



Article

Machine Learning Approaches for Predicting the Ablation Performance of Ceramic Matrix Composites

Jayanta Bhusan Deb , Jihua Gou ^{*}, Haonan Song and Chiranjit Maiti

Composite Materials and Structures Laboratory, Department of Mechanical and Aerospace Engineering, University of Central Florida, Orlando, FL 32816, USA; ja143279@ucf.edu (J.B.D.); ha781243@ucf.edu (H.S.); chiranjit.maiti@ucf.edu (C.M.)

^{*} Correspondence: jihua.gou@ucf.edu

Abstract: Materials used in aircraft engines, gas turbines, nuclear reactors, re-entry vehicles, and hypersonic structures are subject to severe environmental conditions that present significant challenges. With their remarkable properties, such as high melting temperatures, strong resistance to oxidation, corrosion, and ablation, minimal creep, and advantageous thermal cycling behavior, ceramic matrix composites (CMCs) show great promise as a material to meet the strict requirements in these kinds of environments. Furthermore, the addition of boron nitride nanoparticles with continuous fibers to the CMCs can offer thermal resistivity in harsh conditions, which will improve the composites' strength and fracture toughness. Therefore, in extreme situations, it is crucial to understand the thermal resistivity period of composite materials. To forecast the ablation performance of composites, we developed six machine learning regression methods in this study: decision tree, random forest, support vector machine, gradient boosting, extreme gradient boosting, and adaptive boosting. When evaluating model performance using metrics including R2 score, root mean square error, mean absolute error, and mean absolute percentage error, the gradient boosting and extreme gradient boosting machine learning regression models performed better than the others. The effectiveness of machine learning models as a useful tool for forecasting the ablation behavior of ceramic matrix composites was effectively explained by this study.

Keywords: ceramic matrix composites; oxyacetylene torch test; machine learning; regression models; extreme gradient boosting; gradient boosting



Citation: Deb, J.B.; Gou, J.; Song, H.; Maiti, C. Machine Learning Approaches for Predicting the Ablation Performance of Ceramic Matrix Composites. *J. Compos. Sci.* **2024**, *8*, 96. <https://doi.org/10.3390/jcs8030096>

Academic Editor: Francesco Tornabene

Received: 27 January 2024

Revised: 25 February 2024

Accepted: 1 March 2024

Published: 5 March 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The development of composite materials suitable for ultra-high-temperature (UHT) applications is currently the focus of much attention in the scientific community. HfB₂, ZrB₂, HfC, and ZrC are examples of transition metal borides and carbides that are particularly important because of their exceptionally high melting temperatures—which frequently approach or surpass 3000 °C—and their capacity to produce stable, oxidation-resistant compounds [1–3]. The materials known as ultra-high-temperature ceramics (UHTCs) are being researched in great detail for applications that require resistance to oxidation and erosion at temperatures higher than 2000 °C [4–6]. The application of UHTCs as a potential thermal protective covering has the potential to improve carbon-based composites' ablation resistance in harsh settings when exposure to oxygen and extremely high temperatures are present [7–9]. This is explained by the remarkable qualities of the UHTC, which include good modulus, enhanced hardness, high melting point, and remarkable chemical inertness and ablation resistance [10–13]. Their prospective applications range from cutting edge aerospace vehicle components, such as control surfaces, engine inlets and exits, and hot flow route components, to creative thermal protection systems [14–16]. The shortcomings of single-phase ceramics, on the other hand, like their low resistance to heat shock and defects, have led to research on substitute strategies. These materials do not have the

required high-temperature, thermal shock, or fracture toughness properties, even after adding phases like SiC [17]. This indicated that a fiber-reinforced composite strategy is a promising path to explore. Carbon fiber (Cf) and silicon carbide (SiC) fiber are potential candidates if they can provide adequate protection at high application temperatures.

Traditionally, the melting infiltration process is used to create ceramic matrix composites (CMCs). The brittle structure and excessive porosity of the fabricated CMCs made using this process make them incapable of withstanding high mechanical and thermal demands. As an alternative, carbon fibers are added to a ceramic matrix generated from polymers to create polymer-derived ceramic composites with exceptional fracture toughness. Carbon fiber may be able to tolerate high temperatures without oxidizing with the use of protective coatings made of metallic or ceramic materials, such as nickel [18]. Ultra-high-temperature CMCs have a wide range of industrial applications, including space and aerospace propulsion, turbine engines, and nuclear reactors. They can tolerate high thermal loads in harsh conditions [19]. Therefore, it is important to forecast the ultra-high-temperature CMCs' ablation ability to learn more about how long they will last in harsh conditions.

Machine learning (ML) algorithms are very useful for predicting the mechanical properties of composite materials [20]. To make accurate predictions or respond appropriately to new and unknown inputs, machine learning algorithms must be able to recognize patterns and relationships within datasets [21,22]. One of the key components of machine learning that underpins its many applications in a variety of disciplines is its innate capacity to learn from data, generalize to new cases, and make acceptable outcomes without having explicit instructions [23–25]. Qi et al. [26] utilized a decision tree model to forecast the mechanical properties of plastic reinforced with carbon fiber. Similarly, to predict the mechanical properties of polymer composites with alumina modifiers, Kosicka et al. [27] utilized a decision tree approach. Daghigh et al. [28] forecasted the fracture toughness of bio-nanocomposites at numerous sizes with different particle fillers using decision trees and adaptive boosting (AdaBoost) machine learning techniques. Using a random forest regression technique, Hegde et al. [29] investigated the mechanical properties, in particular the hardness of vacuum-sintered Ti-6Al-4V reinforced with SiCp composites. Zhang et al. [30] similarly used a random forest model to predict the mechanical properties of composite laminates, showing that the model could produce precise predictions in a shorter amount of time. Guo et al. [31] showed that the XGBoost regression algorithm is superior to artificial neural networks, support vector regression, classification, and regression trees when it comes to precisely forecasting the tensile strength, ductility, and compressive strength of high-performance fiber-reinforced cementitious composites under the same conditions.

Based on molecular dynamics datasets, Liu et al. [32] used the AdaBoost regression technique to predict the ultimate tensile strength and Young's modulus of graphene-reinforced aluminum nanocomposites. Using additional material parameters, Karamov et al. [33] predicted the fracture toughness of pultruded composites using an AdaBoost regression algorithm. A gradient-boosting regression approach was used by Pathan et al. [34] to predict the yield strengths and macroscopic elastic stiffness of unidirectional fiber composites. Support vector machine regression (SVR) was used by Bonifácio et al. [35] to forecast the mechanical properties of concrete, such as its compressive strength and static Young's modulus. The results showed that SVR had an acceptable predictive capacity. SVR was used by Hasanzadeh et al. [36] to estimate the compressive, flexural, and tensile strengths of basalt-fiber-reinforced concrete. This conventional machine learning technique was also used to simulate the modulus of elasticity and compressive stress–strain curves. The decision tree, extra tree, XGBoost, and random forest regression models' feature relevance analyses revealed that the magnesium matrix composites' reinforcing particle form had the biggest impact on their mechanical properties [37].

Prediction models based on machine learning are widely used in literature to develop and forecast the mechanical properties of composite materials. The literature does, however, appear to be lacking in information about the prediction of thermal properties, particularly for ceramic matrix composites that are subjected to extremely high temperatures. This is a

critical component to gain understanding of how long ultra-high-temperature ceramics last in harsh situations. To fill the research gap and open the doors of new knowledge, this study introduced six machine learning regression models intended to estimate the ablation performance of continuous fiber-reinforced silicon oxy-carbide ceramic matrix composites. The main contributions of this research were developing the decision tree (DT), random forest (RF), support vector machine (SVM), gradient boost (GB), extreme gradient boosting (XGBoost), and adaptive boosting (AdaBoost) regression models using the CMC's experimental burning through time thermal data collected from the oxy-acetylene torch test of ceramic matrix composites. The boosting ensemble techniques (GB, XGBoost, AdaBoost) showed better performance accuracy than the bagging ensemble techniques (random forest, decision trees) and traditional SVM regression modeling for predicting the thermal performance of ceramic matrix composites. This study presented a novel framework for forecasting the ablation performance of ceramic matrix composites under harsh environmental circumstances by using multiple machine learning regression models.

Top of Form

2. Methods

2.1. Data Preparation for Predictive Modeling

The continuous-fiber-reinforced silicon oxy-carbide ceramic matrix composite (CMC) sample was heated to a maximum temperature of 2200 °C in the center of the fire nozzle during the torch test. Poly-siloxane (PSX) resin and woven carbon fabrics were used in the polymer infiltration and pyrolysis technique (PIP) to create continuous-fiber-reinforced silicon oxy-carbide composite. During the oxyacetylene torch test, when the slidable panel is pushed, the sample, which is clamped to it, will encounter the heat. To detect the heat flow of the flame at different distances from the panels, a heat flux sensor was calibrated. Panels were intended to be slid a specific distance along a rail. The temperature of the back surface was determined using thermocouples. Regulators were used to manage the varied ratios of oxygen and acetylene, allowing for the optimization of both flame type and temperature. The American Society for Testing and Materials (ASTM) 258-08 standard test was followed for calibration and testing.

Using a centrifuge, boron nitride (BN) nanoparticles were combined with PSX resin at a weight ratio of 5%. After that, the woven carbon fabric was penetrated by the mixed resin, and six layers of lamination were applied. The carbon fiber or silicon carbide (Cf/SiC) CMCs supplemented with BN nanoparticles underwent pyrolysis in a tube furnace after being post-cured and cured in an autoclave. The torch test was conducted in the direction of thickness until the samples burned through, providing ablative qualities such erosion rate and insulating index. A total of three similar torch tests were conducted with a 524 W/cm² heat flux and same BN weight ratio was maintained at a 25 mm distance from the nozzle. Then, the CMC's burning through temperature data were collected from each test for developing the machine learning prediction models.

2.2. Machine Learning Predictive Modeling

Python 3.9 was used in this study from google colab for executing machine learning (ML) models. All the machine learning algorithms applied in this study followed an standard method described in scikit learn [38] with Python. Figure 1 is the graphical representation of the methodology followed for developing every machine learning model in this study.

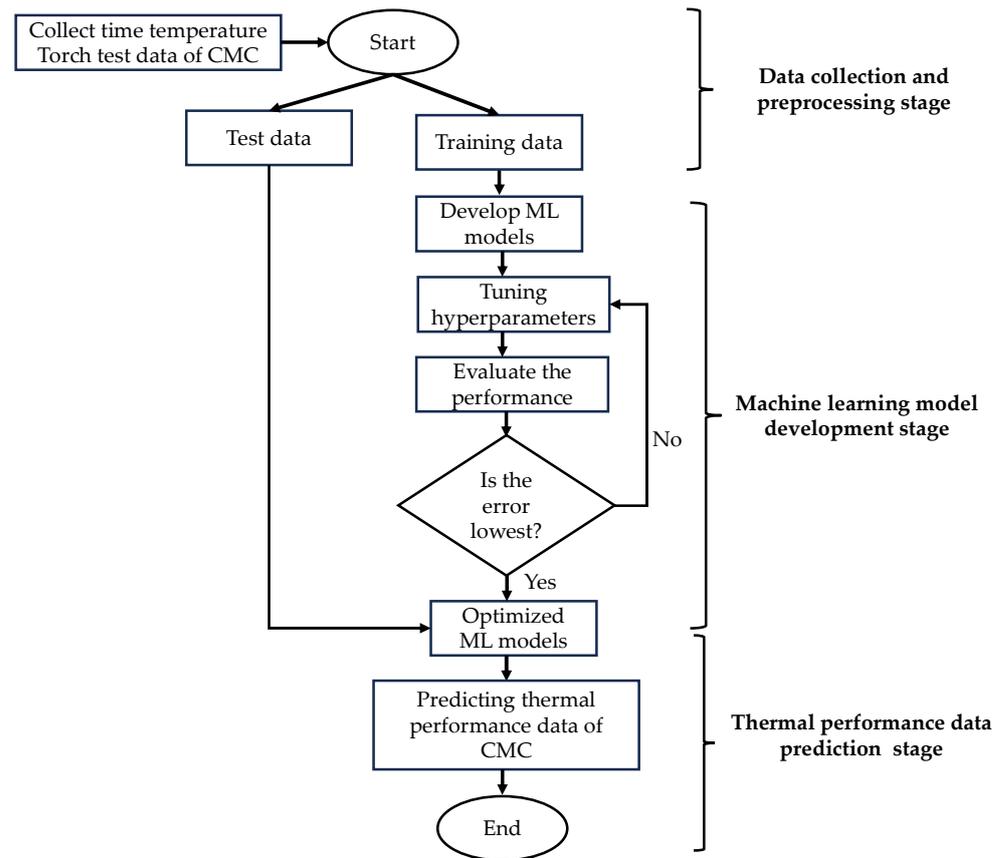


Figure 1. Methodology followed in developing machine learning regression models.

