Preprints.org

Article

A Machine Learning study of the Effect of Process Parameters on Tensile Strength of FDM PLA and PLA-CF

Abdelhamid Ziadia , Mohamed Habibi [*] , Sousso Kelouwani

Article

A Machine Learning Study of the Effect of Process Parameters on Tensile Strength of FDM PLA and PLA-CF

Ziadia Abdelhamid, Habibi Mohamed * and Sousso Kelouwani

Department of Mechanical Engineering, University of Quebec in Trois-Rivieres, 3351 Bd des Forges, Trois-Rivieres, G8Z 4M3, Quebec, Canada

* Correspondence: Mohamed.Habibi@uqtr.ca

Abstract: Material Extrusion is a popular additive manufacturing technology due to its low cost, wide market availability, ability to construct complex parts, safety, and cleanliness. However, optimizing the process parameters to obtain the best possible mechanical properties has not been extensively studied. This paper aims to develop ensemble learning-based models to predict the ultimate tensile strength, Young's modulus, and strain at break of PLA and PLA-CF 3D-printed parts, using printing temperature, printing speed, and layer thickness as process parameters. Additionally, the study investigates the impact of process parameters and material selection on the mechanical properties of the printed parts and uses Genetic Algorithms for multi-objective optimization according to user specifications. The results indicate that process parameters and material selection significantly influence the mechanical properties of the printed parts. The Genetic Algorithm successfully identifies optimal parameter values for the desired mechanical properties. Moreover, this work is the first to model Process-Structure-Properties relationships for an additive manufacturing process and to use a multi-objective optimization approach for multiple mechanical properties, utilizing ensemble learning-based algorithms and Genetic Algorithms.

Keywords: additive manufacturing; material extrusion; machine learning; genetic algorithm; process optimization

1. Introduction

Additive Manufacturing, also known as 3D printing, is revolutionizing the manufacturing and prototyping industries by providing a simpler and more cost-effective product development process than traditional manufacturing techniques. Among various 3D printing technologies, Material Extrusion, or Fused Deposition Modeling (FDM), is the most widely used and rapidly growing technique [1, 2]. Material extrusion was first introduced in 1989 and patented by Scott Crump, the co-founder of Stratasys [3], who also introduced it to the market as Fused Deposition Modeling (FDM). Initially, this technology was also referred to as Fused Filament Fabrication (FFF), which gained popularity after the expiration of the Stratasys FDM patent [1].

The popularity of Material Extrusion can be attributed to its capability to print with a wide range of materials, making it possible to build new technologies with various characteristics. Additionally, its low cost, scalability, and wide market availability make it an attractive option for researchers and industry professionals alike [4].

However, research has shown that parts produced using traditional manufacturing methods exhibit superior mechanical properties compared to those fabricated with Material Extrusion. This is because the parts are printed layer by layer, creating voids that generate internal stress when subjected to mechanical tests. To overcome this obstacle, researchers worldwide are pursuing different strategies to optimize the mechanical properties of 3D-printed parts by selecting the appropriate process parameters [5, 6].

In summary, Material Extrusion, particularly Fused Deposition Modeling, has significantly improved the product development process with its versatility and cost-effectiveness. Despite its

limitations in terms of mechanical properties, ongoing research aims to overcome these challenges and further enhance the capabilities of this innovative technology [4, 5].

Several researchers have employed different approaches to investigate the optimization of process-structure-properties of 3D-printed polymers. Gebisa et al. [5] employed a full factorial design of experiment to study the effect of FDM process parameters on the tensile properties of ULTEM 9085. They investigated five variables, including contours number, contour width, raster width, raster angle, and air gap, and found that raster angle had the greatest influence on mechanical properties. Claver et al. [6] studied the effect of layer height, infill density, and layer orientation on the tensile strength of PLA and ABS. They found that infill density significantly impacted tensile strength, while layer height had little effect.

To enhance the mechanical properties of polymers, researchers have developed composites that combine polymers with reinforcing materials such as fibers, particles, or fillers. Introducing these reinforcing materials allows the resulting composite material to exhibit improved properties such as increased strength, stiffness, toughness, and wear resistance compared to pure polymers. Several researchers have studied and compared the mechanical performance of polymers and composites using various materials and techniques. Ning et al. [7] evaluated the effects of carbon fibers on the mechanical properties of FDM-fabricated parts made of ABS. They found that carbon fibers enhanced tensile strength and Young's modulus but reduced toughness, yield strength, and ductility. Love et al. [8] observed that combining carbon fibers and polymers increased strength, stiffness, thermal conductivity, and reduced distortion in FDM-fabricated parts. Perez et al. [9] investigated the effects of reinforcing materials such as fibers and TiO_2 on the mechanical properties of 3D-printed ABS matrix composites. They found that ABS reinforced with TiO_2 at a 5% weight ratio showed the highest ultimate tensile strength. Aissa et al. [10] experimented with Reinforced Polyamide (RPA) as the printing material and used printing speed, extrusion temperature, and layer thickness as FDM process parameters. They found that extrusion temperature and layer thickness had a more significant impact on tensile strength than printing speed. Zhang et al. [11] developed a data-driven predictive modeling approach to understand the structure-property relationship of FDM-fabricated continuous carbon fiber-reinforced polymers (CCFRP). The ensemble learning algorithm combined eight base learners: multiple linear regression, Least Absolute Shrinkage and Selection Operator (lasso), Multivariate Adaptive Regression Splines (MARS), Generalized Additive Model (GAM), KNN, SVM, Extra- Trees, and XGBoost. Their study concluded that the predictive model accurately predicted the flexural strength of CCFRP specimens, with a minimum RMSE of 9.87%, a minimum RE of 7.75%, and a maximum R^2 of 96.99%.

The primary focus of this study is to use ensemble learning algorithms to predict the mechanical properties of FDM-printed parts using PLA and PLA-CF as printing materials. Additionally, the study aims to understand the impact of process parameters on the mechanical properties of the parts and perform a multi-objective optimization of process parameters to achieve the desired mechanical properties. The rest of this article will be organized as follows: Section 2 will define the machine learning algorithms and performance metrics used in this study. Section 3 will provide details on the experimental and computational methods adopted, and the results obtained will be discussed. Finally, in section 4, the article will be concluded with a summary and future research perspectives.

2. Machine Learning

The four main categories of machine learning techniques are supervised, unsupervised, semi-supervised, and reinforcement. However, in this paper, the focus is on supervised learning, which is a formalization of the concept of learning from examples. The supervised learning approach involves providing the learner with a training set and a test set of labeled data, which it uses to learn from the labeled data to identify the unlabeled data as accurately as possible [12]. This paper focuses explicitly on regression-supervised models since the variables to predict are discrete.

3

2.1. Machine Learning Regression Models

This section provides an in-depth explanation of the various machine-learning regression models employed in this study.

2.1.1. Multiple Linear Regression

Multiple linear regression is a statistical method that enables the prediction of a target variable based on the values of two or more explanatory variables [13]. This technique builds upon linear regression and is sometimes referred to as multiple regression. In this context, the variable to be predicted is designated as the dependent variable, while the variables employed for prediction are called independent variables. The primary objective of this algorithm is to establish a linear association between the independent variables (x) and the dependent variable (y), as illustrated in Equation 1.

$$y = \beta_0 + \beta_1 x_1 + \cdots + \beta_P x_P \tag{1}$$

where β_0 is the value of y when the independent variables are equal to zero, and $\{\beta_1, \ldots, \beta_p\}$ are the estimated regression coefficients.

2.1.2. Decision Tree Regression

A decision tree is a method that employs a tree-like framework for constructing regression models [14]. This approach progressively develops a corresponding decision tree by dividing a dataset into increasingly smaller subsets. The decision tree consists of three types of nodes. The root node serves as the starting point and represents the entire data sample. Interior nodes symbolize the features of the dataset, with their branches denoting the respective decisions. Lastly, leaf nodes signify the outcomes derived from the model.

2.1.3. Least Absolute Shrinkage and Selection Operator (Lasso)

Lasso, a shrinkage method, applies constraints on the coefficients of the least squares estimates [11]. The objective function for the lasso technique is illustrated in Equation 2:

$$J = \sum_{i=1}^{N} \left(y_i - \beta_0 - \sum_{j=1}^{p} \beta_j x_{ij} \right)^2 + \lambda \sum_{j=1}^{p} \beta_j \tag{2}$$

here $\beta 0$ and βj represent the coefficients for the least squares estimates, and λ denotes the tuning parameter that regulates the penalty effect on the estimation of coefficients. The lasso offers an advantage over traditional least squares approaches as the penalty term facilitates managing the trade-off between variance and bias.

2.1.4. Ridge Regression

Ridge Regression, a linear regression variation, contains a regularization method to avoid overfitting and deal with multicollinearity among predictor variables. It predicts the dependent variable y based on the independent variables $x_1, x_2, x_3, \ldots, x_P$ via the linear relationship:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_p x_p + \lambda \sum_{i=1}^{p} \beta_i^2 \tag{3}$$

where, $\beta_0, \beta_1, \beta_2 \ldots, \beta_p$ are the coefficients estimated to minimize the residual sum of squares plus the penalty term. The presence of the penalty term $\lambda \sum_{i=1}^{p} \beta_i^2$ distinguishes Ridge Regression from ordinary linear regression; it penalizes large coefficients with a tuning parameter λ.

Ridge Regression behaves exactly like normal linear regression when ($\lambda = 0$). As λ increases, the model's coefficients approach zero, simplifying the model. To assist the model predict effectively on new data, the best value for λ is commonly chosen through cross-validation. Ridge Regression

performs better when the predictors are correlated, leading to more stable predictions and preventing calculation issues.

2.2. Ensemble Learning Methods

Ensemble learning is a technique that combines multiple base learners to enhance the final prediction [15]. Individual base models might exhibit suboptimal performance due to high variance or bias. However, when these base learners are connected, they form a more robust learner, as their combination effectively reduces bias or variance, leading to improved model performance. In this study, four distinct types of ensemble learning techniques have been employed.

2.2.1. Bagging and Boosting

The key distinction between bagging and boosting lies in their training methodologies. In bagging, the base models are trained concurrently, whereas boosting necessitates sequential training of the base models [16].

