Environment and Natural Resources Journal

Volume 25
Issue X
2027

Machine Learning-Based Classification of Plastic Combustion Emissions Using Multi-Sensor Data

Arinto Yudi Ponco Wardoyo*, Hari Arief Dharmawan, Unggul Pundjung Juswono, Renetha Salma Myesha A, Maria Pramundhitya Wishnu Wardhani, and Yahya Efendi

Physics Department, Faculty of Science, Technology, and Mathematics, Universitas Brawijaya, East Java, Indonesia

ARTICLE INFO

Received: 13 Oct 2025
Received in revised: 2 Sep 2026
Accepted: 8 Sep 2026
Published online: 10 Nov 2026
DOI: 10.32526/ennrj/25/20250290

Keywords:
Plastic combustion/ Particulate matter/ Gas emissions/ Gradient Boosting/ Machine learning/ Feature selection

*** Corresponding author:**
E-mail: a.wardoyo@ub.ac.id

ABSTRACT

Open burning of plastic waste releases particulate and gaseous pollutants whose profiles vary with polymer composition and combustion behavior. This study evaluated a multi-sensor machine-learning approach for classifying emissions from six common plastics: high-density polyethylene (HDPE), low-density polyethylene (LDPE), polyethylene terephthalate (PET), polypropylene (PP), polystyrene (PS), and polyvinyl chloride (PVC). The dataset comprised 2,160 observations (360 per class). Classification used six emission variables: PM_1, $PM_{2.5}$, PM_{10}, CO_2, CO, and total volatile organic compounds (TVOC), while relative humidity and temperature were analyzed separately as smoke-collection-chamber conditions. Kruskal-Wallis tests showed significant differences among plastic classes for all particulate and gaseous variables. Gradient Boosting (GB), Random Forest (RF), and k-nearest neighbor (k-NN) were optimized using the same randomized hyperparameter-search framework with stratified five-fold cross-validation. GB achieved the highest performance, with a cross-validated macro-F1 of 95.95%, test accuracy of 96.76%, and test macro-F1 of 96.75%. Cross-validated permutation importance identified CO2 as the most influential predictor, followed by PM_{10}, CO, $PM_{2.5}$, PM_1, and TVOC. Exhaustive evaluation of 57 feature combinations identified PM_1, $PM_{2.5}$, and CO_2 as the most parsimonious subset; the corresponding three-feature GB model retained 96.06% test accuracy and 96.05% macro-F1. The learning curve showed a 3.26-percentage-point train-validation gap at the largest training size. These results demonstrate strong discrimination of controlled plastic-combustion emission patterns, although mixed-waste, variable-combustion, and long-term sensor-drift validation is required before field application.

HIGHLIGHTS

- Classified emissions from six plastic types using multi-sensor data
- Gradient Boosting achieved the highest full-feature accuracy of 96.76%.
- CO_2 was the most influential predictor based on permutation importance.
- PM_1, $PM_{2.5}$, and CO_2 formed the most parsimonious three-feature subset.
- The reduced Gradient Boosting model retained 96.06% classification accuracy.

1. INTRODUCTION

Plastic waste is increasingly recognized as an important source of environmental contamination, particularly where collection systems are incomplete and open burning is used as an informal disposal practice (Wang et al., 2023). Open combustion of plastic-containing waste can release fine and coarse particulate matter, volatile organic compounds, carbonaceous aerosols, and toxic degradation products that affect air quality and human health (Velis and Cook, 2021; Pathak et al., 2023; Pathak et al., 2024; Wu et al., 2021). Because combustion conditions and polymer chemistry strongly influence the resulting emissions, the measured pollutant pattern can potentially provide information about the type of plastic being burned (Kucbel et al., 2022).

Citation: Wardoyo AYP, Dharmawan HA, Juswono UP, Myesha ARS, Wardhani MPW, Efendi Y. Machine learning-based classification of plastic combustion emissions using multi-sensor data. Environ. Nat. Resour. J. 2027;25(X):xx-xx.

Common polymers such as HDPE, LDPE, PET, PP, PS, and PVC differ in molecular structure, additives, aromaticity, and heteroatom content (Růžičková et al., 2022). These differences affect thermal degradation, soot formation, volatile-product generation, and residual material after combustion (Furman et al., 2021; Kanellopoulos et al., 2021; Kováts et al., 2023; Mentes et al., 2023). PS, for example, can generate intense particulate and aromatic emissions, whereas chlorine-containing PVC follows a substantially different degradation pathway. In actual waste-burning environments these polymer-specific signals may be further modified by moisture, mixed fuels, oxygen availability, airflow, and combustion temperature (Mentes et al., 2024; Ajay et al., 2022).

Conventional chemical-analysis techniques can provide detailed compositional information, but their cost, size, and analytical delay can limit routine near-real-time screening. Compact sensors provide an alternative by simultaneously recording several pollutant indicators and allowing multivariate pattern analysis. Previous work has demonstrated low-cost sensing systems for particulate matter and volatile compounds, including ZH03A- and CCS811-based measurement platforms (Widhowati et al., 2021; Wishnu Wardhani et al., 2021).

Machine-learning methods are increasingly used to interpret multi-sensor environmental data because nonlinear interactions among sensor channels may contain class information that is difficult to capture using univariate thresholds. Common approaches include distance-based classifiers, support-vector machines, tree ensembles, boosting methods, and neural networks. The present study compares three algorithms with distinct learning mechanisms: k-nearest neighbor (k-NN), which assigns classes according to local feature-space proximity (Cover and Hart, 1967); Random Forest (RF), which reduces variance by aggregating many decision trees (Breiman, 2001); and Gradient Boosting (GB), which sequentially improves weak learners by correcting preceding errors (Friedman, 2001).

Accordingly, this study aimed to: (i) quantify differences in particulate and gaseous emissions among six plastic types using non-parametric statistical tests and effect-size analysis; (ii) characterize relative humidity and temperature as environmental conditions of the smoke-collection chamber; (iii) compare k-NN, RF, and GB using a consistent hyperparameter-optimization framework; and (iv) interpret the highest-performing six-feature classifier and determine whether a reduced emission-feature subset could retain comparable classification performance. The study was designed as a controlled laboratory proof of concept for distinguishing polymer-specific emission patterns rather than for direct identification of unknown mixed wastes under field conditions.

2. METHODOLOGY

2.1 Plastic samples and specimen preparation

Six plastic classes were evaluated: HDPE, LDPE, PET, PP, PS, and PVC. Representative materials were obtained from commonly encountered plastic products and were cleaned before preparation. Each combustion specimen had a target mass of 0.5 g. Representative examples of the six plastic materials are shown in Figure 1. Each specimen, weighing 0.5 g, was prepared with approximate dimensions of 300 mm in length and 300 mm in width, while the thickness was adjusted according to the polymer type to achieve the required specimen mass.

2.2 Combustion and emission-collection system

As shown in Figure 2, the experimental arrangement consisted of a combustion unit, an emission-transfer path, a collection chamber, and a multi-sensor measurement unit. Combustion emissions were transferred into a chamber measuring 400 × 300 × 300 mm, corresponding to a volume of 0.036 m^3. A micro-pump operated at 2 L/min for 1 min, transferring approximately 2 L of gas into the collection chamber. Each 0.5 g plastic specimen was burned for 60 s. The collection chamber was ventilated between experimental runs until the sensor signals returned toward their baseline range, thereby reducing carry-over from the preceding plastic sample. The plastic samples were ignited using a gas lighter and combusted at an approximate temperature of 700±50°C. The combustion temperature was not actively controlled but was continuously monitored using a thermocouple sensor.

