

Prediction of Thermogravimetric Data in the Thermal Recycling of e-waste Using Machine Learning Techniques: A Data-driven Approach

Labeeb Ali, Kaushik Sivaramakrishnan,* Mohamed Shafi Kuttiyathil, Vignesh Chandrasekaran, Oday H. Ahmed, Mohammad Al-Harabsheh, and Mohammednoor Altarawneh



Cite This: *ACS Omega* 2023, 8, 43254–43270



Read Online

ACCESS |



Metrics & More

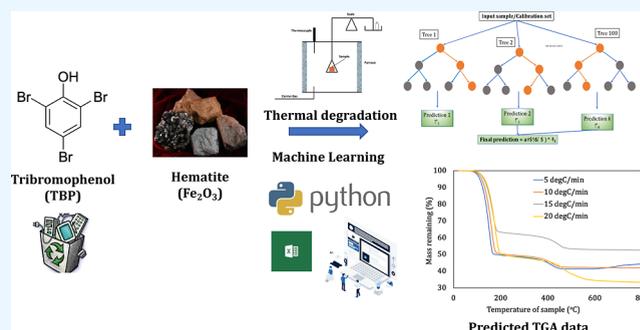


Article Recommendations



Supporting Information

ABSTRACT: The release of bromine-free hydrocarbons and gases is a major challenge faced in the thermal recycling of e-waste due to the corrosive effects of produced HBr. Metal oxides such as Fe_2O_3 (hematite) are excellent debrominating agents, and they are copyrolyzed along with tetrabromophenol (TBP), a lesser used brominated flame retardant that is a constituent of printed circuit boards in electronic equipment. The pyrolytic (N_2) and oxidative (O_2) decomposition of TBP with Fe_2O_3 has been previously investigated with thermogravimetric analysis (TGA) at four different heating rates of 5, 10, 15, and 20 °C/min, and the mass loss data between room temperature and 800 °C were reported. The objective of our paper is to study the effectiveness of machine learning (ML) techniques to reproduce these TGA data so that the use of the instrument can be eliminated to enhance the potential of online monitoring of copyrolysis in e-waste treatment. This will reduce experimental and human errors as well as improve process time significantly. TGA data are both nonlinear and multidimensional, and hence, nonlinear regression techniques such as random forest (RF) and gradient boosting regression (GBR) showed the highest prediction accuracies of 0.999 and lowest prediction errors among all the ML models employed in this work. The large data sets allowed us to explore three different scenarios of model training and validation, where the number of training samples were varied from 10,000 to 40,000 for both TBP and TBP + hematite samples under N_2 (pyrolysis) and O_2 (combustion) environments. The novelty of our study is that ML techniques have not been employed for the copyrolysis of these compounds, while the significance is the excellent potential of enhanced online monitoring of e-waste treatment and extension to other characterization techniques such as spectroscopy and chromatography. Lastly, e-waste recycling could greatly benefit from ML applications since it has the potential to reduce total and operational costs and improve overall process time and efficiency, thereby encouraging more treatment plants to adopt these techniques, resulting in reducing the increasing environmental footprint of e-waste.



1. INTRODUCTION

With the advancement of technology, the increase in the amount of electronic and electrical waste (cumulatively called e-waste) is tremendous and imminent. During the period from 2014 to 2019, the e-waste generated annually increased by ~23% to 55 t (Mt) and is further projected to increase by an alarming ~40% to 75 Mt in the next 7 years.^{1–3} The primary constituents of e-waste are printed circuit boards (PCBs), which consist of both metallic and nonmetallic fractions (NMFs). The metallic constituents such as Au, Ag, Cu, Al, Fe, Ni, and Pb are recovered using leaching techniques through strong inorganic acids, but the efficiency of this process is dependent on the fineness (size of the metallic species) of the e-wastes.^{4–7} Other limitations of this extraction are that it is slow, generates toxic effluents, consumes a lot of concentrated acids, and is expensive. Brominated flame retardants (BFRs) along with resins and ceramics constitute the nonmetallic part of PCBs, where the

BFRs release brominated benzo-dioxins, benzofurans, and most notably toxic hydrogen bromide (HBr) as byproducts during their thermal recycling process due to favorable reaction conditions.⁸ These toxic compounds contaminate the environment and are hazardous to human health.⁹ The Environmental Protection Agency (EPA) of the United States estimated that only ~15% of e-waste is recycled while the rest collects in landfills, which also eventually contaminate subsurface water through the formation of harmful compounds. Thus, in line with

Received: September 20, 2023

Revised: October 12, 2023

Accepted: October 17, 2023

Published: October 30, 2023



the Envision 2030 sustainable development goals of the United Nations (UN), novel techniques are required to be developed to capture these toxic emissions that are emitted during the thermal treatment of e-waste, especially from BFRs.

BFRs aid in delaying or preventing explosions in electronics and electrical equipment when they are subjected to stresses above their design threshold during their lifetime. Their composition makes their thermal treatment complex, leading to aromatic condensation and in the process, releasing HBr, which is thought to arise from 50% of the bromine content in BFRs.¹⁰ The toxicity of HBr is well-known, and it has potential to back-brominate (act as a Br-source for the backward reaction) aliphatic and olefinic hydrocarbons formed during thermal recycling as well as corrode equipment.¹¹ In order to capture HBr and other harmful brominated compounds formed from the pyrolysis of BFRs, a number of different metal oxides have been employed as copyrolytic reactants with BFRs in past research works such as alumina (Al_2O_3),¹⁰ calcium hydroxide ($\text{Ca}(\text{OH})_2$),¹² hematite (Fe_2O_3),^{1,13} copper oxide (CuO),¹⁴ lanthanum oxide (La_2O_3),¹⁴ antimony (Sb_2O_3),¹⁵ and franklinite (ZnFe_2O_4).¹⁶ The effect of these oxides in reducing the HBr yield in the pyrolytic products is quite evident from these works. Pyrolysis alone just distributes the bromine within the pyrolysis products, and hence, a metal oxide is required for bromine removal. Though a specific % removal of bromine from the PCBs was not reported when alumina was used in its copyrolysis in the work by Ali et al.,¹⁰ it was inferred that alumina was less efficient as a debrominating agent as compared to other metal oxides. Wu et al.¹⁵ investigated the bromine removal-effectiveness from high impact polystyrene containing BFRs by copyrolyzing with antimony trioxide (Sb_2O_3). Interestingly, they found that pyrolysis with Sb_2O_3 alone did not have an effect on reducing the Br content in the products. The most effective oxides for removal and capture of bromine from BFRs were reported to be $\text{Ca}(\text{OH})_2$ and iron oxides (Fe_2O_3), showing 100 and >80% removal, respectively.^{1,13,17} The advantage of hematite was seen in the work by Liu et al.,¹⁸ where they reported that iron particles decreased the energy barriers and enhanced the formation of Br-free gases during the pyrolysis of Fe-added WPCBs. In the work by Ma et al.,¹⁷ Fe/zeolite catalysts exhibited a combined catalytic cracking activity as well as excellent debromination performance when used with brominated acrylonitrile–butadiene–styrene (Br-ABS), which is an important resin (an amorphous polymer with fatigue-resistant, rigidity, and hardness properties) constituent of NMF. The ability to switch between different oxidation states in transition metals such as Fe makes it an excellent Lewis acid, thereby assisting in the HBr dissociation and C–Br scission. These mechanisms are prevalent in brominated aliphatics, aromatics, and olefins arising from the remaining 50% of BFR bromine content and subsequently enable the bromine capture and removal from the PCBs. The high debrominating efficiency of Fe is also established from the work by Ma and Kamo,¹⁹ where they noticed an increase of unsubstituted phenols and lower concentration of brominated hydrocarbons in the pyrolytic oil from the NMFs portion of PCBs.

Typical model compounds of BFRs include 2,4,6-tribromophenol (TBP) and tetrabromobisphenol-A (TBBA) since they contain a number of brominated phenolic species and aromatics.²⁰ A plethora of works in literature have focused on the copyrolysis of TBBA with various metal oxides,^{21–25} but very few have focused on TBP.^{1,13} We would like to highlight a recent work by us, where thermogravimetric analysis (TGA) of TBBA

combined with $\text{Ca}(\text{OH})_2$ was investigated and various machine learning (ML) techniques were applied to reproduce the TGA data.²⁶ This work was an extension of a previous work also by Ali et al.,¹² where TGA data were obtained for samples of both pure TBBA and TBBA combined with $\text{Ca}(\text{OH})_2$. Furthermore, these samples were also subjected to pyrolysis under both N_2 and O_2 environments, where the objective was to explore the debromination capability of $\text{Ca}(\text{OH})_2$ from the pyrolytic products of TBBA. TGA is an excellent fingerprinting method to study the behavior of complex samples at different temperatures ranging from 25 to 900 °C, depending on the instrument and the research objectives. The major advantage of thermal analysis is that they allow the variation of parameters such as heating rate (HR), temperature, chemical environment (carried out under inert N_2 or O_2 for combustion reaction), and pressure in order to evaluate their effects on the sample. Though the inferences from TGA provided useful insights such as melting points, phase changes, metal halide formation, stages of decomposition, and valuable information on the nature of interaction between the BFRs and the metal oxides, the production of the mass loss curves itself takes lots of time and is prone to human and instrument errors.^{16,27} ML methods provide a solution for this by directly reproducing the TGA data in much less time using the already existing experimental data. Predicting the TGA data for new input conditions such as HRs, time spent by the sample inside the chamber, heat supplied by the TGA instrument at different times, and temperatures of the sample and the chamber at different times is achieved by establishing linear and nonlinear relationships between these inputs and the output mass loss data. In our previous work,²⁶ we compared the prediction ability of random forest (RF) and support vector (SVR) regression techniques as with ordinary least-squares (OLS)-based multiple linear regression (MLR) for reproducing the TGA data of TBBA copyrolyzed with $\text{Ca}(\text{OH})_2$ under both oxygen and nitrogen environments. These techniques are elaborated in the [ML Methods Section](#) for readers' reference. The RF technique showed the best performance for both the O_2 and N_2 conditions, where all the ML models were trained and predicted using the TGA data at 4 different HRs of 5, 10, 15, and 20 °C/min. The RF regression method outperformed the other ML models when the TGA data at lower HRs were used for training while predicting on the data at higher HRs. Prediction accuracies for the RF and SVR models reached as high as 0.9999 and 0.9973 for the combined TBBA and $\text{Ca}(\text{OH})_2$ samples, whereas linear regression displayed much lower R^2 values of <0.9. In this work, we will interchange the use of prediction accuracies with R^2 values frequently and note that both expressions point to the same intended meaning of the model performance.

ML techniques, such as RF and gradient boosting regression (GBR), are ensemble-based decision trees that work on supervised algorithms. Supervised learning algorithms use labeled data sets to train the ML models with the correct outputs corresponding to the respective inputs. This way, it is easier for the model to recognize patterns in the data compared to unsupervised learning. Their main advantage in the application toward multivariate (with more than one variable feature/attribute) TGA data is that they are able to effectively capture nonlinearity in the data. RF selects multiple subdata points at random from the original data set and trains separate decision trees on each of these data sets, and the final prediction is an average of the individual predictions. The randomness in the selection of input features and observations makes the RF

algorithm robust and versatile and reduces overfitting. GBR differs from the RF in the way the trees are trained, that is, the GBR algorithm trains the trees one after the other sequentially, while in RF, all the trees are trained in parallel. GBR can solve a number of objective functions that are differentiable, which makes it very adaptable. SVR is another supervised learning technique that can be used for both classification and regression of multivariate and complex data sets. Its main advantage is the use of a kernel trick (a mathematical function) to transfer the data to a higher dimension to capture the nonlinearity in the input features.²⁸ In general, decision tree-based techniques require minimal tuning of hyperparameters as compared to SVR or artificial neural networks (ANN). Though SVR requires an extra hyperparameter to tune during model training, it addresses the limitations of ANN such as nonrepeatability of the predictions and converging to local minima.²⁹ Although we could not find any application of GBR to e-waste treatment and bromine capture, it has been applied successfully in other areas of chemical, process, and environmental engineering, and the results were compared with RF and SVR methods. GBR and RF both showed excellent prediction accuracies of >0.97 for the compressive strength of phase change materials (PCMs) integrated with cement composites in the work by Marani and Nehdi.³⁰ RF, SVR, and GBR predicted solar and wind energies from various farms in Spain with equal accuracies in the study by Torres-Barrán et al.³¹ Interestingly, they found it hard to conclude the best performing model since it depended on the farm type as well as the training and testing data sets. In a real-time study, GBR showed superior performance for predicting the yields of oil from plastic pyrolysis when compared with RF, ANN, and linear regression and applied on live plastic samples from the Rhine river.³² In the real-time dynamic temperature prediction of asphalt pavement by Wang et al.,³³ SVR showed better performance than GBR. In order to find the optimal extraction conditions of TBBA—dibromopropyl ether (DBP), which is another BFR constituent, using an organic solvent in an ultrasound instrument, Wan et al.,³⁴ employed SVR. From all these works and our previous work on e-waste treatment, we can see the utility of RF, GBR, and SVR as excellent ML approaches for predicting multivariate data in different academic and industry/real-time scenarios. Moreover, Mousa et al.¹³ utilized TBP for exploring the debromination ability of hematite by copyrolyzing them between 150 and 500 °C, but did not utilize TGA to examine the feed or products postpyrolysis. They suggested that TBP was rarely investigated as a model compound for BFR but noticed that the products of pyrolysis were devoid of brominated hydrocarbons, thus reiterating the utility of hematite as a strong debrominating agent.