2.2.1. Decision Tree Regression Modeling

Decision trees (DTs) are predictive models in supervised learning that are well-known for their robustness, interpretability, and undeniable utility in a variety of applications [39]. The decision tree is a tree-based algorithm that follows a top-down approach to split a series of trees. The main reason for using decision tree technique in our study is the randomization that each feature introduces, which helps to lower the variance in the time temperature dataset. During the splitting process, Friedman mean square error (MSE) was used to calculate variance reduction for every root and leaf node. The principle of each split was to maximize the variance reduction (V_R) function (1) where W_i , T_i , \bar{T}_i , V_{Root} , and V_{Leaf} represent weights in the n^{th} nodes, predicted temperature, average predicted temperature of every feature, and variance in the root and leaf nodes (2), respectively.

$$V_R = V_{Root} - \sum_{i=1}^n W_i \times V_{Leaf} \tag{1}$$

$$V_{Root} = V_{Leaf} = \sum_{i=1}^n (T_i - \bar{T}_i)^2 \tag{2}$$

In our research, BN4, BN5, and BN6 had 4813, 4429, and 4391 datapoints, respectively. To minimize the overfitting and underfitting problem, maximum depth of the tree was selected as 3. Several hyperparameters including splitter (best or random), minimum samples split (2, 3, 4), minimum samples leaf (1, 2, 3), minimum weight fraction leaf (0.0, 0.1, 0.2), maximum features (none, sqrt, log2), random state (none), maximum leaf nodes (none, 10, 20, 30), minimum impurity decrease (0.0, 0.1, 0.2), and minimal cost complexity pruning (0.0, 0.1, 0.2) were tuned to optimize the final model. During tree development, the splitter decides how to choose the split at each node. Whereas ‘random’ selects the best random split, ‘best’ selects the best split. The minimum number of samples needed to separate an internal node is determined using the minimum samples split calculation. It also regulates node size when forming a tree. Minimum samples leaf establishes the bare

minimum of samples needed for a leaf node. It affects how finely detailed the leaves are in the finished tree. Minimal weight fraction leaf is comparable to minimum samples leaf except it is represented as a percentage of the weighted sum of all instances. The maximum features are the highest count of attributes considered while dividing a node. The 'log2' value represents the logarithm base 2 of the total features, 'sqrt' takes the square root of the entire features, and 'none' indicates that all features are considered. Reproducibility is ensured by using a random state as the seed for the random number generator. Maximum leaf nodes restricts the amount of leaf nodes within the tree. A certain number limits the maximum number of leaf nodes, whereas 'none' permits infinite growth. Minimum impurity decrease indicates that if the impurity decrease is greater than this cutoff, the node is split. It restricts the growth of the tree by demanding a minimum improvement in impurity before a split can happen. Minimal cost complexity pruning influences the cost of adding more nodes to the tree during pruning.

The grid search method was used to find the optimized hyperparameters in the decision tree modeling. Figure 2 shows the decision tree modeling for the BN4 sample using 70% training and 30% test data. In this model, 4813 time–temperature datapoints were used for 144.36 s and the value represented the average predicted temperature (\bar{T}_i) of the selected time–temperature samples.

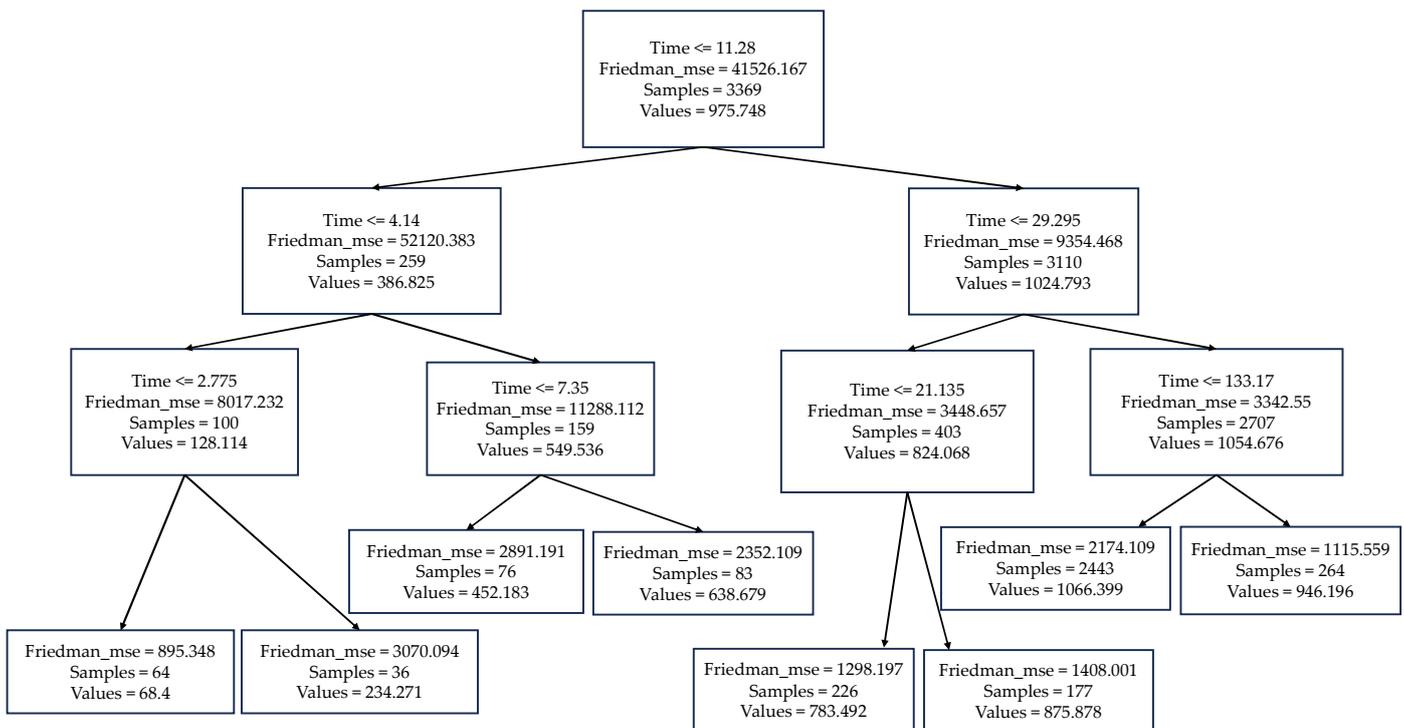


Figure 2. Decision Tree Regression modeling for BN4 sample.

2.2.2. Random Forest Regression Modeling

To develop a more reliable model for regression problems and increasing the performance from decision tree temperature prediction model, random forest (RF) regression was introduced which entailed training an ensemble of decision trees on a random subset of the bootstrap sampled data, and then combining their predictions following bootstrap aggregations [37]. In this study, the row sampling with replacement method was utilized to collect a subset (bootstrap sampled) of torch test time–temperature data (\bar{d}) from the total dataset d , where $d > \bar{d}$. Then, the subset of time–temperature data was passed through the five individual base learner trees having a constant depth of 3 in each tree. Finally, the bootstrap aggregation method was followed to predict the thermal resistivity of the ceramic matrix composites as shown in Figure 3.

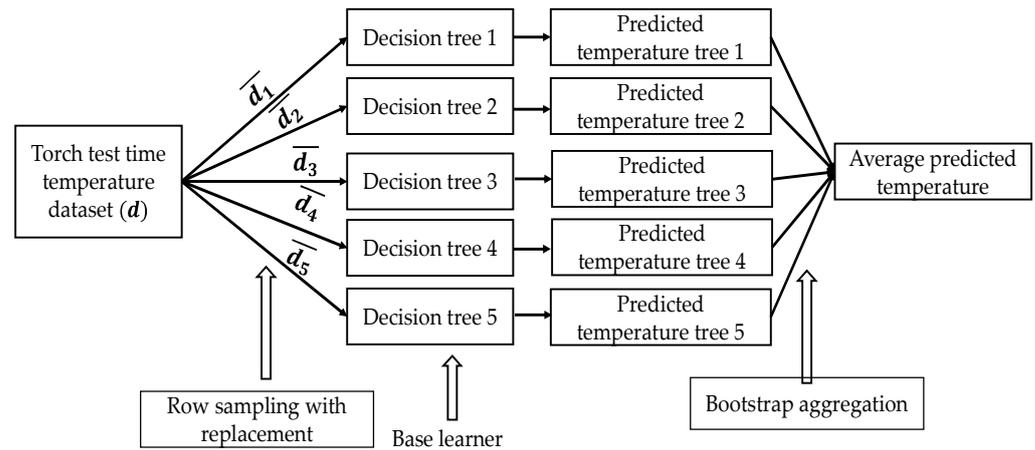


Figure 3. Random Forest Regression modeling.

Various hyperparameters were adjusted to optimize the final model, such as the minimum samples split (2, 3, 4), minimum samples leaf (1, 2, 3), minimum weight fraction leaf (0.0, 0.1, 0.2), maximum features (none, sqrt, log2), random state (none), maximum leaf nodes (none, 10, 20, 30), minimum impurity decrease (0.0, 0.1, 0.2), bootstrap (true), out of bag score (false), and minimal cost complexity pruning (0.0, 0.1, 0.2). Finally, the optimal hyperparameters in the random forest modeling were found using the training data in the grid search method. The bootstrap algorithm determines if each decision tree is constructed using bootstrap samples, which are sampled with replacement. Whether or not to use out-of-bag samples for scoring depends on the out-of-bag score. Samples that are not utilized to train a specific decision tree are known as out-of-bag samples.

2.2.3. Support Vector Machine (SVM) Regression Modeling

The SVM regression method was used in this study because of its ability to find complicated or non-linear relationships between the variables. Choosing an appropriate kernel, training the model by maximizing an objective function, and adjusting hyperparameters to obtain the best possible performance in continuous result prediction are the steps involved in SVM regression [40]. To predict the thermal data, this method created an optimized hyperplane. Then, a marginal hyperplane was introduced on both sides of the hyperplane with equal distance ϵ as shown in Figure 4. To deal with the overfitting problem, we have used 30 datapoints for individual samples and 20 datapoints for all replicates to predict the thermal data outside the marginal plane boundary. The objective function O_f of the SVM model was calculated by the following Equation (3).

$$O_f = \frac{w^2}{2} + C_i \sum_{i=1}^n \zeta_i \tag{3}$$

In the above equation, w represents the slope of the hyperplane, C_i is the number of datapoints outside the marginal planes, and ζ represents their distance from the marginal plane.

To improve the final model, several hyperparameters were adjusted, such as the kernel (linear, polynomial, radial basis function, sigmoid) and epsilon (0.001, 0.01, 0.1, 1). To address issues with non-linearity in the thermal performance data in CMCs, SVM uses a kernel function to map sample data into a high-dimensional space [41]. The kind of decision boundary that the SVM uses is decided by the kernel.

‘Linear’ decision boundaries are used; ‘polynomial’ kernels are used; ‘radial basis function (RBF)’ kernels are used; and ‘sigmoid’ functions are used. The SVM algorithm’s optimization margin of tolerance is determined by epsilon. It establishes a range in which misclassification carries no consequences.