2.3 Multi-sensor measurement and calibration

Particulate matter was measured as PM_1, $PM_{2.5}$, and PM_{10} using the ZH03A optical particle sensor. CO_2 and TVOC were measured using the CCS811 gas-sensing module, while CO was measured using the MiCS-6814 sensor. The particulate and gas-sensing systems had previously been characterized

against reference instruments in earlier development studies (Widhowati et al., 2021; Wishnu Wardhani et al., 2021). Before the combustion experiments, sensor performance was verified using the calibration and reference procedures summarized in Table 1.



Figure 1. Representative samples of HDPE, LDPE, PET, PP, PS, and PVC used in the combustion experiments

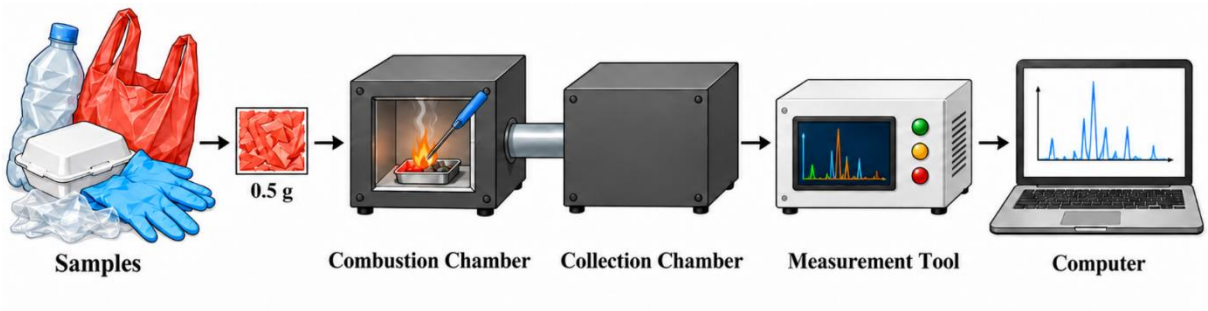


Figure 2. Schematic diagram of the plastic-combustion, smoke-collection, and multi-sensor measurement system

Table 1. Sensor variables, measurement targets, and calibration/reference methods used in this study

Sensor/channel	Measurement target	Calibration/reference method	Reported performance
ZH03A	PM_1, $PM_{2.5}$, PM_{10}	Compared with Kanomax Digital Dust Monitor 3443 and Hinaway CW-HAT200; regression evaluated over 0-1000 µg/m³.	R^2>0.99; reported accuracy ±0.5% indoor and ±1% outdoor.
CCS811	CO_2, TVOC	Compared with a reference TVOC meter (WP6910) in a 0.012 m³ chamber.	R^2>0.99; resolution 0.001 ppm; response time 4 s.
MiCS-6814	CO	Compared with Kanomax 2212	R^2>0.98

2.4 Calculation of total concentration and emission factor

In addition to the instantaneous sensor concentrations used for statistical characterization, total concentration and emission factor were calculated as complementary descriptors of combustion intensity. These quantities were calculated separately from the machine-learning dataset and were not used as classifier input variables, during hyperparameter optimization, for permutation importance, or for feature-subset selection. For each pollutant i, the total emitted amount over the selected measurement interval was obtained

from the area under the concentration-time curve while accounting for the measurement airflow rate, according to Equation 1:

$$M_i = \int_0^T C_i(t)Q(t)dt \quad (1)$$

Where; M_i denotes the integrated pollutant amount, $C_i(t)$ is the measured concentration of pollutant i as a function of time, Q(t) is the airflow rate during measurement, and T is the integration interval. The emission factor was then normalized to the mass of plastic consumed during combustion, as expressed in Equation 2:

$$EF_i = M_i/m_b \quad (2)$$

Where; EF_i is the emission factor of pollutant i and m_b is the mass of plastic burned. These integrated metrics were used to characterize overall emission intensity, whereas the classification analysis was based only on the observation-level PM_1, $PM_{2.5}$, PM_{10}, CO_2, CO, and TVOC measurements.

2.5 Preparation of the machine-learning dataset

The machine-learning dataset comprised 2,160 observations, with 360 observations for each of the six plastic classes. Each observation contained measurements of PM_1, $PM_{2.5}$, PM_{10}, CO_2, CO, TVOC, relative humidity, and temperature. Classification was performed using only the six emission-related variables (PM_1, $PM_{2.5}$, PM_{10}, CO_2, CO, and TVOC). Relative humidity and temperature were analyzed separately as environmental conditions of the smoke-collection chamber and were not included as predictors. The recorded chamber temperature was distinct from the combustion temperature measured during plastic burning.

2.6 Statistical analysis of emission and chamber variables

Descriptive statistics were calculated for all six emission variables by plastic class. Because the distributions contained strong skewness, ceiling values in some particulate channels, and non-uniform dispersion, between-class differences were evaluated using the Kruskal-Wallis test. Epsilon-squared (ε^2) was calculated as an omnibus effect-size measure. Pairwise Mann-Whitney U tests were subsequently performed and adjusted using the Holm procedure to control the family-wise error rate. Statistical significance was defined at $p<0.05$. Relative humidity and temperature were analyzed separately as smoke-collection-chamber variables. Their distributions among plastic groups were tested using the same non-parametric framework. Spearman rank correlation was then used to evaluate associations between chamber humidity/temperature and the six emission variables; the corresponding p-values were Holm-adjusted across the correlation set.

2.7 Machine-learning models and hyperparameter optimization

The 2,160 observations were divided into stratified training (80%) and testing (20%) subsets using a random state of 42. Three supervised classifiers were evaluated: GB, RF, and k-NN. Hyperparameter optimization was conducted exclusively on the training subset using randomized search with a maximum of 24 candidate configurations, stratified five-fold cross-validation, and macro-F1 as the optimization criterion. To ensure direct comparability among the algorithms, the same search budget and cross-validation procedure were applied to all three models. For GB, the search space included 50-300 estimators, learning rates of 0.02-0.20, maximum tree depths of 1-3, subsampling ratios of 0.70-1.00, and minimum leaf sizes of 1-4. For RF, the search included 100-500 trees, unrestricted or bounded tree depth, alternative feature-selection rules, and different node-size constraints. For k-NN, the number of neighbors, distance weighting, distance metric, and leaf size were optimized within a standardized feature pipeline. Performance on the held-out test subset was evaluated using accuracy, macro-F1, macro-precision, macro-recall, and normalized confusion matrices.

2.8 Gradient boosting interpretation and feature reduction

Gradient Boosting was treated as the primary full-feature classifier because it achieved the highest cross-validated and held-out performance under the common hyperparameter-search framework. Model interpretation therefore focused on GB. Cross-validated permutation importance was calculated on validation folds within the training data, and the decrease in macro-F1 after feature permutation was used as the primary feature-importance measure. Built-in model importance and permutation importance on the held-out test subset were used as supporting diagnostics.

To evaluate whether the sensor-input set could be simplified, all 57 possible combinations containing

two to six of the six emission features were evaluated using the optimized GB architecture and five-fold cross-validation on the training data only. The subset with the highest mean macro-F1 was selected without access to the held-out test subset. GB, RF, and k-NN were then re-optimized using the selected subset. A reduced-feature model was interpreted as a parsimonious alternative and not as a replacement for the primary six-feature GB classifier.