In some other works, TGA has been primarily used to obtain the kinetic and thermodynamic parameters for a number of thermo-chemical processes but rarely used for predicting the mass-loss curves themselves.^{16,35} Altarawneh et al.¹⁶ also used density function theory combined with kinetic models for calculating the activation and reaction energies of the conversion of ZnO and franklinite to the corresponding bromide. Active regions of temperatures where the copyrolysis of the zinc oxide with BFRs was postulated to occur as well as the product profiles were also predicted using these models. However, ML approaches would require much less computational effort and provide excellent predictions. Furthermore, the mass loss curves for the reactant mixture were not reproduced in this work. Predicting TGA data poses significant challenges due to its multivariate nature, and except in our previous work on TBBA,²⁶

this area has not been explored in the past for e-waste recycling. Ali et al.³⁵ obtained thermo-kinetic parameters from the TGA data of TBP combined with hematite using model-fitting and iso-conversional methods (model-free determination of kinetic parameters, that is, without the assumption of a reaction pathway), which are effective but time-consuming. In our current work, we intend to reproduce these TGA data at 4 different HRs of 5, 10, 15, and 20 °C/min under N₂ and O₂ conditions using nonlinear ML approaches such as RF, GBR, and SVR. Furthermore, these will be compared with MLR for establishing a baseline prediction accuracy. The significance of our work is that it will help eliminate the use of offline measuring instruments such as TGA, which are generally time-consuming and prone to errors. Our work will lower the process time, significantly reduce human and sampling errors, and successfully be extended to spectroscopy and chromatography as well. This can further help in identifying the optimal conditions for copyrolysis of BFRs with metal oxides with applications in e-waste treatment. The novelty of our work is addressing the literature gap of the lack of predictive models for the TGA data of TBP and hematite combinations. Our versatile and robust models will be applicable across different HRs and dual chemical environments.

The aim and scope of our contribution in this work is three-fold: (i) to predict the mass loss data obtained experimentally through TGA at different HRs based on four ML approaches; (ii) to identify the best-predicting ML model for the TGA data of both TBP and TBP + hematite samples at different HRs under both N₂ and O₂ environments; and (iii) to test the sensitivity of the best-predicting regression technique to variation in the number of training points. This was achieved by training the ML algorithms at 2 or 3 sets of HRs and predicting at the remaining HRs. The ultimate objective is to develop a robust algorithm to generate mass loss curves online before and after the pyrolysis of BFRs in PCBs and minimize the use of physical, offline TGA characterization. Details of the training and validation data sets for the ML models are described in the next section.

2. ML METHODS, DATA SET SPLITTING, PERFORMANCE METRICS, AND WORKFLOW

2.1. Inputs and Outputs Used for the ML Models. For training purposes, the input variables, also known as explanatory variables/regressors/predictors, to the ML models were: (i) temperature of the TGA chamber, (ii) temperature of the sample, (iii) time spent by the sample inside the chamber, (iv) heat flow inside the chamber, and (v) the HRs (5, 10, 15, and 20 °C/min). The masses of the samples remaining in the TGA chamber at each time were used as the outputs. The outputs are also called response variables.

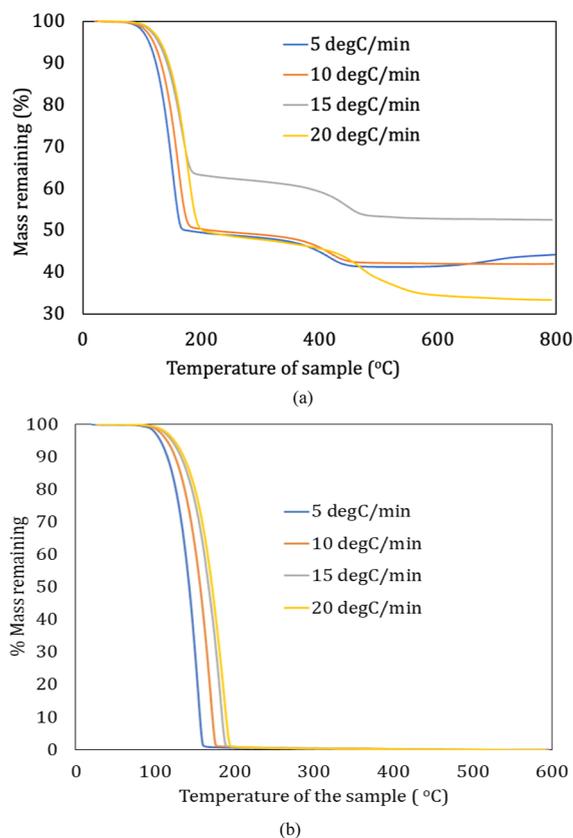
2.2. Available Data Sets and Their Description. From our previous work,¹ we obtained TGA data for TBP separately and the combination of TBP and Fe₂O₃ samples at 4 different HRs of 5, 10, 15, and 20 °C/min under inert pyrolysis (N₂) and combustion (O₂) conditions. The number of observations for samples of TBP alone and TBP + hematite combined is summarized in Table 1. It is to be noted that the number of data points were same under both pyrolysis and combustion environments, and hence the same table is applicable for both.

Figure 1a,b shows plots of the entire TGA data for a sample of TBP separately and TBP + hematite combined, respectively.

We can see that there was only a single decomposition stage in the pure TBP sample (Figure 1b) over all temperatures until 800 °C, whereas at least 2 stages could be seen for the samples of

Table 1. Number of Data Points at Each HR for the Samples Used for TGA under both N₂ and O₂ Environments

HR (°C/min)	sample	number of observations
5	TBP	13,900
	TBP + hematite	18,545
10	TBP	6852
	TBP + hematite	9350
15	TBP	4596
	TBP + hematite	6812
20	TBP	3430
	TBP + hematite	4638

**Figure 1.** Mass loss curves from TGA obtained for (a) TBP + hematite and (b) TBP alone under pyrolysis (N₂) conditions at all four HRs.

TBP mixed with hematite (Figure 1a). There were no major events until ~100 °C for either sample, but devolatilization of TBP was postulated to occur in the first stage of significant mass decrease. For more accurate determinations of mass losses, differential thermogravimetric (DTG) curves are plotted in Ali et al.³⁵ for these samples since DTG curves are derivatives of the parent TGA curves and amplify the regions of decrease in mass losses so that it is easier to identify the effects of major thermal events. The total mass loss is >99%, and Mousa et al.¹³ attributed this to the absence of ash content and other solid residue (Figure 1b). Osman et al.³⁶ found a high mass loss of >85% when polyethylene terephthalate (PET) was subjected to TGA and attributed this to a high volatile matter content. On the other hand, the adding hematite to TBP resulted in residual material that could not be vaporized, as shown in Figure 1a, and this increased with increasing HR. This was due to the formation of FeBr₂ as a result of brominated species from the pyrolysis of TBP reacting with hematite. We also noticed that the onset

temperature of thermal degradation for all HRs was slightly higher by ~30 °C for TBP as compared to hematite-added TBP, and this becomes clearer from the derivative TGA curves as reported in Ali et al.,^{1,35} from where our data were taken for ML training and prediction. Al-Harashsheh³⁷ also noticed a similar onset degradation behavior for the TGA of TBBA combined with electric arc furnace dust. Moreover, they also noticed that an increased TBBA content resulted in an increase in mass loss, and although in our work, we did not vary the ratio of TBP and hematite combination, it still explains why there was higher mass loss for TBP as compared to hematite-added TBP samples. It was also noticed that increasing the HR from 5 to 20 °C/min increased the onset temperature as well as the final temperature of complete degradation for both TBP (Figure 1b) and TBP + hematite (Figure 1a) samples. As seen from these curves, the gap between the curves at the lowest and highest HRs widened to ~10–40 °C as the thermal degradation progressed. Similar observations were also reported for TGA of PET and TBBA samples in other works as well.^{36,37}

For the combined TBP + hematite samples, the first and second regions of major mass losses can be identified as 100–180 °C (a smaller region) and 280–410 °C (a larger region) for 5 °C/min HR. Interestingly, as described previously, these peak regions shifted by 5–10 °C for every 5 °C/min increase in the HR, as seen in Figure 1a. Overall, these peaks shifted by ~30 °C to the right (higher side) as the HRs increased by 4 times, as seen in Figure 1a. The first region corresponded to the melting point of pure TBP (~95.5 °C), where the hematite is known to react with this melted form of BFRs before they evaporate.^{1,12,13,37} Breakage of C–Br bonds and cross-linking reactions facilitate this phase of the reaction.¹¹ The second region of mass-loss from the coprolysis of TBP and hematite was postulated to be a result of three occurrences: (i) removal of oxygen from oxygenated species in the reaction mixture at these higher temperatures, according to Jayashree et al.³⁸ A number of oxygenated species such as esters, dialcohols, and phenols were also previously observed by Mousa et al.¹³ in the condensed products when TBP was coprolyzed with Fe₂O₃ in the temperature range of 150–500 °C. (ii) Char oxidation since the char was formed only when hematite was used and not for TBP alone.³⁵ This char consisted of FeBr₃, carbon ash and other inorganic compounds produced due to reactions of the emitted gases from TBP pyrolysis with Fe₂O₃. Compared to pyrolysis under the N₂ environment, the char produced under combustion conditions is 10% wt lesser and also varied with different HRs according to Ali et al.³⁵ (iii) Last, a sure possibility was the release of HBr since it was reported to occur only at 400 °C for coprolysis of BFRs,³⁷ and 400 °C falls right in between this peak range. Thus, we can conclude that although HRs do not influence the reaction chemistry specifically, they do cause the onset temperatures of different thermal events to be delayed by 20–30 °C, showing as a peak shift to the right in the TGA curves.

2.3. Data Splitting and Workflow Adopted. There has been no consensus for the optimal splitting of the original data set into calibration (also called training) and validation (also called testing) sets when employing ML methods. It is well-known that the calibration set is used for constructing the regression model, and the validation set is used for testing the ML model on “new” and “unseen” input data and test the accuracy of predictions. For a large data set, Larsen and Goutte³⁹ and Dubbs⁴⁰ suggested a split of 90:10 for training/validation while Afendras and Markatou⁴¹ recommended a 50:50 split. On

the other hand, a few other works suggested to train the ML models on lower than half of the data points and validate with the remaining observations.^{42–45} An interesting work by Joseph⁴⁵ found the optimal calibration/validation ratio to be dependent on the number of parameters in the ML model as $\sqrt{\text{number of parameters}}$: 1. This was developed through a closed-form solution, but the clear limitation here was that it was applicable only to linear models.

The two types of scenarios adopted for the prediction of TGA data in our work are depicted in Figure 2. In this figure, R^2 and

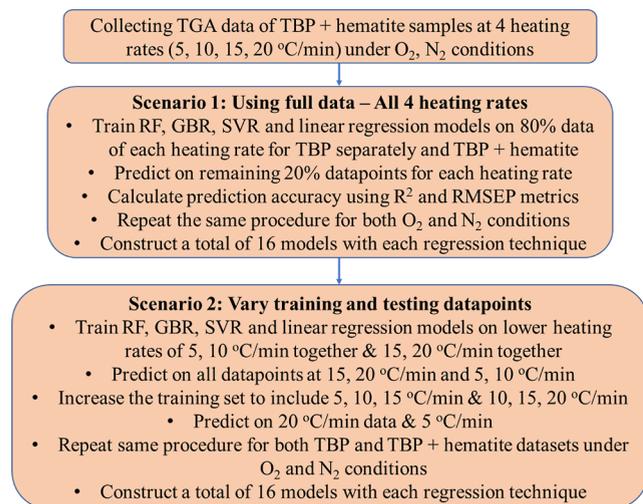


Figure 2. Schematic of the workflow adopted in this work for the use of different ML techniques to predict TGA data for TBP + hematite samples.

RMSEP are the coefficient of determination and root mean squared error (RMSE) of prediction and their calculations are described in the next section (Section 2.4). Keeping in mind the discussion in the previous paragraph, we chose to apply an 80:20 split for the calibration/validation data sets with random sampling (data points chosen at random) for our ML models. This was also the industry standard and was chosen due to the large size of our data set.²⁶ As can be seen in Figure 2, we applied an 80:20 split for calibrating and validating the models in scenario 1, where the ML models were trained by using the TGA data at each HR for both TBP and TBP + hematite samples. However, in scenario 2, two sets of investigations were conducted: (i) the entire data points in the lower HRs considering two sets at a time (5, 10 °C/min) initially and then considering three sets at a time (5, 10, and 15 °C/min) were used as calibration sets for training the ML models and the performance was tested on the data points from the remaining HRs two at a time (15, 20 °C/min) and 20 °C/min alone for the calibration set corresponding to three HRs at a time. Essentially, the validation set here was 15 and 20 °C/min combined for the former and only 20 °C/min for the latter; (ii) the entire data points in the higher HRs consider two at a time (15 and 20 °C/min) and three at a time (10, 15, and 20 °C/min) were used as the calibration set for model construction, while the validation data set consisted of the remaining HRs of 5 and 10 °C/min for 2 at a time and 5 °C/min for 3 at a time calibration, respectively. In this way, 16 separate models were built in both scenarios considering both pyrolysis (N₂) and combustion (O₂) conditions for both TBP and TBP + hematite samples.

Details about data tidying, arrangement, and cleaning are provided in the Supporting Information. Furthermore, detailed information on the software tools used for model building and testing is also provided in the Supporting Information.