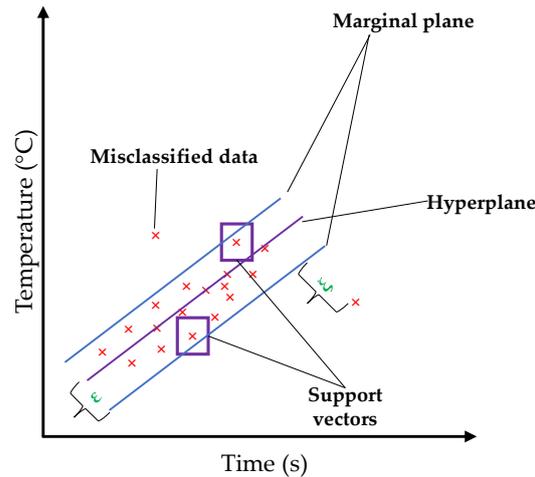


Figure 4. Support Vector Machine Regression modeling.

2.2.4. Gradient Boost Regression (GBR) Modeling

A series of weak learners (decision trees) are constructed in an ensemble using gradient boosted regression, with each model’s goal being to capture the residuals of the preceding models [42]. Firstly, the average ablation performance data of ceramic matrix composites was predicted through a base model. Then, the residuals or errors of the base model were used as an input in the decision tree (DT) model. Finally, the predicted errors were minimized through five serially connected DT models to obtain the finally predicted temperature data as described in Figure 5. The finally predicted temperature (T_p) can be predicted through the following Equation (4). In this equation, $h_0(x)$ represents the predicted temperature of the base model, α represents the learning rate, and $h_i(x)$ is the prediction output of the residuals from the DT models.

$$T_p = h_0(x) + \alpha_1 h_1(x) + \alpha_2 h_2(x) + \alpha_3 h_3(x) + \alpha_4 h_4(x) + \alpha_5 h_5(x)$$

$$T_p = h_0(x) + \sum_{i=1}^5 \alpha_i h_i(x)$$
(4)

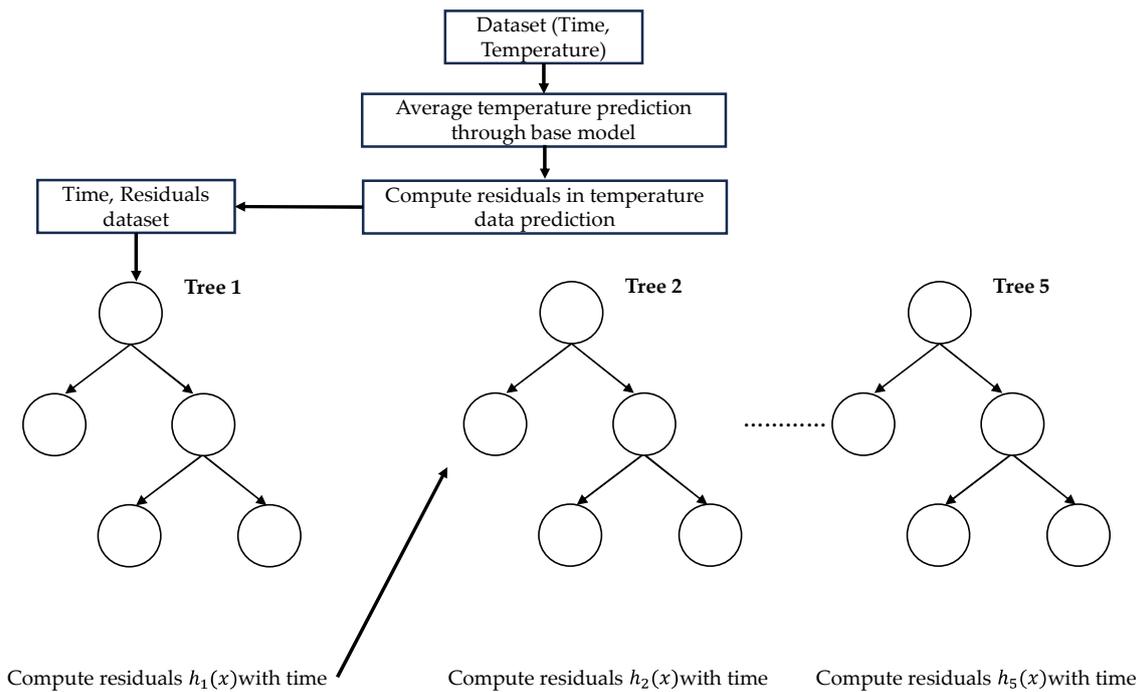


Figure 5. Gradient Boost Regression modeling procedure.

To identify the GBR modeling’s optimum hyperparameters, the grid search approach was applied. The final model was optimized by adjusting a number of hyperparameters, such as the learning rate (0.001, 0.01, 0.1, 0.2, 0.3, 1), minimum samples split (2, 3, 4), minimum samples leaf (1, 2, 3), minimum weight fraction leaf (0.0, 0.1, 0.2), maximum features (none, sqrt, log2), random state (none), maximum leaf nodes (none, 10, 20, 30), minimum impurity decrease (0.0, 0.1, 0.2), and minimal cost complexity pruning (0.0, 0.1, 0.2). The learning rate regulates the contribution of every tree to the final prediction. Although more trees are needed for modeling with smaller values, generalization can be enhanced.

2.2.5. Extreme Gradient Boost Regression (XGBoost) Modeling

To prevent overfitting and improve efficiency, the XGBoost method adds a more regularized gradient tree boosting technique than GBR [43]. To develop the XGBoost model, similarity weight and gain were calculated in the leaf node to predict the residuals from the first tree. Then, residuals or errors were minimized after predicting the temperature through the series of decision trees as depicted in Figure 6. In this study, five DTs were used with the depth of 3 in each tree. In the grid search method, learning rate (0.001, 0.01, 0.1, 0.2, 0.3, 1) was introduced to overcome the overfitting problem.

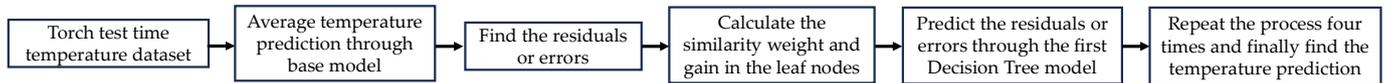


Figure 6. Extreme Gradient Boost Regression modeling procedure.

2.2.6. Adaptive Boosting Regression Modeling

Adaptive boosting (AdaBoost) uses decision trees to train a sequence of weak learners. A weak learner is trained with all datapoints having identical weights at first. Weights are modified in the following cycles in accordance with prediction accuracy. Erroneously predicted points are assigned a higher weight, which directs the subsequent learner towards more difficult cases as shown in Figure 7. This procedure is repeated, and a weighted vote is used to aggregate the guesses made by each learner to create the final prediction [44]. During iterations, AdaBoost prioritizes hard-to-predict cases for enhancing the overall performance of the model. In this study, five DTs were used with the depth of 3 in each tree. In the grid search method, we used linear, square, and exponential loss in the training dataset and learning rate (0.001, 0.01, 0.1, 0.2, 0.3, 1). The absolute disparities between expected and actual values are minimized using linear loss. By minimizing the squared discrepancies between the actual and predicted values, square loss highlights bigger errors. Exponential loss prioritizes mistake reduction by assigning greater weight to misclassified instances.

2.3. Performance Accuracy Measuring Parameters

In this study, four evaluation metrics including coefficient of determination (R^2), mean absolute error (MAE), mean absolute percentage error (MAPE), and root mean square error (RMSE) are analyzed for the test dataset to assess the performance of the machine learning regression models. They are defined as:

$$R^2 = 1 - \frac{\sum_{j=1}^K (e_j - p_j)^2}{\sum_{j=1}^K (e_j - \bar{e})^2} \tag{5}$$

$$MAE = \frac{1}{K} \sum_{j=1}^K |e_j - p_j| \tag{6}$$

$$MAPE = \frac{1}{K} \sum_{j=1}^K \frac{|e_j - p_j|}{|e_j|} \times 100 \tag{7}$$

$$RMSE = \sqrt{\frac{1}{K} \sum_{j=1}^K (e_j - p_j)^2} \tag{8}$$

where K is the number of datasets, e_j and p_j stand for the j^{th} experimental and predicted burning through temperature data of the ceramic matrix composites, respectively, and \bar{e} is the average actual temperature data.

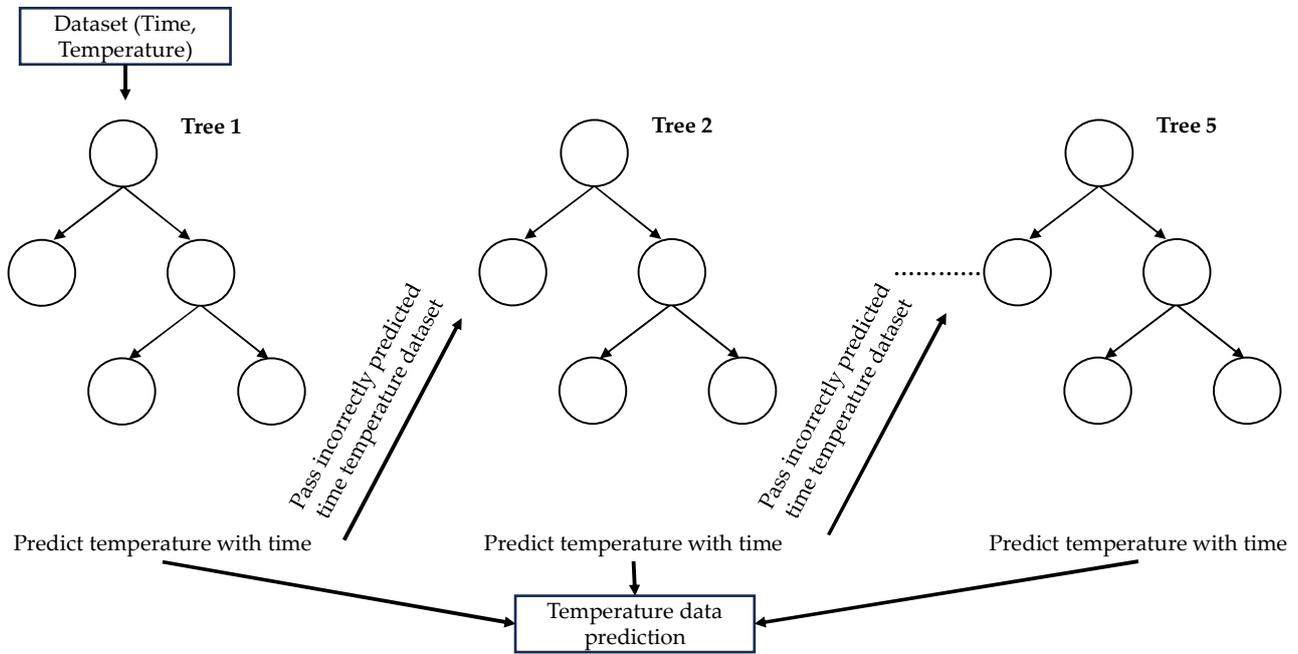


Figure 7. AdaBoost Regression modeling procedure.

3. Results

K-fold cross-validation was used in this study to evaluate the model’s performance and generalizability. First, 50 subsets (folds) of the 50–90% training dataset were created, and then every machine learning model was trained and assessed based on training data size percentage. Then, the optimized training data size was selected for every machine learning model to achieve the highest prediction accuracy. In the cross-validation process, one of the folds was utilized as the validation set and the other folds were used for training in each iteration. Every fold was utilized as the validation set exactly once during this operation. After optimizing the dataset partitioning through different training data size percentages, testing data were used to test all the prediction models to assess the decision tree, random forest, support vector machine, gradient boosting, extreme gradient boosting, and AdaBoost regression algorithms’ accuracy.

