–6. <https://doi.org/10.1016/j.biortech.2008.03.047>.
- [18] Ferrara F, Orsini A, Plaisant A, Pettinai A. Pyrolysis of coal, biomass and their blends: Performance assessment by thermogravimetric analysis. *Bioresour Technol* 2014;171:433–41. <https://doi.org/10.1016/j.biortech.2014.08.104>.
- [19] Peng X, Ma X, Xu Z. Thermogravimetric analysis of co-combustion between microalgae and textile dyeing sludge. *Bioresour Technol* 2015;180:288–95. <https://doi.org/10.1016/j.biortech.2015.01.023>.
- [20] Mallick D, Poddar MK, Mahanta P, Moholkar VS. Discernment of synergism in pyrolysis of biomass blends using thermogravimetric analysis. *Bioresour Technol* 2018;261:294–305. <https://doi.org/10.1016/j.biortech.2018.04.011>.
- [21] Biagini E, Barontini F, Tognotti L. Thermal decomposition of agricultural and food residues: comparison of kinetic models. *Can J Chem Eng* 2017;95(5):913–21. <https://doi.org/10.1002/cjce.v95.10.002>.
- [22] Yıldız Z, Üzüm H, Ceylan S, Topcu Y. Application of artificial neural networks to co-combustion of hazelnut husk–lignite coal blends. *Bioresour Technol* 2016;200:42–7. <https://doi.org/10.1016/j.biortech.2015.09.114>.
- [23] Breiman L. Random Forests. *Mach Learn* 2001;45:5–32. <https://doi.org/10.1023/A:1010933404324>.
- [24] Geurts P, Ernst D, Wehenkel L. Extremely randomized trees. *Mach Learn* 2006;63(1):3–42.
- [25] Luo K, Xing J, Bai Y, Fan J. Prediction of product distributions in coal devolatilization by an artificial neural network model. *Cambust Flame* 2018;193:283–94. <https://doi.org/10.1016/j.combustflame.2018.03.016>.
- [26] Xing J, Luo K, Wang H, Jin T, Fan J. Novel Sensitivity Study for Biomass Directional Devolatilization by Random Forest Models. *Energy Fuels* 2020;34(7):8414–23. <https://doi.org/10.1021/acs.energyfuels.0c00822.s001>.
- [27] Xing J, Luo K, Pitsch H, Wang H, Bai Y, Zhao C, et al. Predicting kinetic parameters for coal devolatilization by means of Artificial Neural Networks. *Proc Combust Inst* 2019;37(3):2943–50. <https://doi.org/10.1016/j.proci.2018.05.148>.
- [28] Xing J, Wang H, Luo K, Wang S, Bai Y, Fan J. Predictive single-step kinetic model of biomass devolatilization for CFD applications: A comparison study of empirical correlations (EC), artificial neural networks (ANN) and random forest (RF). *Renew Energy* 2019;136:104–14. <https://doi.org/10.1016/j.renene.2018.12.088>.
- [29] Xing J, Luo K, Wang H, Fan J. Estimating biomass major chemical constituents from ultimate analysis using a random forest model. *Bioresour Technol* 2019;288:121541. <https://doi.org/10.1016/j.biortech.2019.121541>.
- [30] Noushabadi AS, Dashti A, Ahmadijkani F, Hu J, Mohammadi AH. Estimation of higher heating values (HHVs) of biomass fuels based on ultimate analysis using machine learning techniques and improved equation. *Renew Energy* 2021;179:550–62. <https://doi.org/10.1016/j.renene.2021.07.003>.
- [31] Nitze I, Schultheiss U, Asche H. Comparison Of Machine Learning Algorithms Random Forest, Artificial Neural Network And Support Vector Machine To Maximum Likelihood For Supervised Crop Type Classification n.d.:7.
- [32] Welbl J. Casting Random Forests as Artificial Neural Networks (and Profiting from It). In: Jiang X, Hornegger J, Koch R, editors. *Pattern Recognit.*, Cham: Springer International Publishing; 2014, p. 765–71. [https://doi.org/10.1007/978-3-319-11752-2\\_66](https://doi.org/10.1007/978-3-319-11752-2_66).
- [33] Ullah Z, khan M, Raza Naqvi S, Farooq W, Yang H, Wang S, et al. A comparative study of machine learning methods for bio-oil yield prediction – A genetic algorithm-based features selection. *Bioresour Technol* 2021;335:125292. <https://doi.org/10.1016/j.biortech.2021.125292>.
- [34] Tang Q, Chen Y, Yang H, Liu M, Xiao H, Wang S, et al. 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>.
- [35] Abbas T, Awais MM, Lockwood FC. An artificial intelligence treatment of devolatilization for pulverized coal and biomass in co-fired flames. *Cambust Flame* 2003;132(3):305–18. [https://doi.org/10.1016/S0010-2180\(02\)00482-0](https://doi.org/10.1016/S0010-2180(02)00482-0).
- [36] Sunphorka S, Chalermsinsuwan B, Piomsomboon P. Artificial neural network model for the prediction of kinetic parameters of biomass pyrolysis from its constituents. *Fuel* 2017;193:142–58. <https://doi.org/10.1016/j.fuel.2016.12.046>.