2.4. Random Forest Regression and Its Hyperparameter Tuning. Random forest methods have been applied in a number of subregions of chemical engineering such as quantitative structural–property relationships for predictions of viscosity⁴⁶ and CO₂ solubility in deep eutectic solvents,⁴⁷ composition of particulate matter,⁴⁸ infrared spectroscopy,⁴⁹ gas chromatography-based classification of tobacco,⁵⁰ prediction of air–water interfacial tension in surface-active agents,⁵¹ bio-sorption for dyes,⁵² and bioreactor operations.⁵³ In these works, the performance of RF has been compared against other ML methods such as SVR, ANN, and linear regression, and it was seen that RF outperformed its counterparts comprehensively. Further details of the RF algorithm are provided in the Supporting Information. A schematic of the different trees, their nodes, branches, and the information flow to obtain the final prediction in the RF method is shown in Figure S2.

2.4.1. Advantages of RF Regression.

- The primary advantage is that there is no need for data transformation of the inputs during training of the model.
- Correlations present in input features do not affect the accuracy of predictions. This is due to the randomness in selecting the features and data points from the original data set and the final selection eliminates redundant information. This has been indicated in our previous work as well.²⁶
- Random selection of input data points with replacement and averaging out the tree predictions also decouples the errors in predictions.
- Production of stable predictions and less overfitting tendency as compared to neural networks (ANN). Overfitting is avoided by resampling the subdata sets repeatedly and averaging the predictions of all the individual trees at the end.
- Requirement of lesser hyperparameters, training time, memory, storage space, and computational power than ANN.
- Another benefit of RF algorithm is its ability to handle multidimensional/multivariate data sets very well.
- Excellent handling of nonlinear relationships between inputs and outputs from complex data sets such as TGA, spectroscopy, and chromatography makes it a suitable approach for our work.

2.4.2. Hyperparameter Tuning of the RF Model. The two main parameters to be tuned and optimized for the RF model are “*n_estimators*” and “*max_depth*”. “*n_estimators*” relates to the total number of random and independent trees that make up the model, while “*max_depth*” is the number of horizontal levels that each tree is divided into, which is quite self-explanatory. A couple of examples of the same tree with “*max_depth*” of 2 and 3 are provided in Figures S3 and S4, respectively, in the Supporting Information. These trees were part of the RF algorithm modeled on the TGA data from pyrolysis conditions of TBP + hematite sample at HRs of 5 and 10 °C/min combined for scenario 2 (Figure 2). Each box in the tree is a node, and we can infer that with each increasing horizontal level of the tree, the number of nodes also increases exponentially in the order of 2^{*n*}. For example, the number of nodes (boxes) in Figure S3 is 7, which can be calculated as 2⁰ + 2¹ + 2², while that in Figure S4 is

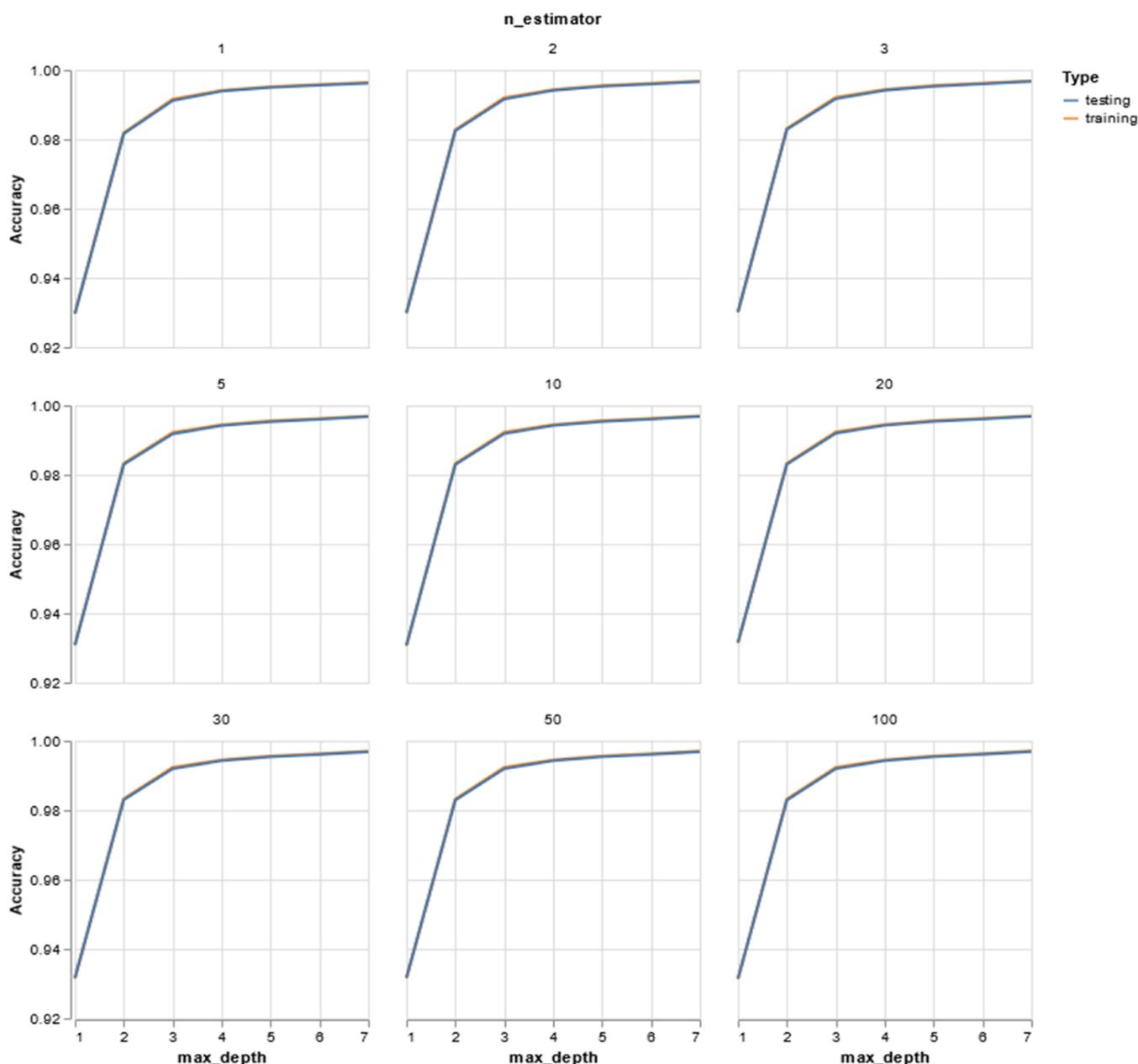


Figure 3. Prediction accuracy vs “ max_depth ” and “ $n_estimators$ ” for the pyrolysis data set of TBP + hematite for RF regression with TGA data at 5 and 10 °C/min used for training.

15, which is in turn calculated as $2^0 + 2^1 + 2^2 + 2^3$. Hence, for a “ max_depth ” of n , the number of nodes in that tree will be $2^0 + 2^1 + 2^2 + \dots + 2^{n-1} + 2^n$. Thus, it can be stated that “ max_depth ” is a measure of the depth of the RF model. On the other hand, “ $n_estimators$ ” indicates how wide the algorithm and can be visualized as the breadth of the model.

Both temperature and time were used as the inputs for the decision making in the nodes, as seen in Figures S3 and S4. The mean squared error (MSE) is the square of the RMSE which has been defined in Section 2.7 of ML Methods titled “Performance Metrics”. This value decreases as we go down the horizontal levels, while “values” represents the predicted weights of the remaining mass of the sample in the TGA at that node. The final predicted value will be the average of all the values of the nodes in the bottom-most level of the trees. The “samples” represent the number of data points that fall within this condition for HRs

considered in this training set. The variation of the prediction accuracies with the number of trees and the depth of the trees as they are varied from 1 to 100 and 1 to 7 respectively, is shown in Figure 3. It can be seen that the prediction accuracies increased sharply and then slowed down only with “ max_depth ”, while they did not depend hugely on the “ $n_estimators$ ”. The nodes are the decision points, and the next tree could predict the final mass based on time alone or temperature alone or a combination of heat flow in the chamber or any combination of these input variables. Obviously the more the number of trees, the higher the training time needed, and this is further multiplied with an increase in the “ max_depth ”. Moreover, we could clearly see that the influence of the “ $n_estimators$ ” was much lesser than that of “ max_depth ” from Figure 3, and hence values of 10 and 7 were chosen for former and latter parameters, respectively, in both the scenarios. Though the curve starts to flatten out from

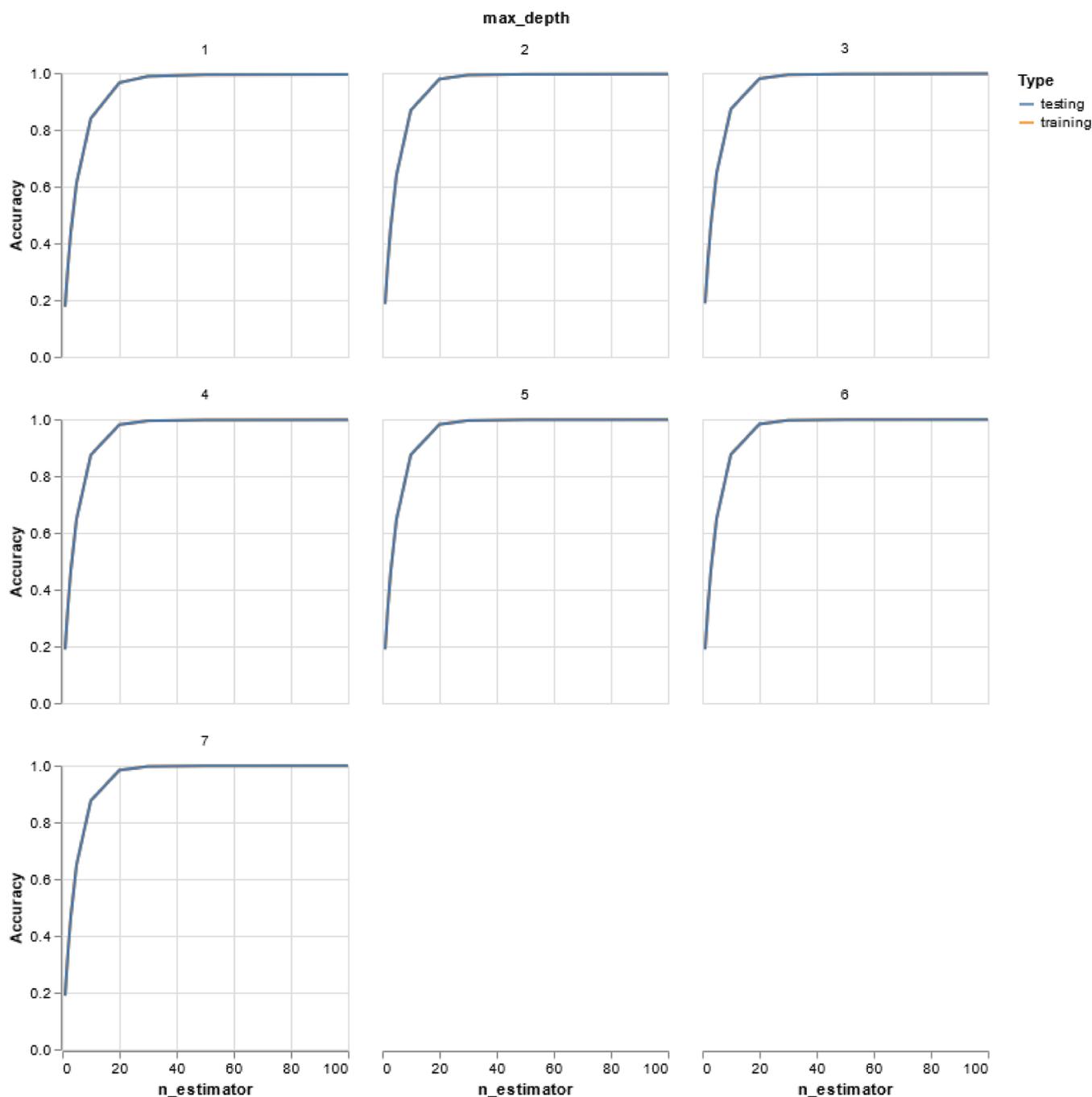


Figure 4. Prediction accuracy vs “*max_depth*” and “*n_estimators*” for the pyrolysis data set of TBP + hematite for GBR model with TGA data at 5 and 10 °C/min used as calibration set for training. This is for scenario 2, where the validation set was combined at 15 and 20 °C/min, on which the model was tested on.

“*max_depth*” values of 3, the maximum prediction accuracies were always obtained at 7, and hence, we chose this for model building.

2.5. Gradient Boosting Regression and Its Hyperparameter Tuning. **2.5.1. Overview and Applications.** GBR is another ensemble learning technique which operates on weak learners one at a time in order to improve on the previous learners in the model.⁵⁴ Each decision tree is built and trained one at a time, and these trees do not go beyond a “*max_depth*” of 3. However, in random forests, there is no restriction on this parameter since all the trees are independently built and grown fully. This reduces the bias, and the final error of prediction is

decreased by reducing the variance. In GBR, the shallower decision trees, which are also called weak learners, result in low variance but higher bias than RF. Hence, the boosting algorithm in GBR decreases the error by reducing the bias. Smaller trees are sufficient in GBR since their construction considers the already built trees before the current one. Unlike bootstrap sampling that is used in the RF approach, the trees are built on a modified version of the original data set. Overfitting is tackled in GBR by correcting the errors of the previous trees while in RF, the predictions are averaged out. In both these cases, the overfitting tendency of the individual trees is greatly avoided. However, the number of trees needs to be chosen with caution

so that an overemphasis on outliers can be avoided. Unlike in RF where random subsets are chosen from the original data set and trained with independent trees, in GBR, the main full data set is supplied to the first tree and subsequent trees work on the residuals of this tree.

