Bao *et al. Green Manuf Open* 2024;2:13
DOI: 10.20517/gmo.2024.052101

Green Manufacturing Open

Research Article

Open Access

Energy efficiency prediction model in fused deposition modeling based on Bayesian optimized random forest

Hong Bao[1,2], Yapeng Li[1], Qingdi Ke[1], Jing Yang[1], Yongshuai Zhou[1]

[1]School of Mechanical Engineering, Hefei University of Technology, Hefei 230009, Anhui, China.
[2]Institute of Intelligent Manufacturing Technology, Hefei University of Technology, Hefei 230051, Anhui, China.

Correspondence to: Prof. Hong Bao, School of Mechanical Engineering, Hefei University of Technology, No. 193 Tunxi Road, Hefei 230009, Anhui, China. E-mail: bhseva7@sina.com

How to cite this article: Bao H, Li Y, Ke Q, Yang J, Zhou Y. Energy efficiency prediction model in fused deposition modeling based on Bayesian optimized random forest. *Green Manuf Open* 2024;2:13. https://dx.doi.org/10.20517/gmo.2024.052101

Received: 21 May 2024 **First Decision:** 29 Jul 2024 **Revised:** 19 Aug 2024 **Accepted:** 24 Aug 2024 **Published:** 1 Sep 2024

Academic Editors: Swee Leong Sing, Zhichao Liu **Copy Editor:** Pei-Yun Wang **Production Editor:** Pei-Yun Wang

Abstract

As the manufacturing industry expands in both scale and energy consumption, the challenge of achieving green and sustainable development becomes more prominent. One effective approach to this challenge is reducing product energy consumption by selecting appropriate process parameters. Process parameters in fused deposition modeling (FDM) play a crucial role in determining the energy consumption during the manufacturing process. Accurately forecasting how these parameters affect energy consumption is essential for realizing green manufacturing in FDM. This paper proposes a method for predicting energy consumption in FDM using Bayesian-optimized random forests (BO-RF). First, response surface methodology (RSM) is utilized to design experiments, obtain the sample datasets, and identify four significant factors affecting processing energy consumption. Then, the sample datasets are partitioned into training and testing subsets; Bayesian optimization is used to optimize the model parameters of the RF, thereby obtaining the prediction model. Finally, using mean square error (MSE), mean absolute percentage error (MAPE), mean percentage error, and coefficient of determination as evaluation metrics, the BO-RF model's predictive performance is evaluated against nine other prediction models. The results demonstrate the BO-RF prediction model's superiority over nine other models.

Keywords: FDM, energy consumption, Bayesian optimization, random forests

INTRODUCTION

Additive manufacturing (AM), often referred to as 3D printing, is a technique for fabricating solid parts by layering materials based on a 3D computer-aided design (CAD) model. Unlike traditional subtractive manufacturing (cutting and machining), AM is a "bottom-up" material accumulation method[1]. Fused deposition modeling (FDM) 3D printing primarily uses filament materials such as polylactic acid (PLA), acrylonitrile butadiene styrene (ABS), and fiber-reinforced composites. The filament is heated and melted, extruded through a nozzle, and then deposited layer by layer onto a preheated substrate, ultimately forming parts with specific shapes and structures. Depending on the characteristics of the part's shape and structure, supports may be added as necessary. Based on its process principles, FDM 3D printing demonstrates high material utilization, aligning with the requirements of green manufacturing. However, this technique still faces challenges such as long printing times and high energy consumption[2]. Due to the unique process characteristics of FDM, numerous printing parameters such as printing speed, infill speed, nozzle temperature, bed temperature, and layer height exist. It is intuitively understood that there is a certain relationship between these parameters and energy consumption. To obtain a more comprehensive understanding of how printing parameters influence energy consumption, researchers have conducted a series of studies using various qualitative and quantitative analysis methods.

In qualitative analysis, the impact of printing parameters and their interactions on energy consumption is studied using experimental design and statistical analysis methods. Griffiths *et al.* employed the design of experiments (DOE) method to investigate the effects of part orientation, number of perimeters, infill rate, and layer thickness on energy consumption[3]. Their results indicated that layer thickness was the most influential factor in determining energy consumption, and the interactions between the studied parameters also had a notable impact on energy consumption. Additionally, using main effects plots and contour plots of energy consumption, they determined the optimal range of printing parameter combinations. Enemuoh *et al.* applied Taguchi orthogonal design, signal-to-noise ratio analysis, and variance analysis to study the effects of five printing parameters: infill rate, infill pattern, layer thickness, printing speed, and number of perimeters on energy consumption[4]. Their findings showed that layer thickness exerted the greatest influence on energy consumption, followed by printing speed, infill rate, infill pattern, and number of perimeters. They identified optimized parameter levels based on the signal-to-noise ratio analysis. Galetto *et al.* used experimental design and variance analysis to examine the impact of layer height, infill rate, nozzle temperature, printing orientation, number of perimeters, printing speed, and retraction speed, as well as the interactions between these parameters, on FDM energy consumption[5]. They found that low infill rate, large layer thickness, and fewer perimeters were conducive to reducing energy consumption. Furthermore, based on the experimental design analysis results, they fitted a polynomial to approximate the relationship between printing parameters and energy consumption, identifying the corresponding optimized parameter combinations for different optimization needs. Elkaseer *et al.* utilized Taguchi orthogonal design, variance analysis, and polynomial fitting methods to establish a quadratic regression model for the relationships between infill rate, layer height, printing speed, printing temperature, and printing orientation with energy consumption[6]. They used this model to predict how printing parameters affect energy consumption, finding that printing speed and layer thickness exerted an important influence on energy consumption. Camposeco-Negrete *et al.* employed Taguchi design and variance analysis to study the effects of layer thickness, infill pattern, printing orientation, and printing position on energy consumption[7]. They further determined the optimized combination of printing parameters through desirability analysis. Al-Ghamdi *et al.* used experimental design and variance analysis to investigate the impact of layer thickness, infill rate, printing speed, and wall thickness on FDM energy consumption[8]. They found strong interactions between the printing parameters and developed a cubic nonlinear regression model. Hassan *et al.* used full factorial design and variance analysis to study how six different infill patterns and four infill rates influence FDM

energy consumption[9]. They established a polynomial regression model discovering that both of them had a notable impact.

In quantitative analysis, it is rare to find studies that utilize physics-based functions to clearly describe how printing parameters affect energy consumption. This is mainly due to the complexity of the relationship, making it challenging to establish an expression using physics-based mathematical models. Therefore, polynomial fitting methods based on quadratic regression are primarily used to establish the relation between printing parameters and energy consumption. Feng *et al.*, based on experimental design results, employed the group method of data handling (GMDH) to establish second-order nonlinear relationship models between three printing parameters and FDM printing energy consumption[10]. Tian *et al.* used a linear regression model to predict the influence of printing parameters on the FDM energy consumption, and proposed a nonlinear optimization approach to identify the optimal combination of printing parameters that minimizes energy consumption while meeting part quality requirements[11]. Alizadeh *et al.* used polynomial fitting to establish a quadratic polynomial regression model for the relationship between layer thickness, printing speed, nozzle temperature, and energy consumption[12]. They then used the nondominated sorting genetic algorithm II (NSGA-II) and technique for order preference by similarity to ideal solution optimization (TOPSIS) methods to solve a multi-objective optimization model for energy efficiency and geometric accuracy of the parts, obtaining optimized combinations of printing parameters based on the quadratic polynomial regression model.