The underlying principle of bagging ensemble learning is straightforward; it depends on samples generated using the bootstrap statistical method, which is based on an original dataset [15]. Given N bootstrap samples, each with a size of S (as shown in Equation 4), bagging proceeds accordingly:

$$\{(r_1^1, r_2^1, \dots, r_S^1), (r_1^2, r_2^2, \dots, r_S^2), \dots, (r_1^N, r_2^N, \dots, r_S^N)\} \tag{4}$$

These samples are subsequently utilized to train multiple base models, as demonstrated in Equation 5:

$$\{m_1, m_2, \dots, m_N\} \tag{5}$$

In a regression scenario, as is the case in this study, the average of the predictions from the base models is computed to generate a final prediction with reduced variance, as depicted in Equation 6:

$$fm_i(x) = \frac{1}{N} \Sigma_{i=1}^N m_i(x) \tag{6}$$

In this study, the bagging method employed is **Random Forest (RF)**, which utilizes deep decision trees as weak learners, fitted to bootstrap samples derived from an initial dataset [17]. In RF, feature sampling also occurs, ensuring that each decision tree trains on a random subset of features. Boosting shares similarities with bagging, as both techniques use multiple base models to achieve a better-performing model [15].

The primary difference between the two is that boosting trains base models sequentially, with each model focusing on the data poorly addressed by its predecessor. Upon completing the process, a learner with reduced bias is obtained. The boosting methods utilized in this study include **Gradient Boosting** and **Extreme Gradient Boosting** (XGB).

Gradient Boosting is an approach wherein the ensemble model is constructed as the sum of weighted weak learners, as illustrated in Equation 7:

$$fm_N(.) = \Sigma_{i=1}^N c_i \times m_i(.) \tag{7}$$

This algorithm transforms the problem into a Gradient Descent-based one. During each sequential iteration, a weak learner is fitted to the negative of the current fitting error in relation to the existing ensemble model [18], as demonstrated in Equation 8:

$$fm_i(.) = fm_{i-1}(.) - c_i \times \nabla_{fm_{i-1}} L(fm_{i-1})(.) \tag{8}$$

where L(.) represents the fitting error of the model, c_i is the step size, and $-\nabla_{fm_{i-1}} L(fm_{i-1})(.)$ denotes the opposite of the current fitting error relative to the existing model.

Extreme Gradient Boosting (XGB) is constructed similarly as a sum of weighted weak learners (refer to Equation 6), where the first weak learner is trained on the entire input data. Subsequent models are trained on the residuals to address the limitations of the previous training, continuing until the stopping criterion is achieved.

XGB strives to prevent overfitting while optimizing computational resources [19, 20]. To accomplish this, an objective function (shown in Equation 9) that assesses both loss and regularization should be minimized:

$$Obj^{(t)} = \sum_{k=1}^{n} l(y_i, \hat{y}_i) + \sum_{k=1}^{t} \Omega(f_i) \qquad (9)$$

In this case, n is the number of values, l denotes the loss function, y_i and $\hat{y}_i$ are the actual label and the predicted label, respectively, f_i is the week learner, and Ω is the regulation term (Equation 10) which is defined as:

$$\Omega(f_i) = \gamma T + \frac{1}{2}\lambda \parallel w \parallel^2 \qquad (10)$$

where γ represents the minimum loss necessary to further partition the leaf node, T denotes the number of leaves in the tree, λ is the regularization parameter, and w refers to the weight assigned to each leaf.

2.2.2. Stacking

Unlike bagging and boosting methods, which rely on basic models utilizing the same algorithm, stacking models employ a hierarchical ensemble framework. The stacking model enhances modeling accuracy by combining different classifiers or regressors [21-23]. The simplest stacking model typically comprises two levels: level 0, which includes basic models, and level 1, which features the meta-learner. The basic models (level-0) utilize various algorithms to learn from the original dataset, generating the meta-feature dataset, while the meta-learner (level-1) processes the meta-feature dataset to produce the final results. K-fold cross-validation is commonly employed during the training process of a single basic model to prevent overfitting and ensure that all original datasets contribute to training for generating new meta-feature datasets.

2.2.3. Blending

The primary distinction between blending and stacking lies in the manner in which basic models generate the meta-feature dataset. Instead of using the K-fold cross-validation method, blending reserves a portion of the training dataset (typically 10% or 20%) as the validation set [23]. Figure 1 provides a summary of the ensemble learning blending method.
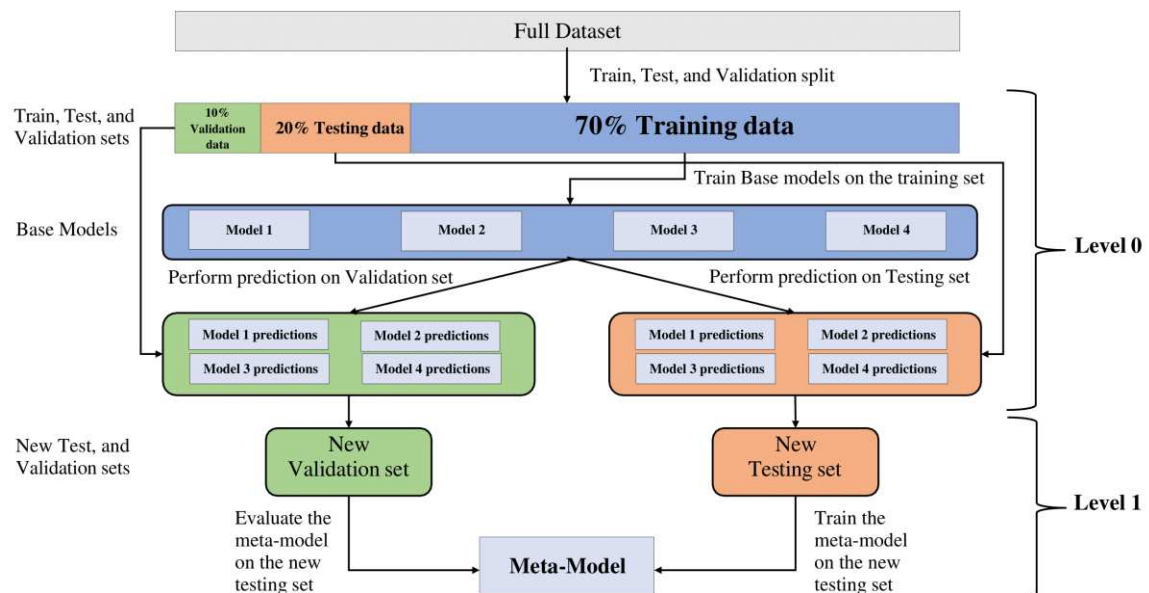
Figure 1. Blending method.

2.3. Performance Metrics

A critical aspect of developing a robust machine learning model is assessing its performance. Various metrics are employed to evaluate the quality of the model. These performance indicators facilitate the measurement of the model's effectiveness based on the available data. By fine-tuning the hyperparameters, it is possible to improve the model's performance.

2.3.1. Root Mean Squared Error (RMSE)

Root Mean Square Error (RMSE) represents the standard deviation of the prediction errors, indicating how closely the data is clustered around the best-fit line [24, 25]. The formula to calculate the RMSE is shown in Equation 11:

$$RMSE = \sqrt{\frac{1}{n}\Sigma_{i=1}^{n}(y_i - \hat{y}_i)^2} \tag{11}$$

In this equation, y represents the actual value, $\hat{y}$ is the predicted value, n denotes the number of values, and i refers to the index of each value.

2.3.2. Coefficient of Determination (R^2)

R^2, also known as the coefficient of determination, represents the difference between 1 and the ratio of the sum of residual squares to the total sum of squares [25], as shown in Equation 12:

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \tag{12}$$

In this equation, y denotes the actual value, $\hat{y}$ is the predicted value, $\bar{y}$ represents the mean value, n refers to the number of values, and i indicates the index of each value.

An R^2 score of 1.0 signifies a perfect fit of the model to the data, while an R^2 value of 0.0 suggests that the predicted values are constant and equal to the mean value of the training data. A negative R^2 score implies that the model's performance is exceptionally poor.

3. Methodology

The CAD models of the test specimens are designed using SolidWorks and saved as STL files. These files are then processed using Ultimaker Cura software, which slices them and generates the G-code for printing the models with the Material Extrusion printer, Ultimaker S5. The test specimens are prepared for tensile strength testing in accordance with the ASTM D638-14 Type I [26] standard method. This test method aims to determine the tensile properties of unreinforced and reinforced plastics in the form of standard dumbbell-shaped test specimens when tested under specified conditions. The test specimens have an overall width of 19mm and an overall length of 165mm.

3.1. Choice of Material

PLA, or polylactic acid, is a biodegradable and bioactive polyester that is derived from renewable resources, such as cornstarch, sugarcane, or cassava roots. It has become a popular material for 3D printing due to its environmental friendliness, ease of printing, and low toxicity. However, its mechanical properties are not as strong as some other materials, such as ABS.

To improve the properties of PLA, researchers have turned to composite materials, such as PLA-CF. Carbon fiber is a strong and lightweight material that can be used to reinforce plastic materials. PLA-CF is a composite material made by adding carbon fiber to PLA. It has improved mechanical properties, such as strength and stiffness, compared to pure PLA.

The use of bioplastics, such as PLA, is becoming increasingly important as concerns about the environment and sustainability grow. Petroleum-based plastics, such as ABS, are non-renewable and

contribute to pollution and waste. The availability of these materials is also decreasing, making it necessary to find alternative materials that are more sustainable. By using PLA and PLA-CF, this study contributes to the development of more environmentally friendly 3D-printing materials.

3.2. Design of the Experiment

The optimization of process-structure-properties of materials is a crucial aspect of 3D printing. In this study, three key parameters were considered for optimizing the 3D printing process of PLA and PLA-CF: Printing Temperature, Layer Thickness, and Printing Speed.

Printing Temperature refers to the temperature at which the material is extruded from the printer nozzle. It is a critical parameter that can significantly impact the final mechanical properties of the printed part. If the temperature is too low, the material may not bond well between layers, leading to a weak and brittle part. On the other hand, if the temperature is too high, the material may burn or degrade, affecting its structural integrity. *Layer Thickness* refers to the thickness of each layer of material that is deposited on top of the previous layer. This parameter affects the resolution and quality of the final printed part. A thinner layer thickness results in higher resolution and smoother surface finish, but it also increases printing time. A thicker layer thickness results in faster printing but may compromise the quality and mechanical properties of the final part. *Printing Speed* refers to the speed at which the printer moves the nozzle across the printing bed. This parameter affects the overall printing time and the quality of the final part. A slower printing speed allows for better adhesion between layers, resulting in a stronger part with a better surface finish. However, it also increases printing time. A faster printing speed reduces printing time but may result in weaker parts with lower quality surface finish.