As mentioned in the introduction, GBR has been used in different sectors of environmental, chemical engineering and biomass disciplines such as compressive strength prediction of PCMs,³⁰ forecasting energy consumptions for residential buildings,⁵⁵ prediction of renewable energy values in wind farms,³¹ oil yield prediction from plastic pyrolysis,³² bio-oil yields from hydrothermal liquefaction of algae,⁵⁶ and even mechanical applications such as estimating the onset of cracks due to fatigue in the design of pavements.⁵⁷ It is to be noted that certain variants of GBR technique such as CatBoost, XGBoost, and LightGBM have been used recently by other researchers but the number of parameters to be optimized in their algorithms are far more than GBR. The limitation of GBR is that it takes more time to train the model, and adding more hyperparameters will only increase the training time further and hence was avoided in this work. Another well-known disadvantage of GBR is that it does not work well with noisy data. However, the TGA data in our work were devoid of any noise, and hence GBR was well-suited. Further details about the objective function and equations involved in formulating the GBR model are given in the Supporting Information document. Figure S5 also shows the tree-framework of GBR depicting the flow of information from the input data set to the predictions.

2.5.2. Advantages.

- Ability to capture nonlinearity in the data sets efficiently.
- Local minima are avoided by the gradient descent algorithms adopted by GBR. Gradient descent is a powerful algorithm that helps the cost function (or objective function) of the GBR to reach its minimum point through a series of steps so that the predicted value is as close as possible to the actual experimental value.
- As in the case of RF, no data transformation is required.
- Overfitting is greatly reduced by progressive error correction in the trees.
- Extremely adaptable and versatile algorithm as GBR can optimize on a variety of loss functions in addition to that given in eq S4.

2.5.3. Hyperparameter Tuning. The tuning parameters for GBR are the same as those for the RF algorithm. Figure 4 shows the variation of prediction accuracy with “*n_estimators*” and “*max_depth*” (for TBP + hematite in scenario 2 using the TGA data at HRs of 5 and 10 °C/min combined as the calibration set), the only difference being that GBR performance depended more on “*n_estimators*” than on “*max_depth*”. This made sense because the GBR algorithm trains one tree at a time rather than all the trees parallelly, and adding more trees would affect the performance to higher degrees. On the other hand, “*max_depth*” affects the performance of RF model to a higher degree since the effect of “*n_estimators*” is nullified by training all the trees simultaneously. It should be kept in mind that the TGA data at HRs of 15 and 20 °C/min were used as the validation set for model prediction using the tuned hyperparameters. However, after ~20 trees, the prediction accuracy flattens out close to 1 and this tells us that too many trees above the optimal value are also a problem, and we should be careful in choosing the number of trees. Moreover, the training time also increases when a higher number of trees are chosen for GBR due to the sequential order

in which they are trained. Hence, we choose 20 for “*n_estimators*” and 3 for “*max_depth*”. There is a third parameter called “learning rate” that is defined in eq S6 in the Supporting Information, and this parameter is randomly initialized. Since the aim of the ML method is to find the global minima and avoid local minima, a lower value of learning rate (0.01) is chosen so that irregular terrain of the cost function path is tackled. The learning rate is essentially used in the gradient descent algorithm and represents the step size for finding the minima in the cost function. Furthermore, we employed stochastic gradient descent algorithm, where the gradient is calculated using each data point separately as opposed to the whole data set for batch gradient descent algorithm.⁵⁸ Choosing one observation at a time at random avoids selecting redundant samples and helps reach the global minima at a faster rate. Batch gradient descent is much slower since the whole data set is used to compute the gradient, which might result in data overload as well.

2.6. Support Vector Regression. **2.6.1. Applications of SVR.** SVM is another ML method that is based on supervised learning and was first introduced as a classification technique by Vapnik.⁵⁹ Over the past 2 decades, SVM has gained immense popularity in different areas of chemical engineering such as catalysis of alkylation reactions⁶⁰ and olefin conversions,⁶¹ predictions of natural gas hydrate temperatures,⁶² concentrations of chemical oxygen demand in aerobic granular reactors for wastewater treatment,⁶³ and pressure drops in slurry flows,⁶⁴ as well as drug discovery through compound identification.⁶⁵ The potential of SVR to predict complex, nonlinear, and multidimensional data obtained from characterization instruments such as TGA²⁶ and infrared spectrometers⁶¹ have also been explored by our coauthors with applications in e-waste treatment and acid-catalyzed olefin oligomerization.

2.6.2. Kernel Trick. Though the initial algorithm of SVR was catered toward data classification, regression and parameter estimation functions can also be performed by tuning the objective function (eq S8) and the constraints (eq S9) accordingly. If the data in the calibration set is nonlinear in nature, a kernel trick transforms the data to a higher dimension, where it becomes linear. Following this, an optimal hyperplane performs the regression or classification function according to the objective function definition.^{28,66,67} Further details about the discriminant function, minimization of the weight term, the added constraints, and the regularization parameter are given in the Supporting Information. In eq S7, $\phi(x_i)$ refers to the kernel function that can be either linear, polynomial, sigmoid, or radial basis function (RBF). Some previous works have indicated that RBF and polynomial kernels are able to capture the nonlinearity in the data much better than linear or sigmoid.^{68,69} Out of RBF and polynomial kernels, RBF has been shown to give the best results in the works of different researchers such as prediction of CO concentrations in urban areas⁷⁰ and monitoring of product concentrations in propylene oligomerization.⁶¹ RBF has also been shown to perform well over a range of varied data sets depicting the physical and chemical properties of gasoline, diesel, ethanol, macromolecules in crude oil, and resins using vibrational spectroscopy.²⁸ Furthermore, there is no limit to the range of a polynomial kernel, whereas the RBF kernel falls between 0 and 1. Moreover, the tuning of a polynomial kernel requires two parameters, whereas RBF kernel requires only a single parameter to be tuned, called the width, σ (eq S10 in Supporting Information document). All of these factors prompted us to choose the RBF kernel for the SVR model in our work to compare with GBR and RF regression. In this way,

the best possible outcomes from each ML model were compared to arrive at the best-performing method.

Since the number of variables/columns in the TGA data set is less, the training time is also less when compared to a spectral data set from infrared spectroscopy. Dimension reduction techniques such as principal component analysis (PCA) would be needed to reduce redundant information from a FTIR data set.⁶¹ The RBF kernel given by eq S10 provides information on the similarity or dissimilarity of the samples in the original data set. The value of the kernel is closer to 1 if the samples are similar and is closer to 0 and vice versa. The kernel width, after tuning, determines whether the relationship between the inputs and outputs is linear or nonlinear.⁶¹ Nonlinear data will lower the value of σ and vice versa. The tuning process is described in Supporting Information, and the regularization parameter and kernel width values obtained for scenarios 1 and 2 are provided in Section 2.6.4.

2.6.3. Advantages of SVR.

- Robust to outliers in the data.
- Much better than principal component regression and partial least-squares regression (PLSR) techniques in establishing nonlinear relationships between inputs and outputs.²⁸
- Gradient descent algorithm used in the SVR approach is similar to the mountaineering problem, where it produces a stable error gradient and converges to global minima.^{29,71} Moreover, stochastic gradient helps the algorithm to jump out of the local minima and reach the global minima.
- A perfect balance is achieved between the complexity of the SVR model and the minimization of the discriminant function (eq S7) through structural risk minimization principle (SRM). This principle essentially utilized a regularization parameter (described in the following section) to reduce the gap between prediction errors in the calibration and validation data sets. This reduces overfitting, which is further explained in Section 2.8. Balabin and Lomakina²⁸ suggested that empirical risk minimization (ERM) on which ANNs are based are not as robust as SRM. This makes SVRs more advantageous than ANNs. ERM is basically parameter selection.

2.6.4. Tuned Hyperparameters in SVR. Table 2 provides the values of the hyperparameters, namely, the regularization parameter (C) and kernel width (σ). As mentioned in the Supporting Information, the role of C (eq S8) is to magnify the penalty of errors in the model predictions and increase the tolerance of the model toward outliers. It helps in maximizing

Table 2. Values of the Regularization Parameter and RBF Kernel Width in SVR Models for Scenarios 1 and 2^a

scenario #	HR (°C/min)	σ	C
1	5	2	4
	10	4	5
	15	3	8
	20	1	3
2	5 and 10	3.5	6
	15 and 20	2.5	6
	5, 10, 15	3	5
	10, 15, 20	5	3

^aThe HRs for scenario 2 indicates the calibration set, i.e., on which the SVR model was trained.

the margin and balancing this with minimizing the error, which is the overall goal of SVR itself. It helps us to achieve a trade-off between model complexity and error in both the calibration and validation data sets. We also notice that the kernel width values are <5 , which meant that the degree of nonlinearity between the inputs and outputs is very high in the data set. The values of regularization parameter tell us that the predictions did not deviate too much from the experimental test values of the validation data set since the penalty for the error was low. This also meant that the SVR model exhibited low overfitting tendencies.

2.7. MLR and PLSR. We have compared the results of the nonlinear regression methods with a couple of linear regression techniques such as OLS-based MLR and PLSR. Linear regression is one of the simplest ML techniques where the regression line is straight, and the model residuals (prediction errors) are calculated based on the vertical distance from this regression line. Hence, curved regions such as those exhibited during thermal events in TGA (Figure 1a,b) would not be predicted well. MLRs mean that there are multiple inputs trained on a single output variable. Although it is not able to handle nonlinear relationships well, MLR has been applied in the prediction of selectivities and conversion rates in microfluidic hydrocarbon oxidation⁷² as well as to predict oil yields from plastic pyrolysis.³² PLSR is an improvement over the MLR since it has the ability to handle correlations between the input variables themselves. PLSR is based on the principle of maximizing the covariance between the scores of the matrices obtained upon applying PCA on the original data set. The only parameter that needed to be chosen in PLSR was the number of latent variables, which was chosen as 3 in our work. Three latent variables explained $\sim 96\%$ of the variance in the output variable (mass remaining) for both TBP and TBP + hematite samples under both N_2 and O_2 environments. The single major drawback of PLSR is that it is still a linear regression technique that is unable to handle nonlinear relationships between inputs and outputs, and this overrides its ability to handle multiple correlations within the inputs. The theory of PLSR is well-established in literature,^{73–76} and hence, a detailed description is not provided here. Lastly, there was no parameter to be tuned for the MLR method.

2.8. Model Evaluation Metrics. Two statistical parameters, namely, RMSE and coefficient of determination (R^2), were used to measure the model performances in our work for the reproduction of TGA data at different HRs. The RMSE is given in eq 1 and R^2 is given by eq 2, respectively. There are many ways to calculate R^2 , and one of the ways is to obtain a direct linear bivariate correlation between the experimental values in the testing data set and the corresponding predicted values. R^2 also represents the explained variance, which is the variance of experimental data points (SST in eq 2) minus sum of squared residuals of the predicted model (SSE in eq 2). Ultimately, the higher the value of R^2 , the higher the prediction accuracy of the model. Second, the RMSE is the square root of the MSE, which is the sum of all the residuals (errors) divided by $n - k - 1$, where n and k are the number of data points (rows in the data set) and input variables (number of columns in the data set except the output variable), respectively. It is to be noted that as the number of data points increases, $n - k - 1$ will be very close to $n - 1$ or n , which is the case in our work.

Table 3. Comparison of the R^2 Values Indicating Predictive Performances of All the ML Methods Applied to Train and Test the TGA Data for Pyrolysis of TBP, and Its Combination with Hematite at the four HRs Used^a

sample	HR (°C/min)	RF	GBR	SVR	linear	PLSR ^b
TBP	5	0.999	0.999	0.909	0.549	0.544
	10	0.999	0.999	0.923	0.578	0.577
	15	0.999	0.999	0.928	0.620	0.623
	20	0.999	0.999	0.922	0.629	0.622
	all 4 combined	0.999	0.999	0.917	0.581	0.583
TBP + hematite	5	0.999	0.998	0.916	0.521	0.520
	10	0.983	0.972	0.828	0.600	0.601
	15	0.996	0.990	0.843	0.376	0.376
	20	0.999	0.995	0.944	0.642	0.642
	all 4 combined	0.995	0.989	0.888	0.586	0.584

^aThese values are for the validation sets. ^bPLSR is partial least-squares regression.

$$\text{RMSE} = (\text{MSE})^{1/2} = \left(\frac{1}{n - k - 1} \sum_{i=1}^n |y_i - y_i^*|^2 \right)^{1/2} \quad (1)$$

$$\begin{aligned} R^2 &= 1 - \frac{\sum_{i=1}^n (y_i - y_i^*)^2}{\sum_{i=1}^n (y_i - y_{\text{mean}})^2} \\ &= 1 - \frac{\text{SSE}}{\text{SST}} \\ &= \frac{\text{explained variance}}{\text{total variance}} \end{aligned} \quad (2)$$

where y_i^* and y_i are the model-predicted and experimental values, respectively. y_{mean} is the mean of all the data points in that output variable.

RMSE is generally defined for both the calibration and validation data sets. The RMSE for the calibration data set is referred to as RMSEC (error of calibration), and for the validation data set, it is called RMSEP (error of prediction). If a model is said to be ideally predicted, these two measures should be equal. However, this does not hold true for real-world data sets. The lower the gap between RMSEP and RMSEC, the higher the generalizability, versatility, and adaptability of the model to newer and “unseen” data sets.⁷⁷ In this case, the overfitting tendency would be lower. RMSEP usually tends to be higher than RMSEC since the former measures the error in newer data sets, which tend to be predicted with lower accuracy. On the other hand, the higher the gap between RMSEP and RMSEC, the higher the overfitting tendency of the model. Both these values will be reported in this work in the **Results and Discussion** section.