In recent years, with the swift advancements in artificial intelligence, machine learning (ML) has found extensive applications in the mechanical manufacturing sector. Applications such as fault diagnosis[13] and defect monitoring in part manufacturing[14] have shown promising results. In terms of regression prediction, ML techniques are frequently employed to forecast the correlation between design parameters and performance metrics in aerospace optimization design[15], establishing surrogate models for optimization, and achieving effective results. As a data-driven modeling and prediction method, ML has shown unique advantages in multi-input single-output regression prediction and has achieved broader adoption in the field of AM. Xia *et al.* employed three ML algorithms, adaptive neuro fuzzy inference system (ANFIS), extreme learning machine (ELM), and support vector regression (SVR), to establish predictive models for the surface roughness of parts formed by arc AM[16]. The study showed that the genetic algorithm (GA)-ANFIS model, optimized by a GA, had relatively better predictive performance. Li *et al.*, aiming to study the influence of process parameters on the performance of parts formed by selective laser melting (SLM), utilized Kriging, radial basis function (RBF), and SVR to build a composite surrogate model based on local error for predicting material utilization, energy consumption, and tensile strength of SLM parts[17]. The predictive method's effectiveness was verified through experiments. Baturynska *et al.*, addressing the low prediction accuracy of linear regression methods in predicting geometric errors of AM parts, proposed applying ML algorithms such as SVR, decision tree (DT) regression, and multilayer perceptron to predict geometric errors[18]. The comparison with linear regression models indicated that ML algorithms had better prediction performance for geometric errors in AM parts. Cai *et al.*, studying the relationship between dynamic strength and printing parameters of 3D-printed polypropylene composites, used six ML algorithms to establish predictive models[19]. The study found that artificial neural networks (ANNs) had the best predictive performance but required longer prediction times, while SVR performed well in both predictive performance and prediction time. Gor *et al.* used ML algorithms such as ANN, K-nearest neighbors (KNNs), and support vector machines (SVMs) to predict the density of parts produced by powder bed fusion AM[20]. Their results indicated that among the ML algorithms used, SVM exhibited relatively the best predictive performance. Kumar *et al.* applied the KNN algorithm to predict the influence of process parameters on the surface roughness of parts formed by microplasma arc metal AM[21].

Experimental validation showed that the prediction error of surface roughness confirmed the effectiveness of the KNN algorithm in predicting surface roughness for microplasma arc metal AM. Ranjan *et al.* used the random forest (RF) algorithm to predict how process parameters influence mechanical properties, such as bending performance and fracture strength of Al-reinforced ABS composite parts fabricated by fused filament fabrication (FFF) 3D printing[22]. Wu *et al.* employed RF, SVR, ridge regression (RR), and least absolute shrinkage and selection operator (LASSO) regression algorithms to establish predictive models to link FDM printing parameters with the surface roughness of formed parts[23]. The experimental validation results demonstrated the effectiveness of these methods in predicting the surface roughness of FDM parts. Furthermore, Li *et al.* utilized an ensemble learning method to create a ML model comprising six heterogeneous base learners to predict the influence of FFF printing parameters on the surface roughness of formed parts[24]. Their study showed that this predictive method achieved high accuracy in forecasting the surface roughness of FFF parts. Feng *et al.* utilized the GMDH to establish an energy consumption predictive model[10]. They further applied the differential evolution (DE) algorithm to optimize printing parameters for low energy consumption and conducted variance analysis to study the impact of three printing parameters on energy consumption. Zhang *et al.* used a back-propagation neural network (BPNN) to establish an energy efficiency predictive model[25]. Based on the model, they employed an adaptive niche GA to solve an energy efficiency optimization model and obtain optimized printing parameter combinations.

Based on the above analysis, it is evident that ML algorithms have realized broader adoption in AM and predicted the correlation between printing parameters and the quality of formed parts, including surface roughness and mechanical properties, resulting in accurate predictive outcomes. However, for different research problems, such as the regression modeling prediction between FDM printing parameters and energy consumption, the accuracy of predictive results varies depending on the ML algorithm used. In the current research on the correlation between FDM printing parameters and energy consumption, whether using polynomial fitting methods or ML methods for regression prediction, there is a lack of in-depth analysis regarding the predictive accuracy of the methods employed. This leads to an insufficient understanding of the most effective ML methods. Furthermore, it is clear that research on energy efficiency optimization in FDM based on ML methods is limited. Consequently, exploring various ML predictive approaches for FDM energy consumption is essential.

METHODS

Experimental design

Experimental design uses statistical methods to study how to design reasonable experimental schemes, obtain experimental data, and analyze experimental results to understand the effects of experimental factors and their interactions on responses. Experimental design primarily consists of the following three parts:

(1) Designing the Experiment Scheme: Determine the objectives of the experiment, specifically the response indicators. Identify the experimental factors to be studied and their levels, and design a reasonable experimental scheme.
(2) Measuring Experimental Results: Conduct the experiments according to the designed scheme, and record the experimental response results.
(3) Analyzing Experimental Results: Use statistical methods to analyze the experimental results. This involves assessing the significance of the effects of experimental factors on the response and analyzing the impact of interactions between experimental factors on the response.

Response surface methodology experimental design

Response surface methodology (RSM) is a technique that uses appropriate experimental design methods to study how to approximate the functional correlation between design variables, constraints, and target values in complex situations. Its essence lies in approximating the relationship between the inputs x_1, x_2, …, x_n and the output y of actual complex structural systems using response surface fitting. For the response surface function that fits the input-output relationship, in general, lower-order polynomials are sufficient to meet accuracy requirements. The determination of the polynomial order depends on the nonlinearities between the inputs and outputs in the actual structural system. When designing the response surface function, the following principles need to be considered: to minimize the number of experiments as much as possible, lower-order and simple expression polynomial functions should be chosen to meet the fitting accuracy requirements. Based on the previous analysis, selecting a reasonable experimental design method, which involves choosing sample points to fit the response surface, is one of the keys to the effective application of RSM. For response surfaces in the form of second-order polynomials, commonly used experimental design methods include:

(1) Full factorial design

Full factorial design is a fundamental experimental design method in RSM. This method requires running all possible experimental combinations. For a full factorial design with m factors each at n levels, the number of experiments required is n^m. As can be seen, the number of experiments increases exponentially with the number of factors. Therefore, considering the cost and time of experiments, full factorial design is not suitable for situations with a large number of factors.

(2) Central composite design

Central composite design (CCD) is an extension of the two-level full factorial design. It involves adding axial points and center points to ensure that the design requirements are met. For an m-factor CCD, the number of experiments needed is typically $2^m + 2m + 1$. Compared to full factorial design, CCD requires fewer experiments, thus improving experimental efficiency.