3.1. Decision Tree Modeling Results

The optimized best hyperparameters of the decision tree algorithms were found using the grid search method as shown in Table 1. The prediction accuracy measurements after utilizing the cross-validation process are listed in Table 2.

Table 1. Optimized hyperparameters in decision tree regression models of BN4, BN5, BN6 and all the replicates.

Optimized Hyperparameters	50% Test Data for BN4	30% Test Data for BN5	20% Test Data for BN6
Splitter	Best	Best	Best
Minimum samples split	2	2	2
Minimum samples leaf	1	1	1
Minimum weight fraction leaf	0.0	0.0	0.0
Maximum features	None	None	None
Random state	None	None	None
Maximum leaf nodes	None	None	None
Minimum impurity decrease	0.0	0.0	0.0
Cost complexity pruning	0.0	0.0	0.0

The high potential of this model is demonstrated by the fact that the mean absolute percentage error is always less than 8%, the R2 scores are higher than 94% for each replicate, and the root mean square error is less than 48 for individual replicates. When the MAPE is continuously less than 5%, it means that the model is accurate since, on average, the predictions are close to the actual values. A high degree of goodness of fit is indicated by R2 values that often exceed 94%, demonstrating the model’s capacity to capture and explain the variation in the torch test data. When the RMSE is continuously less than 48, it indicates that the model’s average prediction mistakes are small.

The expected versus actual ablation behavior utilizing 30% of test datasets for BN4, BN5, and BN6 samples is depicted in Figure 8.

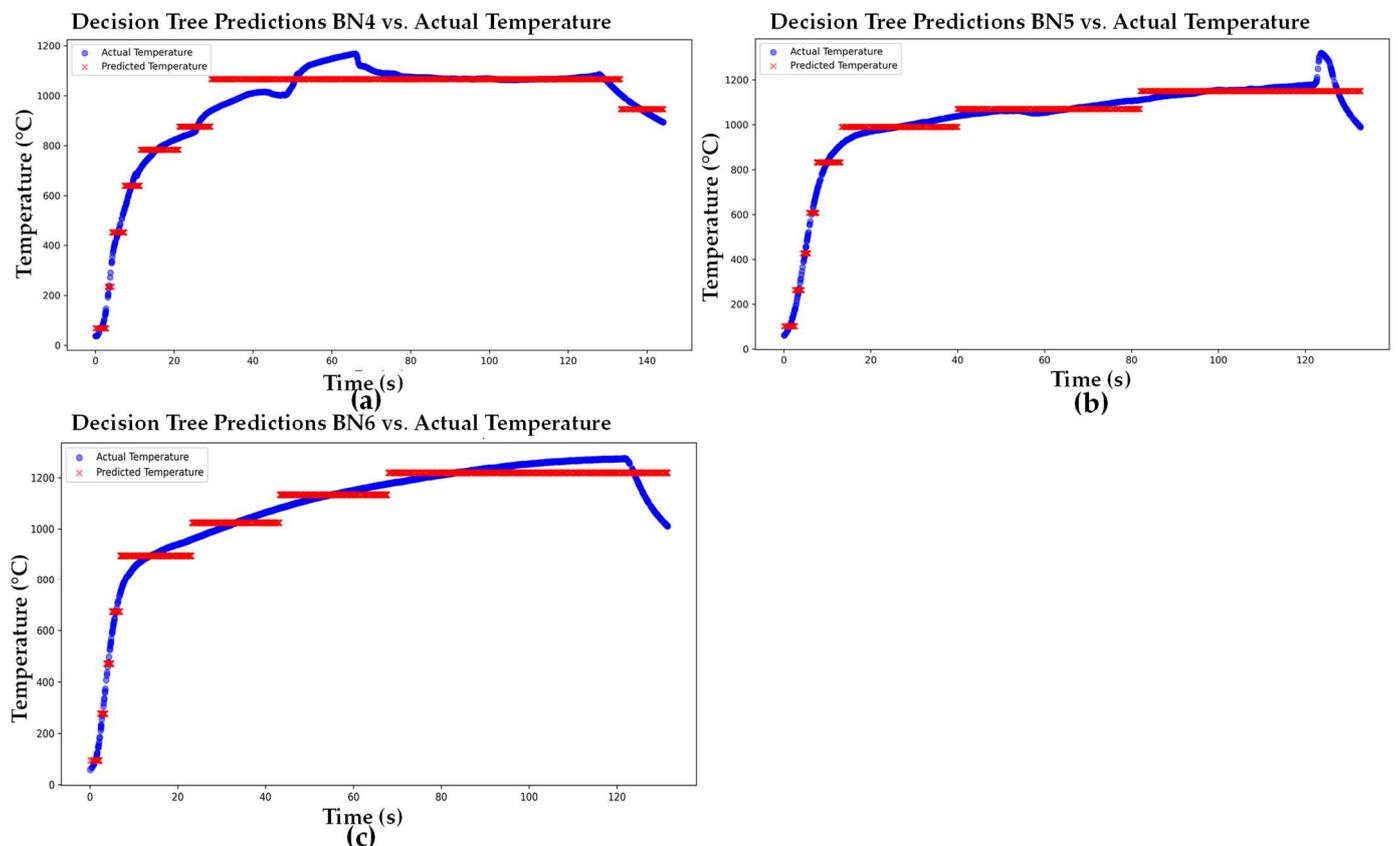


Figure 8. Decision Tree Regression results for (a) BN4, (b) BN5, and (c) BN6 samples.

The results are shown in Figure 8, which indicates that the final decision tree has a maximum depth of 3 which resulted in a configuration with eight leaf nodes. Every leaf

node predicted the average constant burning through temperature of the ceramic matrix composites. The greatest average temperature that was forecasted was 1066.399 °C, while the lowest average temperature that was predicted was 68.4 °C for the BN4 sample using 70% training and 30% test data.

Table 2. Ablation performance results of boron nitride ceramic matrix composites in Decision Tree model.

Performance Parameters	50/50 Split for BN4	70/30 Split for BN5	80/20 Split for BN6	70/30 Split for All Replicates
R-squared (R2) score	0.9501	0.9596	0.9469	0.8845
Root mean squared error (RMSE)	45.28	39.87	47.84	73.24
Mean absolute error (MAE)	32.84	26.68	36.94	60.52
Mean absolute percentage error (MAPE)	4.33%	3.31%	4.03%	7.10%
Simulation execution time	18 min 11 s	18 min 24 s	28 min 49 s	23 min 24 s

3.2. Random Forest Modeling Results

The grid search approach was used to find the random forest algorithm’s optimal hyperparameters, as Table 3 illustrates.

Table 3. Optimized hyperparameters in Random Forest Regression models of BN4, BN5, and BN6 samples.

Optimized Hyperparameters	50% Test Data for BN4	10% Test Data for BN5	20% Test Data for BN6
Minimum samples split	2	3	3
Minimum samples leaf	3	1	1
Minimum weight fraction leaf	0.0	0.0	0.0
Maximum features	None	sqrt	log2
Random state	None	None	None
Maximum leaf nodes	None	10	30
Minimum impurity decrease	0.2	0.2	0.1
Cost complexity pruning	0.0	0.0	0.2
Bootstrap	True	True	True
Out-of-bag score	False	False	False

According to Table 4, the mean absolute percentage error of the random forest model is 1% lower than that of decision tree modeling, and the root mean square error is less than 46 for individual replicates, indicating the model’s outstanding prospects. Figure 9 shows the expected versus actual thermal performance curve for the BN4, BN5, and BN6 samples using 30% test datasets.

Table 4. Ablation performance results of boron nitride ceramic matrix composites in Random Forest model.

Performance Parameters	50/50 Split for BN4	90/10 Split for BN5	80/20 Split for BN6	80/20 Split for All Replicates
R-squared (R2) score	0.9597	0.9604	0.9521	0.8882
Root mean squared error (RMSE)	40.66	33.11	45.41	70.95
Mean absolute error (MAE)	27.84	20.36	34.23	58.72
Mean absolute percentage error (MAPE)	3.65%	2.37%	3.63%	6.66%
Simulation execution time	2 min 18 s	2 min 33 s	2 min 24 s	3 min 40 s

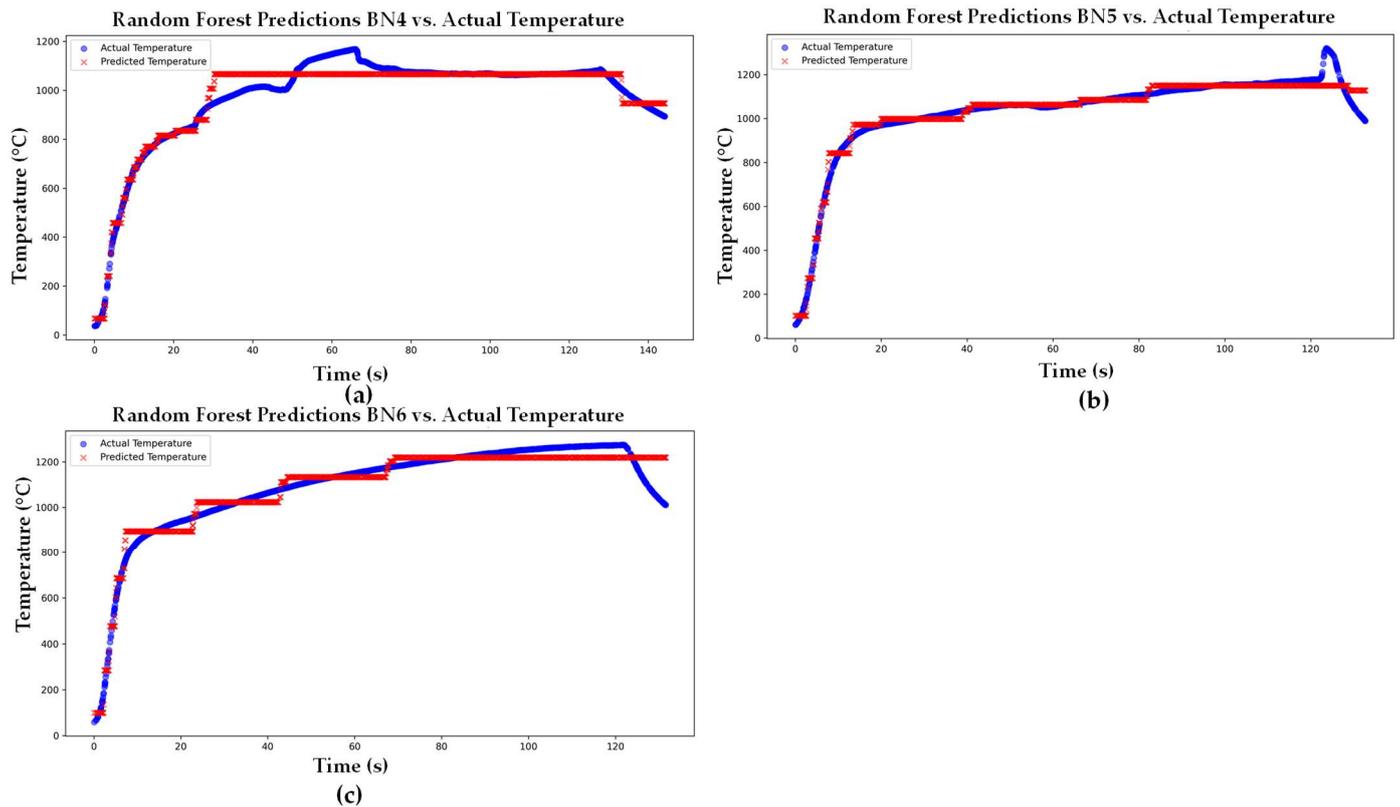


Figure 9. Random Forest Regression results for (a) BN4, (b) BN5, and (c) BN6 samples.

The findings are displayed in Figure 9, which shows that there are five decision trees in the final random forest (RF) model, each having a maximum depth of three. The final RF model configuration contained forty leaf nodes in total from five trees. Forty leaf nodes corresponded to forty expected constant thermal values of composites. For the BN4 sample, which consisted of 70% training and 30% test data, the highest average temperature predicted was 1067.001 °C, while the lowest average temperature predicted was 58.555 °C.