2.9 Learning-curve analysis

Learning-curve analysis was performed to evaluate model generalization and potential overfitting. Training and validation macro-F1 were calculated across increasing training-set sizes under five-fold stratified cross-validation. The final train-validation gap and the direction of the validation curve were used to assess whether model variance persisted as the amount of training data increased and whether additional observations could further improve generalization.

3. RESULTS AND DISCUSSION

3.1 Emission characteristics of the six plastic classes

The six plastic classes produced markedly different particulate profiles (Figure 4). PS generated the highest mean particulate concentrations, reaching 805.63±272.28 for PM_1, 847.87±239.46 for $PM_{2.5}$, and 879.41±212.56 for PM_{10}. PVC showed the lowest corresponding means (27.91±13.70, 34.91±16.71, and 39.09±18.86), followed by HDPE. LDPE occupied an intermediate range, whereas PET and PP displayed broad distributions with individual observations reaching the upper recorded value of 1,000. This ceiling behavior is important because saturation can compress the apparent magnitude of differences among strongly emitting samples.

Figure 3 presents the residual ash and unburned material observed after combustion, illustrating visible differences in post-combustion residues among the plastic types.

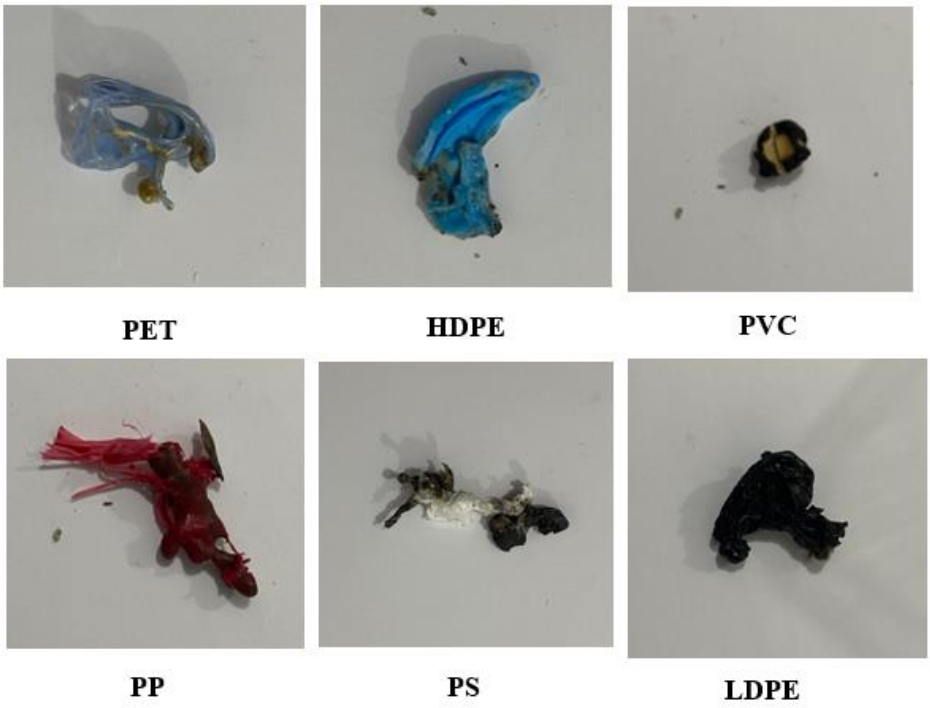


Figure 3. Residual ash and unburned material observed after controlled combustion of the plastic samples

The particulate differences were statistically strong. Kruskal-Wallis tests were significant for PM_1, $PM_{2.5}$, and PM_{10} (all $p<0.001$), with epsilon-squared values of 0.661, 0.662, and 0.663, respectively. These large effect sizes indicate substantial between-class variation in the measured particulate channels. The strong particulate contrast is consistent with polymer-dependent soot and aerosol formation reported in controlled plastic-combustion studies. Figure 4 summarizes the distributions of all six emission variables and highlights between-class variability and outlying observations.

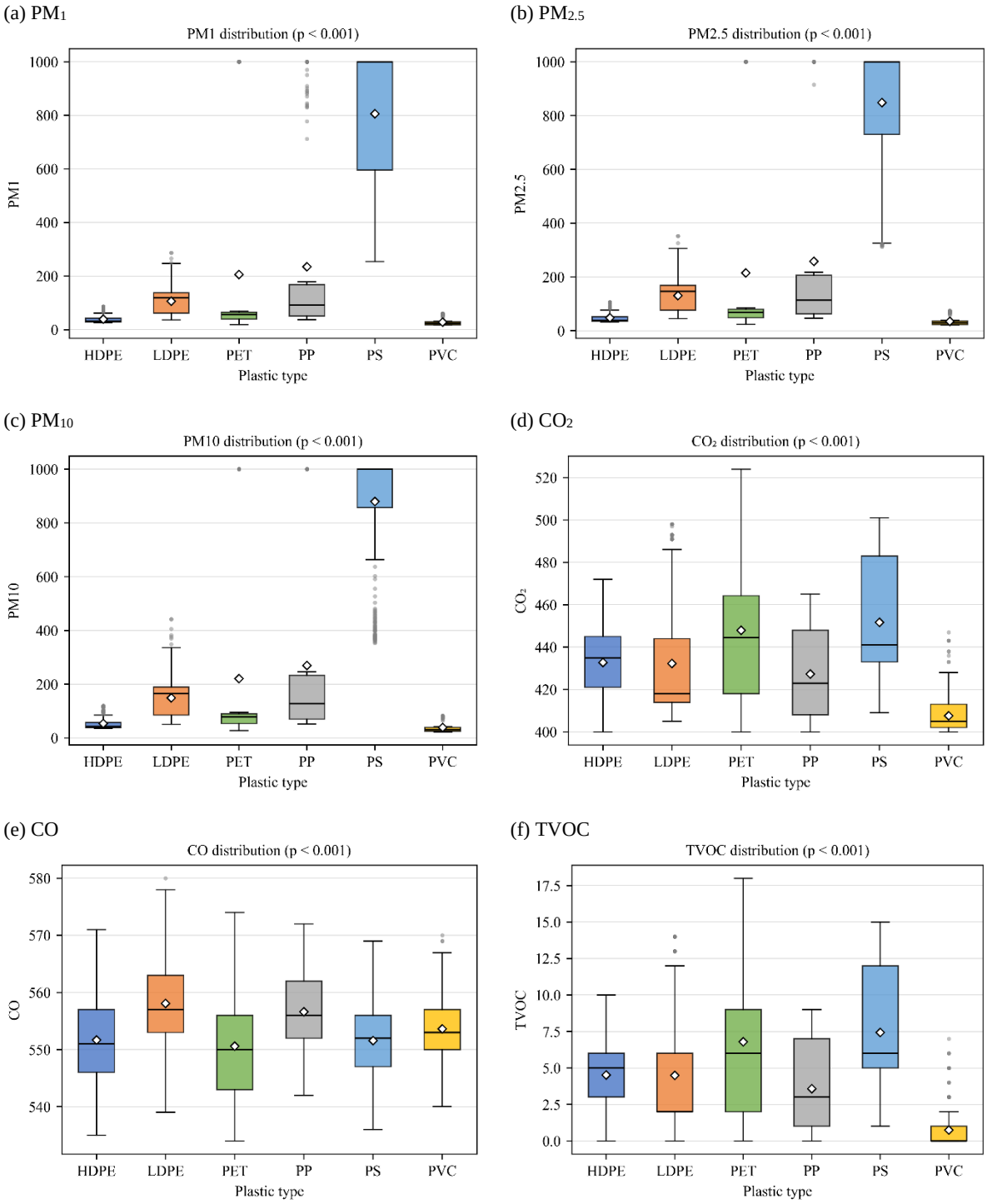


Figure 4. Distributions of (a) PM_1, (b) $PM_{2.5}$, (c) PM_{10}, (d) CO_2, (e) CO, and (f) TVOC across the six plastic classes. Diamonds indicate means; horizontal lines indicate medians.