Other parameters, such as infill density and pattern, and bed temperature, can also affect the final properties of the printed part. However, for this study, these parameters were set as indicated in Table 1 and were not considered in the optimization process.

Table 1. Table of the fixed FDM process parameters and their description.

Factors	Description	Value
Bed Temperature (°C)	Used to heat the build platform	60
Infill density %	The amount of material used in the inside of the print	100
Infill pattern	The form or structure of the material within the component	Lines
Number of contours	The number of contours surrounding the part	1
Number of contours	The number of contours surrounding the part	1

Full Factorial design of the experiment involves testing all possible combinations of the chosen parameter levels to determine their effect on the final properties of the printed part. As depicted in Table 2, three levels were selected for each parameter: Printing Temperature, Layer Thickness, and Printing Speed.

Table 2. Table of the selected factors and their levels.

Factors	Level 1	Level 2	Level 3
Printing Temperature (°C)	200	215	230
Layer Thickness (mm)	0.25	0.35	0.45
Printing Speed (mm/s)	40	50	60

For Printing Temperature, the levels chosen were 200, 215, and 230°C. These temperatures were selected based on the melting point of the PLA and PLA-CF materials and their recommended printing temperature range.

For Layer Thickness, the levels chosen were 0.25, 0.35, and 0.45 mm. This parameter affects the resolution and quality of the final printed part, as well as printing time.

Finally, for Printing Speed, the levels chosen were 40, 50, and 60 mm/s. This parameter affects printing time and the overall quality of the printed part, as well as its mechanical properties.

By applying a Full Factorial design of the experiment, this study can determine the effect of each parameter and their interactions on the final properties of the printed parts. This approach allows for the optimization of the printing process and the production of high-quality parts with the desired mechanical properties.

The full factorial design of the experiment yielded 27 specimens for each material, resulting in a total of 54 specimens for the study. Of these, 27 specimens were printed using pure PLA, and the remaining 27 specimens were printed using PLA-CF. Each specimen represents a unique combination of the chosen parameter levels, allowing for the determination of how each parameter affects the final mechanical properties of the printed part. By testing multiple specimens for each material and parameter combination, the study can determine the consistency and repeatability of the printing process and ensure that the results are statistically significant.

Overall, the use of a full factorial design of the experiment and multiple specimens allows for a comprehensive analysis of the 3D printing process and the properties of the printed parts. This approach can help to optimize the printing process, improve the quality of the final parts, and contribute to the development of more sustainable and environmentally friendly 3D-printing materials.

3.3. Data Collection and Modeling

Tensile tests are a standard method for determining the mechanical properties of materials, including their strength, stiffness, and ductility. In this study, the specimens created using the full factorial design of the experiment were used for tensile tests to evaluate the mechanical properties of PLA and PLA-CF.

The tensile tests were performed using a universal tensile material testing system MTS, which is a commonly used equipment for material testing. The crosshead speed was set to 2 mm/min, which is a typical speed for tensile tests on 3D printed parts. The experimental setup for the tensile test is shown in Figure 2, which includes the MTS equipment, the specimen holder, and the strain gauge for measuring the deformation of the specimen. After the tensile tests were performed, the obtained data were manually preprocessed and stored in CSV files. Preprocessing involves removing any outliers or errors in the data and ensuring that the data is formatted correctly for analysis. The results of data collection and preprocessing are shown in Table A1, which includes the tensile strength, modulus of elasticity, and elongation at break for each specimen.

Figure 2. The experimental setup of the tensile testing.

The tensile test results are used to evaluate the effect of the chosen parameters on the mechanical properties of the printed parts. By analyzing the data and comparing the results for different parameter combinations, the optimal printing parameters can be determined for producing parts with the desired mechanical properties.

In this study, the collected data from tensile tests were used to train ensemble learning models to predict the ultimate tensile strength, Young's modulus, and strain at the break of the printed parts. The blending ensemble learning method was used, which combines multiple weak learners to create a stronger, more accurate predictive model.

The weak learners employed were the bagging method Random Forest, the boosting methods Xgboost and Gradient Boosting, and the regression models Decision Tree Regressor, the Multiple Linear Regression, lasso, and Ridge regression. The Multiple Linear Regression was used as the meta-learner, which combines the predictions of the weak learners to create the final prediction.

To train and evaluate the ensemble learning models, the data was randomly split into training data (70%), testing data (20%), and validation data (10%). The weak learners were trained using the training data, and their predictions were combined to create the data that the meta-learner was trained on. The prediction results of the validation set of the weak learners were combined with the validation set to evaluate the model.

The results of the ensemble learning models (Table 3) showed high accuracy in predicting the mechanical properties of the printed parts. The R^2 score for the prediction of ultimate tensile strength was 91.75%, the R^2 score for Young's modulus was 94.08%, and the R^2 score for strain at break was 88.54% (Figure 3). The RMSE values were also relatively low, indicating that the models have good predictive accuracy.

These results prove that the models trained are capable of predicting with high accuracy the mechanical properties studied. Therefore, the optimization of the process parameters within the range studied became possible. Consequently, it is necessary to analyze the influence of the process parameters used on these mechanical properties.

(a)

(b)

(c)
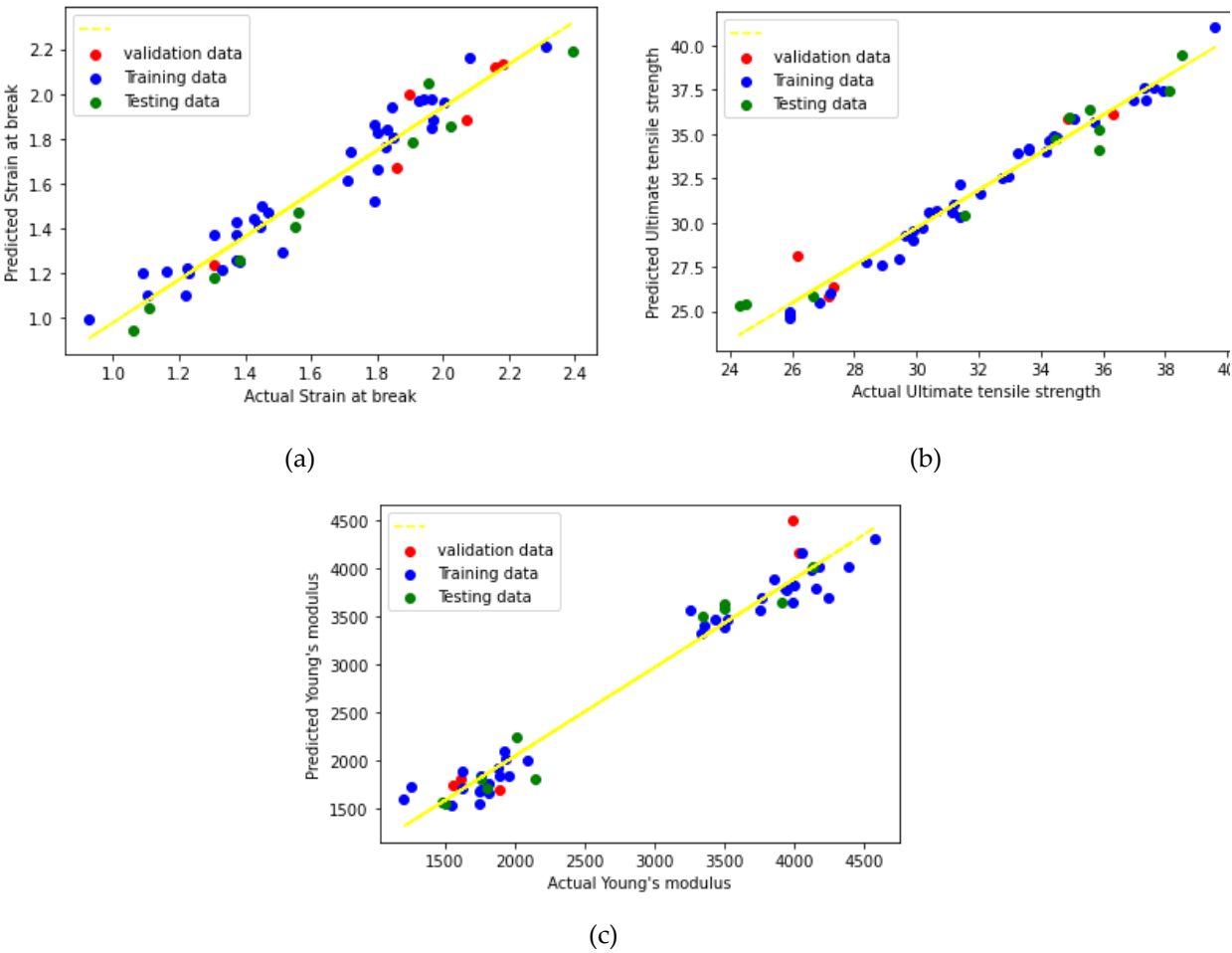
Figure 3. Observed versus predicted mechanical properties: (a) Strain at break, (b) Ultimate tensile strength, (c) Young's modulus.

Table 3. Prediction accuracy of the blending methods.

Property predicted	R² (%)	RMSE	Mean of actual values
Ultimate Tensile strength (σ)	91.75%	1.23	33.87
Young's Modulus (E)	94.08%	278.00	3233.74
Strain at break (ϵ)	88.54%	0.09	1.91

3.4. Influence of the Features Studied on the Mechanical Properties of the Part

Table A1 provides data on the mechanical properties of the PLA and PLA-CF specimens. There is a clear difference between the properties of the two materials, which confirms the effect of carbon fiber on PLA. The addition of carbon fiber improves the mechanical properties of PLA, including its strength, stiffness, and toughness.