- [37] Tokmurzin D, Kus pangaliyeva B, Aimbetov B, Abylkhan B, Inglezakis V, Anthony EJ, et al. Characterization of solid char produced from pyrolysis of the organic fraction of municipal solid waste, high volatile coal and their blends. *Energy* 2020;191:116562. <https://doi.org/10.1016/j.energy.2019.116562>.
- [38] Guan Y, Ma Y, Zhang K, Chen H, Xu G, Liu W, et al. Co-pyrolysis behaviors of energy grass and lignite. *Energy Convers Manag* 2015;93:132–40. <https://doi.org/10.1016/j.enconman.2015.01.006>.
- [39] Quan C, Xu S, An Yi, Liu X. Co-pyrolysis of biomass and coal blend by TG and in a free fall reactor. *J Therm Anal Calorim* 2014;117(2):817–23. <https://doi.org/10.1007/s10973-014-3774-7>.
- [40] Yangali P, Celaya AM, Goldfarb JL. Co-pyrolysis reaction rates and activation energies of West Virginia coal and cherry pit blends. *J Anal Appl Pyrolysis* 2014;108:203–11. <https://doi.org/10.1016/j.jaap.2014.04.015>.
- [41] Zhao H, Sun T, Sun C, Song Q, Li Y, Wang X, et al. Effects of coal pretreatment on the products of co-pyrolysis of caking bituminous coal and corn stalks mixed in equal proportion. *Appl Therm Eng* 2017;125:470–9. <https://doi.org/10.1016/j.applthermaleng.2017.07.012>.
- [42] Wang J, Yan Q, Zhao J, Wang Z, Huang J, Gao S, et al. Fast co-pyrolysis of coal and biomass in a fluidized-bed reactor. *J Therm Anal Calorim* 2014;118(3):1663–73. <https://doi.org/10.1007/s10973-014-4043-5>.
- [43] Aboyade AO, Carrier M, Meyer EL, Knoetze JH, Görgens JF. Model fitting kinetic analysis and characterisation of the devolatilization of coal blends with corn and sugarcane residues. *Thermochim Acta* 2012;530:95–106. <https://doi.org/10.1016/j.tca.2011.12.007>.
- [44] Meng H, Wang S, Wu Z, Zhao J, Chen L, Li J. Thermochemical behavior and kinetic analysis during co-pyrolysis of starch biomass model compound and lignite. *Energy Procedia* 2019;158:400–5. <https://doi.org/10.1016/j,egypro.2019.01.123>.
- [45] Yang F, Zhou A, Hao W, Yang Z, Li H. Thermochemical behaviors, kinetics and gas emission analyses during co-pyrolysis of walnut shell and coal. *Thermochim Acta* 2019;673:26–33. <https://doi.org/10.1016/j.tca.2019.01.004>.
- [46] Lu K-M, Lee W-J, Chen W-H, Lin T-C. Thermogravimetric analysis and kinetics of co-pyrolysis of raw/torrefied wood and coal blends. *Appl Energy* 2013;105:57–65. <https://doi.org/10.1016/j.apenergy.2012.12.050>.
- [47] Chen X, Liu Li, Zhang L, Zhao Y, Zhang Z, Xie X, et al. Thermogravimetric analysis and kinetics of the co-pyrolysis of coal blends with corn stalks. *Thermochim Acta* 2018;659:59–65. <https://doi.org/10.1016/j.tca.2017.11.005>.
- [48] Ulloa CA, Gordon AL, García XA. Thermogravimetric study of interactions in the pyrolysis of blends of coal with radiata pine sawdust. *Fuel Process Technol* 2009;90(4):583–90. <https://doi.org/10.1016/j.fuproc.2008.12.015>.
- [49] Aboyade AO, Görgens JF, Carrier M, Meyer EL, Knoetze JH. Thermogravimetric study of the pyrolysis characteristics and kinetics of coal blends with corn and sugarcane residues. *Fuel Process Technol* 2013;106:310–20. <https://doi.org/10.1016/j.fuproc.2012.08.014>.
- [50] Liu X, Chen M, Wei Y. Combustion behavior of corncob/bituminous coal and hardwood/bituminous coal. *Renew Energy* 2015;81:355–65. <https://doi.org/10.1016/j.renene.2015.03.021>.
- [51] Tian H, Jiao H, Cai J, Wang J, Yang Y, Bridgwater AV. Co-pyrolysis of Miscanthus Sacchariflorus and coals: A systematic study on the synergies in thermal decomposition, kinetics and vapour phase products. *Fuel* 2020;262:116603. <https://doi.org/10.1016/j.fuel.2019.116603>.
- [52] Breiman L, Friedman J, Stone CJ, Olshen RA. *Classification and regression trees*. CRC Press; 1984.
- [53] Louppe G. *Understanding Random Forests: From Theory to Practice*. Université de Liège, Liège, Belgique 2014.
- [54] Xing J, Luo K, Wang H, Gao Z, Fan J. A comprehensive study on estimating higher heating value of biomass from proximate and ultimate analysis with machine learning approaches. *Energy* 2019;188:116077. <https://doi.org/10.1016/j.energy.2019.116077>.
- [55] McKendry P. Energy production from biomass (part 1): overview of biomass. *Bioresour Technol* 2002;83(1):37–46. [https://doi.org/10.1016/S0960-8524\(01\)00118-3](https://doi.org/10.1016/S0960-8524(01)00118-3).