Asian Journal of Research in Computer Science

4(4): 1-14, 2019; Article no.AJRCOS.53584
ISSN: 2581-8260

Novel Approaches to Modelling Flammability Characteristics of Polymethyl Methacrylate (PMMA) via Multivariate Adaptive Regression Splines and Random Forest Methods

Mohammed Okoe Alhassan[1*], Lin Jiang[1], Rhoda Afriyie Mensah[1], Qiang Xu[1] and Michael Boakye Osei[2]

[1]School of Mechanical Engineering, Nanjing University of Science and Technology, 210094 Nanjing, China.
[2]School of Computer Science and Technology, Nanjing University of Science and Technology, 210094 Nanjing, China.

Authors' contributions

This work was carried out in collaboration between all authors. Author MOA designed the study, managed the literature searches and wrote the first draft of the manuscript. Authors LJ and QU supervised the work. Author RAM managed the data collection and analysis. Author MBO performed and analyzed the proposed models. All authors read and approved the final manuscript.

Article Information

DOI: 10.9734/AJRCOS/2019/v4i430121
Editor(s):
(1) Dr. Stephen Mugisha Akandwanaho, Department of Information Systems and Technology, University of KwaZulu-Natal, South Africa.
Reviewers:
(1) Tabe Franklin Nyenty, Yaoundé I University, Cameroon.
(2) Ananda Kumar Karunakaran, Swamy Vivekanandha College of Nursing and Pharmacy Namakkal, India.
(3) David Lizcano, Madrid Open University, Spain.
Complete Peer review History: http://www.sdiarticle4.com/review-history/53584

Original Research Article

Received 26 October 2019
Accepted 31 December 2019
Published 07 January 2020

ABSTRACT

Soft-computing techniques for fire safety parameter predictions in flammability studies are essential for describing a material fire behaviour. This study proposed, two novel Artificial Intelligence developed models, Multivariate Adaptive Regression Splines (MARS) and Random Forest (RF) methods, to model and predict peak heat release rate (pHRR) of Polymethyl methacrylate (PMMA) from Microscale Combustion Calorimetry (MCC) experiment. From the statistical analysis, MARS presented the highest coefficient of determination (R^2) values of (0.9998) and (0.9996) for training and testing respectively, with low MAD, MAPE and RMSE values. Comparatively, MARS

**Corresponding author: E-mail: mohaokoe2020@njust.edu.cn;*

outperformed RF in the predictions of pHRR, through its model algorithms that generated optimized equations for pHRR predictions, covering all non-linearity points of the experimental data. Amongst the input variables (sample mass, THR, HRC, pTemp and pTime), heating rate (β), highly influenced pHRR outcome predictions from MARS and RF models. However, to validate the performance and applicability of the proposed models. Results of MARS and RF were benchmarked with that from Artificial Neural Network (ANN) methods. The MARS and RF models observed the least error deviation when compared with pHRR results for PMMA from the ANN models. This study therefore, recommends the adoption of MARS and RF in the predictions of flammability characteristics of polymeric materials.

Keywords: Flammability; Multivariate Adaptive Regression Splines (MARS); Random Forest (RF); Microscale Combustion Calorimetry (MCC); PMMA.

1. INTRODUCTION

Polymethyl methacrylate (PMMA) is an amorphous molding transparent thermoplastic material with excellent physical and chemical properties. These properties allow for its processing and application for several domestic, industrial and aviation purposes. Amid other materials, PMMA is one of the most successfully explored and evaluated polymers by the thermal and fire safety industries when estimating a material fire behaviour, making it a benchmark for comparison with other materials. However, it is very flammable. That is the reason why continues improvements of its thermal and combustion stability with advanced fire modelling techniques is still important [1,2]. So far, various standard fire tests or calorimetry techniques from small-scale to large-scale to measure the flammability properties and parameters of PMMA materials have been developed and used [1]. Largely, PMMA flammability properties and parameters are quantified with Microscale combustion calorimeter (MCC) apparatus [1,3,4].

The microscale combustion calorimetry (MCC), also termed as pyrolysis combustion flow calorimetry (PCFC) is a small-scale device that uses milli-gram sized mass samples with practices of analytical pyrolysis, combustion gas analysis, and flow calorimetry [5]. From the combustion outcomes, innumerable flammability characteristics and material properties are measured with ASTM D7309-13 standards, which includes, heat release rate (HRR), peak pyrolysis temperature (Tp), total heat released (THR), and heat release capacity (HRC) [6,7,2,8]. Several researchers [9-12,13] have performed and used MCC experimental results to screen and evaluate the flammability of pure polymers. It is worth affirming that the database obtained from MCC experiments can be modelled to predict the flammability and fire safety parameters with developed estimation methods.

Previous researches from literature works used empirical methods with chemistry-based models like additive molar group contribution methods, quantitative structure-property relationship (QSPR), and quantitative structure-activity relationship (QSAR) to estimate flammability properties from measured MCC experiments, with good and acceptable results [14-16]. However, these chemistry-based models required the material molecular moiety/ functional group and chemical structure to be available in the database to estimate a material property. This presented Limitations for measuring material properties, as materials without functional groups or chemical structures in their database could not be estimated [17]. In addition, the models were not able to fully account for the intrinsically non-linear and complex nature of experimental datasets. This presented large errors in the model predictions [17,18]. This created huge gabs in fire parameter predictions that needed to be solved.

In recently times, novel soft-computing systems with Artificial Intelligence (AI) approaches like Artificial Neural Networks (ANN) methods, Support Vector Machines (SVM) systems etc., with machine learning algorithms have been used to model and predict thermal and flammability characteristics of materials, gaining significance importance with low marginal prediction errors [17,18]. One significant advantage of this AI methods is that, they do not require a materials molecular moiety or functional group to be available in the database before evaluating the material properties.