3. RESULTS AND DISCUSSION

3.1. Results for Scenario 1 under Pyrolysis Conditions.

Table 3 provides the R^2 values of all three nonlinear ML models (RF, GBR, and SVR) for the prediction of TGA data in pyrolysis conditions under a nitrogen environment for both TBP and TBP + hematite samples at all four HRs employed in this work. It is reiterated that a comparison with linear regression and PLSR has also been made so as to establish a baseline for the prediction accuracies. Note that these displayed values are for the validation data set. The hyperparameters (“ $n_{\text{estimators}}$ ” and “ max_{depth} ”) for RF and GBR models were chosen as (10 and 7) and (20 and 3), respectively, as reasoned in the **ML methods section** (2.4 and 2.5) previously. Furthermore, the corresponding RMSEP values of the different ML models at all 4

HRs for TBP and TBP + hematite samples are provided in **Figure 5a,b**, respectively.

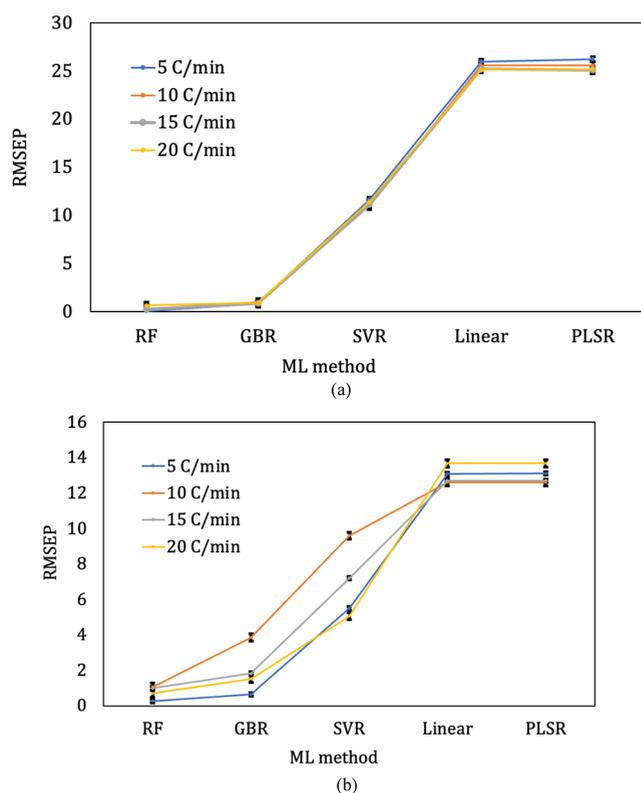


Figure 5. Comparison of the RMSEP values of all ML models at all four HRs applied on the TGA data for the pyrolysis of (a) TBP samples and (b) TBP + hematite samples. Each data point was predicted three times, and hence, the respective standard errors are shown as error bars for each point.

It is very clear from **Table 3** that RF and GBR showed the best performance among all of the ML techniques employed in this work. Exceptional prediction accuracies of >0.99 were obtained for the TGA data of both TBP and TBP + hematite samples at all HRs. It is important to note that the training and testing accuracies for the calibration and validation data sets of all the ML models were similar up to the third decimal for both TBP and TBP + hematite samples, signifying the generalizability of the ML models for TGA data. For TBP alone, both RF and GBR gave the highest R^2 of 0.999 for all HRs individually. This same prediction accuracy was shown for the overall model as

well, which was constructed using the data set combining all four HRs. For this master data set, the model was trained with 80:20 split for the calibration/validation sets. The R^2 values of testing for these overall models combining all 4 HRs are indicated in bold and are highlighted in green in Table 3. For the TBP + hematite samples, RF showed a prediction accuracy of 0.999 for 5 and 20 °C/min but the R^2 values of the other two HRs still remained >0.983. This was slightly higher than the R^2 values shown by GBR models on TBP + hematite samples when the third decimal was considered. For 5 and 20 °C/min, GBR showed R^2 values of 0.998 and 0.995, respectively, which were much higher than SVR and linear regression models. Though the performances of GBR and RF were the same for TBP, RF clearly outperformed GBR for TBP + hematite samples. Overall, both the individual models on the HRs and their combined models indicated a superior performance of RF regression over all the other ML methods employed in this work. We also notice that when hematite was added to TBP, the prediction accuracy of the best performing ML models such as RF, GBR, and SVR slightly decreased. This can be attributed to the complexity in composition introduced by Fe_2O_3 and its interactions with TBP. This is clearly reflected in the TGA mass loss curves for TBP + hematite in Figure 1a, where multiple thermal events were seen as opposed to a single decomposition event in Figure 1b for TBP alone. This was confirmed in the previous works on the copyrolysis of TBP and hematite, which led to a complex mixture of products such as aldehydes, alcohols, esters, aliphatic alkanes, aliphatic alkenes, substituted aromatics, and phenolic derivatives. On the other hand, pyrolysis of pure TBP consisted of only 1,3,5-tribromo benzene and 2,4-dibromophenol, which indicated the removal of -OH and 1 Br atom from TBP, respectively. Similar prediction results using ML techniques for TGA data were obtained in another work by our coauthors Ali et al.,²⁶ when $\text{Ca}(\text{OH})_2$ was copyrolyzed with TBBA, which is another model compound for BFRs consisting of 2 units of TBP.

Though the performance of SVR was lower than those of RF and GBR, it was still much higher than the linear regression techniques such as MLR and PLSR. The performances of these linear regression techniques were poor, with R^2 values of ~0.59 for both TBP and TBP + hematite samples at all HRs (Table 3). This was because of the nonlinearity in the TGA data that can be clearly seen in the mass loss curves in Figure 1a,b for TBP and TBP + hematite, respectively. This implied that linear regression techniques could not capture nonlinearity in the TGA data effectively. It was also very interesting to note that the performances shown by both MLR and PLSR were almost identical in terms of R^2 and this is reflected further in Figure 5a,b, which depict the RMSEP values across all ML methods for both types of samples used in this work. The trend of the performances of all the ML models is clear from the RMSEP values, as there is a gradual rise in the lines for all HRs for both types of samples, as seen in Figure 5. This translates to the fact that as we move from RF and GBR regression trees to the linear MLR and PLSR models, the prediction errors increase significantly from <1 to ~25 in TBP samples and <1 to ~14 in TBP + hematite samples. Moreover, the errors shown by SVR and linear regression models at all HRs were higher for the TBP samples (Figure 5a) while those for the regression trees were opposite, showing marginally higher prediction errors for the hematite-added TBP samples (Figure 5b). Overall, it was clear that RF outperformed all the other ML models at all HRs for the TGA data on both types of samples obtained under a N_2 environment. The order of performance for scenario 1 where

the ML models were trained with TGA data at both individual and combined HRs consisting of 80% data points in the calibration set and remaining 20% data points in the validation set under pyrolysis (N_2) conditions was as follows: RF > GBR > SVR > MLR ~ PLSR. This was confirmed based on both the performance metrics used in this work, that is, R^2 and RMSE.

3.2. Results for Scenario 2 under Pyrolysis Conditions: Effect of Varying the Number of Data Points for Model Training. As given in Figure 2, we have implemented all of the ML models consisting of nonlinear and linear regression on the TGA data of TBP and TBP + hematite separately for scenario 2. This was different from scenario 1 in the number of data points used for ML model training and validation, as follows:

- (i) In scenario 2a, the ML models were trained using 2 sets of HRs (for example, 5 and 10 °C/min with ~28,000 and ~21,000 observations for TBP + hematite and TBP, respectively) and predicted on 15 and 20 °C/min with ~11,000 and ~8000 observations for TBP + hematite and TBP, respectively. Following this, the calibration set was changed to 15 and 20 °C/min and predicted on 5 and 10 °C/min for both sets of samples as well. In summary, the calibration and validation sets are interchanged between 5 and 10 and 15 and 20 °C/min to evaluate the prediction ability of the model in different situations. Table 1 provides the number of observations of TGA data available for each sample at each of the 4 HRs.
- (ii) In scenario 2b, the ML models were trained using 3 sets of HRs (for example, 5, 10, and 15 °C/min with ~34,000 and ~24,000 observations for TBP + hematite and TBP, respectively), while predicting on 20 °C/min. Here, there were ~5000 and ~3500 observations in the validation set for TBP + hematite and TBP samples, respectively. Similar to scenario 2a, this was also repeated with data from 10, 15, and 20 °C/min HRs as the calibration set and 5 °C/min as the validation set. Here, there were 19,000 and 14,000 observations in the validation set for TBP + hematite and TBP samples, respectively. This ensured that the whole range of samples were covered for training and testing, and the ML models were tested for robustness by varying the calibration and validation data sets with the ultimate objective of making the ML models more versatile and generalizable.

3.2.1. Overall Best Predicting Method. The plots of R^2 and RMSEP for both TBP and TBP + hematite samples for scenario 2a and scenario 2b are provided in Figure 6 (TBP) and Figure 7 (TBP + hematite), respectively. In these figures, the HRs correspond to the validation data sets on which the model was predicted on. For example, the green line (scenario 2a—HR 15, 20 °C/min) corresponds to the R^2 and RMSEP values when the models were trained using 5 and 10 °C/min and tested on 15 and 20 °C/min HRs.

A number of interesting observations can be made from these results. First of all, it is confirmed that RF was the best predicting model, with GBR being a close second in most of the cases. However, there was one exception in scenario 2a, for which GBR marginally outperformed RF by 0.02 units in R^2 for both TBP (Figure 6a) and TBP + hematite (Figure 7a) samples. The GBR showed a slightly better prediction accuracy and lower error when both 5 and 10 and 15 and 20 °C/min HRs were used as the validation sets in scenario 2a. The overall trend in the performance of the ML models corroborated with the results shown in Section 3.1 for scenario 1. It can be clearly seen from

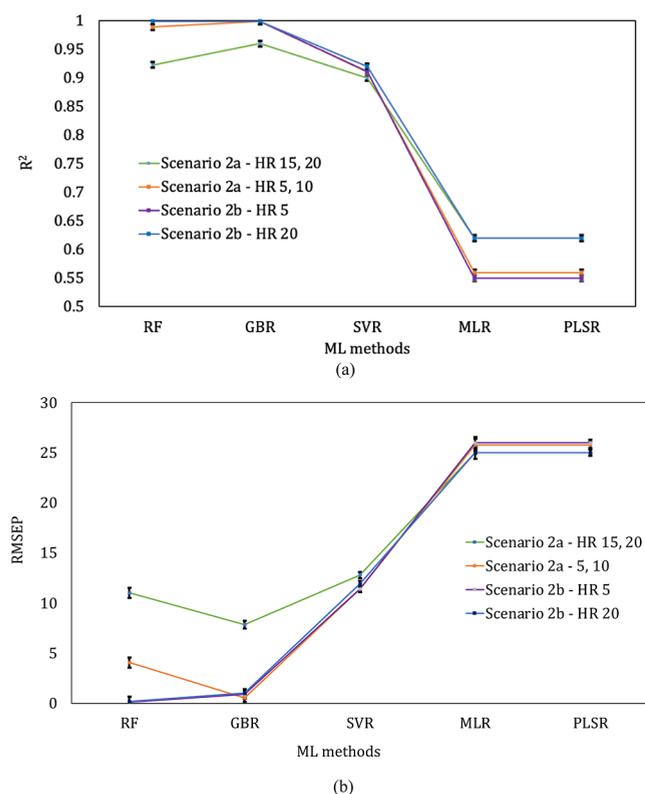


Figure 6. Comparison of (a) R^2 and (b) RMSEP values of all ML models on samples of TBP for scenario 2 with varying number of training data points. Scenario 2a corresponds to training on two sets of HRs and predicting on the other 2 sets, while scenario 2b corresponds to training on three sets of HRs and predicting on the remaining single set.

Figures 6 and 7 that as the R^2 values decrease from the decision trees to the linear regression models, the RMSEP values correspondingly increase. Furthermore, the prediction errors of SVR in scenario 2b increased by 8-fold and 6-fold for both TBP (Figure 6b) and TBP + hematite (Figure 7b) samples, respectively, in comparison to RF and GBR techniques. The RMSEP values of the linear regression models (MLR and PLSR) were 13–14-fold and 9–10-fold higher than those of RF and GBR for TBP (Figure 6b) and TBP + hematite (Figure 7b) samples, respectively, in scenario 2b.

Hence, for both scenarios 2a and 2b, the order of performance in terms of highest to lowest R^2 and lowest to highest RMSEP was RF \sim GBR $>$ SVR $>$ MLR \sim PLSR with prediction accuracies of RF and GBR reaching 0.999 for both types of samples. Moreover, scenario 2a showed slightly worse results than scenario 1 for RF and GBR, in terms of lower R^2 and higher prediction errors (RMSEP). However, scenario 2b showed equal or better performance as compared to scenario 1. These observations were specifically due to the lower number of training data points in the calibration set for the ML models in scenario 2a as compared to scenario 2b and scenario 1. Further comparison of scenario 2a and 2b is provided in Section 3.2.3.

