ARTICLE

Machine learning classification of THz spectra of commercial rubber

Mark Lexter D. De Lara*[1], Allen L. Nazareno[1], Chiquee U. Barcarse[1], Elmer S. Estacio[2], Arvin Lester C. Jusi[3], and Alvin Karlo G. Tapia[3]

[1]Institute of Mathematical Sciences, College of Arts and Sciences, University of the Philippines Los Baños, College, Laguna 4031, Philippines
[2]National Institute of Physics, College of Science, University of the Philippines Diliman, Quezon City 1101, Philippines
[3]Institute of Physics, College of Arts and Sciences, University of the Philippines Los Baños, College, Laguna 4031, Philippines

ABSTRACT

Rubber is a widely used industrial material whose performance and durability are strongly influenced by thermal processing, making reliable quality assessment essential. This study investigated machine learning techniques for classifying commercial rubber samples using terahertz time-domain spectroscopy (THz-TDS) absorption coefficient and refractive index spectra obtained after controlled thermal treatments at 150°C, 175°C, and 200°C. Each spectral representation formed a separate 70-feature dataset. Five supervised classifiers were trained on the multiclass spectral data: logistic regression (LR), k-nearest neighbors (KNN), random forest (RF), support vector machine (SVM), and neural network (NN). Models were evaluated on an independent held-out test set (20% of the data), with five-fold cross-validation performed on the training set (80%) during hyperparameter optimization. Performance was assessed using accuracy, precision, recall, and F1-score, while uncertainty was quantified using nonparametric bootstrap confidence intervals.

Absorption coefficient spectra consistently provided stronger discriminatory information than refractive index spectra. For the absorption dataset, SVM achieved the highest performance, with an accuracy of 0.93 (95% CI: 0.82–1.00) and F1-score of 0.92 (95% CI: 0.80–1.00), followed closely by RF and NN. In contrast, refractive index models showed lower predictive performance and greater uncertainty, with RF achieving the highest F1-score of 0.77 (95% CI: 0.60–0.93). These results demonstrate that THz absorption is more sensitive to thermally induced molecular and microstructural changes in rubber. Overall, integrating THz-TDS absorption spectroscopy with machine learning offers a non-destructive and automated framework for rubber quality evaluation, industrial process monitoring, and intelligent quality-control applications.

*Corresponding author
Email Address: **mddelara@up.edu.ph**
Date received: 30 March 2026
Date revised: 11 August 2026; 22 August 2026
Date accepted: 24 August 2026
DOI: https://doi.org/10.54645/2026192CRD-25

INTRODUCTION

Rubber is a commonly used material known for its unique properties, including elasticity and durability, which are essential for a variety of applications, from automotive components to consumer products. Maintaining high performance and safety standards for rubber products requires reliable quality evaluation methods that can detect even subtle changes in material properties. One important approach is to assess how rubber reacts to thermal stress, as exposure to heat can induce molecular and structural changes that directly affect durability and functionality, especially in demanding service conditions where thermal and mechanical loads are high. (Tanaka et al., 2024).

Tanaka et al. (2024) demonstrated that strained natural rubber undergoes distinct anisotropic structural reorganization when subjected to high temperatures. Specifically, the main polymer chain direction (c-axis) contracts while the perpendicular directions expand, a behavior attributed to thermally activated torsional and rotational movements of the polymer chains. Evaluating rubber samples by analyzing their thermal behavior thus offers valuable insights into material integrity. Thermal analysis can reveal variations in physical properties that might compromise performance, thereby guiding manufacturers in quality control and process optimization. Various thermal analysis techniques, such as thermogravimetric analysis (TGA), differential scanning calorimetry (DSC), thermomechanical analysis (TMA), and dynamic mechanical analysis (DMA), have been widely used to characterize rubber's composition, oxidative stability, glass-transition temperature, and viscoelastic properties (Ginic-Markovic et al., 1998). These methods provide critical insights into the structural and mechanical behavior of elastomers under different temperature conditions, making these methods valuable tools for assessing rubber quality and performance in industrial applications. In this context, non-destructive techniques that

KEYWORDS

Terahertz Time-Domain Spectroscopy, THz-TDS, Non-Destructive Evaluation, Rubber Quality Control

capture these thermally induced chnages with high sensitivity are highly desirable.

Terahertz Time-Domain Spectroscopy (THz-TDS) has emerged as a powerful tool for material characterization and process monitoring. Amenabar et al. (2013) emphasized that recent breakthroughs in Terahertz (THz) technology have transformed non-destructive testing (NDT) by enabling robust, high-resolution imaging and spectroscopy methods. Their review details how THz systems offer non-invasive and non-contact inspection capabilities with high penetration depth and sub-millimeter resolution, which are crucial for detecting defects in composite materials. These advances have paved the way for automated and integrated inspection solutions that can capture both surface and internal features of complex materials. Although their work primarily focuses on composites, the underlying principles of THz NDT are highly relevant to rubber evaluation, where the ability to detect subtle molecular and structural changes induced by thermal stress can significantly enhance quality control. By measuring parameters such as THz refractive index and absorption coefficient, THz-TDS provides detailed information about the intrinsic properties of rubber. The THz refractive index measures how much the material slows the propagation of terahertz waves, reflecting the material's optical density and dielectric properties. The THz absorption coefficient describes how strongly a material absorbs terahertz radiation as it passes through the sample, providing information about the material's molecular composition and vibrational characteristics. Together, these complementary parameters provide valuable information for distinguishing and characterizing different materials. Jusi et al. (2023) demonstrated that THz-TDS and THz imaging effectively differentiate between thermally treated and non-thermally treated rubber samples. The capability of THz-TDS to detect subtle variations in material structure makes it an ideal method for assessing thermally induced quality changes.