Figure 4 presents the behavior of runs 1, 10, 12, and 13 of PLA and PLA-CF. These runs represent different combinations of the printing parameters and their effect on the mechanical properties of the printed parts. The plots show the stress-strain curves for each run, which provide information on the mechanical behavior of the printed parts under tensile load.
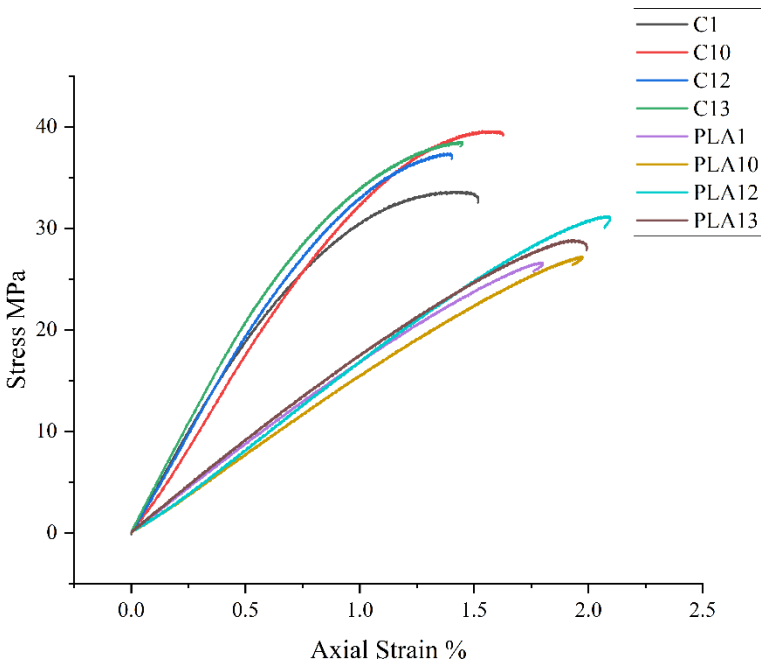
Figure 4. Tensile test behavior of PLA and PLA-CF parts.

3.4.1. Feature Importance.

The Decision Tree Regressor algorithm was used to study the importance of the features, as it provides good interpretability and the ability to output the importance of the features. The importance of the features indicates the extent to which a given input feature affects the outcome, which can be used to evaluate the relative influence of the input features on the mechanical properties of the printed parts.

The results presented in Table 4 show that the material has the greatest importance on all the mechanical properties studied, which confirms the significant effect of carbon fiber on the mechanical properties of PLA. The study of feature importance also confirms that the printing temperature has the most significant influence on the ultimate tensile strength of the printed parts. For Young's modulus and strain at break, the layer thickness is the process parameter with the most significant influence, as shown in Figure 5.
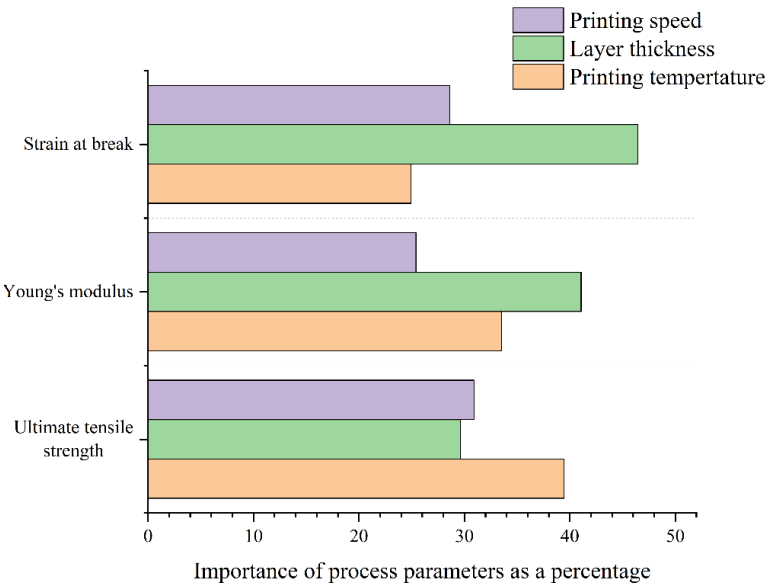
Figure 5. Importance of the process parameters on the mechanical properties.

Table 4. Features importance for each mechanical property studied using decision tree.

	Ultimate tensile strength	Young's Modulus	Strain at break
Material	67.30%	92.70%	71.60%
Printing Temperature	12.89%	2.44%	7.08%
Layer thickness	9.68%	2.99%	13.19%
Printing speed	10.11%	1.85%	8.12%

These findings can be used to optimize the printing process and improve the mechanical properties of the printed parts. By selecting the appropriate material and optimizing the printing temperature and layer thickness, it is possible to achieve the desired mechanical properties for 3D printed parts, contributing to the development of more sustainable and environmentally friendly materials.

Moreover, thanks to the models created to predict the mechanical properties studied, 3d plots of the response surfaces of these properties are created in Figures A1, A2, A3, A4, A5, and A6. These plots can approve the study of the feature importance.

3.4.2. Analysis of the Response Surfaces

The response surfaces presented in Figures B1-B6 provide additional insights into the influence of the printing parameters on the mechanical properties of the PLA and PLA-CF specimens.

For PLA-CF, the response surfaces show that the ultimate tensile strength decreases with increasing layer thickness and printing speed. The printing temperature has a more complex influence on the ultimate tensile strength, with the highest values observed in a range of temperatures between 210°C and 225°C. Young's modulus increases with increasing layer thickness at high printing speeds and low temperatures. The strain at break has its maximum values at a range of layer thicknesses between 0.375mm and 0.425mm, and increases with increasing printing temperature and decreasing printing speed.

For PLA, the response surfaces show that printing temperature has the highest influence on the ultimate tensile strength, with an increase in strength observed with increasing temperature. Layer thickness has a significant influence on Young's modulus, with an increase in modulus observed with increasing layer thickness. The increase in printing speed also produces a higher Young's modulus. The maximum values of strain at break for PLA specimens appear at a range of layer thicknesses between 0.375mm and 0.425mm, a range of printing speed between 45mm/s and 55mm/s, and at low temperatures.

These findings confirm the importance of feature importance analysis, presented in the previous section, and the need for optimizing the printing parameters to achieve the desired mechanical properties.

3.5. Optimization of the Process Parameters

After the analysis presented in Section 3.4, it is concluded that the material used has the most significant influence on the mechanical properties studied. Once the appropriate material is selected, the choice of the process parameters depends on the desired mechanical properties. Therefore, to optimize these relationships and find the optimal set of printing parameters for a given material, a Genetic Algorithm was used.
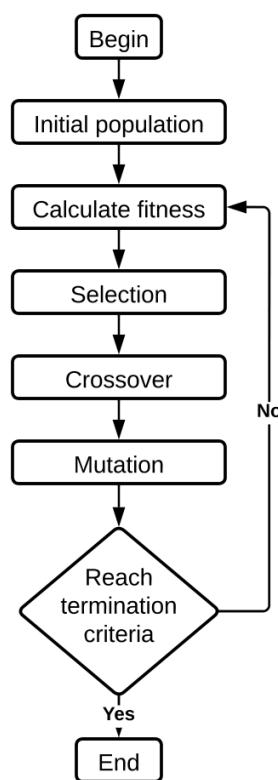
3.5.1. Genetic Algorithm.

The Genetic Algorithm (GA) is a population-based stochastic optimization algorithm that was inspired by Charles Darwin's theory of evolution. The GA algorithm begins by creating an initial population of multiple solutions that represent the chromosomes of the individuals. The fitness of

each individual in the population is then evaluated by calculating a fitness score [28]. Natural selection is applied by selecting the fittest individuals from the population. Crossover and mutation are then applied to produce offspring that descend from the fittest individuals.

The crossover operator takes two parents and produces two offspring by exchanging parts of their chromosomes. The mutation operator introduces small random changes to the offspring's chromosomes to promote diversity in the population. These operators are used to produce a new population, and the process of selection, crossover, and mutation is repeated until the termination criteria are achieved.

The use of GA in 3D printing optimization has become increasingly popular in recent years due to its ability to find the optimal set of printing parameters efficiently. By using GA, it is possible to search through a large parameter space and find the best set of parameters that result in the desired mechanical properties of the printed parts. Figure 6 summarizes the GA process, showing the steps involved in generating the initial population, evaluating fitness, selecting the fittest individuals, and applying crossover and mutation to produce offspring for the next generation.

Figure 6. Flow chart of the Genetic Algorithm.

3.5.2. Optimization of the Process Parameters

The fitness function used in the genetic algorithm plays a crucial role in determining the optimal set of process parameters. In this study, the fitness function is defined as a weighted sum of the predicted mechanical properties using the process parameters. The weights of each mechanical property are determined by the user, based on the application of the printed part and the importance of each property in the design criteria.

The fitness function (Equation 13) includes the ultimate tensile strength, Young's modulus, and strain at break as predicted mechanical properties. The process parameters to optimize are the material (m) used, printing temperature (t), layer thickness (h), and printing speed (s). The predicted mechanical properties, $\hat{\sigma}$, $\hat{E}$, and $\hat{\epsilon}$, are calculated using the ensemble learning models trained in Section 3.3.

$$f(m,t,h,s) = \alpha \times \frac{\hat{\sigma}}{\bar{\sigma}} + \frac{\beta \times \hat{E}}{\bar{E}} + \frac{\gamma \times \hat{\epsilon}}{\bar{\epsilon}} \tag{13}$$

To normalize the mechanical properties, the average values of the mechanical properties in the collected data, $\bar{\sigma}$, $\bar{E}$, and $\bar{\epsilon}$, are used in the fitness function. This normalization is important to ensure that the optimization process is not biased toward a specific mechanical property. The weights of the mechanical properties, α, β, and γ, are assigned by the user. These weights determine the importance of each mechanical property in the optimization process. The user can assign higher weights to the properties that are more critical to the application of the printed part.

By using the fitness function in the genetic algorithm, it is possible to efficiently search for the optimal set of process parameters that result in the desired mechanical properties of the printed parts. This approach can significantly reduce the time and cost required to find the optimal printing parameters for a given material and application.