3.2 Statistical differences in gaseous emissions

At the descriptive level, PS exhibited the highest mean CO_2 (451.63) and TVOC (7.44), whereas PVC showed the lowest mean CO_2 (407.61) and TVOC (0.74). CO concentrations occupied a narrower range, from 550.58 for PET to 558.08 for LDPE. Holm-adjusted pairwise analysis identified numerous significant contrasts, particularly comparisons involving PVC and PS for CO_2 and TVOC (Table 2). For example, the PS-PVC contrast showed very large rank-biserial effects for both CO_2 and TVOC. These results provide inferential evidence that gaseous emission profiles differed among the plastic classes.

All three gaseous variables differed significantly among plastic classes (Table 3). CO_2

showed H=670.65 ($p<0.001$, $\varepsilon^2=0.309$), TVOC showed H=677.99 ($p<0.001$, $\varepsilon^2=0.312$), and CO showed H=258.27 ($p<0.001$, $\varepsilon^2=0.118$). Thus, CO_2 and TVOC produced moderate between-class effects, whereas CO exhibited a smaller but still statistically detectable class effect. This distinction is important: the large number of observations can make relatively small differences statistically significant, so the effect size provides a more meaningful indication of practical separation.

Table 2. Kruskal-Wallis statistics for particulate and gaseous emission variables

Variable	H	df	p-value	ε^2	Interpretation
PM_1	1,428.75	5	<0.001	0.661	Large/moderate-to-large class effect
$PM_{2.5}$	1,431.36	5	<0.001	0.662	Large/moderate-to-large class effect
PM_{10}	1,432.47	5	<0.001	0.663	Large/moderate-to-large class effect
CO_2	670.65	5	<0.001	0.309	Large/moderate-to-large class effect
CO	258.27	5	<0.001	0.118	Small-to-moderate class effect
TVOC	677.99	5	<0.001	0.312	Large/moderate-to-large class effect

Table 3. Smoke-collection-chamber environmental statistics and selected correlations with emission variables

Variable/relation	Value	p-value	Effect/correlation	Interpretation
Humidity, overall	45.10±3.67	-	-	Chamber environment; not ML input
Temperature, overall	40.36±1.36	-	-	Chamber environment; not furnace temperature
Humidity among classes	H=27.20	<0.001	$\varepsilon^2=0.010$	Very small class effect
Temperature among classes	H=208.04	<0.001	$\varepsilon^2=0.094$	Small-to-moderate class effect
Humidity vs CO_2	$\rho=-0.232$	Holm p<0.001	Modest negative	Significant
Humidity vs TVOC	$\rho=-0.228$	Holm p<0.001	Modest negative	Significant
Temperature vs CO	$\rho=-0.315$	Holm p<0.001	Moderate negative	Strongest chamber correlation
Temperature vs CO_2	$\rho=0.259$	Holm p<0.001	Modest positive	Significant

3.3 Smoke-collection-chamber humidity and temperature

Humidity and temperature were evaluated separately from the classification features because they describe the collection-chamber measurement environment rather than the combustion source itself. Across all 2,160 observations, mean relative humidity was 45.10±3.67%RH (Figure 5(a)), while mean chamber temperature was 40.36±1.36 °C as shown in Figure 5(b). These variables were used only to characterize the chamber environment and were not included in the machine-learning predictor matrix.

Both chamber variables showed statistically significant differences among plastic groups (humidity: H=27.20, $p<0.001$; temperature: H=208.04, $p<0.001$) (Table 3), but their effect sizes were substantially smaller than those observed for the particulate variables. Humidity had $\varepsilon^2=0.010$, indicating a very small group effect, whereas chamber temperature had $\varepsilon^2=0.094$, indicating a small-to-moderate effect. Therefore, statistical significance alone should not be interpreted as evidence that chamber conditions dominated the observed plastic-emission patterns.

Spearman analysis further clarified these relationships, as shown in Figure 5(c). Humidity was essentially unrelated to PM_1, $PM_{2.5}$, PM_{10}, and CO, but showed modest negative correlations with CO_2 ($\rho=-0.232$) and TVOC ($\rho=-0.228$). Chamber temperature showed modest negative associations with PM1-PM10 ($\rho\approx-0.209$), a positive association with CO_2 ($\rho=-0.259$) and TVOC ($\rho=0.259$), and the strongest chamber relationship with CO ($\rho=-0.315$). Because humidity and temperature were excluded from the classifier, the reported machine-learning performance is attributable to the six emission channels rather than direct use of these environmental variables.

3.4 Total concentration and emission factor analysis

Total concentration and emission factor provided complementary physical measures of particulate and gaseous emission intensity for each plastic type. These integrated quantities were analyzed

separately from the observation-level machine-learning features and were used to support interpretation of the overall combustion-emission profiles.