3.3. Support Vector Machine Modeling Results

The grid search approach was used to find the support vector machine (SVM) regression modeling's optimal hyperparameters. The SVM model's poor performance is demonstrated by consistently high mean absolute percentage errors, low R2 values, and high root mean square errors in Figure 10. Cross-validation was used to evaluate the model's performance. The epsilon (ϵ) value was set to 1 for the BN4 test sample (20%) and the BN5 and BN6 test samples (10%). To obtain better prediction results, however, a slightly lower epsilon value of 0.1 was used for all replicate models. Figure 10 shows the expected versus real ablation performance using 30% test datasets for the BN4, BN5, and BN6 samples.

The radial basis function (RBF), which has been found to be the optimal kernel in this work, is used to calculate the ceramic matrix composites' ablation performance for each sample. But this SVM model performs worse than other machine learning models because of the limitations of marginal plane distance from the hyperplane as depicted in Table 5.

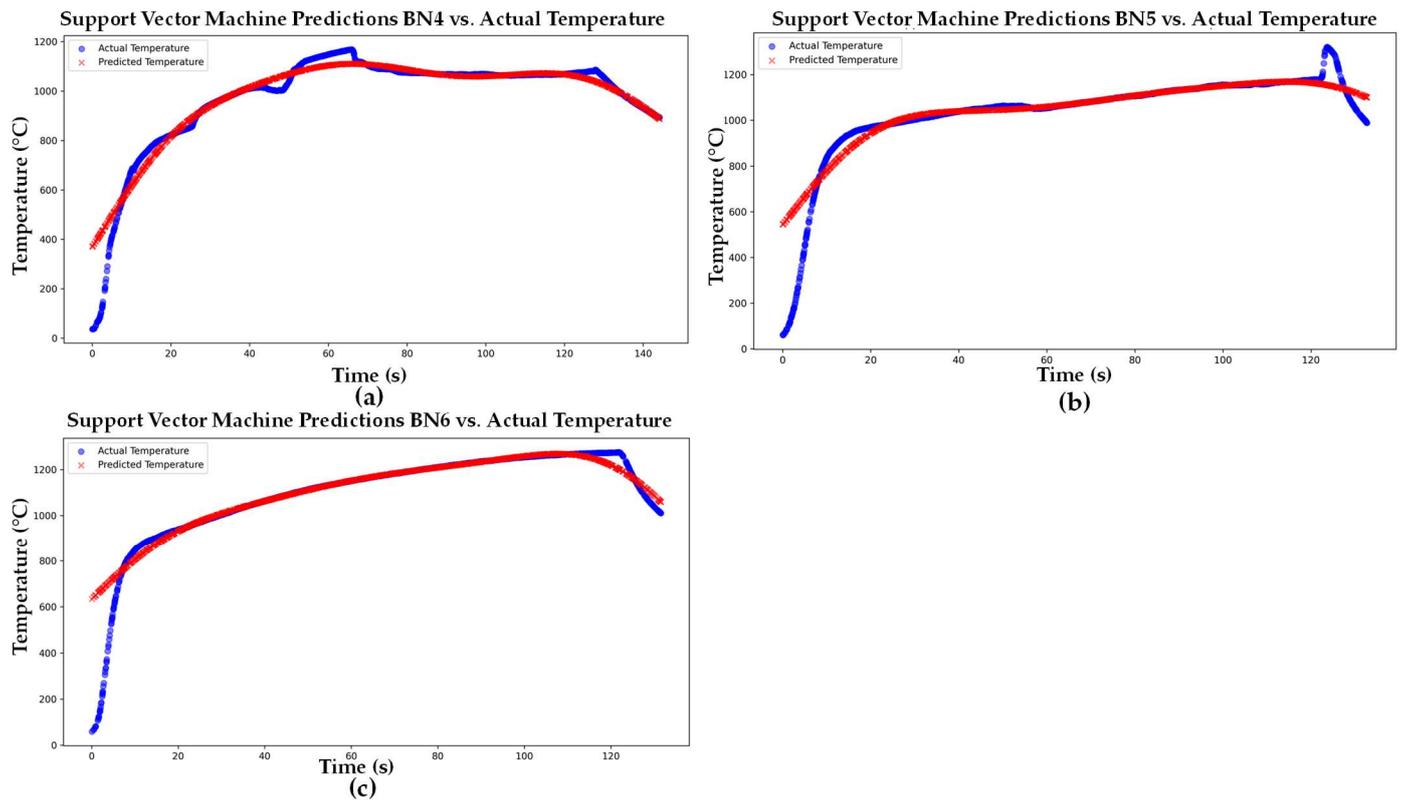


Figure 10. SVM Regression results for (a) BN4, (b) BN5, and (c) BN6 samples.

Table 5. Ablation performance results of boron nitride ceramic matrix composites in Support Vector Machine model.

Performance Parameters	80/20 Split for BN4	90/10 Split for BN5	90/10 Split for BN6	90/10 Split for All Replicates
R-squared (R2) score	0.9289	0.8516	0.8853	0.8704
Root mean squared error (RMSE)	50.63	64.11	61.29	77.06
Mean absolute error (MAE)	23.87	25.87	17.59	56.47
Mean absolute percentage error (MAPE)	10.15%	7.74%	6.80%	10.69%
Simulation execution time	6 min 21 s	11 min 50 s	8 min 31 s	35 min 54 s

3.4. Gradient Boost Modeling Results

Every decision tree was designed in the gradient boost regression (GBR) model with maximum depth 3 to address the overfitting issue. For BN4, the hyperparameters hold true when 50% of the test data are used and 10% of the test data for BN5 and BN6. With a minimum of two sample splits and one leaf, the learning rate is set to one. Consistently set to 0.0, 'none', 'none', 'none', 0.0, and 0.0, respectively, are the minimum weight fraction leaf, maximum features, random state, maximum leaf nodes, minimum impurity decrease, and cost complexity pruning. The design highlights an optimized set of hyperparameters that are customized for specific test data subsets after the cross-validation process, guaranteeing excellent model performance on a variety of datasets used in the investigation.

Table 6 presents the prediction accuracy measurements. The last decision tree's eight leaf nodes in the GBR model were able to estimate the final residuals or errors in the ceramic matrix composites' burning through temperature. After deducting the final residual value from the base model forecast, the final temperature of the CMC was estimated.

Table 6. Ablation performance results of boron nitride ceramic matrix composites in Gradient Boosting model.

Performance Parameters	50/50 Split for BN4	90/10 Split for BN5	90/10 Split for BN6	90/10 Split for All Replicates
R-squared (R2) score	0.9943	0.9915	0.9928	0.9145
Root mean squared error (RMSE)	15.32	15.38	15.41	62.59
Mean absolute error (MAE)	9.35	8.76	9.56	50.82
Mean absolute percentage error (MAPE)	1.35%	1.06%	1.36%	5.63%
Simulation execution time	8 min 44 s	12 min 45 s	13 min 32 s	25 min 20 s

The gradient boosting regression (GBR) model’s outstanding performance is demonstrated by consistently low mean absolute percentage errors (never more than 2%), very high R2 values (never less than 99%), and root mean square errors (never more than 16) for each individual replicate. Collectively, these indications show how accurately, precisely, and potently the GBR model functions as an explanatory tool and how well it captures and explains the variance in the data.

Figure 11 compares the expected and actual ablation behavior for the BN4, BN5, and BN6 samples using 30% test datasets.

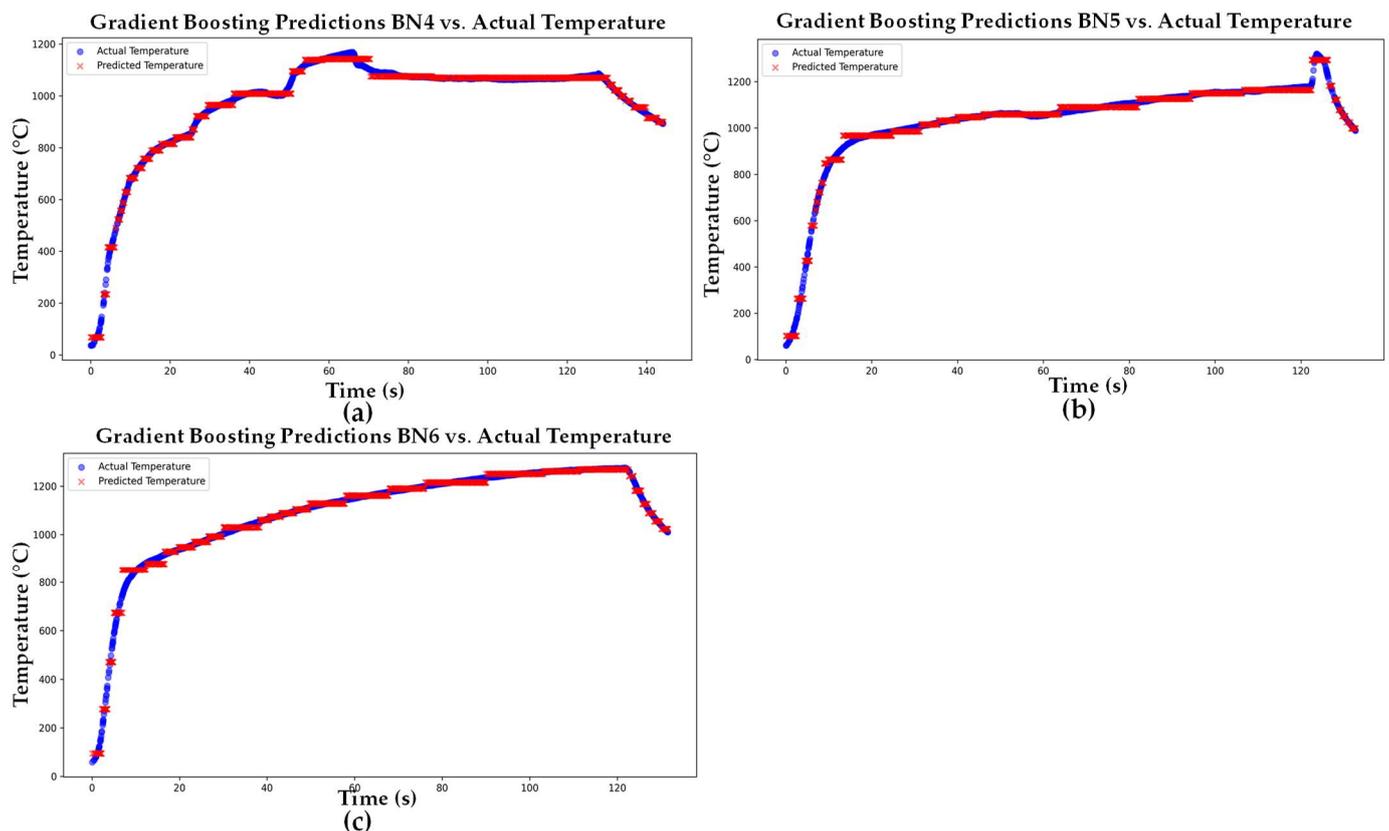


Figure 11. Gradient Boost Regression results for (a) BN4, (b) BN5, and (c) BN6 samples.

As we have shown in our study, a rigorous cross-validation procedure has been used to tune the hyperparameters of the GBR model, which include the maximum depth of every decision tree. Through this procedure, the model is guaranteed to adjust to the subtle differences between various materials, offering a predictive capacity that is material specific. Capturing non-linear interactions in the data is an area where the GBR model excels. It is common for materials to show complex thermal reactions in the event of material-specific ablation, which may not follow straightforward linear patterns. These

non-linear interactions were navigated and captured by the GBR model, which enables us to identify intricate linkages between ablation behavior and material characteristics.