To further improve the evaluation process, machine learning techniques have been incorporated with THz spectroscopy data. Machine learning algorithms such as logistic regression (LR), k-nearest neighbors (KNN), random forest (RF), support vector machine (SVM), and neural network (NN) can be used in processing complex datasets and identifying intricate patterns with minimal human intervention. In the rubber domain, Hu et al. (2024) demonstrated that integrating absorption and refractive index THz-TDS with data-fusion and supervised learning can yield high classification accuracy. Optimization-assisted pipelines have also been reported. Yin et al. (2018) combined absorption THz spectroscopy with principal component analysis (PCA) and a metaheuristic-tuned SVM (CS-SVM) for nondestructive identification of three rubber types. Moreover, Yin et al. (2022) used THz absorption spectra with a particle-swarm-optimized extreme learning machine to classify visually similar rubber vulcanizing accelerators, underscoring the applicability of THz-ML frameworks to both rubber identification and manufacturing quality control. Consistent with these findings, other studies have further validated the robustness of THz-based ML classification across different material systems and spectral datasets (Yilmaz et al., 2023; Cielecki et al., 2023). More recently, Yin et al. (2024) used Savitzky-Golay preprocessing, PCA, and an Improved Honey Badger Algorithm (IHBA)-optimized SVM to identify eight rubber types from THz absorption spectra, reporting 98.96% accuracy.

The rubber industry is an important sector in the Philippines, particularly in Mindanao, where continued efforts are directed toward improving processing efficiency and product quality (Castillo et al., 2014). The quality and consistency of commercially available rubber products are important to both producers and downstream users (Yin et al., 2024). While Yin et al. (2024) focused on differentiating rubber types from absorption spectra using an optimized SVM pipeline, the present study investigated the ability of THz spectroscopy and imaging to distinguish the thermal-treatment states of commercial rubber. Heating temperature was selected as the target class because thermal treatment can induce oxidation, chain scission, and crosslinking, thereby modifying the molecular structure and physical properties of rubber and potentially altering its terahertz spectral and imaging characteristics. Distinguishing such thermal-treatment states can therefore provide a basis for rapid screening of the processing history of commercial rubber materials. This study developed an automated classification framework that integrates THz-TDS with machine learning techniques. The approach exploits the sensitivity of THz radiation to low-frequency molecular and collective excitations, which can be reflected in changes in the absorption coefficient and refractive index of thermally treated rubber. The predictive performance of five machine learning classifiers was evaluated using these THz-derived parameters to determine their capability to distinguish different thermal-treatment states. The resulting framework demonstrates the potential of THz-TDS combined with machine learning as a rapid, non-contact, and non-destructive approach for screening the thermal history of commercial rubber.

MATERIALS AND METHODS

The overall workflow is summarized in Figure 1. Rubber samples subjected to three heating temperatures were measured by THz-TDS using the Menlo Systems TeraK15, from which the refractive index and absorption coefficient were derived. The two were treated as separate datasets thoughout rather that combined, and five classifiers were trained on each and evaluated on a held-out test set. Each stage is detailed in the subsections that follow.

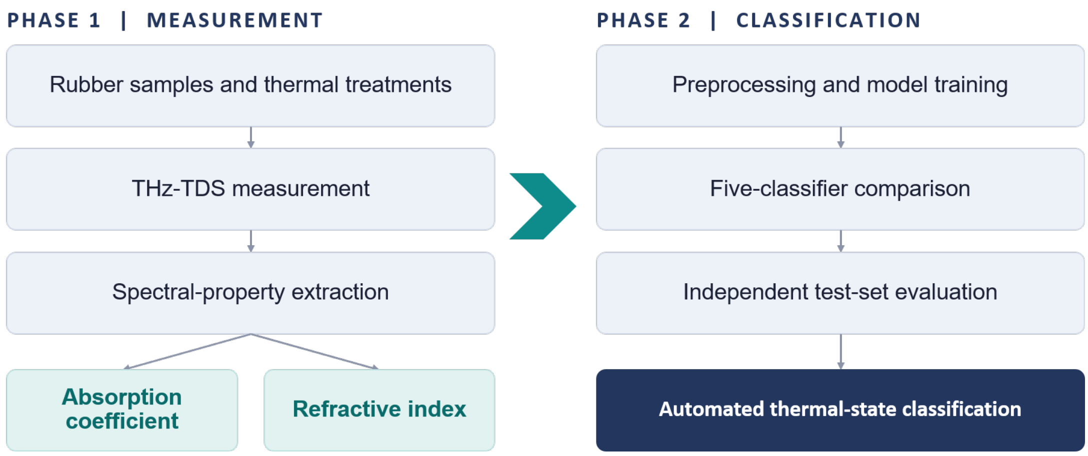

Figure 1: Study workflow. Rubber samples heated at three temperatures were measured by Terahertz Time-Domain Spectroscopy (THz-TDS), and the absorption coefficient and refractive index were analyzed as two separate datasets. Five classifiers were compared on each and assessed on a held-out test set.

Data Collection

The molecular structure of the commercially available rubber sample was verified by FTIR spectra using a Shimadzu IRspirit to confirm the molecular composition of the rubber. Figure 2 shows the FTIR spectra of the rubber sample. This region contains overlapping contributions from C–C–C skeletal twisting, out-of-plane bending, and backbone deformation vibrations (Chen et al., 2013). These vibrational modes are important for identifying alkene substitution patterns and backbone deformation in polyisoprene. Around 1350 cm⁻¹, the scissoring vibration of CH_3 groups supports the presence of methyl groups in polyisoprene. At approximately 1400 cm⁻¹, bending vibrations further confirmed the aliphatic nature of the polymer chains. The band near 1600 cm⁻¹ was attributed to C=C stretching, indicating unsaturation in the polyisoprene backbone (Chen et al., 2013). In the 2800–2900 cm⁻¹ region, absorption features associated with C–H stretching further supported the presence of double bonds and unsaturation in the rubber backbone.