Several successful research works in literature has reported excellent results achieved using soft-computing techniques for parameter

predictions in thermal, combustion and flammability studies. Asante-Okyere, et al. [17] modelled and predicted flammability characteristics and parameters measured from microscale combustion calorimetry (MCC) of polymethyl methacrylate (PMMA) using generalized regression neural Networks (GRNN) and feed forward back propagation Neural networks (FFBPNN). In their study, it was observed that GRNN outperformed FFBP in predicting HRC datasets. Similarly, GRNN attained better THR, pTemp and pHRR predictions during training but generated a relatively poor correlation when estimating the testing data. Also, in the work of Burgaz, et al. [19] Artificial Neural Networks (ANN) techniques were used to model and estimate the thermal stability, crystallinity and thermomechanical properties of poly (ethylene-oxide)/clay nanocomposites. They obtained very good parameter predictions with minimal prediction errors. Hybrid based methods of Genetic Function Approximation and Least Squares Support Vector Machine were applied by Ma, et al. [20] to Predict the ecotoxicity of ionic liquids towards Vibrio fischeri under different temperatures. In addition, comparative evaluation research works by Mensah, et al. [19] undertaken, when they used Group Method of Data Handling Neural networks (GMDHNN) to predict the flammability characteristics of extruded polystyrene from microscale combustion calorimetry (MCC) experiments. With reference from the current-related research works on flammability and thermal analysis studies reviewed, it can be stated that ANN methods have been the most successfully and largely used, making it a benchmark for comparison with other techniques. However, problems of overfitting and slow model training time exist, largely due to the iterative tuning parameters of the models, and the training algorithms used [17,21,19].

It is worth noting that, Multivariate Adaptive Regression Splines (MARS), is a novel soft-computing technique, introduced by Friedman [21], has positioned itself as one of the most significant artificial intelligence (AI) methods in recent times. An advancement on traditional multiple linear regression approach for solving regression function approximation problems with the main purpose to predict the values of a continuous dependent or outcome variable from a set of independent or predictor variables of complex systems. A piecewise basis function with non-linearity between dependent and independent variables. Advantages includes the addition of relative importance of the independent variables in the prediction of the dependent variable in an easily interpretable manner. The occurrence of split sub-regions and discontinuity of the approximating function at the boundaries of the intervals as binary recursive partitioning are usually eliminated. Hence, avoids overfitting and flexible but robust in predictions [22]. This algorithm has been applied to solve several engineering problems, for parameter predictions, including Thermal, Geosciences engineering, and combustion studies. Ziggah and Laari [23] successfully applied MARS technique in their study to transform two-dimensional (2D) coordinates between two nationals geodetic datums (Accra 1929 and Leigon 1977) used in Ghana. They further observed that the MARS technique has the ability to overcome the “black box” nature of artificial neural network (ANN) by given out functional relationship for predicting the parameters output. In addition, Qing-Song et al. [24], adopted two-step multivariate adaptive regression splines in their study to model and evaluate the quantitative relationship between gas chromatography retention indices and the molecular descriptors in the database. Amongst other works done in literature using MARS model for thermal and combustion studies include; Balshi, et al. [25], Durmaz, et al. [26] and Roy, et al. [27].

Another novel non-parametric regression model, Random Forest (RF), is an ensemble learning technique introduced by Breiman [28], as a combination of tree predictors that are robust and hardly overfits; yielding high accurate predictions. This method is widely used as an enhanced prediction model providing variable importance measures that recognizes significant predictor variables. RF encompasses the advantages of low bias, ease of interpretation of variables in decision trees. Overfitting problem in decision trees are avoided by averaging the outcomes across different decision trees [29]. This model has been used to successfully predict and estimate parameters for thermal and combustion analysis applications. Palmer, et al., predicted aqueous solubility using random forest methods [30]. An estimation of coal gross calorific values has been analyzed based on various RF methods [31]. Furthermore, Random Forest, have been applied as a Classification and Regression Tool for Compound Classification and QSAR Modeling tool for compound classification with QSAR modeling [32].

Henceforth, literature works involving RF and MARS demonstrate novelty approaches and are data integrated algorithms that identifies features contributing most to a model prediction using importance input variable contributions. In spite of the novelty of MARS and RF model from literature works, there is no comparative study using both modelling techniques to estimate the flammability characteristics of PMMA from MCC parameters. This present research therefore explores the alternatives to model and predict flammability characteristics and parameters measured from MCC experiments of polymethyl methacrylate (PMMA with MARS and RF models by comparing their predictability performance. Furthermore, the proposed MARS and RF model performance results will be compared with results of predicted pHRR from ANN methods [17] to validate MARS and RF applicability and reliability for engineering parameter predictions. The paper is structured as follows: Firstly, the experimental sample material, procedures, and methods used in this study are described. Secondly, we illustrate and discuss conducted MCC experimental results and obtained MARS-RF model results. Finally, the main conclusions drawn from the results are discussed.

2. MATERIALS AND METHODS

2.1 Materials

The referenced PMMA material and related properties are listed in Table 1. The pellet PMMA material was Milled into powdered form before it was used for the MCC experiment.

2.2 Microscale Combustion Calorimetry (MCC) Measurement

The MCC-2 device [33], was used with guidance from ASTM D7309-13 and "Method A procedure [6,4]. Nine (9) constant heating rates, ranging from 0.1-3.5Ks^{-1} was considered for this experiment. However, 31 sample masses of PMMA in the range of (1-3.5 mg) was also used with 3-groupings. Heating of PMMA samples were carried-out and the temperatures of the Pyrolyzer and combustor was 75-600ºC and 900 ^{0}C, respectively. Peak heat release rate (pHRR), heat release capacity (HRC), peak heat release temperature (pTemp), peak heat release time (pTime) and total heat released (THR) were measured and recorded using "Method A" procedure [6].

2.3 Random Forest (RF)

This algorithm works by growing an ensemble of decorrelated decision trees under bagging technique, based on binary recursive partitioning with a number of bootstrap samples (ntrees), producing a regression tree in a modified operation. Subsequently, random sampling of the numerous predictors (mtry) begin and the algorithm chooses the best split from among those sampled variables rather than considering all variables. Thus, Random Forest is useful in handling thousands of input variables without input variable deletion, with assigned percentage variable importance [34]. Specification of model parameters in RF model has less influence on model output obtaining a default (m_{try}) value as the square root of the total number of variables [35].