Equation 14 represents the range of values for the parameters that will be optimized by the Genetic Algorithm. The optimization process aims to maximize the fitness function while keeping the printing temperature, layer thickness, printing speed, and printing material within the specified range. By setting these boundaries, the GA algorithm can efficiently search for the optimal set of parameters that result in the desired mechanical properties of the printed parts. The range of values for each parameter is based on the experimental results obtained in the previous sections and represents the feasible and optimal values for the printing parameters.

$$\begin{cases} \text{maximize } f \\ m \in \{0,1\}; 0 \text{ attributed to PLA and 1 attributed to PLA-CF} \\ 200°\text{C} < t < 230°\text{C} \\ 0.25\text{mm} < h < 0.45\text{mm} \\ 40\text{mm/s} < s < 60\text{mm/s} \end{cases} \tag{14}$$

During each generation, the fitness score of each individual is calculated using the weighted sum of predicted mechanical properties based on the parameters assigned by the user. The top five fittest individuals are then selected to create the next generation through crossover and mutation. This process is repeated until the termination criteria, which is set to a maximum of 250 generations, is reached.

Figure 7 shows the evolution of fitness scores over the 250 generations. As can be seen, the fitness score improves significantly from the initial population to the later generations. This indicates that the genetic algorithm is effective in finding the optimal set of parameters that maximize the desired mechanical properties. The best individual found by the algorithm provides the optimal combination of printing parameters to achieve the desired mechanical properties for the printed parts.
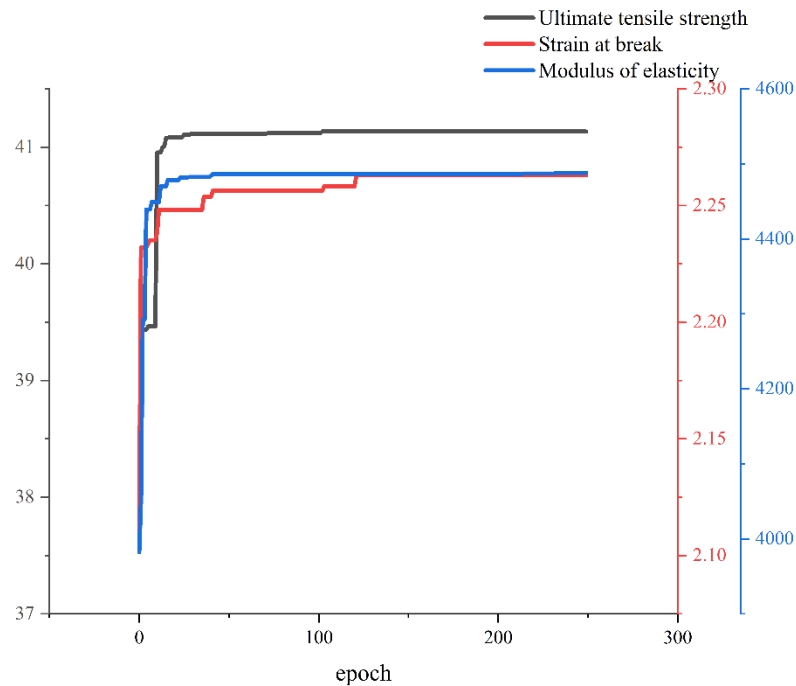
Figure 7. The optimal solution for each iteration of the optimization process for the mechanical properties.

The results summarized in Table 5 show the optimal solutions for different mechanical properties using Fused Deposition Modeling (FDM) 3D printing with PLA-CF (Polylactic Acid - Carbon Fiber) and PLA (Polylactic Acid) materials. The optimal values and parameters for each property are as follows:

1. Ultimate Tensile Strength (UTS) - The optimum solution has a value of 41.129 MPa. The optimal parameters for achieving this value are:

- Printing temperature: 222.28°C
- Layer thickness: 0.261mm
- Printing speed: 40.03mm/s
- Material: PLA-CF

2. Young's Modulus - The optimal value is 4423.63 MPa, and the optimal parameters are:

- Printing temperature: 200.01°C
- Layer thickness: 0.388mm
- Printing speed: 40.038mm/s
- Material: PLA-CF

Note: In this case, the goal was to maximize Young's modulus, which is unusual. To search for the minimum, one could change the weight attributed to Young's modulus in the fitness function by changing its value to -1, for example.

3. Strain at Break - The optimal solution has a value of 2.249%. The parameters for achieving this value are:

- Printing temperature: 200.34°C
- Layer thickness: 0.39mm
- Printing speed: 45.30mm/s
- Material: PLA

These results are confirmed by the response surfaces presented in Figures 9, 10, 11, 12, 13, and 14. These figures demonstrate the relationships between the process parameters and the mechanical properties of the printed materials.

Table 5. Solutions to the optimization process.

Mechanical property	Value	Material	Printing temperature	Layer thickness	Printing speed
Ultimate tensile strength	41.129MPa	PLA-CF	222.28°C	0.261mm	40.30mm/s
Young's Modulus	4423.63MPa	PLA-CF	200.01°C	0.388mm	40.38mm/s
Strain at break	2.249%	PLA	200.34°C	0.390mm	45.30mm/s

4. Conclusion and Future Work

This article presents a data-driven modeling approach to predict the Ultimate Tensile Strength (UTS), Young's Modulus, and Strain at Break of PLA and PLA-CF dog-bone specimens. The specimens were manufactured according to ASTM D638-14 Type I and tested using a universal tensile material testing system (MTS). The study investigated the influence of material type and process parameters such as printing temperature, printing speed, and layer thickness on the mechanical properties. A Genetic Algorithm was employed to optimize the process parameters for specific conditions. Ensemble learning algorithms, including XGBoost, Gradient Boosting Regressor, Random Forest, Decision Tree, Multiple Linear Regression, Lasso, and Ridge Regression, were used to predict the mechanical properties with high accuracy.

This study has shown that PLA-CF specimens exhibit higher Ultimate Tensile Strength (UTS) and Young's Modulus compared to PLA specimens. For PLA-CF specimens, the following trends were observed in relation to process parameters:

1. UTS increases when layer thickness and printing speed decrease, with maximum values at printing temperatures between 210°C and 225°C.

2. Young's Modulus increases with increasing layer thickness and printing speed and decreasing printing temperature.

3. Strain at Break increases with increasing printing temperature and decreasing printing speed, with maximum values for layer thicknesses between 0.375mm and 0.425mm.

For PLA specimens, the following trends were observed:

1. Both UTS and Young's Modulus increase when all studied process parameters (printing temperature, printing speed, and layer thickness) increase.

2. Strain at Break increases when printing temperature decreases, with maximum values for layer thicknesses between 0.375mm and 0.425mm and printing speeds between 45mm/s and 55mm/s.

The Genetic Algorithm used in the study produced values of 41.129 MPa for UTS, 4423.63 MPa for Young's Modulus, and 2.249% for Strain at Break.

The study focused on the effect of printing materials (PLA and PLA-CF) and process parameters (printing speed, printing temperature, and layer thickness) on the mechanical properties of 3D printed parts. Future research aims to investigate the influence of additional process parameters, such as infill density and bed temperature, on the mechanical properties of 3D printed parts, as well as explore more materials to broaden the scope of the study.

Conflicts of Interests: The authors certify that they have NO affiliations with or involvement in any organization or entity with any financial interest or non-financial interest in the subject matter or materials discussed in this manuscript.

Appendix A

**Table A1.** The preprocessed collected data from the mechanical test.

| PLA Run	Ultimate tensile strength (σ)	Modulus of elasticity (E)	Strain at break (ϵ)	PLA-CF Run	Ultimate tensile strength (σ)	Young's Modulus (E)	Strain at break (ϵ)
1	26.66	1746.67	1.79	1	33.61	3754.25	1.42
2	25.92	1744.65	1.84	2	34.17	3356.07	1.44
3	27.34	1505.78	2.00	3	34.44	3497.39	1.55
4	24.50	1543.07	1.95	4	37.93	4391.66	1.16
5	26.85	1202.43	2.31	5	38.14	4174.61	1.30
6	25.93	1482.00	1.96	6	37.61	4581.69	1.10
7	31.41	1812.83	2.08	7	35.85	4125.07	1.08
8	29.42	1891.01	1.89	8	35.74	4132.72	1.10
9	24.30	1258.48	2.18	9	31.13	3984.64	0.92
10	27.23	1556.10	1.96	10	39.60	3505.53	1.56
11	30.20	2143.89	1.97	11	37.34	3986.06	1.30
12	31.21	1629.61	2.06	12	37.39	3853.34	1.38
13	28.87	1809.15	1.92	13	38.53	4153.85	1.44
14	28.39	1762.37	1.94	14	36.99	4000.50	1.30
15	27.16	1617.90	2.024	15	36.31	4060.63	1.22
16	33.59	1801.88	2.15	16	35.57	3768.16	1.37
17	29.87	1756.94	1.83	17	35.88	3939.81	1.22
18	25.90	1625.23	1.717	18	34.43	4034.85	1.06
19	30.21	1895.71	2.054	19	31.40	3497.64	1.37
20	32.03	1755.01	2.39	20	34.65	4250.31	1.38
21	32.95	2094.29	1.90	21	34.88	3941.19	1.51
22	30.64	1958.45	1.82	22	35.05	3339.23	1.70
23	30.38	1881.98	1.80	23	34.24	3518.94	1.47
24	29.65	1922.06	1.79	24	34.53	3256.32	1.79
25	26.15	1897.19	1.45	25	32.76	3343.46	1.37
26	29.91	1931.47	1.85	26	33.24	3432.23	1.33
27	31.54	2011.35	1.85	27	34.89	3912.19	1.22