Though it is a machine learning approach, the GBR model’s capacity to identify intricate patterns and correlations in the data allows it to indirectly represent the system’s physics. Together, the ensemble’s decision trees provide predictions that help the final model output by learning from the data. The GBR model is highly proficient in detecting and exploiting complex thermal patterns that are not readily apparent using conventional analytical techniques when it comes to non-linear time–temperature behavior in CMCs. Various thermodynamic and heat transport processes that control the thermal response of CMCs are reflected in the model’s predictions, even though the model itself does not include explicit physical equations.

3.5. Extreme Gradient Boost Modeling Results

The extreme gradient boosting (XGBoost) model consistently shows a mean absolute percentage error (MAPE) below 2%, demonstrating a high degree of accuracy since its predictions are, on average, quite near to the actual values. When the R2 values are near to 1, it means that the model fits the data well and has a remarkable capacity to explain variation. Furthermore, each replicate’s mean absolute error is reliably less than 10, indicating accurate predictions, and the root mean square error never exceeds 16, highlighting the model’s overall accuracy and dependability in capturing the diversity of the ablation performance data. Based on 30% test datasets, Figure 12 shows the difference between the actual and predicted ablation behavior for the BN4, BN5, and BN6 samples. Most notably, a learning rate of one was established by the XGBoost modeling procedure. Prediction accuracy metrics are summarized in Table 7, which also shows the model’s performance with training data sizes ranging from 50% to 90%.

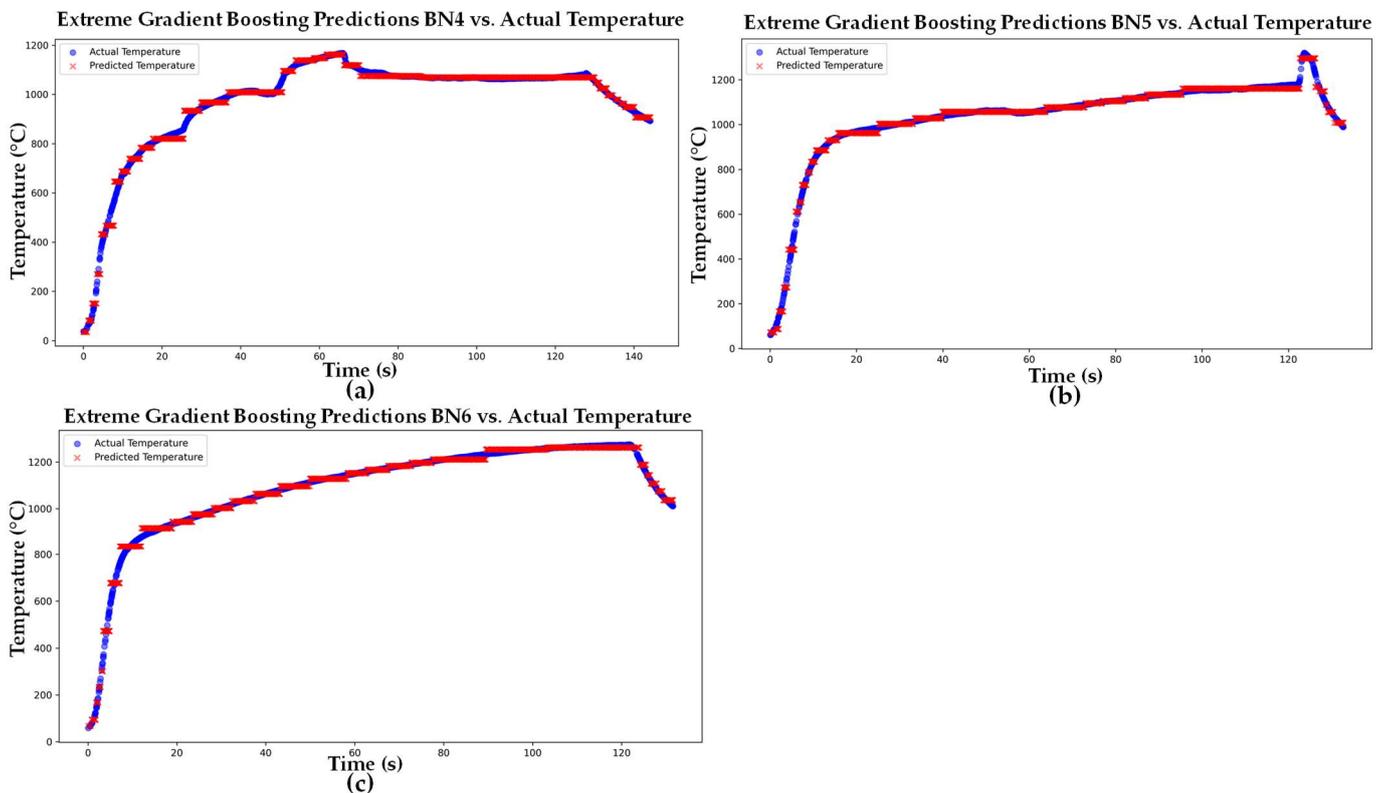


Figure 12. Extreme Gradient Boost Regression results for (a) BN4, (b) BN5, and (c) BN6 samples.

Table 7. Ablation performance results of boron nitride ceramic matrix composites in Extreme Gradient Boosting model.

Performance Parameters	70/30 Split for BN4	70/30 Split for BN5	80/20 Split for BN6	80/20 Split for All Replicates
R-squared (R2) score	0.9940	0.9949	0.9947	0.9174
Root mean squared error (RMSE)	15.59	14.18	15.16	60.97
Mean absolute error (MAE)	9.45	9.43	9.70	48.97
Mean absolute percentage error (MAPE)	1.52%	1.37%	1.20%	5.41%
Simulation execution time	1 s	1 s	1 s	2 s

3.6. Adaptive Boosting (AdaBoost) Modeling Results

The adaptive boosting (AdaBoost) model consistently keeps the root mean square error for individual replicates below 37, attains R2 values greater than 98%, and maintains a mean absolute percentage error below 4%. Using 30% test datasets, Figure 13 compares the actual and expected ablation behavior for the BN4, BN5, and BN6 samples. The AdaBoost process used square loss for the time–temperature data and established a learning rate of one in most prediction models. The prediction accuracy values are presented in Table 8.

It was found from our study that all the performance parameters including RMSE, MAE, MAPE, and R2 score in the combined prediction ML models using all the replicates performed more poorly than the single-sample prediction models. The main reason for poor performance is the difficulty faced by all the above ML models to find the non-linearity in the torch test temperature data in three replicates at the CMC’s burning through time. The RF model took less simulation time than DT models for predicting the thermal performance of composites with total samples of data divided into five trees from one tree. It was also noticeable that XGBoost and AdaBoost regression models took less simulation time than other techniques for temperature prediction.

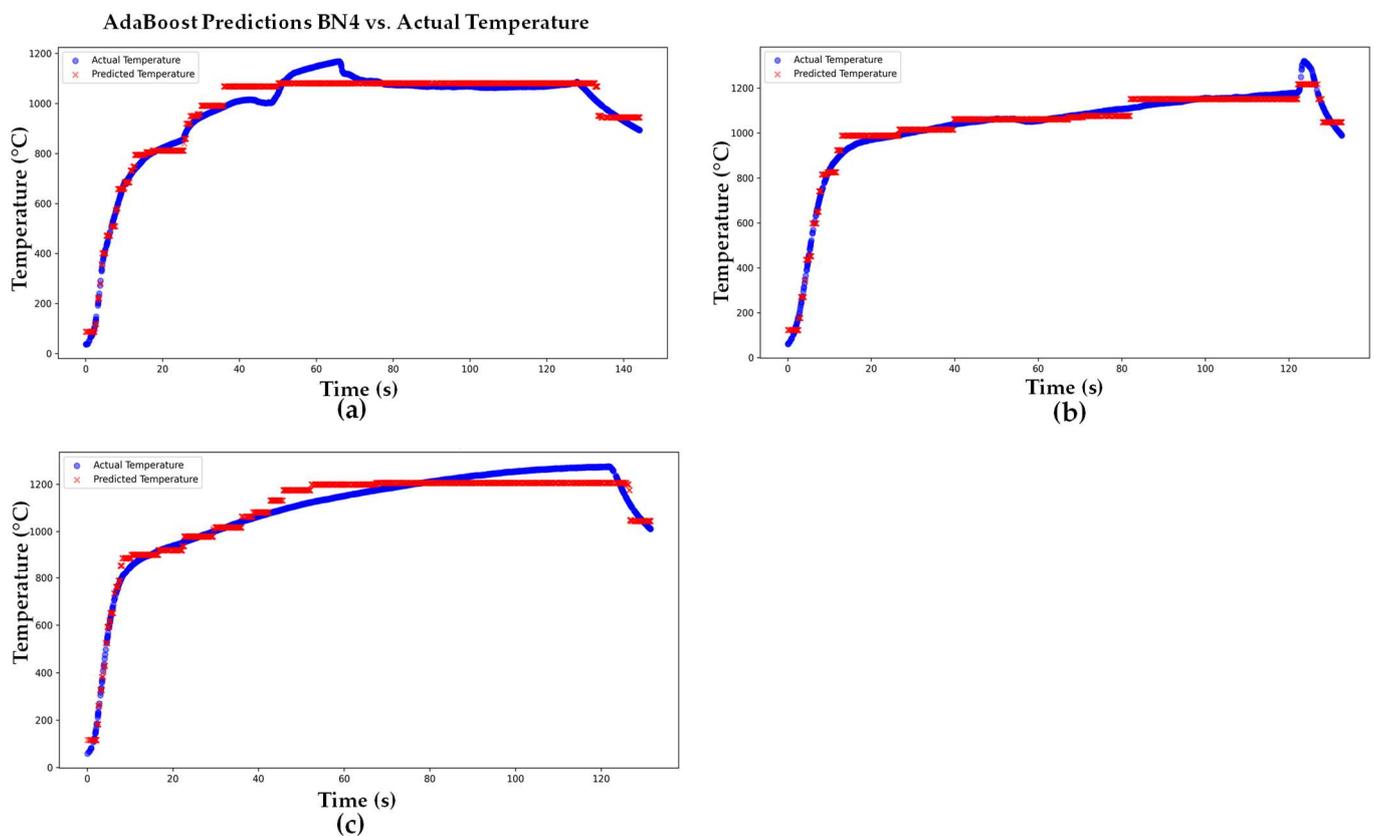


Figure 13. AdaBoost Regression results for (a) BN4, (b) BN5, and (c) BN6 samples.

Table 8. Ablation performance results of boron nitride ceramic matrix composites in Adaptive Boosting model.

Performance Parameters	80/20 Split for BN4	70/30 Split for BN5	80/20 Split for BN6	70/30 Split for All Replicates
R-squared (R2) score	0.9798	0.9846	0.9686	0.9040
Root mean squared error (RMSE)	27.00	24.63	36.79	66.77
Mean absolute error (MAE)	20.88	17.88	30.81	56.12
Mean absolute percentage error (MAPE)	3.19%	2.45%	3.33%	6.92%
Simulation execution time	2 s	1 s	2 s	3 s

4. Comparing the Performance of ML Models

To the best of our knowledge, this work is the first attempt to forecast ceramic matrix composites’ ablation performance data. The coefficient of determination, often known as the R-squared (R2) score, was used as a performance indicator to assess the effectiveness of the regression models analyzed for this study. For each model, the R2 score—which represents the proportion of variance in temperature data explained across time—was estimated. Notably, the R2 scores of the GBR and XGBoost models were nearly 1, indicating an incredible explanation of the consistency between the observed and anticipated torch test thermal data. All the machine learning models developed in this study have an R2 score higher than 0.85, indicating that all the models can accurately predict the torch test data as shown in Figure 14a. The R2 score was considered together with other performance indicators, such as mean absolute error (MAE), mean absolute percentage error (MAPE), and root mean square error, for a thorough evaluation of the model’s performance.