Therefore, the number of trees (ntree) needs to be set sufficiently high to avoid overfitting but a limited generalization error is generally produced. The final predictions are presented as mean value of individual predictions obtained by each decision tree.

The algorithm works by considering a dataset, *D* with *n* number of variables, $D = \{D_1, \ldots, D_n\}$ with $D_i = \{x_i, y_i\}$, given an independent test case D_0 with operator x_0, random forests (RF) regression model can be deduced as follows [28,36]:

i. The data set *D* is sampled with replacement to generate bootstrap resamples $E_1, \ldots, E_M$
ii. For each resample $E_m, m = 1, \ldots, M$, grow a regression tree T_m.
iii. In predicting the test case D_0 with covariate x_0, the combined results given by individual trees gives the predicted value obtained by the whole RF.

Let's Assume $\hat{f}_m^*(x_0)$ denote the prediction of D_0 by *mth* tree, the random forest prediction for regression problems is expressed [37] as:

$$\hat{f}_m^*(x_0) = \frac{1}{M}\sum_{m=1}^{M}\hat{f}_m^*(x_0) \quad (1)$$

For each bootstrap sample taken from the training data, there will be samples left behind

that were not included. These samples are called Out-of-Bag samples or "OOB". Considering a default $m_{try} = \frac{p}{3}$ (where *p* is the number of predictor variables*)* an estimate of the error rate can be obtained, based on the training data, through the following steps to [35];

i. At each bootstrap iteration, the data not in the bootstrap sample is predicted ("out-of-bag", or OOB, data) using the tree grown with the bootstrap sample.
ii. Aggregate of the OOB predictions error rate calculated, and termed "OOB" estimate of error rate. The "mean of squared residuals" is computed as Equation. (2);

$$MSE_{OOB} = n^{-1}\sum_{1}^{n}\{y_i - \hat{y}_i^{OOB}\}^2 \quad (2)$$

Where $\hat{y}_i^{OOB}$ is the average of the OOB predictions for *ith* observations "Percent variance'' as; $1-\frac{MSE_{OOB}}{\hat{\sigma}_y^2}$, where $\hat{\sigma}_y^2$ is computed with *n* as divisor (instead of n-*1*). These performance measures are reliable test error estimate and correlate well with cross validation estimates.

We operated RF with default settings for the number of variables used at each split with an m_{try} of 2 to develop the models.

2.4 Multivariate Adaptive Regression Splines (MARS)

MARS is a non-parametric statistical regression modelling technique for fitting the relationship between a set of input and target variables of systems. No specific assumption about the underlying functional relationship between the input variables and the output is required. A piecewise curve also known as basis functions (BF's) provides superior flexibility to the model's curves with better thresholds, and other linear modelling functions [25]. The MARS generated BF's can then be searched in the stepwise manner. MARS models are built in a two-phase (forward and backward) procedure. In forward phase, it usually results in an overfit model, adds functions and finds potential knots or the forward stepping threshold (*d)*, to improve the performance. Backward phase involves pruning or deleting the least effective terms (N_{prune}). The MARS model is mathematically presented as Equations. (3), [24,38].

$$y = c_0 + \sum_{i=1}^{N} c_i \prod_{j=1}^{K_i} b_{ji}(x_{v(j,i)}) \quad (3)$$

Where; *y* is the output target variable, c_0 is a constant, c_i is the vector coefficients of the non-linear basis functions, $b_{ji}(x_{v(j,i)})$ is the truncated power basis function with $v(j,i)$as the index of the independent variable used in the ith term of the jth product. K_i is a parameter that limits the order of interactions. In this study, the input variables that were considered are mass, heating rate (β), HRC, THR, pTemp, and pTime, and the output variable considered was pHRR.

$$y = pHRR,\ x = (m,\ \beta, \mathrm{HRC},\ \mathrm{THR},\ \mathrm{pTemp},\ \mathrm{pTime})$$

The spline b_{ji} can further be defined as:

$$b_{ji}(x) = (x - t_{ji})_+^q = \begin{cases} (x-t_{ji})^q & if\ \ x < t_{ji} \\ 0 & Otherwise \end{cases} \quad (4)$$

$$b_{ji+1}(x) = (t_{ji} - x)_+^q = \begin{cases} (t_{ji}-x)^q & if\ \ x < t_{ji} \\ 0 & Otherwise \end{cases} \quad (5)$$

Where; t_{ji} represents the "knots" of the spline, $q(q \geq 0)$ denotes the power of the splines in the smoothness degree of the resultant function approximation. The splines are linear functions, when *q = 1*. In the forward step, the BF's selection conforms to Eq. (3), while in the backward procedure, the least contributive BF's are deleted based on the generalized cross-validation (GCV) criterion. The GCV criterion is an adjusted residual sum of squares, expressed as in Equation. (6) [21,38].

$$GCV = \frac{1/N \sum_{i=1}^{N}[y_i - f(x_i)]^2}{\left[1 - \frac{M + \partial(M-1)/2}{N}\right]^2} \quad (6)$$

where *M* is the number of non-constant BF's, ∂ is the penalizing parameter (a default value of 3 is assigned to ∂), the higher the cost of ∂ approaches, the more basis functions will be

eliminated. ∂ is practically only increased during pruning process in order to obtain smaller models. *N* is the maximum number of observations ($N_{\max}$), y_i is the *ith* measured element and $f(x_i)$ denotes the *ith* predicted value of the model. The numerator is the mean squared error of the evaluated model in the training data, penalized by the denominator. The denominator accounts for the increasing variance when the model complexity augments. $(M-1)/2$ is the number of hinge function knots. Minimization of Eq. (3) is done as a basis function is removed from each deletion step until an adequately fitted model is found. BF's and the variable knot locations are data driven procedures and specific to the problem at hand.