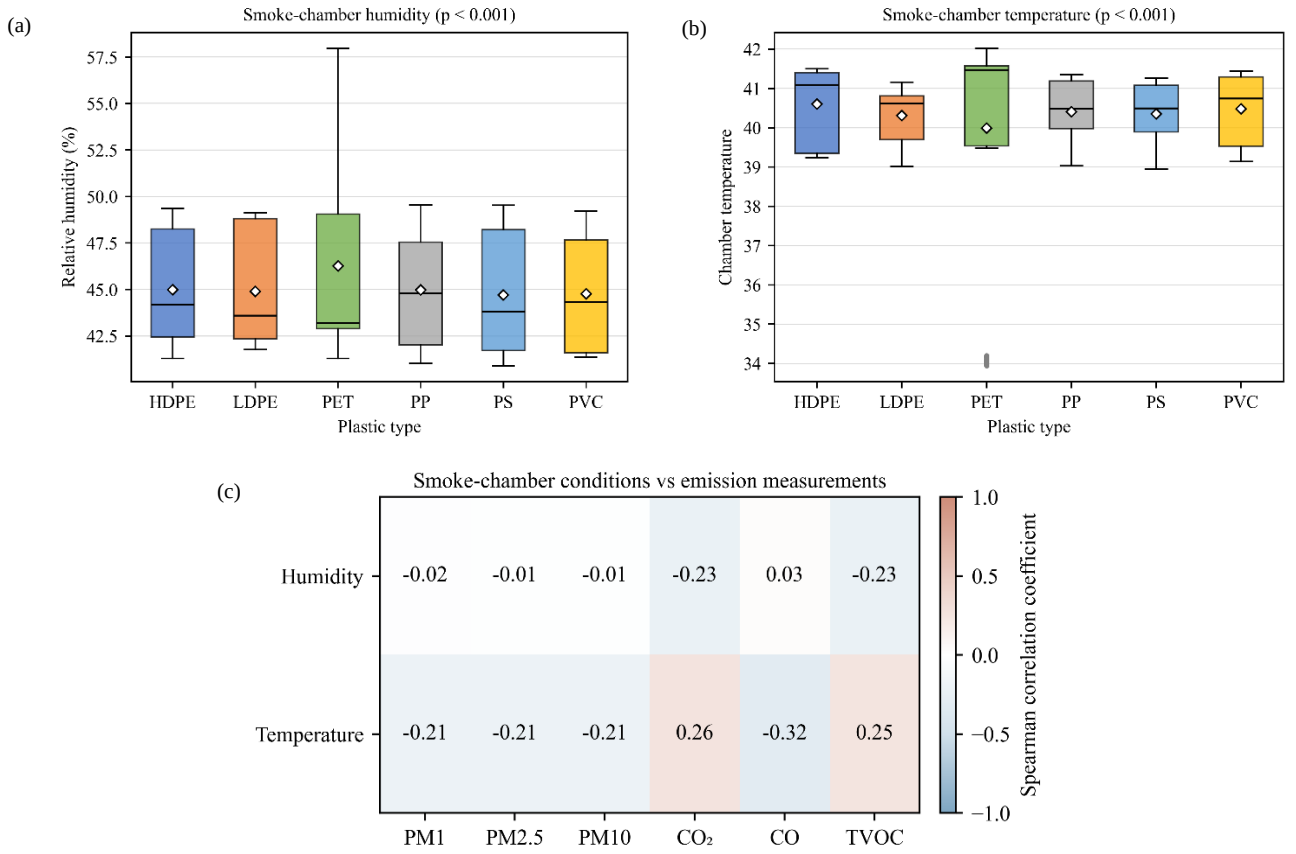


Figure 5. Smoke-collection-chamber analysis: (a) relative-humidity distribution, (b) chamber-temperature distribution, and (c) Spearman correlations between chamber conditions and the six emission measurements.

Figure 6(a) shows the integrated particulate amount for PM_1, $PM_{2.5}$, and PM_{10}. PS produced the largest total particulate values, reaching approximately 2 µg across the three particulate fractions, whereas the remaining plastics were substantially lower. PP and LDPE formed the next highest group, while PET showed the lowest integrated particulate values. The corresponding emission factors in Figure 6(b) preserve the same overall ordering: PS approached approximately 4 µg/g, while the other plastics remained markedly lower. The agreement between total concentration and EF indicates that the high particulate burden observed for PS is not solely a consequence of the initial sample mass but remains evident after normalization to the burned material.

The integrated gas-emission results showed a different scale of emission behavior (Figure 7). CO_2 dominated the integrated gas values for all six plastics at approximately 2,000 micrograms, followed by CO at approximately 1,000-1,200 µg, whereas TVOC remained much smaller. The gas emission factors in Figure 7(b) showed the same general ordering, with CO_2 as the dominant component followed by CO and TVOC. These integrated metrics provide physical context for comparing combustion intensity but were not used to train, interpret, or select the GB, RF, or k-NN classifiers.

3.5 Comparison of the three machine-learning classifiers

Following consistent hyperparameter optimization, GB achieved the highest performance among the three six-feature classifiers (Table 4 and Figure 8). GB obtained a cross-validated macro-F1 of 95.95%, a test accuracy of 96.76%, and a test macro-F1 of 96.75%. RF showed comparable but slightly lower performance, with a cross-validated macro-F1 of 95.49%, test accuracy of 95.60%, and test macro-F1 of 95.52%. k-NN produced lower classification performance, with a cross-validated macro-F1 of 88.26% and test accuracy of 88.66%. These results identify GB as the strongest classifier for the complete six-feature emission dataset.

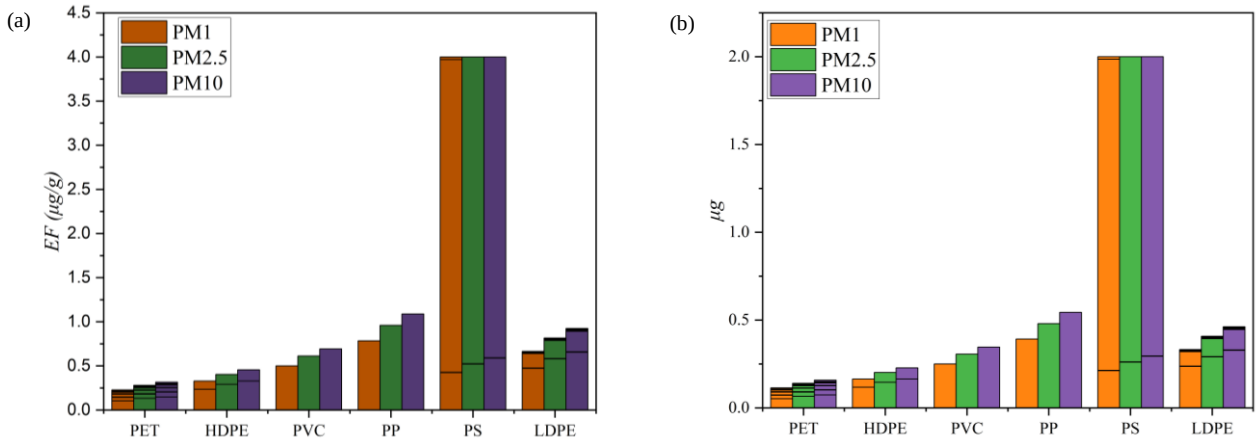


Figure 6. Distribution of (a) total particulate concentration/amount for PM_1, $PM_{2.5}$, and $PM_{10}$and (b) corresponding particulate EF from the combustion of six plastic types.

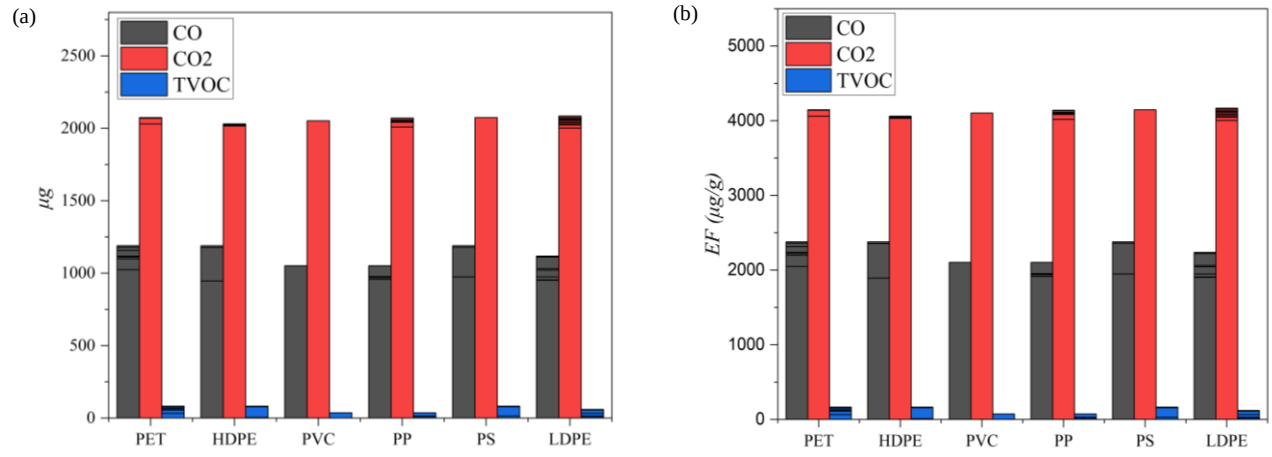


Figure 7. Distribution of (a) total gas concentration for CO, CO_2, and TVOC and (b) corresponding gas emission factor (EF) from the combustion of six plastic types.