3.2.2. Influence of Hematite Addition to TBP on the Performance of ML Methods. In Section 3.1 for scenario 1, we saw that the prediction accuracy slightly decreased for TBP + hematite samples compared to that of TBP alone. This was attributed to the interactions between TBP and hematite, leading to a mixture of products, and this further affected the TGA curve at different temperatures. A direct result of this

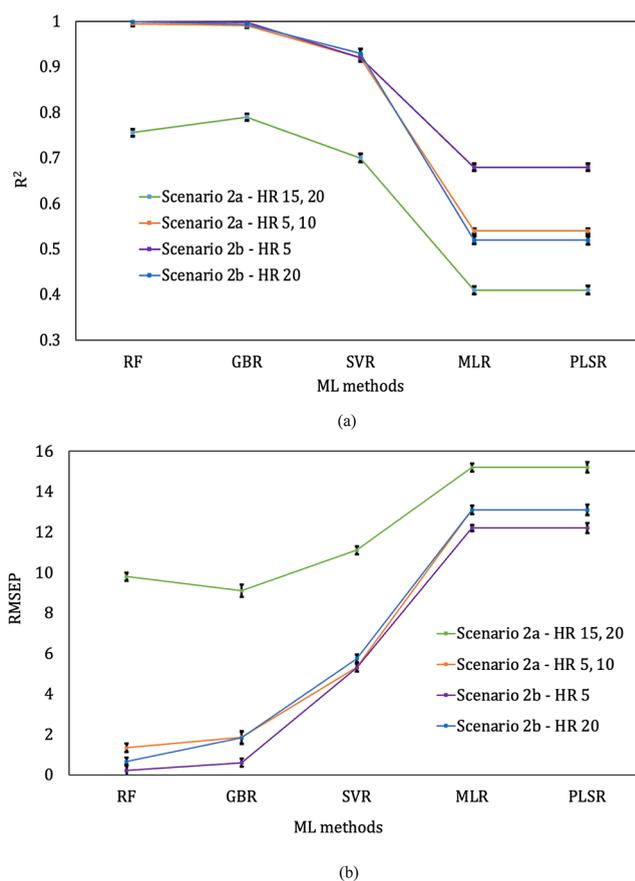


Figure 7. Comparison of: (a) R^2 and (b) RMSEP values of all ML models on samples of TBP + hematite for scenario 2 with varying number of training data points. Scenario 2a corresponds to training on two sets of HRs and predicting on the other 2 sets while scenario 2b corresponds to training on three sets of HRs and predicting on the remaining single set.

interaction is the nonlinearity in the TGA curve, which arises due to the two major thermal events, as detailed in Section 2.2. The spread of these events over a range of temperatures (100–180 and 280–410 °C) and the minor shifts in these temperature ranges as the HRs varied might have affected the performances of the ML models. Though this difference was lesser in scenario 1, this was more prominent in scenario 2a, especially when the models were predicted on 15 and 20 °C/min (green line in Figures 6 and 7). However, this did not have any effect for scenario 2b, where 3 HRs were used as the calibration set and predicted on 1 set. In fact, scenario 2b showed better performance than scenario 1 for TBP + hematite samples for all of the ML models. RF and GBR showed highest R^2 values of 0.999, while MLR and PLSR showed R^2 values of 0.7 for scenario 2b. In scenario 1, the R^2 values for the overall linear regression model were lesser by 0.11 units (Table 3). Even the RMSEP values of scenario 2b for all models on TBP + hematite were less than scenario 1 by 1 or 2 units. It was also clear that the performance of the ML models in scenario 2b was better for TBP + hematite than for TBP alone.

On the other hand, it was very interesting to note that the range of prediction errors (difference in RMSEP values between linear regression and RF models) was doubled for TBP as compared to that of TBP + hematite samples (Figures 6b and 7b, respectively). This was also true for scenario 1, where the RMSEP values for SVR and linear regression models for TBP

were much higher than that of TBP + hematite samples. This trend was also reflected in the R^2 values as the linear regression showed a steep decrease to 0.55 for TBP but only reduced to 0.7 for the TBP + hematite samples in both scenarios (Figures 6a and 7a). This meant that though the performances for RF and GBR were only slightly affected on the negative side by adding hematite to the TBP samples, the performances of SVR and linear regression improved for TBP + hematite. This implied that the ML models could be used successfully for TBP + hematite samples and were also adaptive to a varied number of training points, apart from showing high prediction accuracies. All this was in-line with the central objective of this paper, implying that the use of the offline-TGA instrument could be reduced and eventually eliminated through applying ML approaches, leading to faster process times and lesser human and experimental errors.

3.2.3. Comparison of Scenario 2a Versus 2b: Influence of Number of Training Points. It can be clearly seen from Figures 7b and 6b that the green and orange lines are above the violet and blue lines for all ML models. As indicated in the figures, the former (green and orange) lines represent the RMSEP values of scenario 2a, where the ML models were trained using TGA data at two HRs and predicted on the remaining 2 HRs. The latter (violet and blue) lines correspond to scenario 2b, where the ML models were trained using 3 HRs and predicted on 1 HR data. The larger the number of data points in the calibration set used for model training, the lower the prediction error (RMSEP) and higher the prediction accuracy (R^2). The prediction accuracies shown by RF and GBR for TBP + hematite samples were not overly dependent on the sample sizes, as they showed R^2 values of >0.99 in scenarios 2b and 2a except for the case where data at 15 and 20 °C/min were used for prediction (green line in Figure 7a). Though RF and GBR showed the best performance for scenario 2a on 15 and 20 °C/min as the validation set (green line in Figure 7a), the R^2 value dipped below 0.8 for TBP + hematite. On the other hand, the prediction accuracies for TBP samples in scenario 2a were lower than in 2b (Figure 6a) but still remained >0.92 for the best-performing RF and GBR methods. For the linear regression and SVR models in TBP, it was difficult to clearly distinguish between the performances of scenarios 2a and 2b since the R^2 and RMSEP values were very close to each other (Figure 6a,b). The performances of the linear regression and SVR models in TBP + hematite gave a clear picture that scenario 2b was better than scenario 2a. In summary, it can be stated that all the ML models in scenario 2b outperformed those in scenario 2a for both types of samples, indicating that the number of calibration data points can be a major factor in ML regression model performance for TGA data in e-waste treatment.

3.3. Advantages, Limitations, and Challenges in the ML Approach. Through the methods described in scenario 1 and scenario 2, the full range of our models were tested for their prediction ability on various data sets of lower and higher HRs. An 80:20 split was used in scenario 1 when the entire data set in each HR was considered for calibration and validation. In scenario 2, the number of calibration data points were varied and the ML models were applied to predict on different “unseen” data sets. The performances of the linear and nonlinear regression models were compared in the following ways: (i) in scenario 1, the performances of the ML models were compared among themselves for each type of sample as well as between TBP and TBP + hematite; (ii) in scenario 2, the performances of the ML models were compared with each other for different calibration and validation sets considering 2 and 3 HRs together;

and (iii) the performances of all the ML models were compared between scenarios 1 and 2 and also within scenario 2a and 2b so that the total effect of the number of data points used for training the ML models is seen.

Our results show that the decision tree-based models such as RF and GBR work best for TGA data, and they are also able to predict the entire data for different HRs than the ones they were trained on. This suggests that the constructed models in this work are robust, versatile, and can be generalized for application in TGA data sets in order to remove the use of offline measuring instruments. The advantages of our adopted approach in this work are that it (i) is much quicker generation of mass loss curves for different HRs than from the TGA instrument itself ($\sim 1/20$ th of the time taken for collecting experimental data), (ii) has the potential for online monitoring of the copyrolysis of BFRs (in PCBs present in e-waste) with various metal oxides for the efficient removal and capture of bromine, (iii) will reduce the environmental impact of e-waste and reduce the toxic emissions of HBr, (iv) uses less physical equipment, which reduces instrumental and human errors, (v) has adaptability of the best-predicting ML models to large data sets of sizes $>10,000$ data points and ability to deal with multidimensional and nonlinear data, and (vi) has the adaptability of the best-predicting ML models to different chemical environments such as inert N_2 and combustion conditions with O_2 . Section 3.4 reports the prediction accuracies of the RF and GBR models along with other techniques for the TGA data obtained under the O_2 conditions. It could be seen that the performances of these models were as good as the N_2 conditions, and this showed the versatility of the models to different chemical environments as well. This work has indeed shown the potential to be extended to complex, nonlinear data from other characterization techniques such as spectroscopy and chromatography. The best-predicting ML models in this work are based on decision trees and require some hyperparameters to be tuned. However, the advantage is that this takes very less time (<15 s) and only requires parameters to be tuned. SVR requires an additional parameter and did not perform as well as RF and GBR.

However, there are certain limitations and challenges associated with the ML approach for data prediction. First of all, the process of data arrangement and tidying is the most crucial part of the entire process. The detailed process of tidying and the functions used in Python are given in the Supporting Information. Data from the laboratory equipment are never perfect. It almost always contains noise, which needs to be filtered out during the preprocessing step. Next, the raw data from the instrument are messy and can contain missing values. These must be corrected by “tidying” the data, which essentially involves identifying the dependent (output) and independent (input) variables, ensuring that each column is a unique identifier, each row is a different trial, and the rows in the input column correctly match with those in the output variables. In essence, a tidy data set means that every row is an observation, every column is a unique variable, and every cell has a single value. All this was done completely using Python in our work and is detailed in the Supporting Information document.

The next limitation is the dependence of the performances of the ML models on the number of data points used for training. The simple rule is that the more the data points, the better. The data points correspond to the number of samples or observations (rows) in the data set. TGA consists of numerous data points since the HR is much smaller than the range of temperatures explored. The higher the number of training points

in the calibration set, the better the ML model is equipped to predict the output for a new data set. We could see that RF and GBR showed excellent prediction accuracies of >0.99 and even reached 0.999 for both types of samples. The number of data points used for training was seen as a major challenge in the work by Siddiquee et al.⁷² for predicting product selectivity for hydrocarbon oxidation. They had only 10 observations in total, which they split into calibration and validation sets equally and also varied the data points for training similar to what we did here. Their objective was not only predicting product selectivity but also comparing the relative importance of a number of input variables that influenced the output. In general, it is recommended to have a minimum of 1000 data points to train the ML model.^{34,69,78–80} Anything below that is seen as a major limitation and will not give accurate results, or the results cannot be trusted.

Another limitation of our work is the absence of real-time testing of the best-predicting RF and GBR models, which will be carried out in future. It will be a challenge to build a sensor that will be integrated with the pyrolysis process in e-waste treatment and will give the mass loss curves at different HRs. Real-time testing also involves generating the TGA data online as the HR and the chemical environment (N₂ or O₂ conditions) are specified within the temperature range. Our work has provided high confidence that this will be possible and will be successfully carried out in the near future. A further challenge here will be the interpretation of the mass loss curves and regions of thermal degradation without human intervention. At this point, chemical analysis and interpretation are not possible without a chemical engineer or an expert in TGA analysis. However, a mapping framework can be built using another ML approach, in which the mass loss temperatures and the corresponding chemical interpretations are fed as inputs and outputs, respectively.

Choosing which ML models are suitable for a particular data set is also another challenge in our approach. This requires extensive literature review, knowledge of the ML models and their principles of working, and sufficient trial and error as well. The types of ML models that best suit a particular data set are chosen based on the number of data points available, the type of instrument, the expected nonlinearity that may be present in the data, and the number of input and output variables. Ultimately, we go with models that are easy to develop and implement, robust, versatile, and efficient, have a history of being successful, are able to establish nonlinear relationships, and take less time to train. Models that are able to deal with nonlinear data also model linear data very well, but the reverse is not true. Lastly, though we saw that the ML models performed very well for the TGA data under both N₂ and O₂ conditions, the challenge will significantly increase when there is a reacting medium in a real-time scenario. This will complicate the mass loss data even further, bringing more nonlinearity to the curves and making prediction difficult. The only solution to this is higher number of data points for calibrating the ML models, and this will improve prediction accuracy.

3.4. ML Performance for TGA Data under Combustion Conditions. The performances of the ML models when trained on the TGA data of TBP and TBP + hematite samples obtained under the O₂ environment were observed to be very similar to that of the N₂ environment, as previously discussed. A detailed discussion for combustion data prediction is not carried out here due to similarity of the results with pyrolysis conditions. However, the prediction accuracies of TBP + hematite samples for scenario 1 and scenario 2a, 2b with all 5 ML models are given

in Table 4. Again, it can be seen that the order of performance of the ML models was: RF > GBR > SVR > MLR = PLSR. When

Table 4. R² Values of all ML Models for All Scenarios for TGA Data Prediction of TBP + Hematite Samples under Combustion Conditions

scenario #	HR (°C/min)	RF	GBR	SVR	MLR	PLSR
scenario 1	5	0.999	0.996	0.932	0.621	0.621
	10	0.995	0.994	0.943	0.593	0.593
	15	0.996	0.993	0.944	0.593	0.593
	20	0.999	0.995	0.951	0.602	0.602
scenario 2a	5, 10	0.991	0.993	0.912	0.576	0.576
	15, 20	0.982	0.982	0.889	0.555	0.555
scenario 2b	5	0.999	0.999	0.892	0.543	0.543
	20	0.999	0.998	0.902	0.571	0.571

compared to the N₂ environment, the performances of SVR and linear regression were slightly better for combustion conditions, and no particular reason could be attributed to this since the nature of the data was similar. An important inference from this observation would be that the best-predicting ML models (RF and GBR) were not only able to capture nonlinearity and multidimensionality in the TGA data but could also handle interactions between the sample and the chemical gaseous environment under which TGA was performed. This was because the combustion data were also predicted with high accuracy by the decision tree-based models, which showed excellent R² values of 0.999 for scenario 1 and scenario 2 (Table 4). We also noticed that the prediction accuracies in scenarios 2a were slightly lower than those in scenario 1, and this was purely because of the lower number of training data points used in scenario 2a than in scenario 1 (Table 1). Interestingly, scenario 2b showed equivalent prediction accuracies for the RF and GBR models with scenario 1 under the O₂ conditions, as was the case under the N₂ conditions. Lastly, the performances of the ML models in scenario 2b were higher than in scenario 2a since the models were calibrated using TGA data from three sets of HRs in scenario 2b as opposed to 2 sets in scenario 2a (Table 4). In conclusion, we can confidently say that our ML models built in this work are versatile and robust and are generalizable for “unseen” data sets.