Appendix B
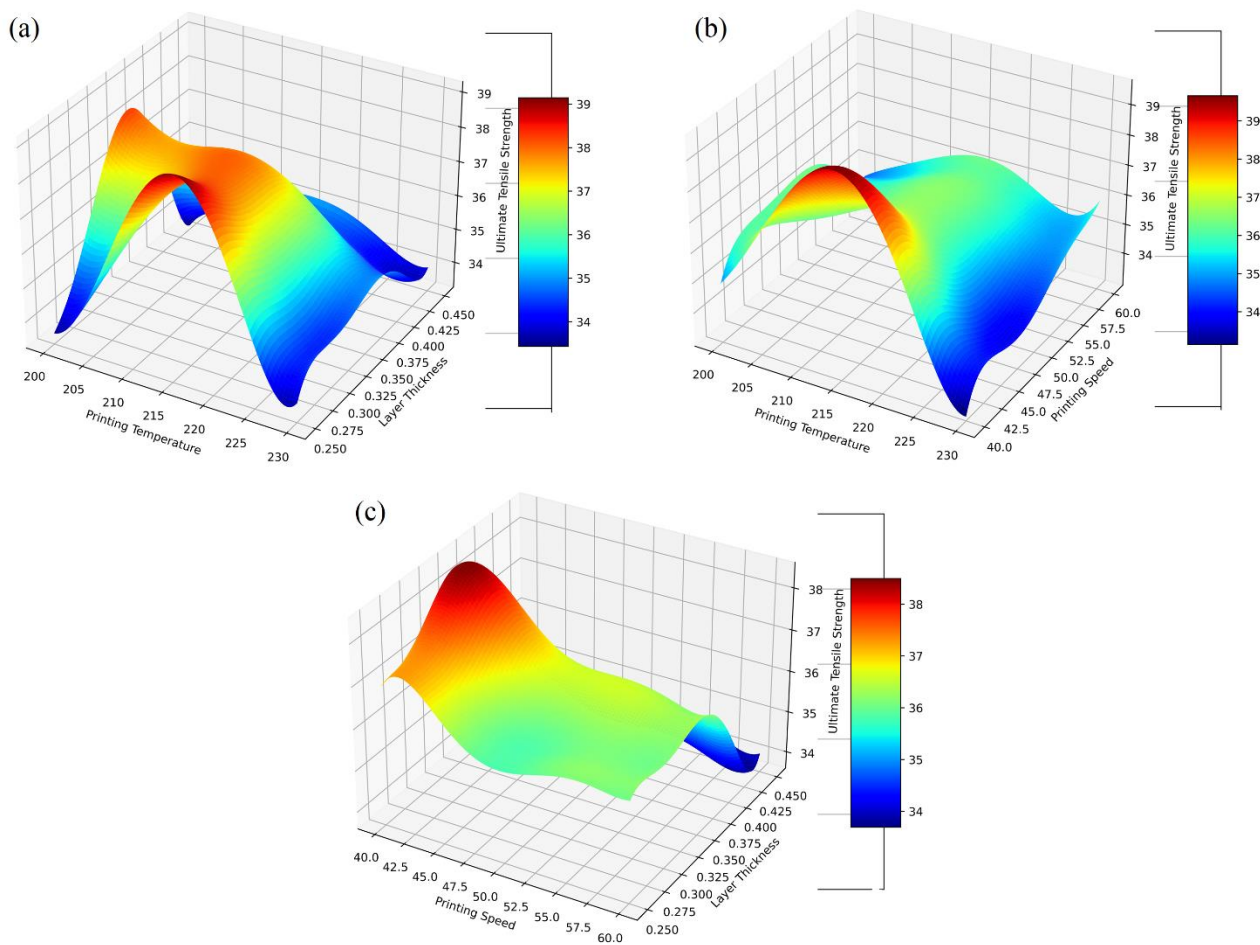
Figure A1. Response surface 3D plots of the predicted ultimate tensile strength of PLA-CF:(a) Printing Temperature vs Layer Thickness, (b) Printing Temperature vs Printing Speed, (c) Printing Speed vs Layer Thickness.
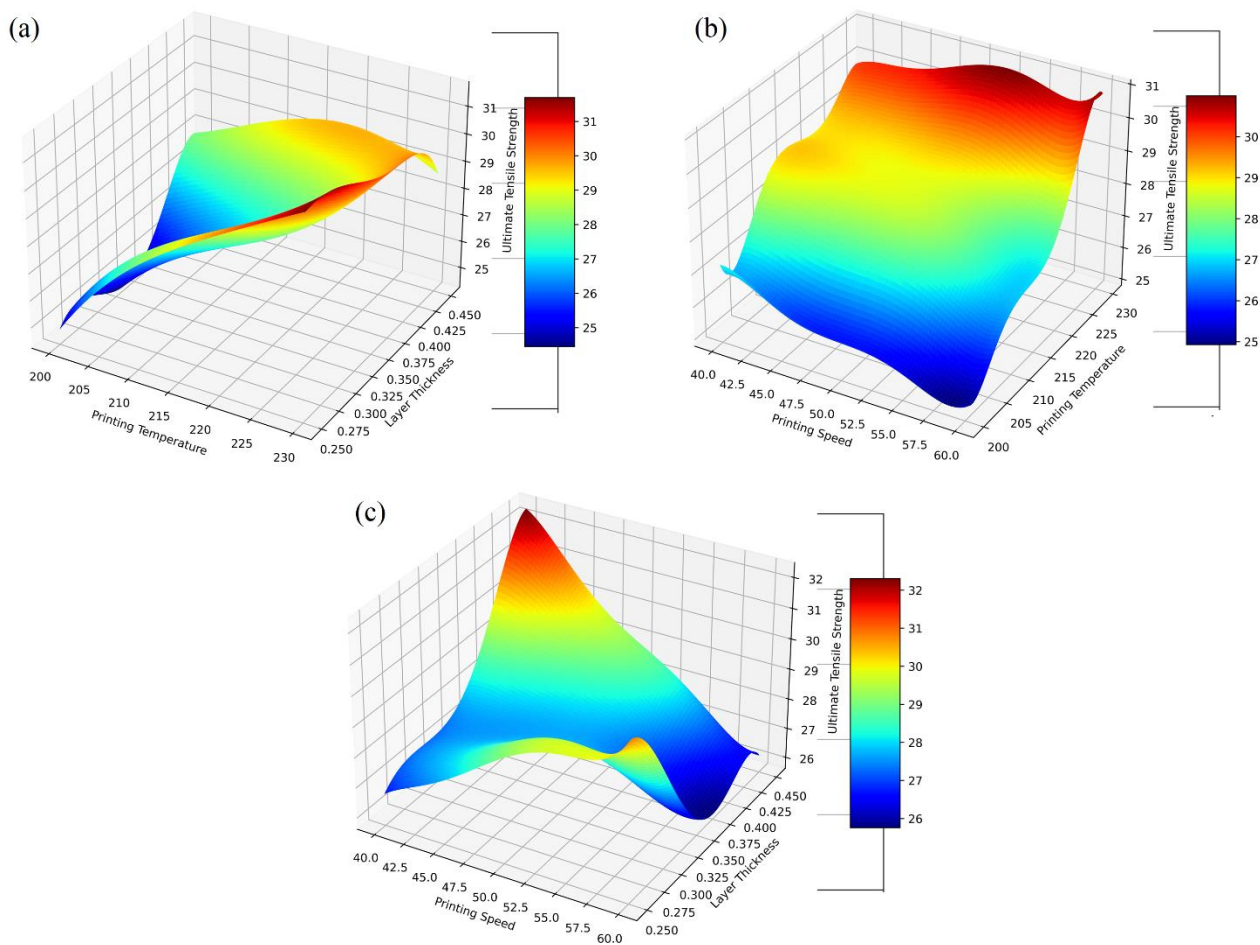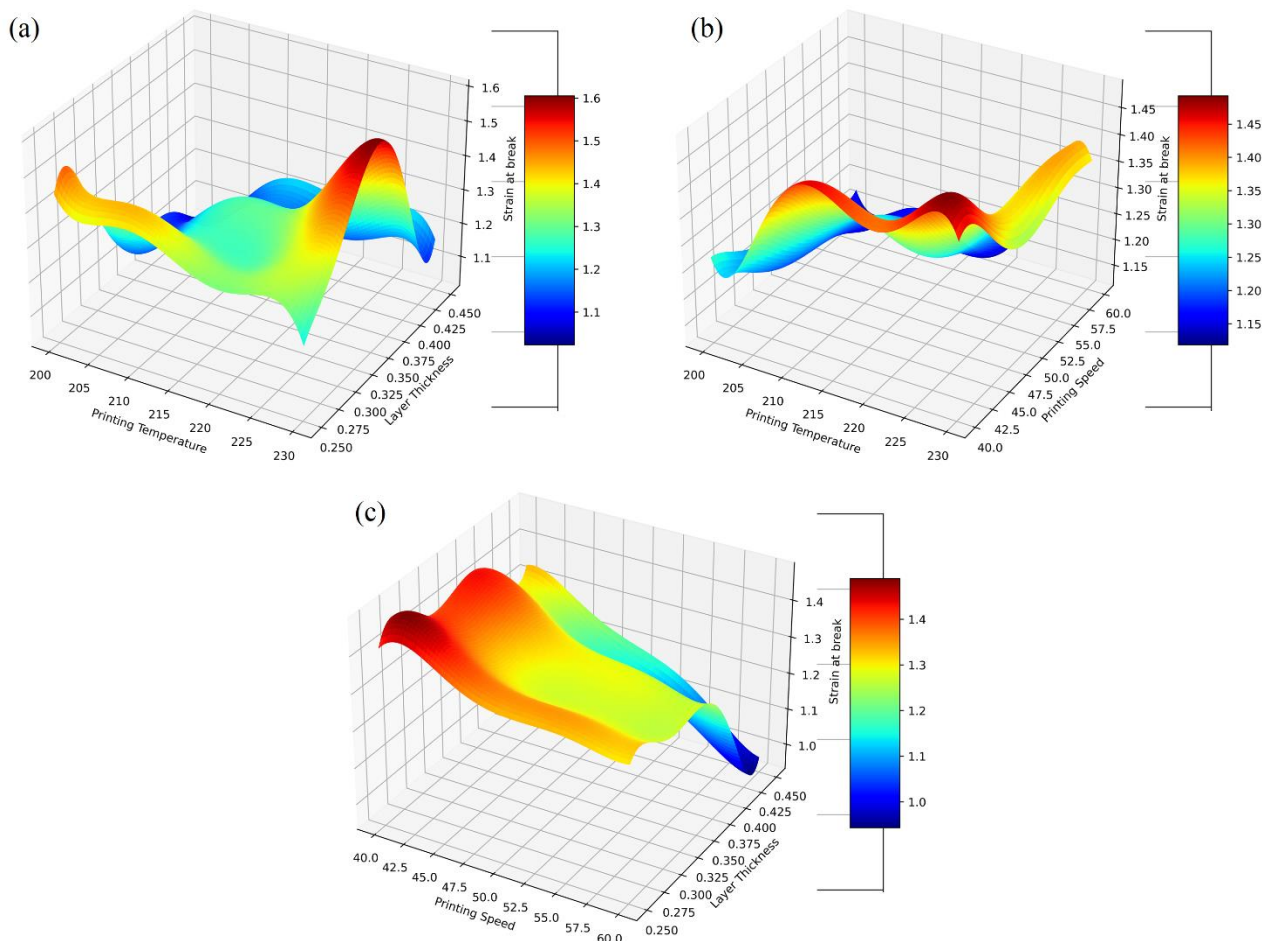
Figure A2. Response surface 3D plots of the predicted ultimate tensile strength of PLA:(a) Printing Temperature vs Layer Thickness, (b) Printing Speed vs Printing Temperature, (c) Printing Speed vs Layer Thickness.