Table 1. Properties of "black" PMMA

Property	Value	
Molecular formula	$([C_5H_8O_2]_n)$	[4]
Density	1180 kg/m^3	[4,17]
Thermal conductivity	0.185 W/Mk	[4,17]
Specific heat	1.510 J/gK	[4,17]

2.5 Model Performance Indicators

In this study, various statistical indices were computed to assess the performance of the developed MARS and RF models including coefficient of determination (R^2), mean absolute deviation (MAD), mean absolute percentage error (MAPE), and root mean squared error (RMSE), and are mathematically expressed in Eq. (7-10) respectively. Generally, with (R^2) values close to 1 indicates high degree of similarity between predicted and measured parameter values. Low MAPE, MAD and RMSE values indicate high confidence in model-predicted values. The MARS and RF models were developed using Salford Predictive Modeler (SPM) [39], Software Suite (version 8.3, Salford Systems, San Diego, CA) which can be accessed online.

$$R^2 = \left(\frac{\sum_{i=1}^{n}(N_i-\bar{N})(P_i-\bar{P})}{\sqrt{\sum_{i=1}^{n}(N_i-\bar{N})^2 \times \sqrt{\sum_{i=1}^{n}(P_i-\bar{P})^2}} \right)^2 \quad (7)$$

$$MAD = \frac{\sum_{i=1}^{n}|N_i-P_i|}{n} \quad (8)$$

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\frac{|N_i-P_i|}{N_i} \times 100 \quad (9)$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(N_i-P_i)^2} \quad (10)$$

Given,

$$\bar{N} = \frac{1}{n}\sum_{i=1}^{n} N_i \quad (11)$$

Where *n* is the total sample number, *p* is the total number of regressors in the training model, N_i is the measured parameter value, P_i is the predicted parameter value, $\bar{N}$ is the mean measured parameter value and $\bar{P}$ is the mean predicted parameter value [40].

3. RESULTS AND DISCUSSION

3.1 MCC Experimental Results

It is remarkable to know that under forced flaming combustion, heat release rate (HRR) is the principal indicator of fire hazard, therefore it is easy to obtain the fire hazard from measuring the HRR of a material [41]. Fig. 1, demonstrates a plot of specific heat release rate against temperatures of MCC test results from PMMA. Indications from Fig. 1 shows an increase in heat releasing rate as temperature increases [4,13,17].

The statistical description of the obtained MCC parameters from the measured experimental data of thirty-one (31) sample masses ranging from (1-3.5 mg) with the nine (9) corresponding heating rates considered for this study are summarized and listed in Table 2, showing the minimum, maximum, mean and standard deviation values. These results are to show the measure of dispersion of the parameters determined around its mean values.

3.2 Performance of MARS Model

The main objective was to develop an optimized MARS model that can predict flammability characteristics of PMMA from MCC parameters. These parameters (mass, β, THR, HRC, pTemp, pTime) were considered as inputs variables to predict peak heat release rate (pHRR) as the output variable. Out of the total dataset

(31samples), Six (6) sample masses representing (20%) were used to test the generalization capability of model. The remaining 80% (25 samples) were used to train the datasets for accuracy. MARS model is generally affected by various parameters. Parameters were then varied within limits to obtain the required number of combinations to build and train the model. This includes; the max. degree of self-interactions, Max functions, threshold, prune, generalized cross-validation penalty per knot, as given in Table 3 [6].

To build the MARS model, 10 basis functions were selected during the forward phase stepwise knot placement. Subsequently, a model generated piecewise linear GCV of 647.25529 was used to prune and eliminate the least significance basis functions by deletion, during the backward phase. For the prediction of pHRR, piecewise interaction produced the lowest mean squared errors (MSE) in both training and testing stage as shown in Fig. 2. The variation between the predicted pHRR and the model data input variables indicates better performance during the training of the model as presented in Fig. 3.

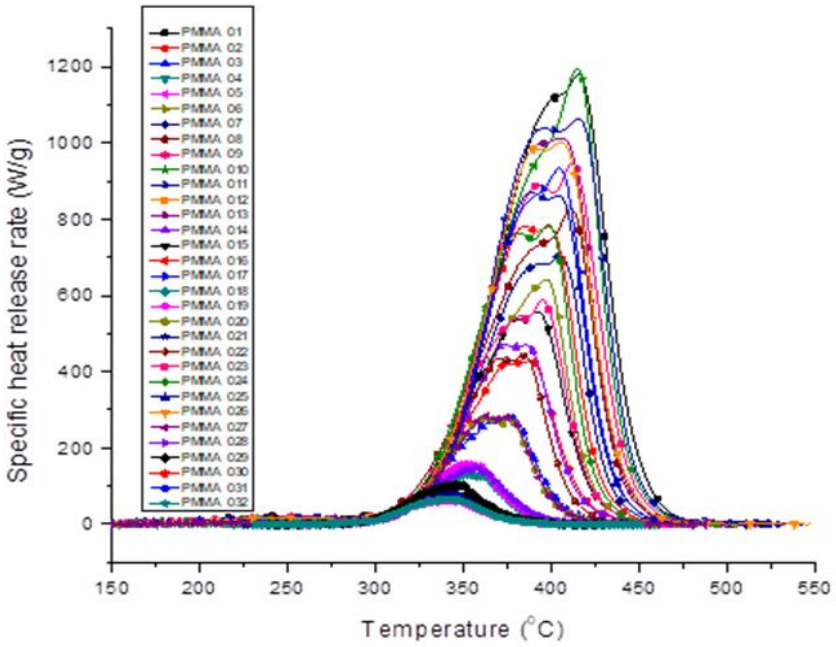


Fig. 1. Curves of plotted specific heat release rate versus temperature from MCC experiment of PMMA

Table 2. Statistical description of the results obtained from the MCC experimental datasets

	pHRR (W/g)	HRC (J/g °C)	THR (kJ/g)	pTemp (°C)	pTime (s)
Mean	639.9	506.8	26.6	376.2	607.1
S. D	418.5	164.4	5.9	21.1	774.9
Max	1278.8	918.5	31.8	406.9	2479.2
Min	53.8	320.3	10.2	338.1	120.7