(3) Box-Behnken design

Box-Behnken design sets three levels for each experimental factor and is commonly used to study the nonlinear interactions between response variables and factors. For the same number of factors, Box-Behnken design requires fewer experiments compared to CCD, offering higher experimental efficiency.

Establishment of RF model based on Bayesian optimization

DT

DT is a commonly used ML algorithm, and its model training process is based on a tree structure, conducting a series of attribute tests according to the sample datasets. Typically, a single DT model consists of a root node, several internal nodes, and leaf nodes, where the leaf nodes correspond to the results obtained by the model's decision, and the root node and internal nodes correspond to specific attribute judgments. Based on the results of attribute judgments, the sample datasets contained in each node are divided into subsets and assigned to corresponding child nodes. The root node serves as the initial attribute judgment node and contains the complete sample datasets.

For the training sample dataset:

$$D = \{(x_1, y_1), (x_2, y_2), ..., (x_m, y_m)\} \; y_i \in R \tag{1}$$

Attribute set:

$$A = \{a_1, a_2, ..., a_d\} \tag{2}$$

In the case of a classification problem, the training process of a DT model is illustrated in Figure 1, showing that the training of the DT model is a recursive process.

The key issue in achieving effective modeling of DTs is how to divide the optimal attributes. In DT learning, the division of optimal attributes is mainly achieved through the following three criteria:

(1) Information Gain:

$$Gain(D, a) = Ent(D) - \sum_{v=1}^{V} \frac{|D^v|}{|D|} Ent(D^v) \tag{3}$$

$$Ent(D) = \sum_{k=1}^{|y|} p_k \log_2 p_k \tag{4}$$

Where Ent (D) represents the information entropy, and P_k denotes the proportion of samples belonging to the k class in the training sample datasets D, k = 1, 2, …, |y|. D^v indicates the number of samples containing the value of the attribute corresponding to the internal node in the sample datasets D. A larger information gain for a particular attribute indicates that the attribute is more favorable for splitting.

(2) Gain Ratio:

$$Gain_{ratio(D,a)} = \frac{Gain(D, a)}{IV(a)} \tag{5}$$

$$IV(a) = \sum_{v=1}^{V} \frac{|D^v|}{|D|} \log_2 \frac{|D^v|}{|D|} \tag{6}$$

The gain ratio helps address the issue of attribute preference when using information gain as a criterion. Specifically, for attributes with a wider range of values, a higher information gain might occur, potentially leading to erroneous attribute partitioning, such as using sample ID as the optimal splitting attribute.

(3) Gini Index:

$$Gini(D) = \sum_{k=1}^{|y|} \sum_{k' \neq k} p_k p_{k'} = 1 - \sum_{k=1}^{|y|} p_k^2 \tag{7}$$

$$Gini_{index(D,a)} = \sum_{v=1}^{V} \frac{|D^v|}{|D|} Gini(D^v) \tag{8}$$

The aforementioned optimal attribute partitioning criteria correspond to three classic DT algorithms: the ID3 DT algorithm, the C4.5 DT algorithm, and the classification and regression tree (CART) DT algorithm. The CART algorithm is applicable to both classification and regression tasks. Therefore, in the predictive modeling of energy consumption in FDM 3D printing, a DT model based on the CART algorithm is used.

RF model

RF is an ensemble learning algorithm based on the Bagging algorithm, using DTs as base learners. Its primary modeling principle involves establishing several sample datasets through bootstrapping, constructing CART regression tree models for each dataset, and ultimately computing the arithmetic

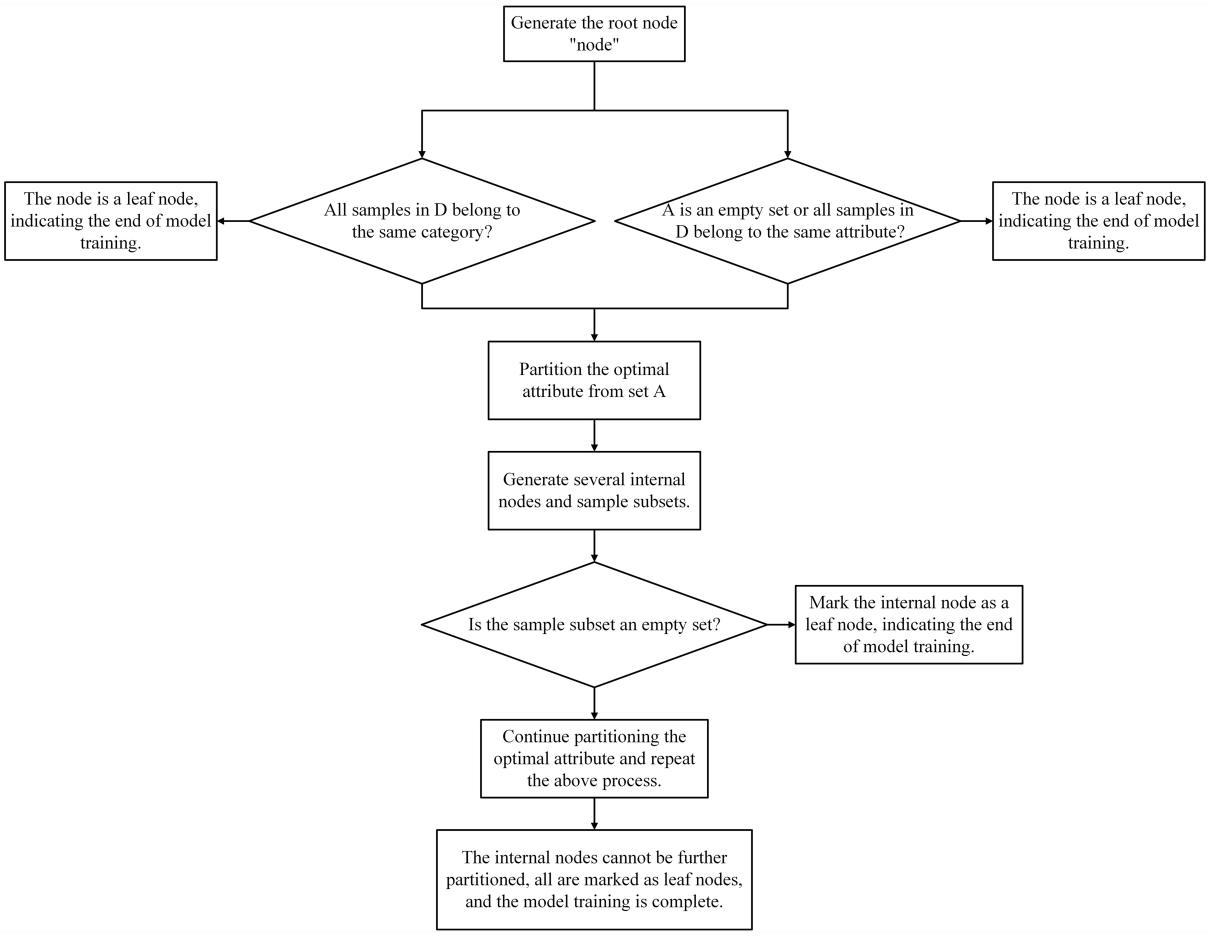

Figure 1. DT model training process. DT: Decision tree.

average of the predictive results from these regression tree models to yield the prediction of the RF model.