4. CONCLUSIONS

In this work, nonlinear and linear regression ML techniques were applied on the data from TGA obtained for the coprolysis of TBP and hematite (Fe₂O₃). The data were obtained at four different HRs of 5, 10, 15, and 20 °C/min under both pyrolysis and combustion environments. Out of all the ML techniques employed, decision tree-based RF regression showed the best performance for predicting the TGA data under different scenarios, that is, when all the HRs were combined in the calibration set as well as when the number of data points used for training were varied. GBR was a close second in terms of the prediction ability for the TGA data in all of the scenarios explored in this work. R² values reached as high as 0.999 for both RF and GBR models, which meant they could be successfully used to replace the TGA instrument itself. Another key finding was that the TGA data for the TBP + hematite samples were predicted with slightly higher accuracy than those for TBP samples alone. This showed the ability of RF and GBR in capturing nonlinearity of the data effectively. SVR performed

much better than the linear regression models but could not reach the level of RF or GBR. The performance of RF was strongly established under both pyrolysis and combustion environments, thus confirming the versatility and robustness of the ML models under various conditions. Our work has shown immense potential for online monitoring of the pyrolysis process in e-waste treatment and recycling. Furthermore, the generalizability of the ML models could be extended to other types of characterization techniques such as infrared spectroscopy and chromatography. Our ML models have shown that the nonlinear and complex relationships in the data from these techniques can be predicted very well. Moreover, the removal of offline measuring instruments facilitated by the ML models could be useful in predicting product concentrations in a number of chemical processes such as biomass conversions, catalytical conversions of hydrocarbons and in sludge treatment. Online predictions quicken the process and also reduce human and instrument errors.

■ ASSOCIATED CONTENT

Data Availability Statement

In order to access all the predicted and experimental data along with the python codes for all the nonlinear and linear ML regression models applied on the TGA data for construction and validation, the readers are directed toward the [Supporting Information](#) document, where we have provided the link to the GitHub repository that hosts all the input, output data and codes for tidying, hyperparameter tuning and prediction.

Supporting Information

The Supporting Information is available free of charge at <https://pubs.acs.org/doi/10.1021/acsomega.3c07228>.

Data cleaning description, software used for the regression models which also contains the link to the public repository (GitHub) for access to the codes and the predicted TGA data for all the ML techniques used in this work, and theory of RF, GBR, and SVR along with details of their hyperparameter tuning ([PDF](#))

■ AUTHOR INFORMATION

Corresponding Author

Kausik Sivaramakrishnan – Department of Chemical and Petroleum Engineering, United Arab Emirates University, Al-Ain 15551, United Arab Emirates; orcid.org/0000-0002-6590-3265; Email: kausiva@uaeu.ac.ae

Authors

Labeeb Ali – Department of Chemical and Petroleum Engineering, United Arab Emirates University, Al-Ain 15551, United Arab Emirates

Mohamed Shafi Kuttiyathil – Department of Chemical and Petroleum Engineering, United Arab Emirates University, Al-Ain 15551, United Arab Emirates; orcid.org/0000-0001-5624-3907

Vignesh Chandrasekaran – Department of Computer Science, University of British Columbia, Vancouver V6T 1Z4, Canada

Oday H. Ahmed – Department of Physics, College of Education, Al-Iraqia University, Baghdad 10071, Iraq

Mohammad Al-Harashsheh – Chemical Engineering Department, Jordan University of Science and Technology, Irbid 22110, Jordan

Mohammednoor Altarawneh – Department of Chemical and Petroleum Engineering, United Arab Emirates University, Al-

Ain 15551, United Arab Emirates; orcid.org/0000-0002-2832-3886

Complete contact information is available at:

<https://pubs.acs.org/10.1021/acsomega.3c07228>

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This study has been supported by the following grants: (i) Startup Grant from United Arab Emirates University (UAEU)—grant number 12N104, (ii) 2019 Abu Dhabi Award for Research Excellence (AARE)—(by ASPIRE, part of the Advanced Technology Research Council (ATRC, Abu Dhabi, United Arab Emirates), grant number: 21N225-AARE2019—ADEK—103).

■ REFERENCES

- (1) Ali, L.; Shafi Kuttiyathil, M.; Altarawneh, M. Oxidative and Pyrolytic Decomposition of an Evaporated Stream of 2,4,6-Tribromophenol over Hematite: A Prevailing Scenario during Thermal Recycling of e-Waste. *Waste Manage.* **2022**, *154*, 283–292.
- (2) Zhao, L.; Lu, Y.; Zhu, H.; Cheng, Z.; Wang, Y.; Chen, H.; Yao, Y.; Zhang, J.; Li, X.; Sun, Z.; Zhang, C.; Sun, H. E-Waste Dismantling-Related Occupational and Routine Exposure to Melamine and Its Derivatives: Estimating Exposure via Dust Ingestion and Hand-to-Mouth Contact. *Environ. Int.* **2022**, *165*, 107299.
- (3) Forti, V.; Baldé, C. P.; Kuehr, R.; Bel, G. *The Global E-Waste Monitor 2020*: Bonn/Geneva/Rotterdam, 2020.
- (4) Sahle-Demessie, E.; Mezgebe, B.; Dietrich, J.; Shan, Y.; Harmon, S.; Lee, C. C. Material Recovery from Electronic Waste Using Pyrolysis: Emissions Measurements and Risk Assessment. *J. Environ. Chem. Eng.* **2021**, *9* (1), 104943.
- (5) Tian, X.; Xie, J.; Hu, L.; Xiao, H.; Liu, Y. Waste LEDs in China: Generation Estimation and Potential Recycling Benefits. *Resour. Conserv. Recycl.* **2022**, *187*, 106640.
- (6) Hsu, E.; Durning, C. J.; West, A. C.; Park, A.-H. A. Enhanced Extraction of Copper from Electronic Waste via Induced Morphological Changes Using Supercritical CO₂. *Resour. Conserv. Recycl.* **2021**, *168*, 105296.
- (7) Ghimire, H.; Ariya, P. A. E-Wastes: Bridging the Knowledge Gaps in Global Production Budgets, Composition, Recycling and Sustainability Implications. *Sustain. Chem.* **2020**, *1* (2), 154–182.
- (8) Altarawneh, M. Temperature-Dependent Profiles of Dioxin-like Toxicants from Combustion of Brominated Flame Retardants. *J. Hazard. Mater.* **2022**, *422*, 126879.
- (9) Rene, E. R.; Sethurajan, M.; Kumar Ponnusamy, V.; Kumar, G.; Bao Dung, T. N.; Brindhadevi, K.; Pugazhendhi, A. Electronic Waste Generation, Recycling and Resource Recovery: Technological Perspectives and Trends. *J. Hazard. Mater.* **2021**, *416*, 125664.
- (10) Ali, L.; A.Mousa, H.; Al-Harashsheh, M.; Al-Zuhair, S.; Abu-Jdayil, B.; Al-Marzouqi, M.; Altarawneh, M. Removal of Bromine from the Non-Metallic Fraction in Printed Circuit Board via Its Co-Pyrolysis with Alumina. *Waste Manage.* **2022**, *137*, 283–293.
- (11) Altarawneh, M.; Saeed, A.; Al-Harashsheh, M.; Dlugogorski, B. Z. Thermal Decomposition of Brominated Flame Retardants (BFRs): Products and Mechanisms. *Prog. Energy Combust. Sci.* **2019**, *70*, 212–259.
- (12) Ali, L.; Kuttiyathil, M. S.; Ahmed, O. H.; Altarawneh, M. Separation of Bromine and Hydrocarbons from Polymeric Constituents in E-Waste through Thermal Treatment with Calcium Hydroxide. *Sep. Purif. Technol.* **2023**, *307*, 122836.
- (13) Mousa, N. A.; Ali, L.; Kuttiyathil, M. S.; Mousa, H. A.; Altarawneh, M. Exploring the Potential of Hematite as a Debromination Agent for 2,4,6-Tribromophenol. *Chem. Eng. J. Adv.* **2022**, *11*, 100334.

- (14) Terakado, O.; Ohhashi, R.; Hirasawa, M. Bromine Fixation by Metal Oxide in Pyrolysis of Printed Circuit Board Containing Brominated Flame Retardant. *J. Anal. Appl. Pyrolysis* **2013**, *103*, 216–221.
- (15) Wu, H.; Shen, Y.; Harada, N.; An, Q.; Yoshikawa, K. Production of Pyrolysis Oil with Low Bromine and Antimony Contents from Plastic Material Containing Brominated Flame Retardants and Antimony Trioxide. *Energy Environ. Res.* **2014**, *4* (3), 105.
- (16) Altarawneh, M.; Ahmed, O. H.; Al-Harashsheh, M.; Jiang, Z.-T.; Dlugogorski, B. Z. A Kinetic Model for Halogenation of the Zinc Content in Franklinite. *Appl. Surf. Sci.* **2021**, *562*, 150105.
- (17) Ma, C.; Yu, J.; Chen, T.; Yan, Q.; Song, Z.; Wang, B.; Sun, L. Influence of Fe Based ZSM-5 Catalysts on the Vapor Intermediates from the Pyrolysis of Brominated Acrylonitrile-Butadiene-Styrene Copolymer (Br-ABS). *Fuel* **2018**, *230*, 390–396.
- (18) Liu, J.; Wang, H.; Zhang, W.; Wang, T.; Mei, M.; Chen, S.; Li, J. Mechanistic Insights into Catalysis of In-Situ Iron on Pyrolysis of Waste Printed Circuit Boards: Comparative Study of Kinetics, Products, and Reaction Mechanism. *J. Hazard. Mater.* **2022**, *431*, 128612.
- (19) Ma, C.; Kamo, T. Enhanced Debromination by Fe Particles during the Catalytic Pyrolysis of Non-Metallic Fractions of Printed Circuit Boards over ZSM-5 and Ni/SiO₂-Al₂O₃ Catalyst. *J. Anal. Appl. Pyrolysis* **2019**, *138*, 170–177.
- (20) Chang, J.; Pan, W.; Liu, X.; Xue, Q.; Fu, J.; Zhang, A. The Formation of PBDFs from the Ortho-Disubstituted Phenol Precursors: A Comprehensive Theoretical Study on the PBDD/Fs Formation from 2,4,6-Tribromophenol. *Sci. Total Environ.* **2020**, *713*, 136657.
- (21) Oleszek, S.; Grabda, M.; Shibata, E.; Nakamura, T. Study of the Reactions between Tetrabromobisphenol A and PbO and Fe₂O₃ in Inert and Oxidizing Atmospheres by Various Thermal Methods. *Thermochim. Acta* **2013**, *566*, 218–225.
- (22) Terakado, O.; Ohhashi, R.; Hirasawa, M. Thermal Degradation Study of Tetrabromobisphenol A under the Presence Metal Oxide: Comparison of Bromine Fixation Ability. *J. Anal. Appl. Pyrolysis* **2011**, *91* (2), 303–309.
- (23) Rzyman, M.; Grabda, M.; Oleszek-Kudlak, S.; Shibata, E.; Nakamura, T. Studies on Bromination and Evaporation of Antimony Oxide during Thermal Treatment of Tetrabromobisphenol A (TBBPA). *J. Anal. Appl. Pyrolysis* **2010**, *88* (1), 14–21.
- (24) Grabda, M.; Oleszek-Kudlak, S.; Rzyman, M.; Shibata, E.; Nakamura, T. Studies on Bromination and Evaporation of Zinc Oxide during Thermal Treatment with TBBPA. *Environ. Sci. Technol.* **2009**, *43* (4), 1205–1210.
- (25) Terakado, O.; Ohhashi, R.; Hirasawa, M. Bromine Fixation by Metal Oxide in Pyrolysis of Printed Circuit Board Containing Brominated Flame Retardant. *J. Anal. Appl. Pyrolysis* **2013**, *103*, 216–221.
- (26) Ali, L.; Sivaramakrishnan, K.; Kuttiyathil, M. S.; Chandrasekaran, V.; Ahmed, O. H.; Al-Harashsheh, M.; Altarawneh, M. Prediction of Thermogravimetric Data in Bromine Captured from Brominated Flame Retardants (BFRs) in e-Waste Treatment Using Machine Learning Approaches. *J. Chem. Inf. Model.* **2023**, *63* (8), 2305–2320.
- (27) Ahmed, O. H.; Altarawneh, M.; Al-Harashsheh, M.; Jiang, Z.-T.; Dlugogorski, B. Z. Recycling of Zincite (ZnO) via Uptake of Hydrogen Halides. *Phys. Chem. Chem. Phys.* **2018**, *20* (2), 1221–1230.
- (28) Balabin, R. M.; Lomakina, E. I. Support Vector Machine Regression (SVR/LS-SVM)—an Alternative to Neural Networks (ANN) for Analytical Chemistry? Comparison of Nonlinear Methods on near Infrared (NIR) Spectroscopy Data. *Analyst* **2011**, *136*, 1703.
- (29) Pierna, J. A. F.; Baeten, V.; Renier, A. M.; Cogdill, R. P.; Dardenne, P. Combination of Support Vector Machines (SVM) and near-Infrared (NIR) Imaging Spectroscopy for the Detection of Meat and Bone Meal (MBM) in Compound Feeds. *J. Chemom.* **2004**, *18*, 341–349.
- (30) Marani, A.; Nehdi, M. L. Machine Learning Prediction of Compressive Strength for Phase Change Materials Integrated Cementitious Composites. *Constr. Build. Mater.* **2020**, *265*, 120286.
- (31) Torres-Barrán, A.; Alonso, C.; Dorransoro, J. R. Regression Tree Ensembles for Wind Energy and Solar Radiation Prediction. *Neuro-computing* **2019**, *326–327*, 151–160.
- (32) Belden, E. R.; Rando, M.; Ferrara, O. G.; Himebaugh, E. T.; Skangos, C. A.; Kazantzis, N. K.; Paffenroth, R. C.; Timko, M. T. Machine Learning Predictions of Oil Yields Obtained by Plastic Pyrolysis and Application to Thermodynamic Analysis. *ACS Eng. Au* **2023**, *3* (2), 91–101.
- (33) Wang, X.; Pan, P.; Li, J. Real-Time Measurement on Dynamic Temperature Variation of Asphalt Pavement Using Machine Learning. *Measurement* **2023**, *207*, 112413.
- (34) Wan, Y.; Zeng, Q.; Shi, P.; Yoon, Y.-J.; Tay, C. Y.; Lee, J.-M. Machine Learning-Assisted Optimization of TBBPA-Bis-(2,3-Dibromopropyl Ether) Extraction Process from ABS Polymer. *Chemosphere* **2022**, *287*, 132128.
- (35) Ali, L.; Kuttiyathil, M. S.; Al-Harashsheh, M.; Altarawneh, M. Kinetic Parameters Underlying Hematite-Assisted Decomposition of Tribromophenol. *Arab. J. Chem.* **2023**, *16* (3), 104540.
- (36) Osman, A. I.; Farrell, C.; Al-Muhtaseb, A. H.; Al-Fatesh, A. S.; Harrison, J.; Rooney, D. W. Pyrolysis Kinetic Modelling of Abundant Plastic Waste (PET) and in-Situ Emission Monitoring. *Environ. Sci. Eur.* **2020**, *32* (1), 112.
- (37) Al-Harashsheh, M.; Altarawneh, M.; Aljarrah, M.; Rummanah, F.; Abdel-Latif, K. Bromine Fixing Ability of Electric Arc Furnace Dust during Thermal Degradation of Tetrabromobisphenol: Experimental and Thermodynamic Analysis Study. *J. Anal. Appl. Pyrolysis* **2018**, *134*, 503–509.
- (38) Jayashree, M.; Parthibavarman, M.; Prabhakaran, S. Hydrothermal-Induced α -Fe₂O₃/Graphene Nanocomposite with Ultrahigh Capacitance for Stabilized and Enhanced Supercapacitor Electrodes. *Ionics* **2019**, *25* (7), 3309–3319.
- (39) Larsen, J.; Goutte, C. On Optimal Data Split for Generalization Estimation and Model Selection. *Neural Networks for Signal Processing—Proceedings of the IEEE Workshop*, 1999; pp 225–234.
- (40) Dubbs, A. Test Set Sizing Via Random Matrix Theory. **2021**, arXiv:2112.05977v4.
- (41) Afendras, G.; Markatou, M. Optimality of Training/Test Size and Resampling Effectiveness in Cross-Validation. *J. Stat. Plann. Inference* **2019**, *199*, 286–301.
- (42) Nguyen, Q. H.; Ly, H.-B.; Ho, L. S.; Al-Ansari, N.; Van Le, H.; Tran, V. Q.; Prakash, I.; Pham, B. T. Influence of Data Splitting on Performance of Machine Learning Models in Prediction of Shear Strength of Soil. *Math. Probl. Eng.* **2021**, *2021*, 1–15.
- (43) Pham, B. T.; Prakash, I.; Jaafari, A.; Bui, D. T. Spatial Prediction of Rainfall-Induced Landslides Using Aggregating One-Dependence Estimators Classifier. *J. Indian Soc. Rem. Sens.* **2018**, *46* (9), 1457–1470.
- (44) Dobbin, K. K.; Simon, R. M. Optimally Splitting Cases for Training and Testing High Dimensional Classifiers. *BMC Med. Genom.* **2011**, *4* (1), 31–38.
- (45) Joseph, V. R. Optimal Ratio for Data Splitting. *Stat. Anal. Data Min.* **2022**, *15* (4), 531–538.
- (46) Rajappan, R.; Shingade, P. D.; Natarajan, R.; Jayaraman, V. K. Quantitative Structure-Property Relationship (QSPR) Prediction of Liquid Viscosities of Pure Organic Compounds Employing Random Forest Regression. *Ind. Eng. Chem. Res.* **2009**, *48* (21), 9708–9712.
- (47) Wang, J.; Song, Z.; Chen, L.; Xu, T.; Deng, L.; Qi, Z. Prediction of CO₂ Solubility in Deep Eutectic Solvents Using Random Forest Model Based on COSMO-RS-Derived Descriptors. *Green Chem. Eng.* **2021**, *2* (4), 431–440.
- (48) Chen, J.; De Hoogh, K.; Gulliver, J.; Hoffmann, B.; Hertel, O.; Ketzel, M.; Weinmayr, G.; Bauwelinck, M.; Van Donkelaar, A.; Hvidtfeldt, U. A.; Atkinson, R.; Janssen, N. A. H.; Martin, R. V.; Samoli, E.; Andersen, Z. J.; Oftedal, B. M.; Stafoggia, M.; Bellander, T.; Strak, M.; Wolf, K.; Vienneau, D.; Brunekreef, B.; Hoek, G. Development of Europe-Wide Models for Particle Elemental Composition Using Supervised Linear Regression and Random Forest. *Environ. Sci. Technol.* **2020**, *54* (24), 15698–15709.
- (49) Teh, S. K.; Zheng, W.; Lau, D. P.; Huang, Z. Spectroscopic Diagnosis of Laryngeal Carcinoma Using Near-Infrared Raman