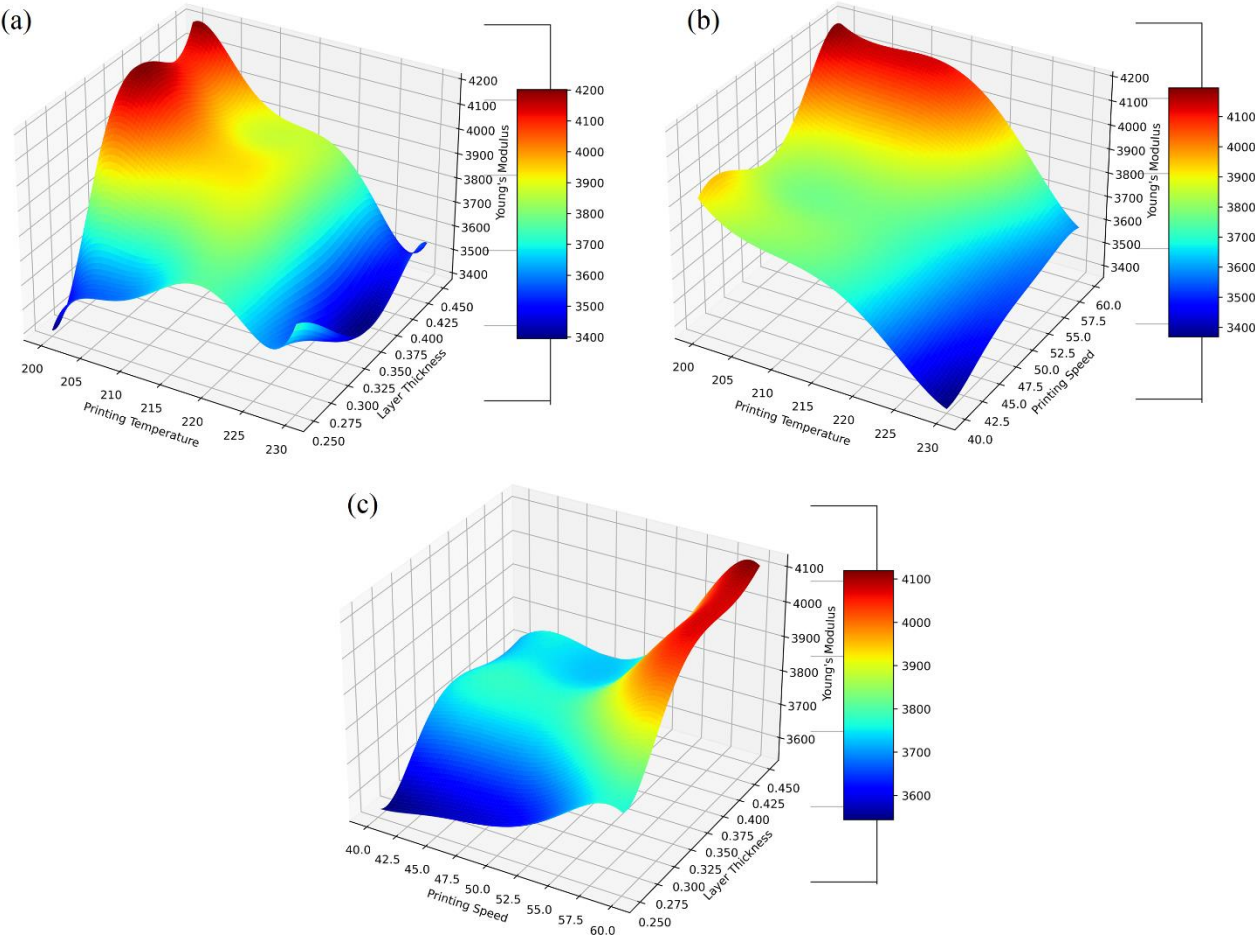
Figure A3. Response surface 3D plots of the predicted Young's modulus of PLA-CF:(a) Printing Temperature vs Layer Thickness, (b) Printing Temperature vs Printing Speed, (c) Printing Speed vs Layer Thickness.
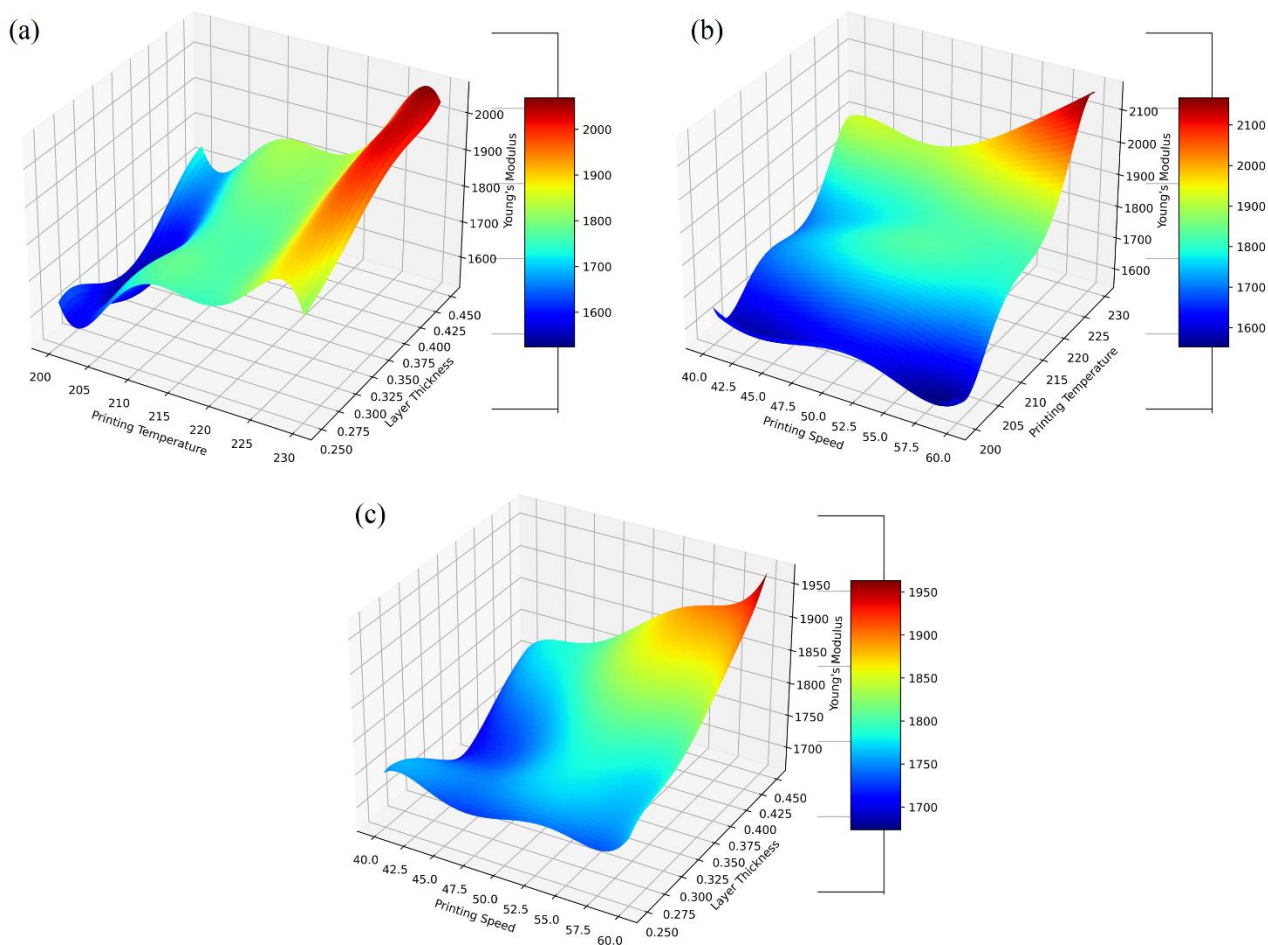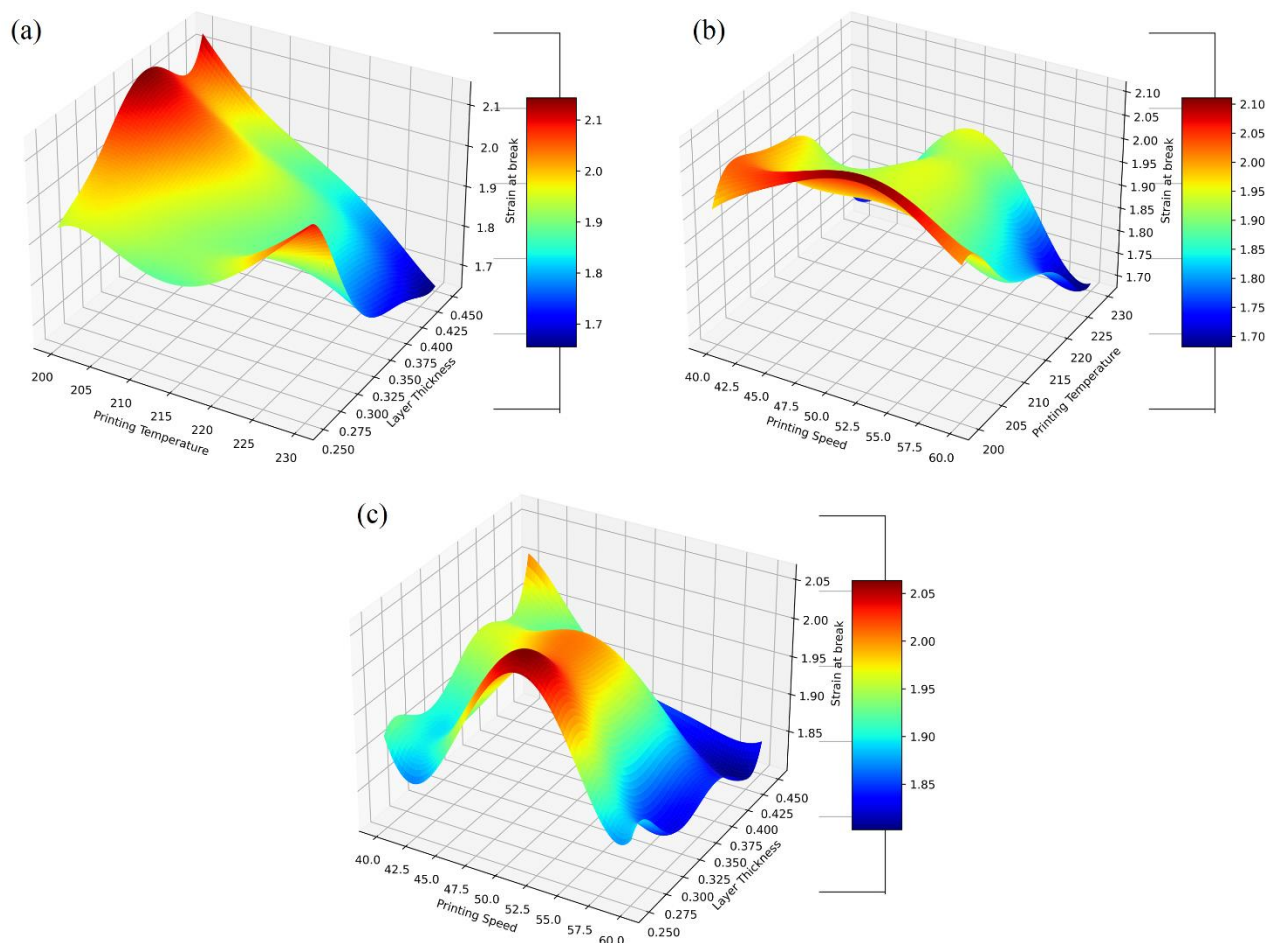
Figure A4. Response surface 3D plots of the predicted Young's modulus of PLA:(a) Printing Temperature vs Layer Thickness, (b) Printing Speed vs Printing Temperature, (c) Printing Speed vs Layer Thickness.