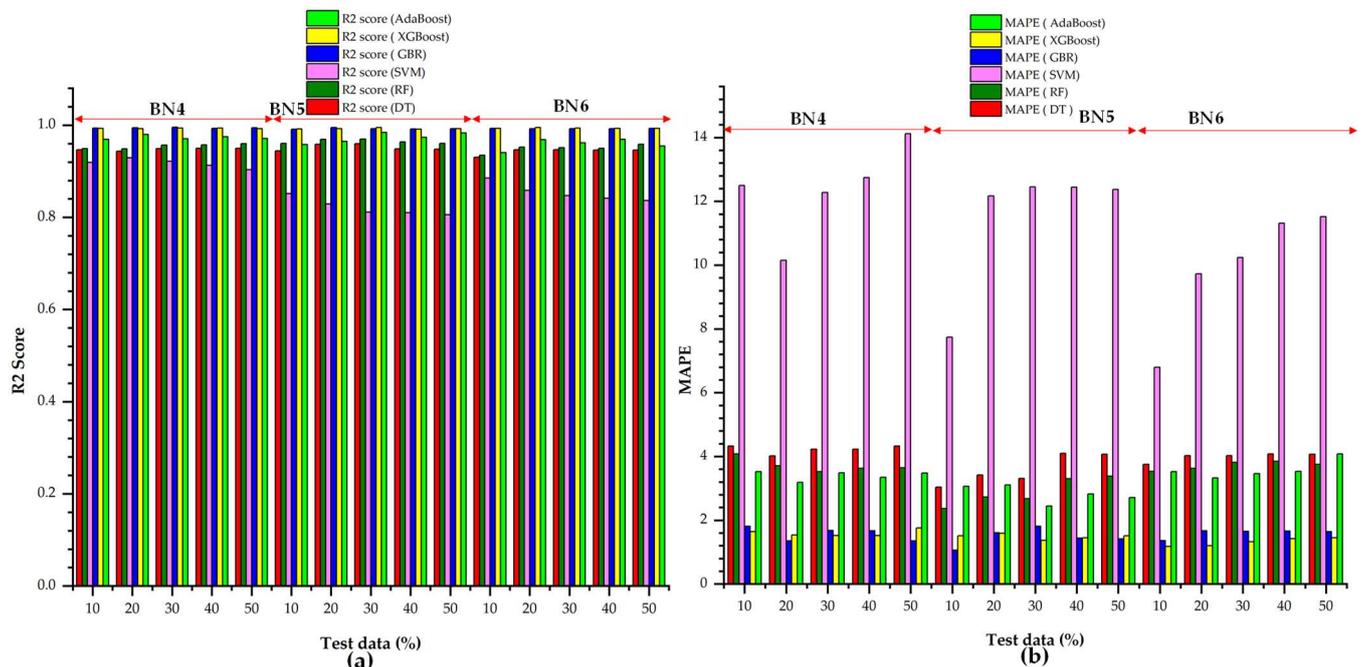


Figure 14. Performance (a) R² score and (b) MAPE results of individual replicates with test data percentage.

According to our research, boosting (XGBoost, AdaBoost, and GBR) regression approaches had a mean absolute percentage error that was higher than that of bagging (random forest), SVM, and decision tree regression techniques. The algorithm’s ability to sequentially learn from the mistakes of prior learners after first learning from weak learners is responsible for the significant gain in performance observed in boosting ensemble techniques. The average percentage difference between the actual and anticipated temperature

data values highlights the fact that GBR and XGBoost were the most accurate in estimating the ablation performance of ceramic matrix composites during the burning phase as described in Figure 14b. On the other hand, the smaller distance between the marginal plane lines and the hyperplane is responsible for the SVM regression model’s poorer performance. This results in incorrect data classification and a rise in the mean absolute percentage error when predicting the burning through temperature data of the CMCs.

The root mean squared error (RMSE) is used to evaluate the predictive machine learning models’ accuracy regarding the thermal behavior of ceramic matrix composites. This measurement is equivalent to taking the square root of the average of the squared discrepancies between the actual and expected temperatures obtained from the torch test. The GBR and XGBoost models beat other machine learning regression models when the RMSE value is less than 19 as shown in Figure 15a. The SVM regression model’s reduced performance can be ascribed to a significant quantity of incorrectly categorized datapoints situated beyond the marginal plane lines.

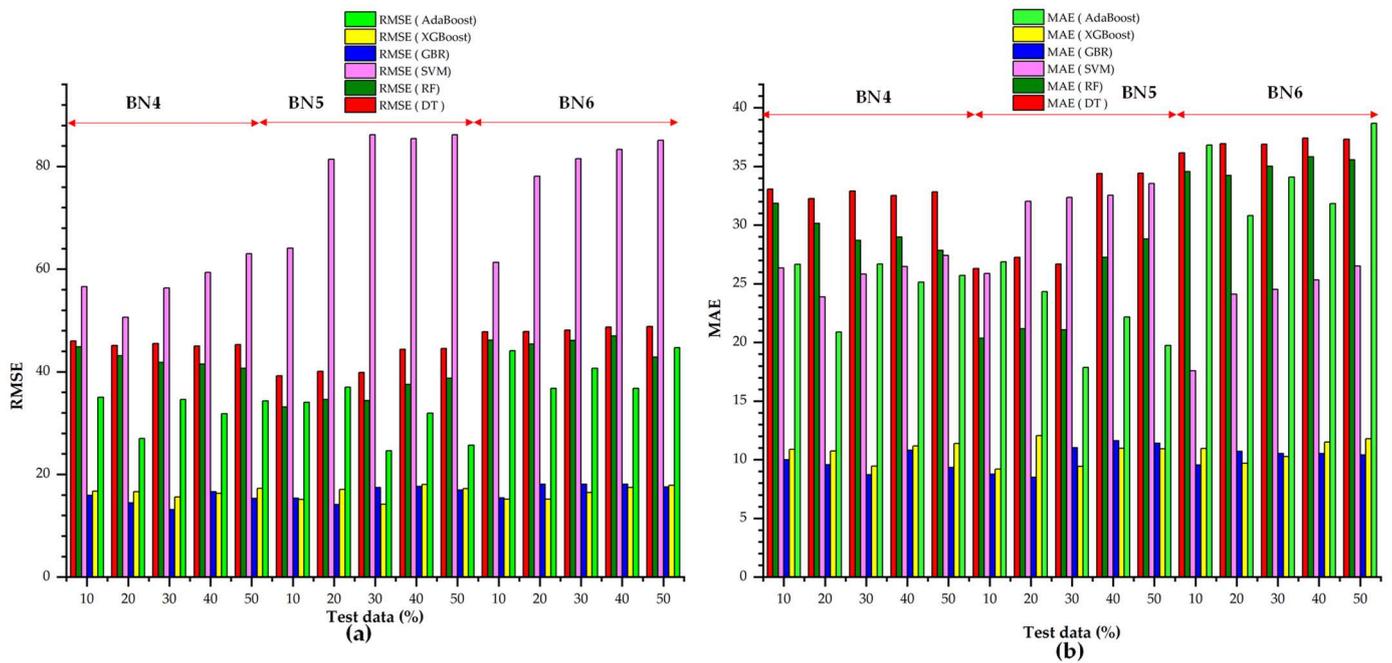


Figure 15. Performance (a) RMSE and (b) MAE results of individual replicates with test data percentage.

The mean absolute error (MAE) is used to assess the temperature forecasting model’s accuracy over the course of ceramic matrix composite (CMC) burning. In the torch test, this measure shows the average absolute difference between the experimental and projected temperature values. Lower MAE values indicate better model performance in comparison to other machine learning models for each sample, especially in the XGBoost and GBR models.

Because various datasets have varying levels of complexity, random forest often has more diverse hyperparameters. The best splitter and fewer hyperparameters are often found in simpler decision trees topologies. AdaBoost trains on time–temperature data using square loss and a learning rate of 1, which is a less complex hyperparameter than that of the decision tree and random forest. The adjustment of hyperparameters for superior model performance, particular to a certain material, is highlighted in this comparison. To ensure flexibility and precision in forecasting the ablation behavior of ceramic matrix composites, each model is customized to fit a specific test data subset.

Across a range of test data splits, the XGBoost and AdaBoost models consistently provide the fastest execution speeds. The execution durations of gradient boosting are modest; it takes more than XGBoost and AdaBoost, but less than SVM and decision tree. In

general, support vector machines take longer to execute than ensemble models. Predictive performance and computational efficiency may have a role in model selection, particularly in the context of big datasets or real-time applications. In this situation, XGBoost and AdaBoost perform exceptionally well.

Although the prediction accuracy of the regression models has been our main emphasis, we recognize the significance of taking computing efficiency into account, particularly in real-world applications. Especially for big datasets or applications where real-time predictions are crucial, computational efficiency is a key component. The XGBoost and GBR models often showed good accuracy in our investigation, but it is vital to remember that these models' computational efficiency might be affected by variables like the quantity and complexity of the dataset.

The XGBoost technique performed efficiently in our study because of its capacity for parallel and distributed computing. Column block encoding and tree pruning are two of its optimization strategies that are responsible for its scalability. However, GBR used sequential boosting which could gradually construct a more intricate and expressive model because of its sequential structure. GBR identified intricate, non-linear patterns in the torch test data since each weak learner adds a component to the overall model. The process of boosting made sure that the model assigned greater weight to hard-to-predict cases, which enhanced its capacity to predict complex interactions in the ablation performance data. Throughout the training phase, GBR allows an adaptive modification of the model's complexity. Hyperparameters such as the number of estimators and the learning rate regulate the inclusion of weak learners. The flexibility of the model guarantees its ability to strike the best possible balance between preventing overfitting and capturing the underlying non-linear patterns.

5. Conclusions

This is the first time in the literature that machine learning approaches have been utilized to predict the thermal performance of boron-nitride-nanoparticle-containing continuous-fiber-reinforced silicon oxy-carbide ceramic matrix composites in the oxy-acetylene torch test. Our study revealed that 70% torch test training data showed better performance accuracy in developing the XGBoost model and 80% torch test training data showed better performance accuracy in developing the GBR model to predict the ablation performance of ceramic matrix composites. The primary shortcoming of the machine learning model developed in this study is its inability to explain the fundamental principles of physics governing the ablation behavior of ceramic matrix composites. This limitation can be addressed by creating machine learning techniques that take physics into account. Future researchers could potentially use the models developed in this study for optimizing the design parameters (fiber volume fraction, resins, boron nitride nanoparticles, fibers, etc.) of continuous-fiber-reinforced silicon oxy-carbide ceramic matrix composites. Future researchers also can use the machine learning models developed in this study for predicting the lifecycle of other thermally resistive composite materials. This research marks the initiation of an effective approach for predicting thermal properties by utilizing oxy-acetylene torch test data, underscoring the significance of machine learning in forecasting the operational longevity of ceramic matrix composites in challenging applications like aeroengines and gas turbines.