The key steps in RF modeling include:
(1) Utilizing bootstrapping to randomly sample the original training datasets, resulting in T training subsets.
(2) Based on the T training subsets obtained in step (1), further establishing T CART regression tree models through random selection of attributes.
(3) Employing each CART regression tree model to predict output results $f_i(x)$.
(4) Calculating the prediction output results $f(x)$ of the RF model through simple averaging, as given in

$$f(x) = \frac{1}{T}\sum_{i=1}^{T} f_i(x) \tag{9}$$

According to the above analysis, the modeling process of RF is illustrated in Figure 2.

Bayesian optimization algorithm

The numbers of regression trees in a RF, leaf nodes, and attributes in the randomly selected subsets all significantly influence the model's prediction performance. Therefore, parameter optimization is crucial for enhancing the performance of predictive models. Commonly used parameter optimization methods include grid search, GAs, particle swarm optimization (PSO), *etc.* In this paper, we adopt the Bayesian optimization

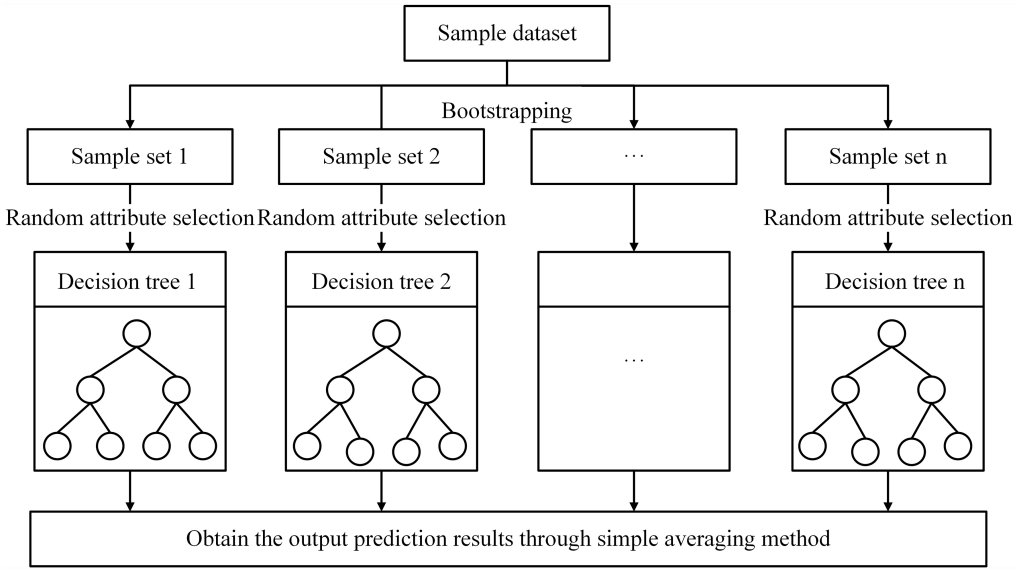

Figure 2. RF modeling process. RF: Random forest.

algorithm for parameter optimization of the RF model. Compared to other optimization algorithms, Bayesian optimization algorithm leverages known sampled information to determine new sampling points, thus possessing the advantage of high optimization efficiency.

The problem of optimizing RF model parameters based on the Bayesian optimization algorithm can be formulated as:

$$x^* = \underset{x \in X}{\arg\min} f(x) \tag{10}$$

Where x^* represents the set of model optimization parameters, x denotes the model parameters to be optimized, X stands for the parameter space for the optimization parameters, and f (x) signifies the objective function to be optimized, which is the prediction error of the RF model.

The solution steps of the Bayesian optimization algorithm mainly include:
(1) Estimating the posterior distribution of the objective function using the probability surrogate model of Gaussian process based on known sampled points and their function values.
(2) Determining the optimal sampling point for the next iteration according to the sampling rule of the acquisition function.

Based on the above analysis, the modeling process of RF based on Bayesian optimization is illustrated in Figure 3.

EXAMPLE VERIFICATION

Energy efficiency function

FDM 3D printing generates energy consumption mainly during the startup, preheating, printing, and reset stages, with the printing stage contributing the highest energy consumption due to its prolonged duration[26]. To consider the feasibility and ease of energy consumption measurement, this study measures the total energy consumption throughout the entire printing process, from the startup to the reset stages of

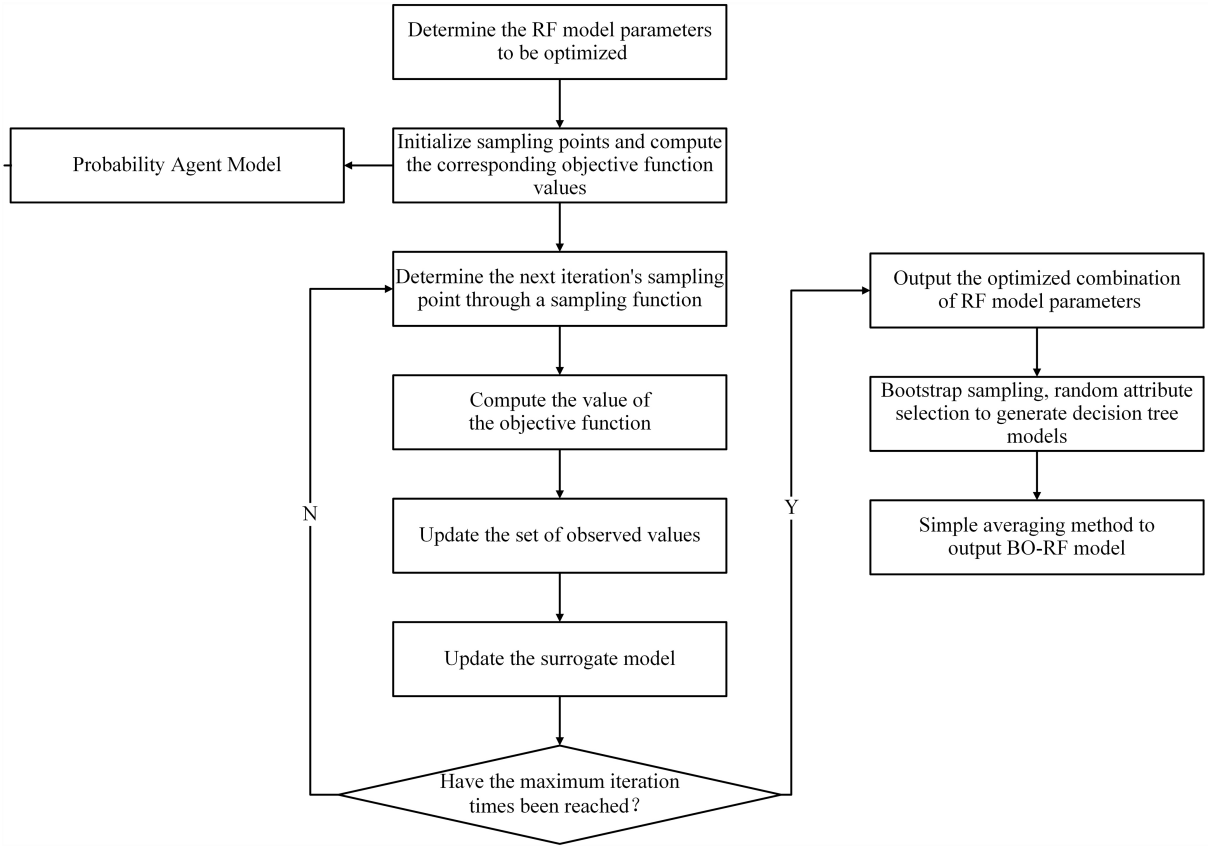

Figure 3. BO-RF modeling process. BO-RF: Bayesian-optimized random forest.