Figure A5. Response surface 3D plots of the predicted Strain at break of PLA-CF:(a) Printing Temperature vs Layer Thickness, (b) Printing Temperature vs Printing Speed, (c) Printing Speed vs Layer Thickness.

Figure A6. Response surface 3D plots of the predicted Strain at break of PLA:(a) Printing Temperature vs Layer Thickness, (b) Printing Speed vs Printing Temperature, (c) Printing Speed vs Layer Thickness.

References

1. Khosravani, M.R., Reinicke, T.: On the environmental impacts of 3D printing technology. Applied Materials Today (2020), **20**, 100689. https://doi.org/10.1016/j.apmt.2020.100689

2. Shaqour, B., Abuabiah, M., Abdel-Fattah, S., Juaidi, A., Abdallah, R., Abuzaina, W., Qarout, M., Verleije, B., Cos, P.: Gaining a better understanding of the extrusion process in fused filament fabrication 3D printing: a review. International Journal of Advanced Manufacturing Technology (2021), **114**(5-6), 1279–1291. https://doi.org/10.1007/ s00170-021-06918-6

3. Radadiya, V.A., Gandhi, A.H.: A Study of Tensile Characteristics for Glass and Carbon Fiber Along with Sandwiched Reinforced ABS Composites. Journal of The Institution of Engineers (India): Series C (2022), **103**(5), 1049–1057. https://doi.org/10.1007/s40032-022-00848-2

4. Farhan Khan, M., Alam, A., Ateeb Siddiqui, M., Saad Alam, M., Rafat, Y., Salik, N., Al-Saidan, I.: Real-time defect detection in 3D printing using machine learning. Materials Today: Proceedings (2020), **42**, 521–528. https://doi.org/10.1016/j.matpr.2020.10.482

5. Gebisa, A.W., Lemu, H.G.: Influence of 3D printing FDM process parameters on tensile property of ultem 9085. Procedia Manufacturing (2019), **30**, 331–338. https://doi.org/10.1016/j.promfg.2019.02.047

6. Rodr´ıguez-Panes, A., Claver, J., Camacho, A.M.: The influence of manufacturing parameters on the mechanical behaviour of PLA and ABS pieces manufactured by FDM: A comparative analysis. Materials (2018), **11**(8). https://doi.org/10.3390/ma11081333

7. Ning, F., Cong, W., Qiu, J., Wei, J., Wang, S.: Additive manufacturing of carbon fiber reinforced thermoplastic composites using fused deposition modeling. Composites Part B: Engineering (2015), **80**, 369–378. https://doi.org/10.1016/j.compositesb.2015.06.013

8. Love, L.J., Kunc, V., Rios, O., Duty, C.E., Elliott, A.M., Post, B.K., Smith, R.J., Blue, C.A.: The importance of carbon fiber to polymer additive manufacturing. Journal of Materials Research (2014), **29**(17), 1893–1898. https://doi.org/10.1557/jmr.2014.212

9. Torrado Perez, A.R., Roberson, D.A., Wicker, R.B.: Fracture surface analysis of 3D-printed tensile specimens of novel ABS-based materials. Journal of Failure Analysis and Prevention (2014), **14**(3), 343–353. https://doi.org/10.1007/s11668-014-9803-9

10. Ouballouch, A., Alaiji, R.E., Ettaqi, S., Bouayad, A., Sallaou, M., Lasri, L.: Evaluation of dimensional accuracy and mechanical behavior of 3D printed reinforced polyamide parts. Procedia Structural Integrity (2019), **19**, 433– 441. https://doi.org/10.1016/j.prostr.2019.12.047

11. Zhang, Z., Shi, J., Yu, T., Santomauro, A., Gordon, A., Gou, J., Wu, D.: Predicting flexural strength of additively manufactured continuous carbon fiber- reinforced polymer composites using machine learning. Journal of Computing and Information Science in Engineering (2020), **20**(6), 1–9. https://doi.org/10.1115/1.4047477

12. Goh GD, Yap YL, Tan HKJ, et al: Process--structure--properties in polymer additive manufacturing via material extrusion: A review. Crit Rev Solid State Mater Sci (2020), 45:113–133

13. Maulud, D., Abdulazeez, A.M.: A Review on Linear Regression Compre- hensive in Machine Learning. Journal of Applied Science and Technology Trends (2020), **1**(4), 140–147. https://doi.org/10.38094/jastt1457

14. Suthaharan, S.: Decision Tree Learning (2016), vol. 1, pp. 237–269. https://doi.org/10.1007/978-1-4899-7641-3 10

15. Seni, G., Elder, J.F.: Ensemble Methods in Data Mining: Improving Accu- racy Through Combining Predictions. Synthesis Lectures on Data Mining and Knowledge Discovery (2010), **2**(1), 1–126. https://doi.org/10.2200/ S00240ED1V01Y200912DMK002

16. Kuncheva, L.I., Skurichina, M., Duin, R.P.W.: An experimental study on diversity for bagging and boosting with linear classifiers. Information Fusion (2002), **3**(4), 245–258. https://doi.org/10.1016/S1566-2535(02)00093-3

17. Biau, G., Scornet, E.: A random forest guided tour. Test (2016), **25**(2), 197–227 https://arxiv.org/abs/1511.05741. https://doi.org/10.1007/ s11749-016-0481-7

18. Hepp, T., Schmid, M., Gefeller, O., Waldmann, E., Mayr, A.: Approaches to regularized regression - A comparison between gradient boosting and the lasso. Methods of Information in Medicine (2016), **55**(5), 422–430. https://doi.org/10.3414/ME16-01-0033

19. Ren, X., Guo, H., Li, S., Wang, S., Li, J.: A novel image classification method with CNN-XGBoost model. Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics) (2017), **10431 LNCS**, 378–390. https://doi.org/10.1007/978-3-319-64185-0 28

20. Dhaliwal, S.S., Nahid, A.A., Abbas, R.: Effective intrusion detection system using XGBoost. Information (Switzerland) (2018), **9**(7). https://doi.org/10.3390/info9070149

21. Wu, T., Zhang, W., Jiao, X., Guo, W., Alhaj Hamoud, Y.: Evaluation of stacking and blending ensemble learning methods for estimating daily reference evapotranspiration. Computers and Electronics in Agriculture (2021), **184**(October 2020), 106039. https://doi.org/10.1016/j.compag.2021.106039

22. Cui, S., Yin, Y., Wang, D., Li, Z., Wang, Y.: A stacking-based ensemble learning method for earthquake casualty prediction. Applied Soft Computing (2021), **101**, 107038. https://doi.org/10.1016/j.asoc.2020.107038

23. Sun, W., Trevor, B.: A stacking ensemble learning framework for annual river ice breakup dates. Journal of Hydrology (2018), **561**(January), 636–650. https://doi.org/10.1016/j.jhydrol.2018.04.008

24. Wang, W., Lu, Y.: Analysis of the Mean Absolute Error (MAE) and the Root Mean Square Error (RMSE) in Assessing Rounding Model. IOP Conference Series: Materials Science and Engineering (2018), **324**(1). https://doi.org/10.1088/1757-899X/324/1/012049

25. Chicco, D., Warrens, M.J., Jurman, G.: The coefficient of determination R-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation. PeerJ Computer Science (2021), **7**, 1–24. https://doi.org/10.7717/PEERJ-CS.623

26. Materials, P., Materials, E.I.: Standard Test Method for Tensile Properties of Plastics (2015), 1, 1–17. https://doi.org/10.1520/D0638-14.1

27. Vinoth Babu, N., Venkateshwaran, N., Rajini, N., Ismail, S.O., Mohammad, F., AlLohedan, H.A., Suchart, S.: Influence of slicing parameters on surface quality and mechanical properties of 3D-printed CF/PLA composites fabricated by FDM technique. Materials Technology (2022), **37**(9), 1008–1025. https://doi.org/10.1080/10667857.2021.1915056

28. Audet, C., Hare, W.: Genetic Algorithms. Springer Series in Operations Research and Financial Engineering, (2017), 57–73. https://doi.org/10. 1007/978-3-319-68913-5 4

29. Author 1, A.B.; Author 2, C.D. Title of the article. *Abbreviated Journal Name* **Year**, *Volume*, page range.