The optimized GB configuration consisted of 300 estimators, a learning rate of 0.20, a maximum tree depth of 3, a minimum leaf size of 4, and a subsampling ratio of 0.85. The optimized RF model used 500 trees with unrestricted depth, $\log_2$ feature sampling, and a minimum leaf size of 1, whereas the optimized k-NN model used five distance-weighted neighbors with Manhattan distance. The strong performance of GB after systematic optimization indicates that its classification capability was sensitive to hyperparameter configuration, highlighting the importance of consistent optimization when comparing different learning algorithms.

Table 4. Optimized full six-feature classification performance

Model	CV macro-F1 (%)	CV train-validation gap (pp)	Test accuracy (%)	Test macro-F1 (%)	Macro-precision (%)	Macro-recall (%)
Gradient boosting	95.95	3.88	96.76	96.75	96.86	96.76
Random forest	95.49	4.34	95.60	95.52	95.77	95.60
k-NN	88.26	11.56	88.66	88.70	89.42	88.66

The normalized confusion matrices in Figure 9 show that GB provided consistently high class-level discrimination and slightly cleaner separation than RF, whereas k-NN produced more cross-class confusion. The performance difference between GB and RF was modest in absolute terms, but GB was superior on both the cross-validation criterion used for model selection and the held-out test metrics. GB was therefore

selected as the primary classifier, with RF and k-NN serving as comparison models.

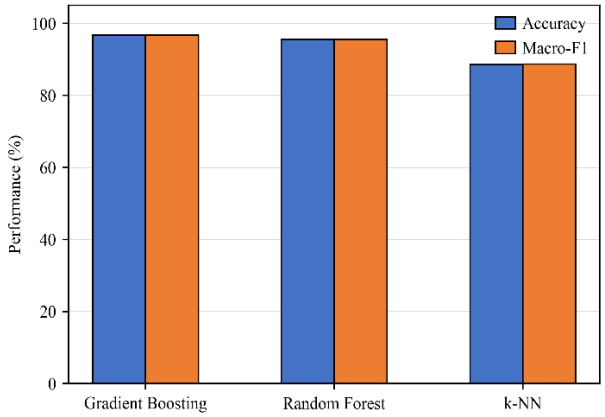


Figure 8. Classification performance of GB, RF, and k-NN using the six emission features

3.6 Interpretation of the primary gradient boosting classifier

Because GB achieved the highest full-feature classification performance, feature interpretation focused on this model. Cross-validated permutation importance ranked CO_2 first (mean decrease in macro-F1=0.334), followed by PM_{10} (0.132), CO (0.069), $PM_{2.5}$ (0.057), PM_1 (0.057), and TVOC (0.005) (Figure 10). This ranking indicates that CO_2 supplied the largest unique predictive contribution within the fitted GB model, while particulate information-particularly PM_{10}-also contributed strongly.

Permutation importance should be interpreted in the context of correlated predictors. PM_1, $PM_{2.5}$, and PM_{10} are physically related size fractions and therefore contain overlapping information. A variable may receive a modest permutation score even when it is physically informative if another correlated feature can compensate after permutation. This is illustrated by the difference between GB built-in importance, which ranked $PM_{2.5}$ highly, and cross-validated permutation importance, which emphasized CO_2 and PM_{10}. For this reason, permutation importance was used as the primary interpretive result, while the built-in tree importance was considered supportive rather than definitive.

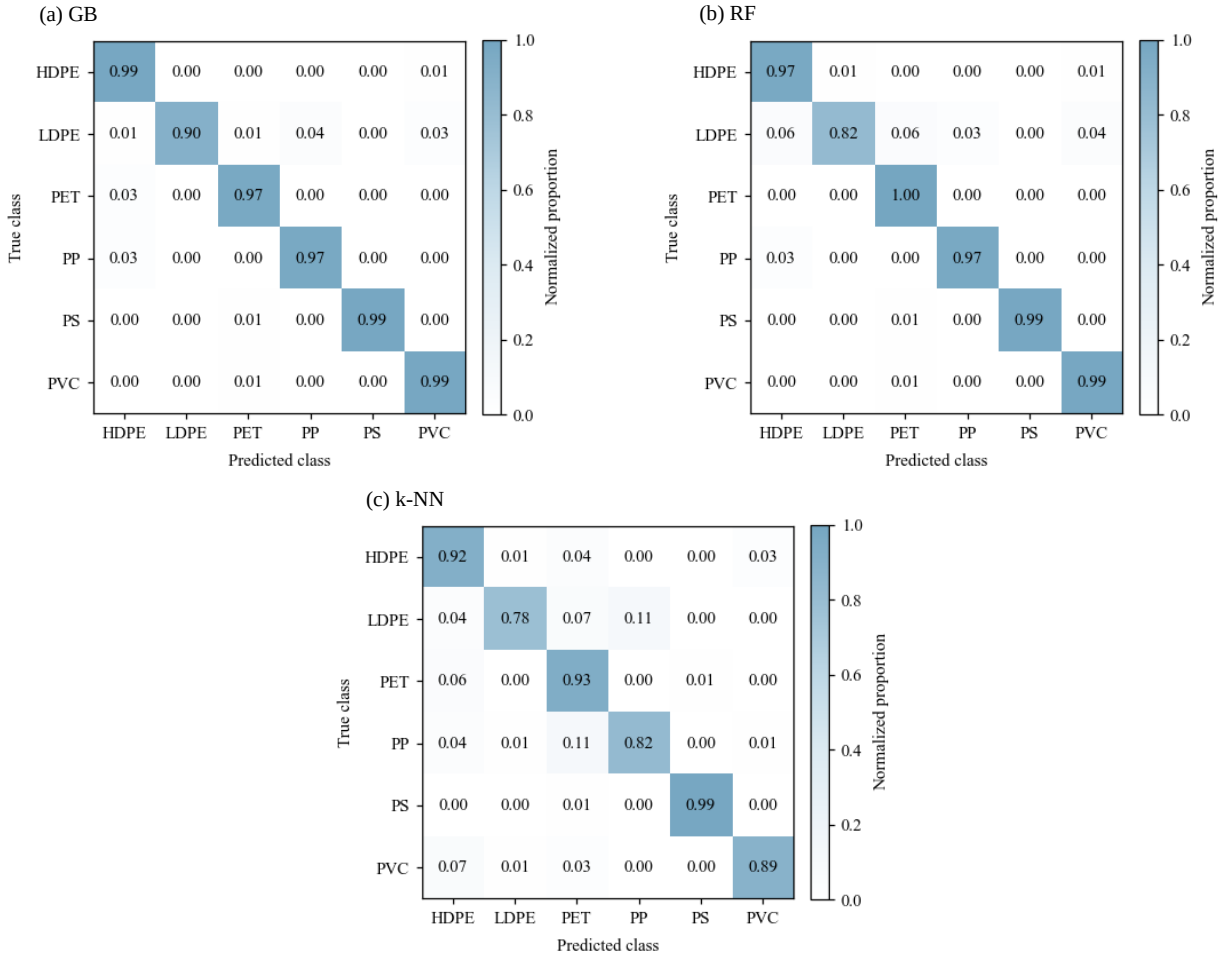


Figure 9. Normalized confusion matrices for (a) GB, (b) RF, and (c) k-NN using six emission features

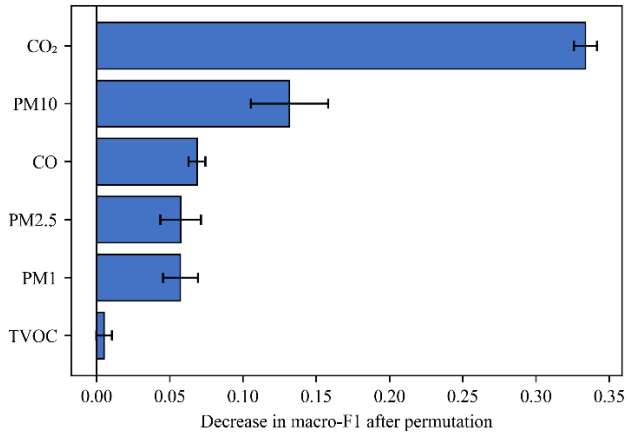


Figure 10. Cross-validated permutation importance of the six emission variables for the primary Gradient Boosting classifier. Error bars show variation across validation folds.