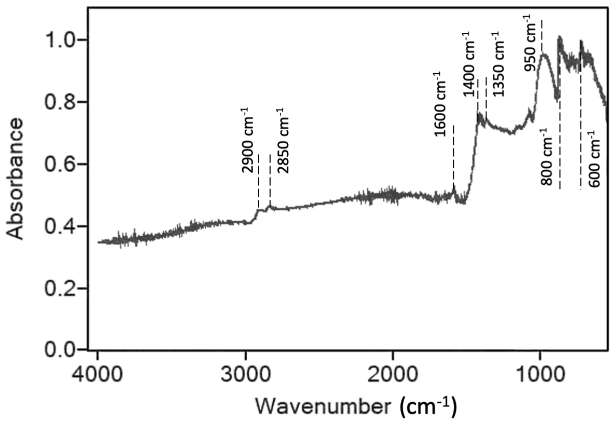

Figure 2: FTIR curve of representative rubber sample showing peak characteristics of rubber.

The samples were measured using THz spectroscopy and imaging. The main constituent, polyisoprene, was shown to exhibit a significantly stronger THz absorption response than those of typical additives and fillers in commercial rubbers (Hirakawa et al., 2020). This suggests that the measured THz response is dominated by the polyisoprene matrix, although contributions from fillers and other additives cannot be completely excluded.

The THz images were generated for the rubber samples heated at 150°C, 175°C, and 200°C. The three temperatures were selected to represent distinct stages of the material's thermal response, from relatively preserved structure to the onset of phase transition and, eventually, more pronounced thermal degradation. This range enables the evaluation of changes in the material's structural and spectral properties resulting from thermal treatment. Heating was done using a laboratory furnace, which was programmed to maintain the target temperature for a duration of ten minutes. For each THz image, 15 points were randomly selected within the region of the THz image corresponding to the rubber sample.

Figure 3 shows a THz scan for a sample heated at 175°C, showing darker patches that may represent areas with physical and/or chemical changes. From these points, the THz spectra, $E_{sam,T}(\nu)$, of the rubber samples subjected to a temperature T (150°C, 175°C, and 200°C) were determined as functions of frequency ν. The THz spectra of a point corresponding to air were used as the reference, $E_{ref,T}(\nu)$. From these spectra at a temperature T, the refractive index $n_T(\nu)$, shown in Figure 4, was calculated using Equation 1, as described in Jusi et al. (2023):

$$n_T(\nu) = 1 + \frac{c}{2\pi\nu d_T}\phi_T(\nu) \tag{1}$$

where c corresponds to the speed of light ($c \approx 3.0 \times 10^8\ m/s$), d_T is the thickness of the sample subjected to a temperature T, and $\phi_T(\nu)$ is the phase difference between $E_{sam,T}(\nu)$ and $E_{ref,T}(\nu)$. The absorption coefficient $\alpha_T(\nu)$ at every temperature T, shown in Figure 5, was determined using Equation 2 (Jusi et al., 2023):

$$\alpha_T(\nu) = -\frac{2}{d_T} ln\left[\frac{E_{sam,T}(\nu)}{E_{ref,T}(\nu)}\frac{(n_T(\nu)+1)^2}{4n_T(\nu)}\right]. \tag{2}$$

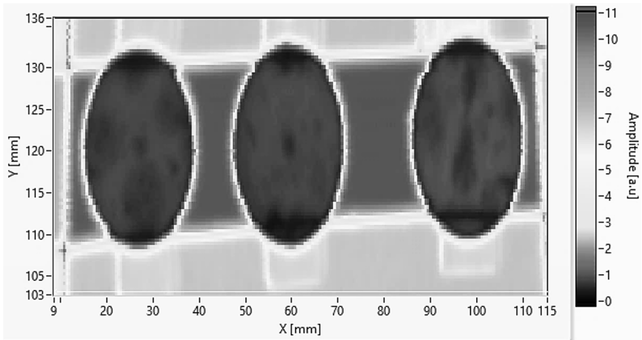

Figure 3: THz image scan for the samples heated at 175°C. For each target heating temperature (150°C, 175°C, and 200°C), three samples were heated. For each sample, 15 spectra were measured from 15 points within each sample, yielding a total of 45 spectra per temperature.

Each spectrum was sampled at 70 distinct frequency points within the range of ν = 0.2–2.0 THz. From every measured spectrum, both the refractive index $n_T(\nu)$ and absorption coefficient $\alpha_T(\nu)$ were calculated; however, they were organized and analyzed as two separate datasets. The refractive index dataset contained 70 $n_T(\nu)$ features per observation, whereas the absorption coefficient dataset contained 70 $\alpha_T(\nu)$ features per observation. Thus, every model input was 70-dimensional and contained either $n_T(\nu)$ values or $\alpha_T(\nu)$ values, not both. The frequency range was selected based on the operational bandwidth and signal-to-noise ratio (SNR) characteristics of the THz-TDS system. Although the instrument provides broadband coverage beyond 6 THz, the SNR in the present configuration remained acceptable only up to 2.0 THz. The analysis was therefore restricted to 0.2–2.0 THz, yielding 70 frequency points with robust SNR. Three rubber samples were prepared at each target temperature, and 15 spectra were obtained from each sample, giving 45 observations per temperature. Consequently, the analysis used two separate data matrices, each with 135 observations and 70 frequency-based features.

Data Preprocessing

The predictor and target variables were identified, and the spectral datasets were preprocessed as described below. The target variable represents the heating temperatures (150°C, 175°C, and 200°C). The two spectral datasets were analyzed separately. For the refractive index dataset, the predictors were the 70 $n_T(\nu)$ values measured at the fixed frequency points from 0.2 to 2.0 THz. For the absorption coefficient dataset, the predictors were the corresponding 70 $\alpha_T(\nu)$ values. Thus, each dataset contained 135 spectral observations, with each observation represented by 70 frequency-based features. The heating-temperature classes were represented by integer labels, with 150°C, 175°C, and 200°C assigned the labels 0, 1, and 2, respectively. These classes refer to specific thermal conditions associated with probing the thermal history of the material. These integers were used only as nominal class identifiers for compatibility with the machine learning algorithms.