the FDM 3D printer. Given that the energy consumption varies depending on the specific part being printed, using energy consumption as the sole research object yields results that lack general applicability. Therefore, this study further investigates the energy efficiency of the printing process, referred to as printing energy efficiency, based on the energy consumption of the printing process.

There are numerous indicators for evaluating the energy efficiency of the manufacturing process[27], with specific energy and energy utilization rate being two commonly used indicators. In this study, specific energy is adopted as the evaluation indicator for the energy efficiency of FDM 3D printing. In FDM, the specific energy can be represented by the ratio of the total energy consumption during the printing process to the total material consumption. A lower specific energy indicates that less energy is consumed per unit volume of material, thereby signifying higher energy efficiency.

Based on the above analysis, the energy efficiency function for FDM 3D printing, in terms of specific energy, can be expressed as

$$E_{SEC} = \frac{E}{V_m} = \frac{E}{Vf_p} \tag{11}$$

Where E_{SEC} is the specific energy consumption; E is the total energy consumption during the manufacturing process; V_m is the total material consumption of the manufactured parts. V is the volume of the molded part; f_p is the filling density of the molded part. Based on the above analysis, it is evident that the smaller the unit energy consumption and specific energy consumption, the higher the energy utilization efficiency.

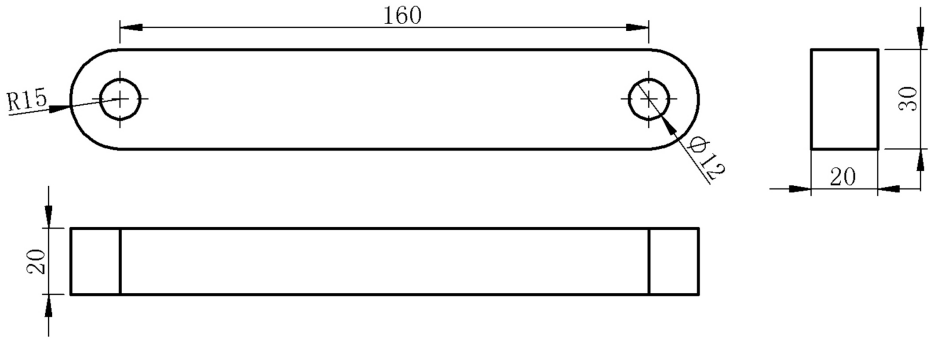

Figure 4. Prototype part drawing.

Experimental materials and equipment

The printing material used in this study is PLA filament, and the selected FDM printing device is the UP PLUS 2. The part diagram of the printed specimen is shown in Figure 4.

After completing the 3D modeling of the part using SolidWorks based on the part drawing, the next step involves setting up the print parameters and performing the slicing operation using Cura software. First, the set values for various process parameters are obtained from the Test design results based on Box-Benhnken. Next, these parameters are configured in the Cura software. Then, the model is exported and sent to the printer for printing. During the printing process, the Aitek AWS2013 power meter is used to measure the energy consumption. Finally, the energy consumption data for FDM 3D printing is obtained. The primary function of the slicing software is to slice the model layer by layer and generate different paths based on the model's shape and the set process parameters, ultimately producing the GCode file for the entire 3D model, which can be exported for offline printing.

Experimental design results

There are numerous printing parameters in FDM 3D printing. These parameters can be categorized into the following three types:

(1) Process-related printing parameters: These include layer height, bed temperature, nozzle temperature, printing speed, infill speed, bottom/top thickness, retraction speed, retraction distance, and so on.

(2) Machine-related printing parameters: These include nozzle diameter, filament width, and other related parameters.

(3) Part geometry-related printing parameters: These include the printing orientation of the part, specific geometric features, and similar attributes.

Based on existing research results from the literature[5] and the controllable parameters of the UP PLUS 2 printer the previously listed parameters were further refined. After analysis, this paper focuses on the following six printing parameters for further study: Layer Height, Nozzle Temperature, Bed Temperature, Print Speed, Travel Speed, and Infill Speed. According to process requirements, the ranges for these parameters are as follows: Layer Height: 0.1-0.3 mm; Nozzle Temperature: 200-220 °C; Bed Temperature: 60-80 °C; Printing Speed: 40-60 mm/s; Travel Speed: 90-110 mm/s; Fill Speed: 60-80 mm/s. Based on these ranges, each experimental factor is set to three levels. The experimental factors and their levels are summarized in Table 1.

Table 1. Print parameters and factors level

Factor Level	Layer height (mm)	Nozzle temperature (°C)	Bed temperature (°C)	Printing speed (mm/s)	Travel speed (mm/s)	Fill speed (mm/s)
1	0.1	200	60	40	90	60
2	0.2	210	70	50	100	70
3	0.3	220	80	60	110	80

Based on the determined experimental factors and levels, the experiment involves six factors with three levels each. Considering the number of experiments and efficiency, this study chooses the Box-Behnken design type. The specific energy consumption (E_{SEC}) is used as the experimental response. Using Design Expert software, an experimental matrix based on the Box-Behnken design type was created, as shown in Table 2, comprising a total of 54 experimental runs.

Analysis of variance

The analysis of variance (ANOVA) significance test method was used to analyze the factors influencing the energy efficiency of FDM printing. When the P-value is ≤ 0.01, the experimental factor has a very significant effect on the response; When $P \leq 0.05$, the effect is significant; When $P > 0.05$, the experimental factor does not significantly affect the response.

The ANOVA results for FDM printing energy efficiency are shown in Table 3. From the results, it can be seen that the model's P-value is < 0.0001 ($P < 0.05$), indicating that the model accurately reflects the real situation. The P-value for the lack of fit is 0.9439 ($P > 0.05$), suggesting that the lack of fit is not significant, and thus, the model has good explanatory power. After validating and confirming the model's effectiveness, it can be further concluded that the factors with a very significant impact on FDM printing energy efficiency are layer height (X1), bed temperature (X3), and printing speed (X4). The factor with a relatively significant impact is fill speed (X6), while the remaining printing parameters have a less significant impact on FDM printing energy efficiency. Regarding the interactions between printing parameters, the interaction between layer height and heated bed temperature has a very significant impact on FDM printing energy efficiency, the interaction between layer height and printing speed has a relatively significant impact, and the interactions between other printing parameters have a less significant impact on FDM printing energy efficiency.