3.7 Best-feature analysis and reduced-input model

Exhaustive evaluation of all 57 possible two-to-six-feature subsets showed that the best three-feature GB combination was PM_1, $PM_{2.5}$, and CO_2, with mean cross-validated macro-F1 of 96.35±0.73% (Table 5 and Figure 11). A two-feature combination consisting of PM_1 and CO_2 already achieved 96.06% macro-F1, while adding $PM_{2.5}$ produced the highest mean performance. Adding TVOC as a fourth feature yielded virtually the same cross-validated macro-F1 (96.35%), whereas five- and six-feature combinations were slightly lower. The three-feature subset was therefore selected on the basis of parsimony and maximum training-CV performance.

The reduced-feature finding is noteworthy because PM_{10} ranked second in permutation importance but was not retained in the optimal three-variable subset. This apparent discrepancy is expected when correlated particulate variables contain redundant information. PM_1 and $PM_{2.5}$ can jointly capture much of the particulate pattern, while CO_2 contributes a complementary gaseous signal. Thus, feature importance and feature selection answer different questions: permutation importance estimates unique predictive contribution within the full model, whereas subset optimization identifies the combination that best preserves performance with fewer inputs.

After the selected three-feature subset was fixed, all three algorithms were retuned. GB remained the strongest reduced-input model according to the cross-validation criterion, with CV macro-F1 of 96.35%, test accuracy of 96.06%, and test macro-F1 of 96.05%. RF achieved the same test accuracy (96.06%) with a nearly identical macro-F1 of 96.04%, while k-NN reached 95.60% accuracy. Thus, GB remained the leading reduced-input model, while RF represented a highly competitive alternative with nearly identical held-out performance.

Table 5. Best Gradient Boosting subset at each feature count and reduced-model comparison

Feature count	Best GB subset	CV macro-F1 (%)	CV SD (%)
2	PM_1, CO_2	96.06	0.69
3	PM_1, $PM_{2.5}$, CO_2	96.35	0.73
4	PM_1, $PM_{2.5}$, CO_2, TVOC	96.35	0.67
5	PM_1, $PM_{2.5}$, CO_2, TVOC	96.01	0.94
6	PM_1, $PM_{2.5}$, CO_2, CO, TVOC	95.95	1.17

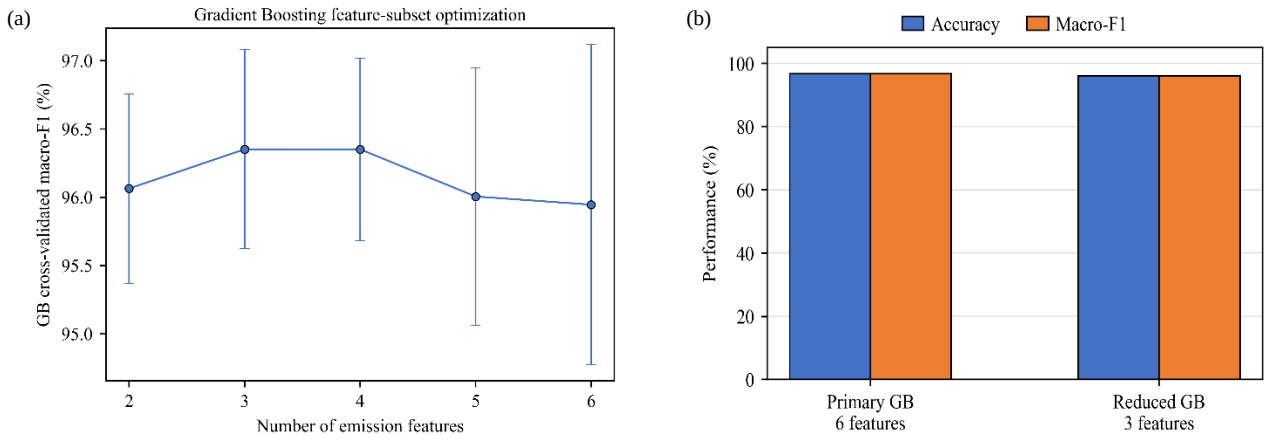


Figure 11. Feature-reduction analysis: (A) best GB cross-validated macro-F1 by number of emission features and (B) comparison between the six-feature primary GB classifier and the selected reduced-feature configuration.

3.8 Dataset sufficiency and overfitting assessment

The full-feature GB learning curve increased steadily as more training observations were introduced (Figure 12). Validation macro-F1 improved from 88.69% at 259 training observations to 96.49% at 1,728 observations, while training macro-F1 decreased slightly from 99.85% to 99.75%. At the largest training size, the train-validation gap was 3.26 percentage points. This residual gap indicates some remaining model variance, but it is considerably smaller than the gap observed at small training sizes and is consistent with useful generalization within the present controlled dataset.

The upward trend of the validation curve at the largest sample size suggests that additional data could still improve robustness. The 2,160-observation dataset provided a substantial basis for the present comparative modeling under controlled laboratory conditions, but it should not be interpreted as representing the full diversity of real plastic-waste combustion. Additional experiments spanning mixed polymers, different combustion temperatures, variable ventilation, moisture, and independent acquisition campaigns remain necessary for external validation.

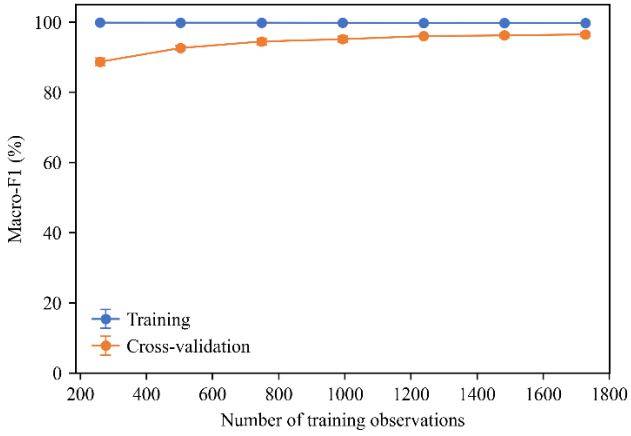


Figure 12. Learning curve of the primary six-feature GB classifier. Training and validation macro-F1 are shown as the number of training observations increases.

3.9 Practical interpretation, limitations, and future work

The results support two complementary conclusions. First, the complete six-feature sensor set provided the highest held-out performance, with GB reaching 96.76% accuracy. Second, nearly the same discrimination was achieved using only PM_1, $PM_{2.5}$, and CO_2, reducing the number of required emission inputs by half. This reduction may simplify future hardware, data acquisition, or embedded implementation while preserving most of the classification capability. However, sensor reduction should be treated as an engineering optimization rather than evidence that the excluded pollutants are environmentally unimportant.