Before model training, the datasets were partitioned into training (80%) and testing sets (20%) using stratified sampling to preserve class proportions. Each feature was standardized via z-score normalization:

$$z_{ij} = \frac{x_{ij} - \mu_j}{\sigma_j},$$

where x_{ij} is the observed value of the jth feature for the ith sample, while μ_j and σ_j denote the mean and the standard deviation of the jth feature, respectively. The scaling parameters (μ_j and σ_j) were fitted on the training set and subsequently applied to the test set to prevent data leakage.

Training of the Classification Models

Five supervised machine learning classifiers were trained separately for the absorption coefficient and refractive index datasets: LR, KNN, RF, SVM, and NN. Hyperparameter optimization was performed exclusively on the training subset using grid search with five-fold cross-validation, with classification accuracy as the optimization criterion. Classification accuracy is defined here as the proportion of samples whose predicted thermal state matches the true thermal state. During grid search, this proportion was computed on each held-out validation fold, and the mean cross-validation accuracy was taken as the arithmetic average of the five resulting values. The optimal hyperparameter configuration for each classifier was selected as the one attaining the highest mean cross-validation accuracy, and this configuration was then used for final model performance evaluation on the held-out test subset. The hyperparameter search spaces explored for each classifier are summarized in Table 1.

Table 1: Hyperparameter search spaces and optimal configurations of the machine learning classifiers for the absorption coefficient and refractive index THz spectral datasets.

Model*	Hyperparameters Explored	Optimal Hyperparameters (Absorption Dataset)	Optimal Hyperparameters (Refractive Index Dataset)
Logistic Regression (LR)	C ∈ {0.001, 0.1, 1, 10, 100, 1000} Solver ∈ {newton-cg, lbfgs, liblinear, sag, saga}	C = 1000; Solver = newton-cg	C = 1000; Solver = newton-cg
k-Nearest Neighbors (KNN)[a]	n_neighbors ∈ {1, 2, …, 50}	n_neighbors = 1	n_neighbors = 1
Random Forest (RF)	n_estimators ∈ {50, 100, 150, 200, 250, 300} max_depth ∈ {None, 10, 20, 30, 40} min_samples_split ∈ {2, 4, 6} max_features ∈ {sqrt, log2} criterion ∈ {gini, entropy}	criterion = gini; max_depth = None; max_features = sqrt; min_samples_split = 6; n_estimators = 200	criterion = gini; max_depth = None; max_features = sqrt; min_samples_split = 4; n_estimators = 100
Support Vector Machine (SVM)	C ∈ {0.1, 1, 10, 100} kernel ∈ {linear, rbf} γ ∈ {0.1, 0.01, 0.001}	C = 10; γ = 0.01; kernel = rbf	C = 100; γ = 0.1; kernel = rbf
Neural Network (NN)	hidden_layer_sizes ∈ {(i), (i, i)}, i = 1,…,20 activation ∈ {tanh, relu} solver = adam α_reg ∈ {0.0001, 0.01}	activation = relu; α_reg = 0.0001; hidden_layer_sizes = (20, 20); solver = adam	activation = tanh; α_reg = 0.0001; hidden_layer_sizes = (15, 15); solver = adam

*Based on grid search using the scikit-learn framework (Pedregosa et al., 2011). Hyperparameter names follow that implementation, and any hyperparameter not listed was retained at its default value. [a]Used the standard Euclidean distance metric. Hyperparameter definitions. *Logistic Regression:* C, inverse of the regularization strength, with smaller values imposing a stronger penalty on the magnitude of the fitted coefficients; Solver, the numerical optimization algorithm used to minimize the regularized loss. *k-Nearest Neighbors:* n_neighbors, the number *k* of nearest training samples consulted when classifying a new sample, whose label is assigned by majority vote among those neighbors. *Random Forest:* n_estimators, the number of decision trees in the ensemble; max_depth, the maximum permitted depth of each tree, where None allows growth until all leaves are pure or contain fewer samples than min_samples_split; min_samples_split, the minimum number of samples an internal node must contain to be eligible for further splitting; max_features, the number of features randomly considered at each split, expressed as the square root or base-2 logarithm of the total number of features; criterion, the impurity measure used to score candidate splits. *Support Vector Machine:* C, the penalty applied to misclassified training samples, controlling the trade-off between a wide decision margin and training-set fidelity; kernel, the function defining the similarity between samples and hence the geometry of the decision boundary; γ, the coefficient of the radial basis function kernel, governing the effective radius of influence of a single training sample. *Neural Network:* hidden_layer_sizes, the hidden-layer architecture, where (*i*) denotes one hidden layer of *i* neurons and (*i*, *i*) two hidden layers of *i* neurons each; activation, the nonlinear activation function applied at the hidden neurons; solver, the algorithm used to optimize the network weights; α_reg, the strength of the L2 penalty applied to the network weights.

Test Set Evaluation and Performance Metrics

After hyperparameter tuning, each optimized classifier was evaluated on the independent held-out test set. Model performance was assessed using accuracy, precision, recall, and F1-score. All metrics were computed using macro-averaging, which assigns equal weight to each heating-temperature class (150°C, 175°C, and 200°C), ensuring that model performance was not dominated by any single class. Accuracy measures overall classification correctness, precision quantifies the reliability of predicted class labels, recall assesses sensitivity to true class instances, and the F1-score provides a balanced summary of precision and recall (Sokolova and Lapalme, 2009).

Bootstrap-Based Statistical Analysis and Confidence Intervals

To quantify uncertainty in the test set performance estimates, nonparametric bootstrap resampling was employed (Efron and Tibshirani, 1993). Each optimized classifier was fitted once on the training set and used to predict the labels of the fixed independent test set. The resulting pairs of observed and predicted labels were then resampled with replacement 1,000 times, and accuracy, precision, recall, and F1-score were recomputed for each resample. Percentile-based 95% confidence intervals were obtained from the 2.5th and 97.5th percentiles of the resulting distributions. The intervals thus reflect variability attributable to test-set composition for a fixed trained model, and do not incorporate variability from retraining or from the train–test partition itself.

To complement the confidence-interval analysis, a Nemenyi post-hoc test was performed to assess pairwise differences in classifier performance across repeated cross-validation runs.