Table 3. MARS model parameters

Parameters	Values
Max functions $10e^4$	10 – 40
Generalized cross-validation penalty per knot	0, 2 – 4
Self-interactions	No
Max interactions	2 – 4
Threshold	$1.0000e^{-4}$
Prune	Yes

The statistical results for predicting pHRR shows significant performance predictability, observing high coefficient of determination (R^2) value of 0.9998 and 0.9996 during training and testing respectively, with low and insignificant RMSE, MAD and MAPE values as listed in Table 4. The results from Table 4, signify high confidence in the model-predicted values agreeing well-with the experimental results, this indicates that the model is good and can generate reliable predictions. The relative important input variables in predicting pHRR are presented in decreasing order of contributions with plotted evidence of this assertion in Figs. (4 (a), (b) and (c). Out of the 6 input variables considered only three (3) had significant impact in the models' prediction. Heating rate (β) was observed to be the input variable with the most significant impact on the model performance. This confirms works done in literature by Q. Xu, et al. [4,13].

The final optimized MARS model for predicting pHRR consisting of 7 generated basis functions after the GCV's least basis functions (BF's) deletion. Related equations are presented in Table 5. This "basis functions" with their respective hinge functions generated covered all points of non-linearity of the experimental data [40]. The optimized model equation is the weighted sum of the "basis functions" as expressed in Equation. (12). A regression equation (13) for predicting pHRR of PMMA was generated as a function of (β, HRC, pTemp) from Table 5, Eq. (3). In using Equations. (13), pHRR can easily be predicted.

Y = 1216.14 + 383.221 * BF1 - 463.686 * BF2 + 1.56806 * BF3 - 2.53986 * BF4 - 2.38165 * BF6 - 1.4712 * BF7 - 2.85618 * BF9; (12)

$$\begin{aligned} pHRR &= 1216.14 + 383.221\,(\beta - 2) \\ &- 463.686\,(2 - \beta) + 1.56806\,(HRC - 534.2) \\ &- 2.53986\,(534.2 - HRC) \\ &- 2.38165\,(385.3 - pTemp) \\ &- 1.4712(HRC - 416.5) - \\ &2.85618\,(pTemp - 367.3) \end{aligned} \quad (13)$$

3.3 Performance of Random Forest (RF)

The random forest (RF) model was developed by setting the number of trees to be built at 200 and number of predictors for each node at 3 with minimum non-terminal node size of 5. The output variable selected was (pHRR), while the input variables considered were mass, β, THR, HRC, pTemp, pTime. 80% of the total dataset (31) representing 25 samples were used to train (OOB) the model and the remaining 20% partitioned to test the model accuracy. In predicting the pHRR, lowest MSE values were attained against number of ensemble trees for both training and testing functions for the model as demonstrated in Fig. 5. The statistical performance evaluation results from the model presents low error of RMSE, MAPE and MAD with high coefficient of determination (R^2) value of 0.9749 and 0.9562 during training (OOB) and testing respectively, signifying better predictive ability of well-fitted RF model as presented in Table 6. Unlike MARS, in Random Forest (RF) all the 6 input variables with their respective contributions in building the model are presented. Heating rate (β) was observed to had more substantial contributive effect to the models' performance with graphical evidence in Fig. 6.

Table 4. Statistical performance of MARS model in predicting pHRR

MCC parameter	pHRR			
Performance indicators	**R^2**	**MAD**	**MAPE**	**RMSE**
Training	0.9998	4.3306	0.0076	4.9717
Testing	0.9996	7.2297	0.0189	8.7518

Table 5. The basis function (BF's) and their related equations in the pHRR MARS model

Basis function	Hinge function
BF1	max (0, HEATING_RATE - 2);
BF2	max (0, 2 - HEATING_RATE);
BF3	max (0, HRC - 534.2);
BF4	max (0, 534.2 - HRC);
BF6	max (0, 385.3 - PTEMP);
BF7	max (0, HRC - 416.5);
BF9	max (0, PTEMP - 367.3);

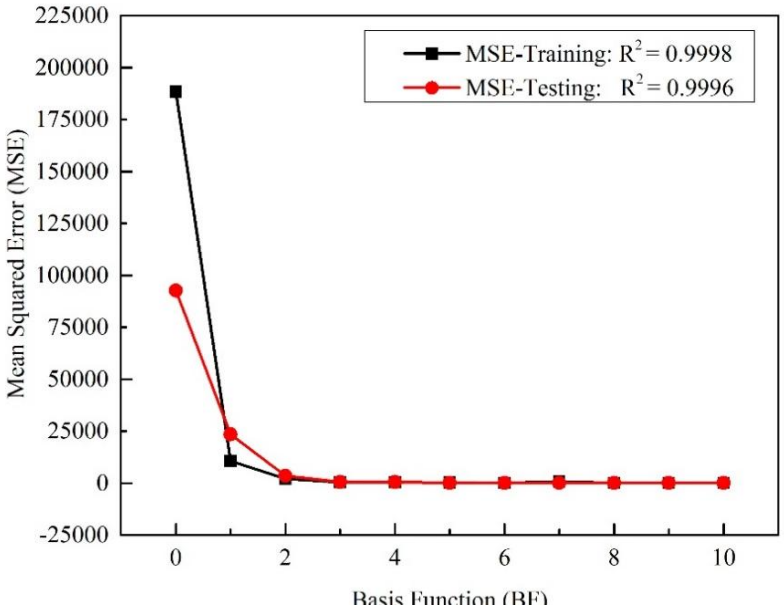


Fig. 2. Plots of MARS Mean Squared Error (MSE) versus basis function based on training and testing dataset