The study remains subject to several limitations. The experiments were conducted using individual, known plastic types in a controlled collection chamber. Real waste burning frequently involves polymer mixtures, food residues, biomass, paper, moisture, and variable airflow. Some particulate observations reached the maximum recorded value of 1,000, indicating possible measurement saturation. Chamber humidity and temperature showed statistically detectable relationships with several emission variables, although these environmental variables were excluded from the classifier. Long-term sensor drift, recalibration frequency, combustion-temperature variability, and between-device variability also require systematic study. The model should therefore be regarded as a laboratory-scale proof of concept rather than a field-ready detector of arbitrary plastic waste. A logical next stage is external validation using an independent experimental campaign, followed by mixed-plastic and mixed-waste combustion, systematic variation of combustion temperature and oxygen supply, and assessment of sensor aging. Periodic reference checks should also be incorporated so that model performance can be distinguished from hardware drift. These steps are required before the approach can be generalized to environmental monitoring or source-identification applications.

4. CONCLUSION

This study evaluated multi-sensor and machine-learning approaches for distinguishing combustion emissions from six plastic types using 2,160 observations. Classification was based on six emission variables (PM_1, $PM_{2.5}$, PM_{10}, CO_2, CO, and TVOC) while relative humidity and chamber temperature were evaluated separately as environmental conditions of the smoke-collection chamber. Significant differences were observed among plastic classes for all particulate and gaseous variables, with particularly strong effects for the particulate fractions. Using a consistent hyperparameter-optimization framework, Gradient Boosting achieved the highest six-feature classification performance, with a test accuracy of 96.76% and macro-F1 of 96.75%. Cross-validated permutation importance identified CO_2 as the most

influential predictor, followed by PM_{10} and CO. Feature-subset optimization further showed that PM_1, $PM_{2.5}$, and CO_2 provided a parsimonious three-feature configuration while retaining 96.06% test accuracy. The learning-curve analysis indicated a relatively small train-validation gap of 3.26 percentage points at the largest training size. Total concentration and emission factor provided complementary measures of overall combustion-emission intensity but were not used as machine-learning predictors. These findings demonstrate the potential of multi-sensor emission patterns for discriminating plastic types under controlled conditions. Further validation using mixed wastes, independent acquisition campaigns, systematic variation in combustion conditions, and long-term sensor-stability assessment is required before field application.

ACKNOWLEDGEMENTS

The authors acknowledge the support of the Physics Department, Universitas Brawijaya, and the laboratory personnel who assisted with the plastic-combustion and sensor measurements. This research was funded by the Professor Grant of the Faculty of Mathematics and Natural Sciences, Universitas Brawijaya, in 2023 under contract number 4158.03/UNl0.F09/PN/2023.

AUTHOR CONTRIBUTIONS

Conceptualization, A.Y.P.W., Y.E., and H.A.D.; Methodology, A.Y.P.W., Y.E. and U.P.J.; Software, Y.E.; Validation, A.Y.P.W., H.A.D., and Y.E.; Formal Analysis, Y.E. and R.S.M.A.; Investigation, M.P.W.W. and R.S.M.A.; Resources, U.P.J.; Data Curation, Y.E.; Writing-Original Draft Preparation, Y.E.; Writing-Review and Editing, A.Y.P.W. and H.A.D.; Visualization, Y.E.; Supervision, A.Y.P.W.; Project Administration, U.P.J.; Funding Acquisition, A.Y.P.W.

DECLARATION OF CONFLICT OF INTEREST

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

REFERENCES

Ajay SV, Kirankumar PS, Sanath K, Prathish KP, Haridas A. An experimental simulation study of conventional waste burning practices in India for the assessment and inventorisation of PCDD/F/dl-PCB emissions. Journal of Environmental Management 2022;303:Article No. 114109.

Breiman L. Random forests. Machine Learning 2001;45:5-32.

Cover TM, Hart PE. Nearest neighbor pattern classification. IEEE Transactions on Information Theory 1967;13(1):21-7.

Friedman JH. Greedy function approximation: A gradient boosting machine. The Annals of Statistics 2001;29(5):1189-232.

Furman P, Styszko K, Skiba A, Zięba D, Zimnoch M, Kistler M, et al. Seasonal variability of PM_{10} chemical composition including 1,3,5-triphenylbenzene, marker of plastic combustion and toxicity in Wadowice, South Poland. Aerosol and Air Quality Research 2021;21(3):Article No. 200223.

Kanellopoulos PG, Verouti E, Chrysochou E, Koukoulakis K, Bakeas E. Primary and secondary organic aerosol in an urban/industrial site: Sources, health implications and the role of plastic enriched waste burning. Journal of Environmental Sciences 2021;99:222-38.

Kováts N, Hubai K, Sainnokhoi TA, Eck-Varanka B, Hoffer A, Tóth Á, et al. Ecotoxicity of PM_{10} emissions generated during controlled burning of waste PET. Environmental Toxicology and Pharmacology 2023;99:Article No. 104118.

Kucbel M, Švédová B, Raclavská H, Růžičková J, Raclavský K, Šafář M, et al. Measurement of organic and elemental carbon in the char deposits from the combustion of permitted and undesirable fuels in domestic boilers. Fuel 2022;319:Article No. 123749.

Mentes D, Nagy G, Szabó TJ, Hornyák-Mester E, Fiser B, Viskolcz B, et al. Combustion behaviour of plastic waste - A case study of PP, HDPE, PET, and mixed PES-EL. Journal of Cleaner Production 2023;402:Article No. 136850.

Mentes D, Jordán A, Farkas L, Muránszky G, Fiser B, Viskolcz B, et al. Evaluating emissions and air quality implications of residential waste incineration. Scientific Reports 2024;14: Article No. 21314.

Pathak G, Nichter M, Hardon A, Moyer E, Latkar A, Simbaya J, et al. Plastic pollution and the open burning of plastic wastes. Global Environmental Change 2023;80:Article No. 102648.

Pathak G, Nichter M, Hardon A, Moyer E. The open burning of plastic wastes is an urgent global health issue. Annals of Global Health 2024;90(1):Article No. 3.

Růžičková J, Raclavská H, Juchelková D, Kucbel M, Raclavský K, Švédová B, et al. Organic compounds in the char deposits characterising the combustion of unauthorised fuels in residential boilers. Energy 2022;257:Article No. 124724.

Velis CA, Cook E. Mismanagement of plastic waste through open burning with emphasis on the Global South: A systematic review of risks to occupational and public health. Environmental Science and Technology 2021;55(11):7186-207.

Wang X, Firouzkouhi H, Chow JC, Watson JG, Carter W, De Vos ASM. Characterization of gas and particle emissions from open burning of household solid waste from South Africa. Atmospheric Chemistry and Physics 2023;23(15):8921-37.

Widhowati AA, Ponco Wardoyo AY, Dharmawan HA, Nurhuda M, Budianto A. Development of a portable volatile organic compounds concentration measurement system using a CCS811 air quality sensor. Proceedings of the 2021 International Symposium on Electronics and Smart Devices (ISESD); 2021 Jun 29-30; Bandung: Indonesia; 2021.

Wishnu Wardhani MP, Ponco Wardoyo AY, Dharmawan HA, Nurhuda M, Budianto A. Design of a concentration measurement system for volcanic particulate matters in different size distribution. Proceedings of the 2021 International Symposium on Electronics and Smart Devices (ISESD); 2021 Jun 29-30; Bandung: Indonesia; 2021.

Wu D, Li Q, Shang X, Liang Y, Ding X, Sun H, et al. Commodity plastic burning as a source of inhaled toxic aerosols. Journal of Hazardous Materials 2021;416:Article No. 125820.