Model Implementation and Reproducibility

All analyses were implemented in Python (version 3.12) using the scikit-learn library (Pedregosa et al., 2011). Numerical computation and data handling were performed using Python libraries NumPy and pandas, while Matplotlib was used for visualization. Bootstrap resampling was implemented using utilities from scikit-learn. A fixed random seed was used for data splitting, model initialization, and bootstrap resampling to ensure full reproducibility of the results. The Python programs used for model training, hyperparameter optimization, performance evaluation, and statistical analysis are publicly available in the GitHub repository at https://github.com/anazareno2988/THz-Rubber-ML.

RESULTS AND DISCUSSION

Spectral Characteristics and Information

Figures 4 and 5 show the refractive index and absorption coefficient, respectively, of the samples heated at different temperatures. The THz refractive index and absorption coefficient characterize the phase delay and energy loss of THz waves in a material, respectively. Visual inspection revealed only modest differences in spectral magnitude and shape across temperature classes, and no well-defined absorption peaks readily associated with specific molecular vibrational modes were apparent. This behavior is a characteristic of rubber and other disordered polymeric materials in the THz frequency range. As noted by Wietzke et al. (2011), THz absorption in polymers typically follows a smooth power law or broadband response rather than exhibiting discrete resonances as seen in Figure 5 from around 0.2 to 1 THz. This behavior arises from the lack of long-range molecular order, resulting in a continuum of low-frequency excitations associated with chain segment motions, hindered rotations, and inter- and intra-molecular fluctuations as discussed in Wietzke et al. (2009). Consequently, discriminative information related to thermal treatment is distributed across the spectral bandwidth and is not readily accessible through simple visual or univariate analysis.

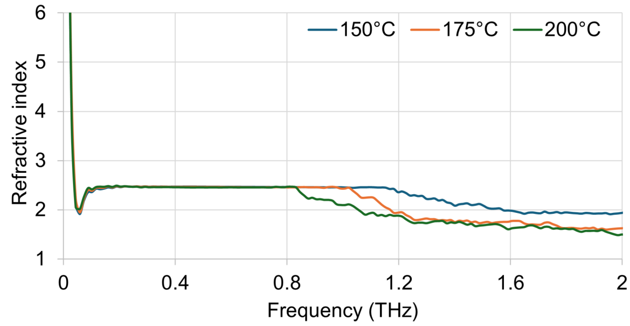

Figure 4: Refractive index of the samples heated at different temperatures. From the 45 spectra collected at each target heating temperature, one representative spectrum was selected to illustrate the corresponding change in the refractive index induced by heating.

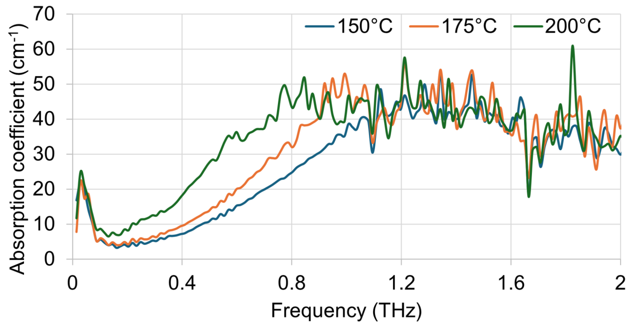

Figure 5: Absorption coefficient of the samples heated at different temperatures. From the 45 spectra collected at each target heating temperature, one representative spectrum was selected to illustrate the change in the absorption of the samples associated with heating.

Table 2: Mean classification performance with 95% bootstrap confidence intervals of five machine learning models on the refractive index THz spectral dataset.

Model	Accuracy (95% CI)	Precision (95% CI)	Recall (95% CI)	F1-score (95% CI)
LR	0.558 (0.370–0.741)	0.556 (0.379–0.737)	0.556 (0.366–0.735)	0.544 (0.374–0.714)
KNN	0.594 (0.407–0.778)	0.599 (0.381–0.806)	0.596 (0.418–0.762)	0.565 (0.386–0.740)
RF	0.780 (0.630–0.926)	0.801 (0.633–0.933)	0.780 (0.620–0.926)	0.770 (0.597–0.925)
SVM	0.668 (0.481–0.815)	0.674 (0.486–0.849)	0.668 (0.489–0.843)	0.652 (0.467–0.824)
NN	0.742 (0.556–0.889)	0.753 (0.571–0.914)	0.741 (0.568–0.900)	0.727 (0.553–0.888)

LR, logistic regression; KNN, k-nearest neighbors; RF, random forest; SVM, support vector machine; NN, neural network; confidence intervals (CIs).

Performance on Refractive Index Data

Classification performance on the refractive index dataset (Table 2) was substantially lower across all models compared with the absorption coefficient results. The RF achieved the highest overall accuracy of 0.780 (95% CI: 0.630, 0.926) and F1-score of 0.770 (95% CI: 0.597, 0.925), followed by the NN with an accuracy of 0.742 (95% CI: 0.556, 0.889) and F1-score of 0.727 (95% CI: 0.553, 0.888). The SVM and KNN models exhibited moderate performance, while LR showed the lowest performance.

The broader overlap of CIs among models reflects reduced class separability when using refractive index features alone. Unlike absorption coefficients, refractive index primarily captures changes in phase velocity and bulk dielectric response, which tend to vary more smoothly with thermal treatment (Wietzke et al., 2011). As a result, differences between temperature classes were less pronounced, leading to greater overlap in feature space and thereby reducing the ability of the models to discriminate among classes.

Performance on Absorption Coefficient Data

For the absorption coefficient dataset (Table 3), all five classifiers achieved relatively strong predictive performance. The SVM attained the highest mean accuracy of 0.926 (95% CI: 0.815–1.000) and F1-score of 0.919 (95% CI: 0.798–1.000), followed closely by the RF model with an accuracy of 0.889 (95% CI: 0.778–1.000) and F1-score of 0.882 (95% CI: 0.739–1.000). The NN also demonstrated comparable performance, achieving an accuracy of 0.886 (95% CI: 0.741–1.000) and an F1-score of 0.879 (95% CI: 0.741–1.000). In contrast, LR and KNN yielded lower metric values, though their performance remained above chance level. The relatively wide CIs for LR and KNN suggest greater sensitivity to test-set sampling variability.