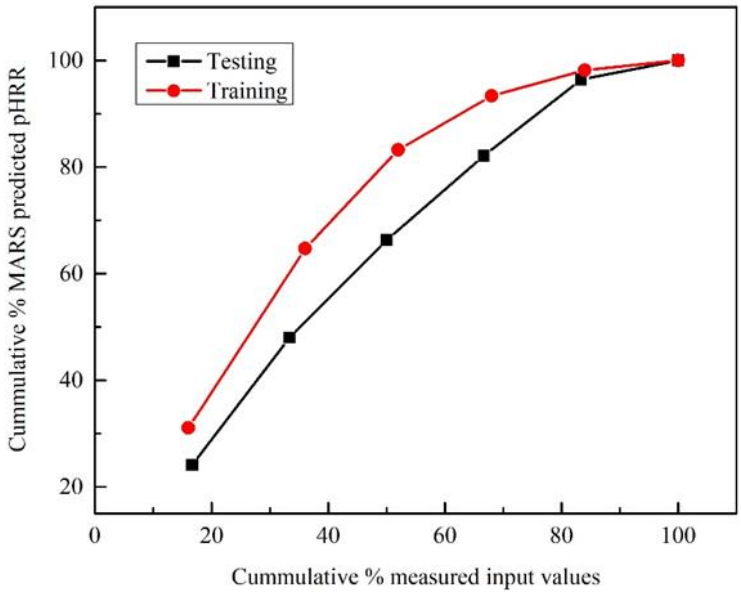
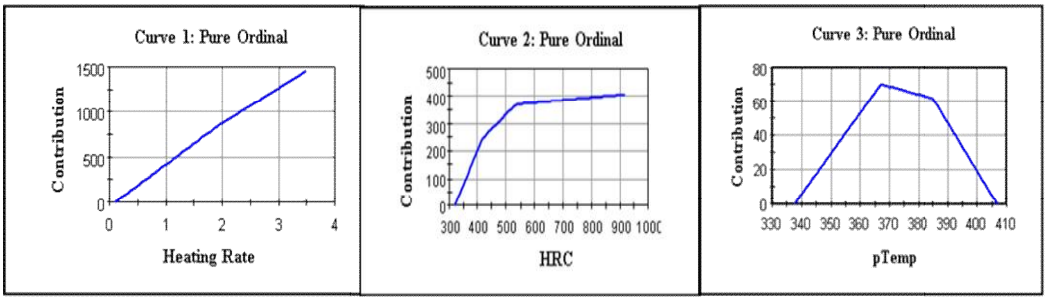


Fig. 3. % MARS output Prediction Versus. % measured input variables during training and testing

Table 6. Statistical performance of RF model in predicting pHRR

MCC parameter	pHRR			
Performance indicators	**R^2**	**MAD**	**MAPE**	**RMSE**
Training (OOB)	0.9749	51.7575	0.2395	73.1342
Testing	0.9562	64.8713	0.0927	76.9207



(a) (b) (c)

Fig. 4. Plots showing important contributions of (a) Heating rate (b) HRC and (c) pTemp

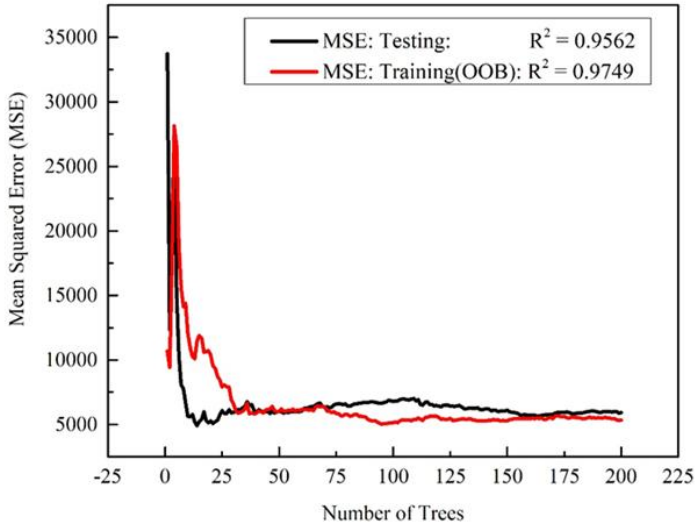


Fig. 5. Plots of MSE versus number of trees based on training and testing data for RF model

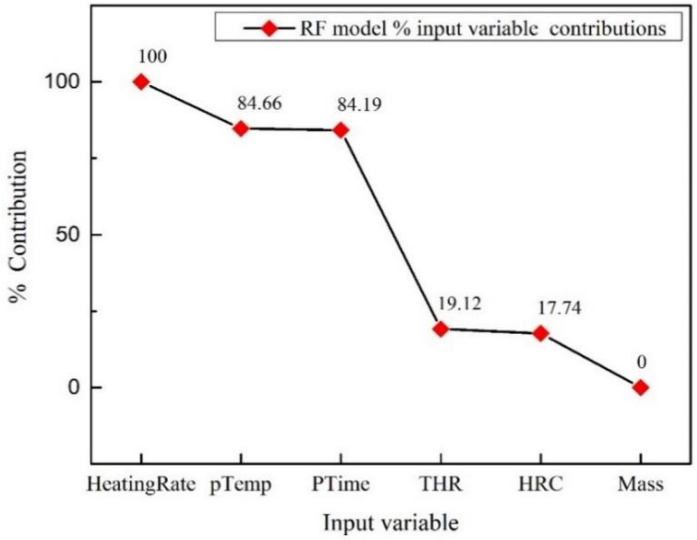


Fig. 6. Plot of input variables percentage contributions to the RF modelling process

Table 7. Comparison of the mean error deviation from MARS and RF and ANN models for predicting pHRR of PMMA

Model	Predicted pHRR (mean) for PMMA (W/g)	Measured pHRR (mean) for PMMA (W/g)	Error deviation
MARS	642.86	639.9	2.96
RF	642.93	639.9	3.03
FFBPNN [17]	643.95	639.9	3.05
GRNN [17]	642.04	639.9	3.14

3.4 pHRR Models Prediction Comparisons

Predicted pHRR mean error deviations observed from MARS and RF models where compared with Artificial Neural network (ANN) methods of feed forward back propagation (FFBP) and generalized regression neural network (GRNN) predicted mean pHRR results of PMMA from Asante-Okyere, et al. [17]. This was done to further validate MARS and RF models performances and applicability. It was observed that, the mean pHRR values of MARS and RF attained error deviations of 2.96 and 3.03 as compared to FFBPNN and GRNN that obtained 3.05 and 3.14 respectively for evaluating PMMA. The MARS and RF models comparatively showed less deviation from the measured pHRR mean, this indicates MARS and RF models capability and reliability for predictions and good agreement with experimental data. The ANN models on the hand exhibited relatively fair deviation errors. This error margin could partly be due to the rigorous operator iterative tuning of ANN model parameters, and the type of training algorithms used [17]. The errors observed from the models are presented in Table 7.