Spectroscopy and Random Recursive Partitioning Ensemble Techniques. *Analyst* **2009**, *134* (6), 1232–1239.

(50) Lin, X.; Sun, L.; Li, Y.; Guo, Z.; Li, Y.; Zhong, K.; Wang, Q.; Lu, X.; Yang, Y.; Xu, G. A Random Forest of Combined Features in the Classification of Cut Tobacco Based on Gas Chromatography Fingerprinting. *Talanta* **2010**, *82* (4), 1571–1575.

(51) Ricardo, F.; Ruiz-Puentes, P.; Reyes, L. H.; Cruz, J. C.; Alvarez, O.; Pradilla, D. Estimation and Prediction of the Air-Water Interfacial Tension in Conventional and Peptide Surface-Active Agents by Random Forest Regression. *Chem. Eng. Sci.* **2023**, *265*, 118208.

(52) De Miranda Ramos Soares, A. P.; De Oliveira Carvalho, F.; De Farias Silva, C. E.; da Silva Gonçalves, A. H.; De Souza Abud, A. K. Random Forest as a Promising Application to Predict Basic-Dye Biosorption Process Using Orange Waste. *J. Environ. Chem. Eng.* **2020**, *8* (4), 103952.

(53) Shrivastava, R.; Mahalingam, H.; Dutta, N. N. Application and Evaluation of Random Forest Classifier Technique for Fault Detection in Bioreactor Operation. *Chem. Eng. Commun.* **2017**, *204* (5), 591–598.

(54) Hastie, T.; Tibshirani, R.; Friedman, J. *Elements of Statistical Learning*; Springer Series in Statistics; Springer: New York, NY, 2001.

(55) Olu-Ajayi, R.; Alaka, H.; Sulaimon, I.; Sunmola, F.; Ajayi, S. Building Energy Consumption Prediction for Residential Buildings Using Deep Learning and Other Machine Learning Techniques. *J. Build. Eng.* **2022**, *45*, 103406.

(56) Zhang, W.; Li, J.; Liu, T.; Leng, S.; Yang, L.; Peng, H.; Jiang, S.; Zhou, W.; Leng, L.; Li, H. Machine Learning Prediction and Optimization of Bio-Oil Production from Hydrothermal Liquefaction of Algae. *Bioresour. Technol.* **2021**, *342*, 126011.

(57) Gong, H.; Sun, Y.; Huang, B. Gradient Boosted Models for Enhancing Fatigue Cracking Prediction in Mechanistic-Empirical Pavement Design Guide. *J. Transp. Eng. B: Pavements* **2019**, *145* (2), 04019014.

(58) Ll, M.; Baxter, J. *Boosting Algorithms as Gradient Descent in Function Space*; NIPS: New Orleans, LA, USA, 1991.

(59) Vapnik, V. N. *Statistical Learning Theory*; Wiley: New York, 1998.

(60) Nandi, S.; Badhe, Y.; Lonari, J.; Sridevi, U.; Rao, B. S.; Tambe, S. S.; Kulkarni, B. D. Hybrid Process Modeling and Optimization Strategies Integrating Neural Networks/Support Vector Regression and Genetic Algorithms: Study of Benzene Isopropylation on Hbeta Catalyst. *Chem. Eng. J.* **2004**, *97* (2–3), 115–129.

(61) Sivaramakrishnan, K.; Nie, J.; de Klerk, A.; Prasad, V. Least Squares-Support Vector Regression for Determining Product Concentrations in Acid-Catalyzed Propylene Oligomerization. *Ind. Eng. Chem. Res.* **2018**, *57*, 13156–13176.

(62) Mesbah, M.; Soroush, E.; Rezakazemi, M. Development of a Least Squares Support Vector Machine Model for Prediction of Natural Gas Hydrate Formation Temperature. *Chin. J. Chem. Eng.* **2017**, *25* (9), 1238–1248.

(63) Ahmad Yasmin, N. S.; Abdul Wahab, N.; Ismail, F. S.; Musa, M. J.; Halim, M. H. A.; Anuar, A. N. Support Vector Regression Modelling of an Aerobic Granular Sludge in Sequential Batch Reactor. *Membranes* **2021**, *11* (8), 554.

(64) Lahiri, S. K.; Ghanta, K. C. Prediction of Pressure Drop of Slurry Flow in Pipeline by Hybrid Support Vector Regression and Genetic Algorithm Model. *Chin. J. Chem. Eng.* **2008**, *16* (6), 841–848.

(65) Jiang, J. L.; Su, X.; Zhang, H.; Zhang, X. H.; Yuan, Y. J. A Novel Approach to Active Compounds Identification Based on Support Vector Regression Model and Mean Impact Value. *Chem. Biol. Drug Des.* **2013**, *81* (5), 650–657.

(66) Cortes, C.; Vapnik, V. Support Vector Networks. *Mach. Learn.* **1995**, *20*, 273–297.

(67) Gunn, S. R. *Support Vector Machines for Classification and Regression*; School of Electronics and Computer Science, University of Southampton, 1998.

(68) Chauchard, F.; Cogdill, R.; Roussel, S.; Roger, J. M.; Bellon-Maurel, V. Application of LS-SVM to Non-Linear Phenomena in NIR Spectroscopy: Development of a Robust and Portable Sensor for Acidity Prediction in Grapes. *Chemom. Intell. Lab. Syst.* **2004**, *71*, 141–150.

(69) Schölkopf, B.; Smola, A. J. *Learning with Kernels: Support Vector Machines, Regularization, Optimization, and Beyond*; MIT Press: Cambridge, MA, 2002.

(70) Yeganeh, B.; Motlagh, M. S. P.; Rashidi, Y.; Kamalan, H. Prediction of CO Concentrations Based on a Hybrid Partial Least Square and Support Vector Machine Model. *Atmos. Environ.* **2012**, *55*, 357–365.

(71) Brereton, R. G.; Lloyd, G. R. Support Vector Machines for Classification and Regression. *Analyst* **2010**, *135* (2), 230–267.

(72) Siddiquee, M. N.; Sivaramakrishnan, K.; Wu, Y.; de Klerk, A.; Nazemifard, N. A Statistical Approach Dealing with Multicollinearity among Predictors in Microfluidic Reactor Operation to Control Liquid-Phase Oxidation Selectivity. *React. Chem. Eng.* **2018**, *3*, 972–990.

(73) Stähle, L.; Wold, S. Partial Least Squares Analysis with Cross-Validation for the Two-Class Problem: A Monte Carlo Study. *J. Chemom.* **1987**, *1*, 185–196.

(74) Wold, S.; Sjöström, M.; Eriksson, L. PLS-Regression: A Basic Tool of Chemometrics. *Chemom. Intell. Lab. Syst.* **2001**, *58*, 109–130.

(75) Martens, H.; Naes, T. Multivariate Calibration. *Chemometrics*; Springer Netherlands: Dordrecht, 1984; pp 147–156.

(76) Sjöström, M.; Wold, S.; Lindberg, W.; Persson, J.-Å.; Martens, H. A Multivariate Calibration Problem in Analytical Chemistry Solved by Partial Least-Squares Models in Latent Variables. *Anal. Chim. Acta* **1983**, *150*, 61–70.

(77) Siddiquee, M. N.; Sivaramakrishnan, K.; Wu, Y.; de Klerk, A.; Nazemifard, N. A Statistical Approach Dealing with Multicollinearity among Predictors in Microfluidic Reactor Operation to Control Liquid-Phase Oxidation Selectivity. *React. Chem. Eng.* **2018**, *3* (6), 972–990.

(78) Mbaabu, O. Introduction to Random Forest in Machine Learning. <https://www.section.io/engineering-education/introduction-to-random-forest-in-machine-learning/#:~:text=Advantages%20of%20random%20forest,over%20the%20decision%20tree%20algorithm> (accessed on 2023-10-15).

(79) Janet, J. P.; Chan, L.; Kulik, H. J. Accelerating Chemical Discovery with Machine Learning: Simulated Evolution of Spin Crossover Complexes with an Artificial Neural Network. *J. Phys. Chem. Lett.* **2018**, *9* (5), 1064–1071.

(80) Goldsmith, B. R.; Esterhuizen, J.; Liu, J. X.; Bartel, C. J.; Sutton, C. Machine Learning for Heterogeneous Catalyst Design and Discovery. *AIChE J.* **2018**, *64*, 2311–2323.