The comparatively high test-set accuracy and F1-scores obtained from the absorption-coefficient dataset indicate that these spectral features contain information useful for distinguishing among the three heating-temperature classes. In particular, SVM, RF, and NN achieved higher performance than KNN on this dataset, with SVM obtaining the highest accuracy and F1-score. These results suggest that these models can use the differences in the absorption-coefficient values for temperature classification. However, the performance metrics alone do not identify the exact physical reasons for the class differences.

Table 3: Mean classification performance with 95% bootstrap confidence intervals of five machine learning models on the absorption coefficient THz spectral dataset.

Model	Accuracy (95% CI)	Precision (95% CI)	Recall (95% CI)	F1-score (95% CI)
LR	0.853 (0.704–0.963)	0.857 (0.701–0.972)	0.853 (0.708–0.970)	0.842 (0.685–0.967)
KNN	0.776 (0.593–0.926)	0.770 (0.594–0.933)	0.777 (0.628–0.923)	0.759 (0.593–0.917)
RF	0.889 (0.778–1.000)	0.903 (0.788–1.000)	0.889 (0.756–1.000)	0.882 (0.739–1.000)
SVM	0.926 (0.815–1.000)	0.933 (0.833–1.000)	0.926 (0.815–1.000)	0.919 (0.798–1.000)
NN	0.886 (0.741–1.000)	0.898 (0.776–1.000)	0.886 (0.762–1.000)	0.879 (0.741–1.000)

LR, logistic regression; KNN, k-nearest neighbors; RF, random forest; SVM, support vector machine; NN, neural network; confidence intervals (CIs).

Comparative Analysis Between Datasets

Absorption coefficient-based models consistently exhibited higher mean accuracy scores and F1-scores, and generally narrower CIs, as shown in Figure 6, indicating greater statistical robustness. In contrast, refractive index-based models showed lower predictive performance and higher uncertainty, particularly for LR and KNN classifiers.

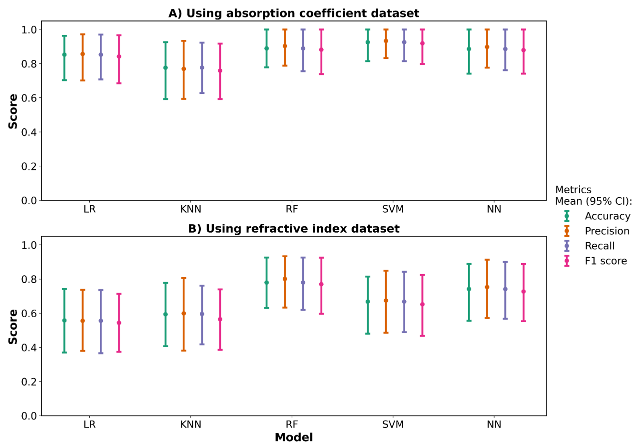

Figure 6: Error-bar plots showing mean classification performance with 95% bootstrap confidence intervals (CIs) for five machine learning models evaluated on (A) absorption coefficient and (B) refractive index THz spectral datasets. Points denote mean metric values, and error bars represent confidence intervals estimated via nonparametric bootstrap resampling of the independent test set. Metrics shown include accuracy, precision, recall, and F1-score, macro-averaged across three heating-temperature classes.

The results of the present study are consistent with earlier work showing that THz measurements can detect differences in rubber materials. Hirakawa et al. (2011) showed that THz absorption is sensitive to differences in rubber compounds and additives, while Peters et al. (2013) demonstrated the use of THz-TDS for monitoring material changes during rubber production. Hu et al. (2024) further showed that absorption, refractive-index, and time-domain data can be used to classify different rubber types. Similarly, Yin et al. (2018, 2024) used absorption-derived THz spectra with optimized SVM-based models for rubber identification. Unlike these studies, the present work focuses on distinguishing the heating conditions of commercial rubber and directly compares absorption coefficient and refractive index as separate inputs across five classifiers. RF performed well in both datasets, consistent with previous studies on the use of ensemble methods for classification, particularly with high-dimensional data (Ahn et al., 2007; Shaik and Srinivasan, 2018). The higher performance of SVM and NN on the absorption-coefficient dataset suggests that these models can use differences across the absorption-coefficient features for temperature classification.

A combined representation was also explored by concatenating the 70 absorption-coefficient features and 70 refractive-index features into a 140-dimensional input vector. However, the resulting classification performance was generally intermediate between those obtained from the two separate datasets and did not exceed the performance of the absorption-coefficient dataset alone. This indicates that the refractive-index features provided limited additional discriminative information beyond that already captured by the absorption coefficients. Consequently, the main analysis focused on the separate spectral representations to more clearly assess their individual predictive value.

Post-hoc Statistical Comparison of Classifiers

The Nemenyi post-hoc results indicated that for the absorption coefficient dataset, most model pairs did not differ significantly, except for RF vs SVM ($p = 0.0059$), SVM vs NN ($p = 0.0006$), and KNN vs NN ($p = 0.0463$), which showed significant differences. For the refractive index-based dataset, nearly all model pairs differed significantly, except for KNN vs SVM, KNN vs NN, and SVM vs NN, which did not differ significantly. These findings indicate that absorption coefficient-based features yielded more stable and consistent performance across different classifiers, while refractive index features accentuated differences between models, making the choice of classifier more critical when using refractive index-based representations of THz spectral data. Detailed pairwise p-values are provided in Tables 4 and 5.

Table 4: Nemenyi post-hoc test results for the absorption coefficient dataset.

	LR	KNN	RF	SVM	NN
LR	1.0000	0.8744	0.2420	0.1138	0.0666
KNN	0.8744	1.0000	0.1841	0.1547	0.0463
RF	0.2420	0.1841	1.0000	0.0059	0.5066
SVM	0.1138	0.1547	0.0059	1.0000	0.0006
NN	0.0666	0.0463	0.5066	0.0006	1.0000

LR, logistic regression; KNN, k-nearest neighbors; RF, random forest; SVM, support vector machine; NN, neural network.