4. CONCLUSION

In this present study, flammability parameters were measured from microscale combustion calorimetry (MCC) experiments from polymethyl methacrylate (PMMA) material. The statistical descriptions from the experimental dataset was then normalized and compiled. Two soft-computing techniques, Random Forest (RF) and Multivariate Adaptive Regression Splines (MARS) models were developed and used to model and predict peak heat release rate (pHRR) of PMMA, considering heating rate (β), mass (m), total heat released (THR), heat release capacity (HRC), heat release temperature (pTemp) and heat release time (pTime) as input variables. The aim was to identify the best model for flammability characteristic predictions.

It was evidence from the statistical results that, MARS model achieved the highest coefficient of determination (R^2) value of (0.9998) during training and (0.9996) for testing respectively, with least standard errors of MAD, MAPE, and RMSE, as compared with RF that achieved R^2 (0.9749) and R^2 (0.9562) for training and testing respectively in evaluating the flammability characteristics of PMMA from MCC parameters. This study therefore indicates, good performance agreement between the experimental and predicted results. Although, both models exhibited excellent prediction capabilities. On the whole, MARS outperformed RF in the prediction of pHRR data.

The study further revealed heating rate (β) as the input variable with more significant impact on the outcome of MARS and RF prediction models performance as compared to mass, total heat released (THR), heat release capacity (HRC), heat release temperature (pTemp) and heat release time (pTime). This affirms from works done in literatures the positive correlation between heating rate (β) and pHRR; an important factor for predicting and describing combustion and fire spread parameters and growth from materials. In fact, MARS model was seen to be computationally more efficient with algorithms employed to construct series of simplified approximation linear regressions to finding the optimal model.

However, to validate the performance and applicability of MARS and RF for assessing flammability properties and MCC parameters, benchmarks from Artificial Neural Network (ANN) methods of feed forward back propagation neural networks (FFBPNN) and generalized regression neural network (GRNN). FFBPNN and GRNN mean values of predicted pHRR results for PMMA from Asante-Okyere, et al. [17] where compared with MARS and RF models predicted pHRR mean results observed from this study. Overall, MARS and RF observed the least deviations of 2.96 and 3.05 followed by FFBNN (3.03) and GRNN (3.14) respectively, from

measured pHRR value for PMMA. This indicates that MARS and RF models are as good as ANN methods for predicting the flammability characteristics of PMMA. This study therefore adds to literature for fire safety parameter predictions of materials with slight superiority and accuracy as compared to the ANN methods. and strongly proposes MARS and RF models as efficient and more reliable technique in predicting the flammability characteristics of PMMA and other polymeric materials.

COMPETING INTERESTS

Authors have declared that no competing interests exist.

REFERENCES

1. Hsinjin EY. Quantitative microscale assessment of polymer *fl*ammability, Plastics Research; 2015.
   Available:https://doi.org/10.2417/spepro.005899
   (Accessed: 3 May 2019)
2. Matala A. Methods and applications of pyrolysis modeling for polymeric materials, Doctor degree dissertation. VTT Science 44; 2013.
   [ISSN: 2242-119x]
3. Lyon RE, Walters RN. Microscale Combustion Calorimeter, US Patent, US5981290A; 1999.
4. Xu Q, Jin C, Griffin GJ, Matala A, Hostikka S. A PMMA flammability analysis using the MCC: Effect of specimen mass, J. Therm. Anal. Calorim. 2016;126:1831–1840.
5. Lyon RE. Heat release rate calorimeter for milligram samples, US Patent, US6464391B2; 2002.
6. Standard test method for determining flammability characteristics of plastics and other solid materials using microscale combustion calorimetry ASTM D 7309-13, American Society for Testing and Materials (International), West Conshohocken, PA; 2013.
7. Lyon RE, Walters RN, Stoliarov SI, and Safronava N. Principles and practice of micro-scale combustion calorimetry, Tech. Rep. DOT/FAA/TC- 12/53, US Department of Transportation, FAA; 2013.
   Available:http://www.tc.faa.gov/its/worldpac/techrpt/tc12-53-r1.pdf
8. Walters RN, Lyon RE. Microscale combustion calorimeter for determining flammability parameters of materials. Evol Technol. Compet. Edge. 1997;42:1335–44.
9. Lyon RE, Walters RN, Stoliarov SI. Screening flame-retardants for plastics using Microscale Combustion Calorimetry. Polym Eng. Sci. 2007;47(10):1501–10.
10. Xu Q, Jin C, Jiang Y. Compare the flammability of two extruded polystyrene foams with micro-scale combustion calorimeter and cone calorimeter tests. Journal of Thermal Analysis and Calorimetry. 2017;127(3):2359-66.
11. Yang CQ, He Q, Lyon RE, Hu Y. Investigation of the flammability of different textile fabrics using micro-scale combustion calorimetry. Polym Degrad Stab. 2010;95(2):108–15.
12. Xu Q, Jin C, Majlingova A, Restas A. Discuss the heat release capacity of polymer derived from microscale combustion calorimeter. Journal of Therm. Anal. Calorim. 2017;133:1.
13. Xu Q, Jin C, Jiang Y. Analysis of the relationship between MCC and thermal analysis results in evaluating flammability of EPS foam, J. Therm. Anal. Calorim. 2014;118(2):687–693.
14. Lyon RE, Takemori MT, Safronava N, Stoliarov SI, Walters RN. A molecular basis for polymer flammability, Polymer 50. 2009;2608–2617.
15. Walters RN, Lyon RE. Molar group contributions to polymer flammability, J. Appl. Polym. Sci. 2003;87:548–563.
16. Parandekar PV, Browning AR, Prakash O. Modeling the flammability characteristics of polymers using quantitative structure–property relationships (QSPR), Polym. Eng. Sci. 2015;55:1553–1559.
17. Asante-Okyere S, Xu Q, Mensah RA, Jin C, Ziggah YY. Generalized regression and feed forward back propagation neural networks in modelling flammability characteristics of polymethyl methacrylate (PMMA). Thermochimica Acta. 2018;667: 79-92.
18. Ma S, Lv M, Deng F, Zhang X, Zhai H, Lv W. Predicting the ecotoxicity of ionic liquids towards Vibrio fischeri using Genetic Function Approximation and Least Squares Support Vector Machine, Journal of Hazardous Materials. 2015;283:591-598.
19. Burgaz E, Yazici M, Kapusuz M, Alisir SH, Ozcan H. Prediction of thermal