Data preprocessing

The quality of the datasets significantly affects the predictive performance of the model. Therefore, it is essential to preprocess the datasets obtained through experimental design, including feature processing, dataset partitioning and normalization.

Feature processing

The number of features or dimensions in the datasets used for training the BO-RF model can influence the predictive performance of the model. When the dimensionality is too high, the predictive accuracy of the model will decrease. Hence, it is necessary to analyze and process the features of the datasets. In this study, a variance analysis of the response surface experimental design results revealed that layer height, bed temperature, printing speed, and infill speed had significant effects on energy consumption. Therefore, dimensionality reduction was performed on the original datasets with six features, and ultimately, four features were selected for training the BO-RF model.

Table 2. Test design results based on Box-Benhnken

Experimental sequence	Printing parameters						Experimental results	
	Layer height (mm)	Nozzle temperature (°C)	Bed temperature (°C)	Printing speed (mm/s)	Travel speed (mm/s)	Fill speed (mm/s)	Equivalent printing time (s^2/mm^3)	E_{SEC} (J/mm^3)
1	0.2	220	60	50	110	70	0.0014	55.9
2	0.3	210	60	50	100	70	0.0009	40.7
3	0.2	210	60	40	100	70	0.0018	60.1
4	0.2	210	80	60	100	80	0.0010	77.7
5	0.1	210	80	50	100	60	0.0033	139.7
6	0.2	220	80	50	110	70	0.0014	67.7
7	0.3	210	70	40	110	60	0.0014	45.5
8	0.1	210	70	60	110	70	0.0024	103.5
9	0.2	200	70	50	110	70	0.0014	62.0
10	0.3	210	70	60	90	70	0.0008	39.3
11	0.2	220	70	50	90	80	0.0013	60.6
12	0.3	210	80	50	100	60	0.0011	51.1
13	0.1	220	70	40	100	70	0.0036	121.2
14	0.2	210	70	50	100	70	0.0014	61.6
15	0.2	210	70	50	100	70	0.0014	59.2
16	0.2	210	70	50	100	70	0.0014	61.1
17	0.1	210	60	50	100	80	0.0025	89.0
18	0.2	220	70	50	90	60	0.0017	64.9
19	0.3	200	70	40	100	70	0.0012	46.4
20	0.2	210	60	60	100	60	0.0014	51.6
21	0.1	200	70	60	100	70	0.0024	106.1
22	0.2	200	60	50	110	70	0.0014	50.2
23	0.2	210	70	50	100	70	0.0014	61.6
24	0.2	200	70	50	90	80	0.0013	55.9
25	0.1	210	60	50	100	60	0.0033	99.0
26	0.2	210	80	60	100	80	0.0010	63.0
27	0.2	210	70	50	100	70	0.0014	60.1
28	0.2	220	70	50	110	60	0.0017	62.5
29	0.2	200	60	50	90	70	0.0014	48.8
30	0.1	200	70	40	100	70	0.0036	121.2
31	0.2	210	80	40	100	80	0.0016	74.8
32	0.3	220	70	40	100	70	0.0012	46.9
33	0.3	210	60	50	100	80	0.0008	35.0
34	0.2	210	80	40	100	60	0.0021	80.5
35	0.1	210	70	60	90	70	0.0024	108.4
36	0.2	220	80	50	90	70	0.0014	75.3
37	0.2	210	60	40	100	80	0.0016	54.0
38	0.1	210	70	40	90	70	0.0036	124.1
39	0.1	210	80	50	100	60	0.0033	133.5
40	0.1	210	70	40	110	70	0.0036	122.2
41	0.2	200	70	50	90	60	0.0017	63.4
42	0.3	210	70	40	90	70	0.0012	49.7
43	0.2	210	70	50	110	80	0.0013	59.2
44	0.2	210	80	50	90	70	0.0014	69.1
45	0.2	210	80	50	110	70	0.0014	67.7
46	0.3	200	70	60	100	70	0.0008	40.2

47	0.2	220	70	50	110	80	0.0013	57.3
48	0.1	220	70	60	100	70	0.0024	106.5
49	0.2	210	70	50	100	70	0.0014	59.2
50	0.3	220	70	60	100	70	0.0008	42.1
51	0.3	210	70	60	110	70	0.0008	40.7
52	0.2	220	60	50	90	70	0.0014	51.6
53	0.2	210	60	60	100	80	0.0010	49.2
54	0.3	210	80	50	100	80	0.0008	48.8

Table 3. Variance analysis results of print energy efficiency

Source of variance	Sum of squares	Degrees of freedom	Mean square	F value	*P* value
Model	38,469.78	27	1,424.81	131.91	< 0.0001
X_1-layer height	27,298.83	1	27,298.83	2,527.32	< 0.0001
X_2-nozzle temperature	3.91	1	3.91	0.36	0.5525
X_3-bed temperature	2,187.90	1	2,187.90	202.56	< 0.0001
X_4-printing speed	542.36	1	542.36	50.21	< 0.0001
X_5-travel speed	8.27	1	8.27	0.77	0.3897
X_6-fill speed	142.47	1	142.47	13.19	0.0012
X_1X_2	0.50	1	0.50	0.046	0.8313
X_1X_3	233.59	1	233.59	21.63	< 0.0001
X_1X_4	97.84	1	97.84	9.06	0.0058
X_1X_5	0.29	1	0.29	0.027	0.8715
X_1X_6	12.47	1	12.47	1.15	0.2925
X_2X_3	10.43	1	10.43	0.97	0.3348
X_2X_4	0.40	1	0.40	0.037	0.8480
X_2X_5	5.41	1	5.41	0.50	0.4855
X_2X_6	0.94	1	0.94	0.087	0.7704
X_3X_4	3.74	1	3.74	0.35	0.5612
X_3X_5	21.31	1	21.31	1.97	0.1720
X_3X_6	3.14	1	3.14	0.29	0.5943
X_4X_5	2.89	1	2.89	0.27	0.6091
X_4X_6	11.91	1	11.91	1.10	0.3033
X_5X_6	2.22	1	2.22	0.21	0.6542
X_1^2	2,128.44	1	2,128.44	197.05	< 0.0001
X_2^2	0.55	1	0.55	0.051	0.8230
X_3^2	14.95	1	14.95	1.38	0.2500
X_4^2	39.72	1	39.72	3.68	0.0662
X_5^2	1.71	1	1.71	0.16	0.6938
X_6^2	2.58	1	2.58	0.24	0.6289
Residual	280.84	26	10.80		
Lack of fit	147.26	19	7.75	0.41	0.9439
Pure error	133.58	7	19.08		
Total deviation	38,750.62	53			

Dataset partitioning

As per the analysis, 54 sets of experimental data were selected as the sample datasets. After removing duplicate sample data, 47 sets of experimental data were retained. To ensure that the training set had enough samples to meet the model's predictive accuracy, 35 sets of data were randomly assigned for training set data, while the remaining 12 sets were allocated for test set data.