Table 5: Nemenyi post-hoc test results for the refractive index dataset.

	LR	KNN	RF	SVM	NN
LR	1.0000	<0.0001	<0.0001	<0.0001	0.0001
KNN	<0.0001	1.0000	<0.0001	1.0000	0.1948
RF	<0.0001	<0.0001	1.0000	<0.0001	<0.0001
SVM	<0.0001	1.0000	<0.0001	1.0000	0.1948
NN	0.0001	0.1948	<0.0001	0.1948	1.0000

LR, logistic regression; KNN, k-nearest neighbors; RF, random forest; SVM, support vector machine; NN, neural network.

Implications for Rubber Quality Assessment

From an application standpoint, these findings have direct implications for non-destructive rubber quality evaluation. The consistently better performance achieved using absorption coefficient data indicates that this spectral parameter more effectively captures thermally induced changes relevant to material integrity. Since thermal exposure is closely linked to changes in mechanical properties, durability, and aging behavior, reliable classification of heating conditions provides a meaningful proxy for quality assessment.

The integration of THz-TDS with machine learning classifiers offers a rapid, contact-free, and automated framework suitable for industrial inspection. In practice, absorption-based models, particularly SVM, RF, and NN, could be deployed for rapid screening of rubber products, enabling early detection of thermal treatment-induced changes or process deviations. While refractive index measurements may still provide complementary information, their lower discriminative power suggests that they are better suited as auxiliary features rather than standalone predictors.

CONCLUSION

In conclusion, this study presents a THz-TDS classification framework, rigorously evaluated using statistical methods, for distinguishing the thermal-treatment states of commercial rubber. The results demonstrate that rubber exposed to 150°C, 175°C, and 200°C can be differentiated using frequency-resolved THz measurements. A controlled, side-by-side comparison further shows that the absorption coefficient provides more reliable discriminatory information than the refractive index for this task. The study also benchmarks five classifier families using independent test evaluation, bootstrap 95% confidence intervals, and post-hoc comparisons, rather than relying solely on the best accuracy obtained from a single optimized pipeline. Among the absorption-based models, SVM achieved the strongest performance, with an accuracy of 0.926 and an F1-score of 0.919, whereas RF performed best on the refractive index data. Rather than proposing a new optimization algorithm, the present work extends THz-based rubber classification to process-state assessment. Instead, it extends THz-based rubber classification to process-state assessment and identifies the spectral representation and model families most suitable for detecting heating-induced changes. Thermal treatment is a critical processing parameter because it influences the crosslink network and polymer structure of rubber. Therefore, accurately distinguishing thermal-treatment states provides an indirect but practical indicator of process quality. While this study did not directly evaluate mechanical properties, the proposed THz-TDS and machine learning framework demonstrates the potential for rapid, non-destructive verification of thermal processing conditions relevant to rubber quality control. Although the dataset is relatively small and limited to five classifiers and three heating conditions, the results support the use of absorption-based THz-TDS with nonlinear machine learning for rapid, non-destructive process monitoring and quality assessment.

ACKNOWLEDGMENTS

The authors would like to thank the Department of Science and Technology-Philippine Council for Industry, Energy, and Emerging Technology Research and Development, University of the Philippines Emerging Interdisciplinary Research Program and the JSPS Core-to-Core Program's Asian Network for Terahertz Molecular Science for the financial support.

CONFLICT OF INTEREST

All authors reported no known potential conflicts of interest.

CONTRIBUTIONS OF INDIVIDUAL AUTHORS

De Lara: Conceptualization, Project administration, Formal analysis, Writing - Review and Editing, Methodology, Software, Visualization. **Nazareno**: Conceptualization, Methodology, Supervision, Writing - Review and Editing, Formal analysis, Software, Validation. **Barcarse**: Conceptualization, Methodology, Writing - Original Draft. **Estacio**: Data Curation, Resources. **Jusi**: Data Curation, Writing - Review and Editing, Visualization, and Validation. **Tapia**: Conceptualization, Data Curation, Methodology, Supervision, Writing - Original Draft, Writing - Review and Editing. All authors discussed the results and contributed to the final manuscript. All authors have read and agreed to the published version of the manuscript.

REFERENCES

Ahn H, Moon H, Fazzari MJ, Lim N, Chen JJ, Kodell RL. Classification by ensembles from random partitions of high-dimensional data. Computational Statistics & Data Analysis 2007; 51(12):6166–6179. https://doi.org/10.1016/j.csda.2006.12.043

Amenabar I, Lopez F, Mendikute A. An introductory review to THz non-destructive testing of composite materials. Journal of Infrared, Millimeter, and Terahertz Waves 2013; 34(2):152–169. https://doi.org/10.1007/s10762-012-9949-z

Castillo AMA, Furoc-Paelmo R, Retes LP, Maloles GM, Sanfuego AD, Carranza CP. Production, processing, and marketing of rubber in Laguna province, Philippines. Ecosyst Dev J. 2014;4(2). https://journals.uplb.edu.ph/index.php/EDJ/article/view/198

Chen D, Shao H, Yao W, Huang B. Fourier transform infrared spectral analysis of polyisoprene of a different microstructure. International Journal of Polymer Science 2013; 2013:1–5. https://doi.org/10.1155/2013/937284

Cielecki PP, Hardenberg M, Amariei G, Henriksen ML, Hinge M, Klarskov P. Identification of black plastics with terahertz time-domain spectroscopy and machine learning. Scientific Reports 2023; 13(1):22399. https://doi.org/10.1038/s41598-023-49765-z

Efron B, Tibshirani RJ. An Introduction to the Bootstrap. New York: Chapman & Hall/CRC, 1993. https://doi.org/10.1201/9780429246593

Ginic-Markovic M, Choudhury NR, Dimopoulos M, Williams DR, Matisons J. Characterization of elastomer compounds by thermal analysis. Thermochimica Acta 1998; 316(1):87–95. https://doi.org/10.1016/S0040-6031(98)00290-1