stability, crystallinity and thermomechanical properties of poly(ethylene-oxide)/clay nanocomposites with Artificial Neural Networks, Thermochi. Acta. 2014; 575:159-166.
Available:https://doi.org/10.1016/j.tca.2013.10.032.

20. Mensah RA, Jiang L, Asante-Okyere S, Xu Q, Jin C. Comparative evaluation of the predictability of neural network methods on the flammability characteristics of extruded polystyrene from microscale combustion calorimetry. J.Therm. Anal. Cal. 2019;133: 1–10.
21. Friedman JH. Multivariate adaptive regression splines. The Annals of Statistics. 1991;19(1):1-141.
22. Friedman JH, Roosen CB. An introduction to multivariate adaptive regression splines", Stat Methods Med Res. 1995; 4:197.
Available:https://doi.org/10.1177/096228029500400303.
23. Ziggah YY, Laari PB. Application of multivariate adaptive regression spline approach for 2D coordinate transformation. Ghana Journal of Tech. 2018;2(2):50-62.
24. Xu QS, Massart DL, Liang YZ, Fang KT. Two-step multivariate adaptive regression splines for modeling a quantitative relationship between gas chromatography retention indices and molecular descriptors, Journal of Chromatography A. 2003;998(1–2):155-167.
Available:https://doi.org/10.1016/S0021-9673(03)00604-6
25. Balshi MS, Mcguire AD, Duffy P, Flannigan M, Walsh J, Melillo J. Assessing the response of area burned to changing climate in western boreal North America using a Multivariate Adaptive Regression Splines (MARS) approach. Global Change Biology. 2009;15:578-600.
DOI:10.1111/j.1365-2486 .2008. 01679.x
26. Durmaz M, Karslioglu MO, Nohutcu M. Regional VTEC modeling with multivariate adaptive regression splines. Advances in Space Research. 2010;46:180-189.
27. Roy SS, Roy R, Balas VE. Estimating heating load in buildings using multivariate adaptive regression splines, extreme learning machine, a hybrid model of MARS and ELM, Renew. Sustain. Energy Rev. 2018;82:4256–4268.
28. Breiman L. Random forests. Machine Learning. 2001;45:5–32.
29. Pramila PV, Mahesh V. Comparison of multivariate adaptive regression splines and random forest regression in predicting forced expiratory volume in one second. International Journal of Bioengineering and Life Sciences. 2015;9 (4):338-342.
30. Palmer DS, O'Boyle NM, Glen CR, Mitchell JBO. Random forest models to predict aqueous solubility. J. of Chemical Inform. and Modeling. 2007;47(1):150-158.
31. Matin. SS, Chelgani SC. Estimation of coal gross calorific value based on various analyses by random forest method, Fuel. 2016;177:274-278.
32. Svetnik V, Liaw A, Tong CJ, Culberson C, Sheridan RP, and Feuston BP. Random forest: A classification and regression tool for compound classification and QSAR modeling. Journal of Chemical Information and Computer Sciences. 2003;43(6): 1947-1958.
Available:https://doi.org/10.1021/ci034160g.
33. Govmark Datasheet of Micro-scale Combustion Calorimeter (MCC2), the Govmark Organization, Inc.
34. Bylander T. Estimating generalization error on two class datasets using out-of-bag estimates. Machine Learning. 2002; 48(18,22):287–297.
35. Liaw A, Wiener M. Classification and regression by random forest. R. News. 2002;2(3):18–22.
36. Nawar S, Mouazen M. Comparison between random forests, Artificial Neural Networks and Gradient Boosted Machines Methods of On-Line Vis-NIR Spectroscopy Measurements of Soil Total Nitrogen and Total Carbon, Sensors. 2017; 2428(17):1-22.
DOI:10.3390/s17102428.
Available: www.mdpi.com/journal/sensors.
37. Segal M, Xiao Y. Multivariate random forests. WIREs Data Mining Knowledge Discovery. 2011;1:80–87.
38. Adoko AC, Jiao YY, Wu L, Wang H, Wang Z-H. Predicting tunnel convergence using multivariate adaptive regression spline and artificial neural network, tunneling and underground space tech. 2013;(38): 368-376.
39. Salford Predictive Modeler (SPM). RF™, MARS™ User Guide. Salford Systems, San Diego, CA, USA; 2001.

40. Dey P, Das AK. Application of multivariate adaptive regression spline-assisted objective function on optimization of heat transfer rate around a cylinder, Nuclear Engineering and Tech. 2016;48:1315-1320. Available:http://dx.doi.org/10.1016/j.net.2016.06.011.

41. Babrauskas V, Peacock RD. Heat release rate: the single most important variable in fire hazard. Fire Safety Journal. 1992;18: 255.

Peer-review history:
The peer review history for this paper can be accessed here:
http://www.sdiarticle4.com/review-history/53584