Normalization

In the sample data of ML, there often exists a disparity in the orders of magnitude between different feature attributes. For example, in this paper, the order of magnitude for layer height is 10^{-1}, while for bed temperature, it is 10^{1}. The differences in orders of magnitude between feature attributes can lead to features with higher magnitudes dominating the ML algorithm, causing the model to overlook information contained in other features, thus resulting in a decrease in prediction accuracy. To prevent the differing magnitudes from affecting the predictive accuracy of the BO-RF model, the datasets need to be normalized[28]:

$$x' = \frac{x - x_{min}}{x_{max} - x_{min}}\left(x'_{max} - x'_{min}\right) + x'_{min} \tag{12}$$

Where x' indicates the processed sample data, x'_{max} stands for the maximum value in the processed sample data, x'_{min} corresponds to the minimum value in the processed sample data, x denotes the original sample datasets, x_{max} points to the maximum value in the original sample datasets, and x_{min} represents the minimum value in the original sample datasets. The normalization range chosen in this study is [0, 1].

Model evaluation metrics

To quantitatively evaluate the modeling and predictive performance of different ML models, four evaluation metrics are chosen: mean square error (MSE), coefficient of determination (R^2), mean absolute percentage error (MAPE), and maximum percentage error (MPE).

(1) MSE

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 \tag{13}$$

Where n represents the number of samples in the test set, y_i denotes the measured value of the output response, and $\hat{y}_i$ means the predicted value of the output response. A smaller MSE indicates a better predictive capability of the model.

(2) R^2

$$R^2 = \frac{(n\sum_{i=1}^{n} y_i\hat{y}_i - \sum_{i=1}^{n}\hat{y}_i\sum_{i=1}^{n} y_i)^2}{(n\sum_{i=1}^{n}\hat{y}_i^2 - (\sum_{i=1}^{n}\hat{y}_i)^2)(n\sum_{i=1}^{n} y_i^2 - (\sum_{i=1}^{n} y_i)^2)} \tag{14}$$

The R^2 characterizes the percentage of variability in a dependent variable that is explained by the model. The optimal value is 1, indicating excellent performance of the regression prediction model. The closer the value of R^2 is to 1, the better the predictive performance of the model.

(3) MAPE

$$MAPE = \frac{1}{N}\sum_{i=1}^{N}\left|\frac{y_i - \hat{y}_i}{y_i}\right| \times 100\% \tag{15}$$

(4) MPE

$$MPE = \max\left|\frac{y_i - \hat{y}_i}{y_i}\right| \times 100\% \tag{16}$$

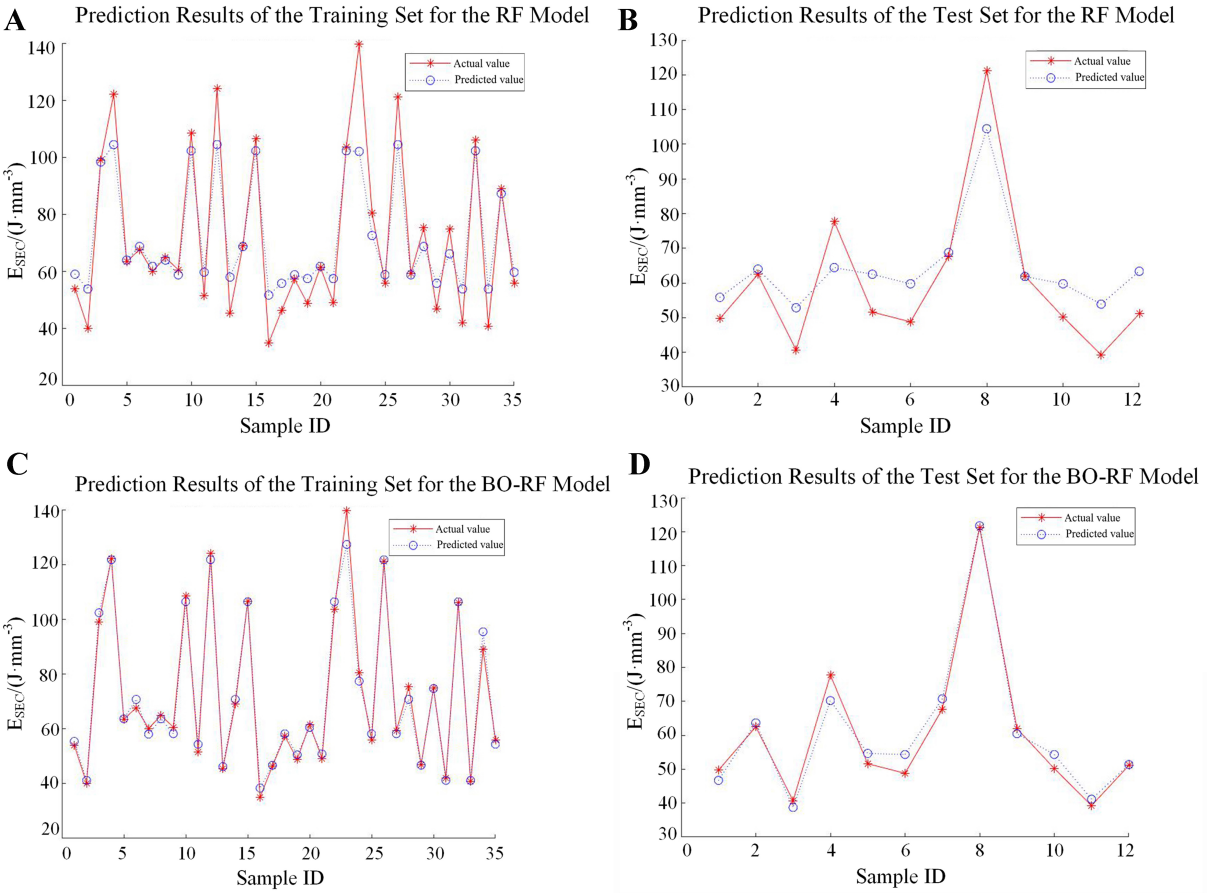

Figure 5. Energy efficiency modeling prediction results of RF and BO-RF models. (A) Prediction results of the training set for the RF model; (B) Prediction results of the test set for the RF model; (C) Prediction results of the training set for the BO-RF model; (D) Prediction results of the test set for the BO-RF model. RF: Random forest; BO-RF: Bayesian-optimized random forest.

In the four model evaluation metrics, MSE is used to assess the dispersion of prediction results, R^2 is used to evaluate the linear correlation between predicted values and actual values, while MAPE and MPE are employed to evaluate the accuracy of prediction results.

RESULTS

In the BO-RF model, the Bayesian optimization algorithm is used to optimize the parameters of the RF model, including the numbers of DTs, leaf nodes, and attributes in the randomly selected subsets. The sample datasets consist of the 47 experimental data sets introduced in the data preprocessing section, with 35 sets used for the training set and the remaining 12 sets for the testing set. Figure 5 shows the energy efficiency prediction results based on the RF and the BO-RF model. Here, E_{SEC} is used to represent energy consumption, which is the energy consumption per unit volume of material processed.