Hirakawa Y, Ohno Y, Gondoh T, Mori T, Takeya K, Tonouchi M, et al. Nondestructive evaluation of rubber compounds by terahertz time-domain spectroscopy. Journal of Infrared, Millimeter, and Terahertz Waves 2011; 32(12):1457–1463. https://doi.org/10.1007/s10762-011-9832-3

Hirakawa Y, Yasumoto Y, Gondo T, Sone R, Morichika T, Minato T, Hojo M. Application of terahertz spectroscopy to rubber products: Evaluation of vulcanization and silica macro dispersion. Electronics 2020; 9(4):669. https://doi.org/10.3390/electronics9040669

Hu J, Liu W, Yang L, Lv H, Zhan C, Qiao P. Research on classification methods for rubber based on terahertz time-domain spectroscopy with data fusion strategy. Infrared Phys Technol. 2024;139:105324. https://doi.org/10.1016/j.infrared.2024.105324

Jusi ALC, Castrosanto MA, Rocha AJD, Ramoso JPA, Estacio E, Tapia AKG. THz time-domain spectroscopy and imaging of heated commercialized rubber. Materials Today: Proceedings 2023. https://doi.org/10.1016/j.matpr.2023.09.016

Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grobler J, Blondel M, Prettenhofer P, Weiss R, Dubourg V, VanderPlas J, Passos A, Cournapeau D, Brucher M, Perrot M, Duchesnay E. Scikit-learn: Machine learning in Python. Journal of Machine Learning Research 2011; 12:2825–2830.

Peters O, Schwerdtfeger M, Wietzke S, Sostmann S, Scheunemann R, Wilk R, et al. Terahertz spectroscopy for rubber production testing. Polymer Testing 2013; 32(5):932–936. https://doi.org/10.1016/j.polymertesting.2013.05.003

Shaik AB, Srinivasan S. A brief survey on random forest ensembles in classification model. In: Lecture Notes in Networks and Systems. 2018:253–260. https://doi.org/10.1007/978-981-13-2354-6_27

Sokolova M, Lapalme G. A systematic analysis of performance measures for classification tasks. Information Processing & Management 2009; 45(4):427–437. https://doi.org/10.1016/j.ipm.2009.03.002

Tanaka R, Yasui T, Kitamura Y, Tsunoda K, Takagi H, Shimizu N, Igarashi N, Masunaga H, Urayama K, Sakurai S. Application of a strained natural rubber at high temperatures. ACS Applied Polymer Materials 2024; 6(5):2799–2806. https://doi.org/10.1021/acsapm.3c02969

Wietzke S, Jansen C, Jung T, Reuter M, Baudrit B, Bastian M, Chatterjee S, Koch M. Terahertz time-domain spectroscopy as a tool to monitor the glass transition in polymers. Optics Express 2009; 17:19006–19014. https://doi.org/10.1364/OE.17.019006

Wietzke S, Jansen C, Reuter M, Jung T, Kraft D, Chatterjee S, et al. Terahertz spectroscopy on polymers: A review of morphological studies. Journal of Molecular Structure 2011; 1006(1–3):41–51. https://doi.org/10.1016/j.molstruc.2011.07.036

Yilmaz VS, Eseller KE, Aslan O, Bayraktar E. Classification of different recycled rubber-epoxy composite based on their hardness using laser-induced breakdown spectroscopy (LIBS) with comparison machine learning algorithms. Inventions 2023; 8(2):54. https://doi.org/10.3390/inventions8020054

Yin X, Mo W, Wang Q, Qin B. A terahertz spectroscopy nondestructive identification method for rubber based on CS-SVM. Advances in Condensed Matter Physics 2018; 2018:1618750. https://doi.org/10.1155/2018/1618750

Yin X, He W, Wang L, Mo W, Li A. Classification of rubber vulcanizing accelerators based on particle swarm optimization extreme learning machine and terahertz spectra. Journal of Applied Spectroscopy 2022; 88:1315–1323. https://doi.org/10.1007/s10812-022-01313-9

Yin X, Zhang F, Luo Y, Mo W. Research on rubber classification and recognition based on terahertz time-domain spectroscopy and improved honey badger algorithm. Optik 2024; 315:172014. https://doi.org/10.1016/j.ijleo.2024.172014

SUPPORTING INFORMATION

Definition of Performance Metrics

Model performance was assessed using accuracy, precision, recall, and F1-score. Let $k = 0, 1, 2$ correspond to the heating-temperature classes 150°C, 175°C, and 200°C, respectively. For each class k, the class was treated as the positive class and the remaining classes as negative, where TP_k, FP_k, and FN_k denote the numbers of true-positive, false-positive, and false-negative predictions, respectively. Accuracy was defined as the proportion of correctly classified observations:

$$Accuracy = \frac{number\ of\ correctly\ classified\ observations}{total\ number\ of\ observations}.$$

The precision and recall for class $k$were defined as

$$P_k = \frac{TP_k}{TP_k + FP_k}\ and\ R_k = \frac{TP_k}{TP_k + FN_k},$$

respectively. Precision measures the proportion of observations predicted as class $k$that truly belong to that class, whereas recall measures the proportion of observations belonging to class $k$that were correctly identified. The class-specific F1-score was calculated as the harmonic mean of precision and recall:

$$F1_k = \frac{2P_kR_k}{P_k + R_k}.$$

Because this study involved multiclass classification, precision, recall, and F1-score were reported using macro-averaging:

$$P_{macro} = \frac{1}{K}\sum_{k=1}^{K} P_k,\ R_{macro} = \frac{1}{K}\sum_{k=1}^{K} R_k, F1_{macro} = \frac{1}{K}\sum_{k=1}^{K} F1_k.$$

Macro-averaging assigns equal importance to the 150°C, 175°C, and 200°C classes, regardless of the number of observations in each class, and therefore prevents the reported performance from being dominated by any single class (Sokolova and Lapalme, 2009).