Author Contributions: Conceptualization, J.B.D.; methodology, J.B.D.; validation, J.B.D.; formal analysis, J.B.D.; investigation, H.S. and J.B.D.; experimental data resources, H.S.; writing—original draft preparation, J.B.D.; writing—review and editing, J.B.D., C.M. and J.G.; result visualization, J.B.D. and J.G.; supervision, J.G.; project administration, J.G.; funding acquisition, J.G. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: Data will be made available upon request.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Wuchina, E.; Opila, E.; Opeka, M.; Fahrenholtz, B.; Talmy, I. UHTCs: Ultra-high temperature ceramic materials for extreme environment applications. *Electrochem. Soc. Interface* **2007**, *16*, 30. [[CrossRef](#)]
2. Monteverde, F.; Bellosi, A.; Scatteia, L. Processing and properties of ultra-high temperature ceramics for space applications. *Mater. Sci. Eng. A* **2008**, *485*, 415–421. [[CrossRef](#)]
3. Fahrenholtz, W.G.; Hilmas, G.E.; Talmy, I.G.; Zaykoski, J.A. Refractory diborides of zirconium and hafnium. *J. Am. Ceram. Soc.* **2007**, *90*, 1347–1364. [[CrossRef](#)]
4. Opeka, M.M.; Talmy, I.G.; Zaykoski, J.A. Oxidation-based materials selection for 2000 °C+ hypersonic aerosurfaces: Theoretical considerations and historical experience. *J. Mater. Sci.* **2004**, *39*, 5887–5904. [[CrossRef](#)]
5. Chamberlain, A.; Fahrenholtz, W.; Hilmas, G.; Ellerby, D. Characterization of zirconium diboride for thermal protection systems. *Key Eng. Mater.* **2004**, *264–268*, 493–496. [[CrossRef](#)]
6. Savino, R.; Fumo, M.D.S.; Paterna, D.; Serpico, M. Aerothermodynamic study of UHTC-based thermal protection systems. *Aerosp. Sci. Technol.* **2005**, *9*, 151–160. [[CrossRef](#)]
7. Fu, Q.; Zhang, P.; Zhuang, L.; Zhou, L.; Zhang, J.; Wang, J.; Hou, X.; Riedel, R.; Li, H. Micro/nano multiscale reinforcing strategies toward extreme high-temperature applications: Take carbon/carbon composites and their coatings as the examples. *J. Mater. Sci. Technol.* **2022**, *96*, 31–68. [[CrossRef](#)]
8. Jin, X.; Fan, X.; Lu, C.; Wang, T. Advances in oxidation and ablation resistance of high and ultra-high temperature ceramics modified or coated carbon/carbon composites. *J. Eur. Ceram. Soc.* **2018**, *38*, 1–28. [[CrossRef](#)]
9. Ni, D.; Cheng, Y.; Zhang, J.; Liu, J.-X.; Zou, J.; Chen, B.; Wu, H.; Li, H.; Dong, S.; Han, J.; et al. Advances in ultra-high temperature ceramics, composites, and coatings. *J. Adv. Ceram.* **2021**, *11*, 1–56. [[CrossRef](#)]
10. Dusza, J.; Švec, P.; Girman, V.; Sedlák, R.; Castle, E.G.; Csanádi, T.; Kovalčíková, A.; Reece, M.J. Microstructure of (Hf-Ta-Zr-Nb) C high-entropy carbide at micro and nano/atomic level. *J. Eur. Ceram. Soc.* **2018**, *38*, 4303–4307. [[CrossRef](#)]
11. Tallarita, G.; Licheri, R.; Garroni, S.; Barbarossa, S.; Orrù, R.; Cao, G. High-entropy transition metal diborides by reactive and non-reactive spark plasma sintering: A comparative investigation. *J. Eur. Ceram. Soc.* **2019**, *40*, 942–952. [[CrossRef](#)]
12. Chen, X.; Ni, D.; Kan, Y.; Jiang, Y.; Zhou, H.; Wang, Z.; Dong, S. Reaction mechanism and microstructure development of ZrSi₂ melt-infiltrated Cf/SiC-ZrC-ZrB₂ composites: The influence of reform pore structures. *J. Mater.* **2018**, *4*, 266–275. [[CrossRef](#)]
13. Zhang, B.; Yin, J.; Wang, Y.; Yu, D.; Liu, H.; Liu, X.; Reece, M.J.; Huang, Z. Low temperature densification mechanism and properties of Ta_{1-x}Hf_xC solid solutions with decarbonization and phase transition of Cr₃C₂. *J. Mater.* **2021**, *7*, 672–682.
14. Levine, S.R.; Opila, E.J.; Halbig, M.C.; Kiser, J.D.; Singh, M.; Salem, J.A. Evaluation of ultra-high temperature ceramics for aero-propulsion use. *J. Eur. Ceram. Soc.* **2002**, *22*, 2757–2767. [[CrossRef](#)]
15. Gasch, M.; Ellerby, D.; Irby, E.; Beckman, S.; Gusman, M.; Johnson, S. Processing, properties and arc jet oxidation of hafnium diboride/silicon carbide ultra high temperature ceramics. *J. Mater. Sci.* **2004**, *39*, 5925–5937. [[CrossRef](#)]
16. Zhang, X.; Hu, P.; Han, J.; Meng, S. Ablation behavior of ZrB₂-SiC ultra high temperature ceramics under simulated atmospheric re-entry conditions. *Compos. Sci. Technol.* **2008**, *68*, 1718–1726. [[CrossRef](#)]
17. Talmy, I.; Zaykoski, J.; Opeka, M. Synthesis, processing and properties of TaC-TaB₂-C ceramics. *J. Eur. Ceram. Soc.* **2010**, *30*, 2253–2263. [[CrossRef](#)]
18. Song, H. Processing and Characterization of Ultra High Temperature and High Conductive Composites. Ph.D. Thesis, University of Central Florida, Orlando, FL, USA, 2022.
19. Guo, L.; Wang, Y.; Liu, B.; Zhang, Y.; Tang, Y.; Li, H.; Sun, J. In-situ phase evolution of multi-component boride to high-entropy ceramic upon ultra-high temperature ablation. *J. Eur. Ceram. Soc.* **2023**, *43*, 1322–1333. [[CrossRef](#)]
20. Shokrollahi, Y.; Nikahd, M.M.; Gholami, K.; Azamirad, G. Deep Learning Techniques for Predicting Stress Fields in Composite Materials: A Superior Alternative to Finite Element Analysis. *J. Compos. Sci.* **2023**, *7*, 311. [[CrossRef](#)]
21. Islam, M.S.; Rahimi, A. A Three-Stage Data-Driven Approach for Determining Reaction Wheels' Remaining Useful Life Using Long Short-Term Memory. *Electronics* **2021**, *10*, 2432. [[CrossRef](#)]
22. Sirajul Islam, M.; Rahimi, A. Fault Prognosis of Satellite Reaction Wheels Using A Two-Step LSTM Network. In Proceedings of the 2021 IEEE International Conference on Prognostics and Health Management (ICPHM), Detroit, MI, USA, 7–9 June 2021; pp. 1–7. [[CrossRef](#)]
23. Deb, J.; Ahsan, N.; Majumder, S. Modeling the Interplay Between Process Parameters and Part Attributes in Additive Manufacturing Process with Artificial Neural Network. In Proceedings of the ASME 2022 International Mechanical Engineering Congress and Exposition, Columbus, OH, USA, 30 October–3 November 2022; Volume 2A. [[CrossRef](#)]
24. Deb, J.B. Data-Driven Prediction Modeling for Part Attributes and Process Monitoring in Additive Manufacturing. Master's Thesis, Western Carolina University, Cullowhee, NC, USA, 2023.
25. Alam, S.; Deb, J.B.; Al Amin, A.; Chowdhury, S. An artificial neural network for predicting air traffic demand based on socio-economic parameters. *Decis. Anal. J.* **2024**, *10*, 100382. [[CrossRef](#)]
26. Qi, Z.; Zhang, N.; Liu, Y.; Chen, W. Prediction of mechanical properties of carbon fiber based on cross-scale FEM and machine learning. *Compos. Struct.* **2019**, *212*, 199–206. [[CrossRef](#)]

27. Kosicka, E.; Krzyzak, A.; Dorobek, M.; Borowiec, M. Prediction of selected mechanical properties of polymer composites with alumina modifiers. *Materials* **2022**, *15*, 882. [[CrossRef](#)] [[PubMed](#)]
28. Daghigh, V.; Lacy, T.E., Jr.; Daghigh, H.; Gu, G.; Baghaei, K.T.; Horstemeyer, M.F.; Pittman, C.U., Jr. Machine learning predictions on fracture toughness of multiscale bio-nano-composites. *J. Reinf. Plast. Compos.* **2020**, *39*, 587–598. [[CrossRef](#)]
29. Hegde, A.L.; Shetty, R.; Chiniwar, D.S.; Naik, N.; Nayak, M. Optimization and Prediction of Mechanical Characteristics on Vacuum Sintered Ti-6Al-4V-SiCp Composites Using Taguchi's Design of Experiments, Response Surface Methodology and Random Forest Regression. *J. Compos. Sci.* **2022**, *6*, 339. [[CrossRef](#)]
30. Zhang, C.; Li, Y.; Jiang, B.; Wang, R.; Liu, Y.; Jia, L. Mechanical properties prediction of composite laminate with FEA and machine learning coupled method. *Compos. Struct.* **2022**, *299*, 116086. [[CrossRef](#)]
31. Guo, P.; Meng, W.; Xu, M.; Li, V.C.; Bao, Y. Predicting mechanical properties of high-performance fiber-reinforced cementitious composites by integrating micromechanics and machine learning. *Materials* **2021**, *14*, 3143. [[CrossRef](#)]
32. Liu, J.; Zhang, Y.; Zhang, Y.; Kitipornchai, S.; Yang, J. Machine learning assisted prediction of mechanical properties of graphene/aluminium nanocomposite based on molecular dynamics simulation. *Mater. Des.* **2022**, *213*, 110334. [[CrossRef](#)]
33. Karamov, R.; Akhatov, I.; Sergeichev, I.V. Prediction of Fracture Toughness of Pultruded Composites Based on Supervised Machine Learning. *Polymers* **2022**, *14*, 3619. [[CrossRef](#)]
34. Pathan, M.V.; Ponnusami, S.A.; Pathan, J.; Pitongsawat, R.; Erice, B.; Petrincic, N.; Tagarielli, V.L. Predictions of the mechanical properties of unidirectional fibre composites by supervised machine learning. *Sci. Rep.* **2019**, *9*, 13964. [[CrossRef](#)]
35. Bonifácio, A.L.; Mendes, J.C.; Farage, M.C.R.; Barbosa, F.S.; Barbosa, C.B.; Beaucour, A.-L. Application of Support Vector Machine and Finite Element Method to predict the mechanical properties of concrete. *Lat. Am. J. Solids Struct.* **2019**, *16*, 7. [[CrossRef](#)]
36. Hasanzadeh, A.; Vatin, N.I.; Hematibahar, M.; Kharun, M.; Shooshpasha, I. Prediction of the mechanical properties of basalt fiber reinforced high-performance concrete using machine learning techniques. *Materials* **2022**, *15*, 7165. [[CrossRef](#)]
37. Huang, S.-J.; Adityawardhana, Y.; Sanjaya, J. Predicting Mechanical Properties of Magnesium Matrix Composites with Regression Models by Machine Learning. *J. Compos. Sci.* **2023**, *7*, 347. [[CrossRef](#)]
38. Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; Dubourg, V.; et al. Scikit-learn: Machine learning in Python. *J. Mach. Learn. Res.* **2011**, *12*, 2825–2830.
39. Costa, V.G.; Pedreira, C.E. Recent advances in decision trees: An updated survey. *Artif. Intell. Rev.* **2023**, *56*, 4765–4800. [[CrossRef](#)]
40. Kibrete, F.; Trzepieciński, T.; Gebremedhen, H.S.; Woldemichael, D.E. Artificial intelligence in predicting mechanical properties of composite materials. *J. Compos. Sci.* **2023**, *7*, 364. [[CrossRef](#)]
41. Chen, S.; Gu, C.; Lin, C.; Zhang, K.; Zhu, Y. Multi-kernel optimized relevance vector machine for probabilistic prediction of concrete dam displacement. *Eng. Comput.* **2020**, *37*, 1943–1959. [[CrossRef](#)]
42. Park, S.; Jung, S.; Lee, J.; Hur, J. A Short-Term Forecasting of Wind Power Outputs Based on Gradient Boosting Regression Tree Algorithms. *Energies* **2023**, *16*, 1132. [[CrossRef](#)]
43. Nguyen, H.; Cao, M.-T.; Tran, X.-L.; Tran, T.-H.; Hoang, N.-D. A novel whale optimization algorithm optimized XGBoost regression for estimating bearing capacity of concrete piles. *Neural Comput. Appl.* **2023**, *35*, 3825–3852. [[CrossRef](#)]
44. Wen, L.; Li, Y.; Zhao, W.; Cao, W.; Zhang, H. Predicting the deformation behaviour of concrete face rockfill dams by combining support vector machine and AdaBoost ensemble algorithm. *Comput. Geotech.* **2023**, *161*, 105611. [[CrossRef](#)]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.