To validate the superiority of the BO-RF model, comparisons are made among the BO-RF energy efficiency prediction model, RF-based energy efficiency prediction model, SVR-based energy efficiency prediction model, and BPNN-based energy efficiency prediction model in terms of the four evaluation metrics: MSE, MPE, MAPE, R^2 [Table 4]. To visually represent the comparison results, a comparison histogram is plotted [Figure 6]. Among these models, GS-SVR stands for grid search-based SVR, GA-SVR for GA-based SVR, PSO-SVR for PSO-based SVR, BO-SVR for Bayesian optimization-based BVR, BO-MK-SVR for Bayesian optimization-based multi-kernel SVR, GA-BPNN for GA-based BPNN, PSO-BPNN for PSO-based BPNN,

Table 4. Comparison of prediction performance of each energy efficiency modeling

	MSE		R^2		MAPE		MPE	
	Training set	Test set	Training set	Test set	Training set	Test set	Training set	Test set
BO-RF	0.0008	0.001	0.9891	0.974	2.8%	5.0%	9.7%	11.4%
RF	0.0105	0.010	0.9352	0.909	11.8%	16.8%	48.0%	37.1%
GS-SVR	0.0013	0.002	0.9852	0.950	3.5%	7.5%	11.4%	9.6%
GA-SVR	0.0003	0.002	0.9966	0.950	1.9%	6.6%	6.8%	21.8%
PSO-SVR	0.0422	0.021	0.8901	0.749	15.8%	14.2%	41.4%	37.8%
BO-SVR	0.0005	0.002	0.9931	0.965	2.3%	6.3%	9.2%	14.2%
BO-MK-SVR	0.0058	0.009	0.9214	0.816	11.5%	16.8%	24.9%	27.2%
GA-BPNN	0.0046	0.004	0.9554	0.927	5.5%	9.3%	53.3%	30.8%
PSO-BPNN	0.0003	0.003	0.9963	0.944	1.5%	7.0%	8.2%	17.1%
BO-BPNN	0.0033	0.017	0.9573	0.646	4.4%	13.0%	45.0%	50.4%

MSE: Mean square error; R^2: coefficient of determination; MAPE: mean absolute percentage error; MPE: maximum percentage error; BO-RF: Bayesian-optimized random forest; RF: random forest; GS-SVR: grid search-based support vector regression; GA-SVR: genetic algorithm-based support vector regression; PSO-SVR: particle swarm optimization-based support vector regression; BO-SVR: Bayesian optimization-based support vector regression; BO-MK-SVR: Bayesian optimization-based multi-kernel support vector regression; GA-BPNN: genetic algorithm-based back-propagation neural network; PSO-BPNN: particle swarm optimization-based back-propagation neural network; BO-BPNN: Bayesian optimization-based back-propagation neural network.

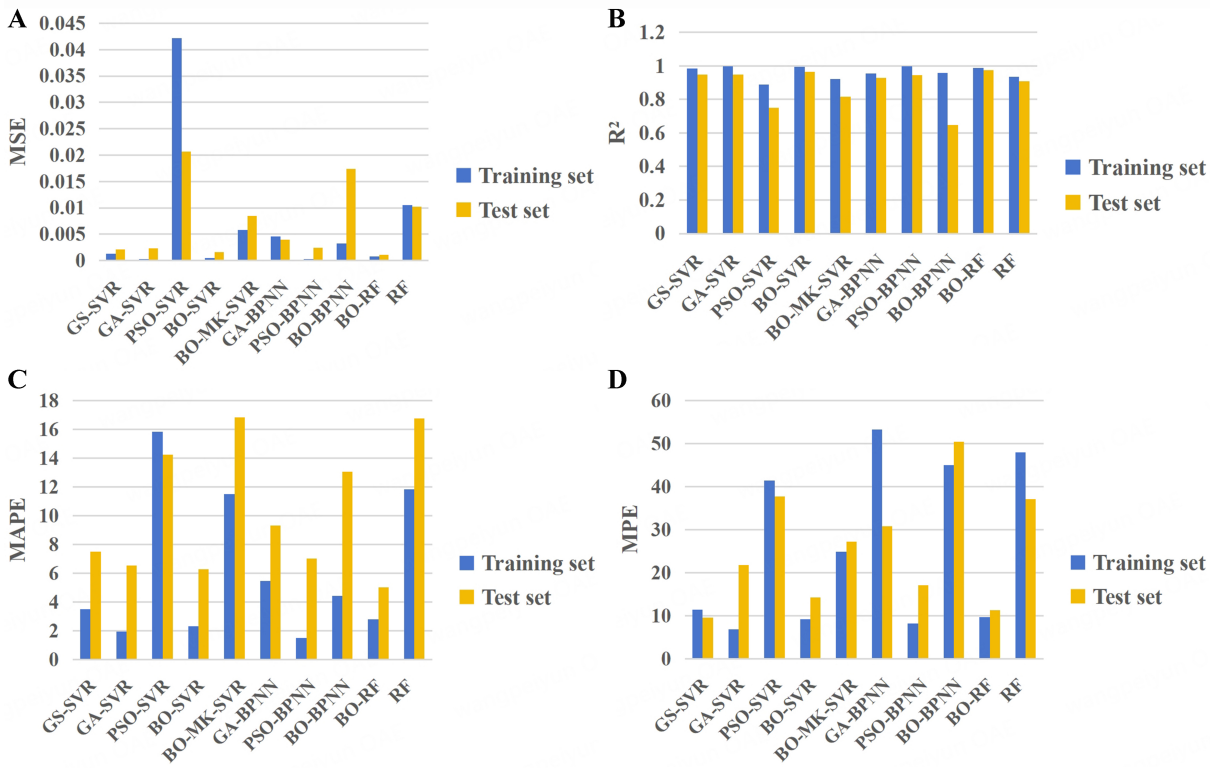

Figure 6. The results of model evaluation metrics for ten models. (A) The comparison of MSE among the prediction results of various models; (B) The comparison of R^2 among the prediction results of various models; (C) The comparison of MAPE among the prediction results of various models; (D) The comparison of MPE among the prediction results of various models. MSE: Mean square error; R^2: coefficient of determination; MAPE: mean absolute percentage error; MPE: maximum percentage error.

and BO-BPNN for Bayesian optimization-based BPNN.

According to the comparison results in Table 4, it is evident that the energy efficiency prediction model based on BO-RF exhibits the best predictive performance. The SVR-based energy efficiency prediction model also demonstrates good predictive performance, albeit slightly inferior to the BO-RF model. However, the predictive performance of the BPNN-based energy efficiency prediction model is relatively poor. Analysis suggests that the BPNN model requires a large amount of sample data, and its predictive performance improves with a larger sample size. When predicting with small sample data, the RF and SVR models perform better.

CONCLUSION AND FUTURE WORK

In this study, we addressed the prediction problem of the relationship between processing parameters and energy consumption in FDM. We proposed a FDM energy consumption prediction model based on BO-RF. Through experimental validation from the perspectives of MSE, MPE, MAPE, and R^2, we compared the prediction results of the BO-RF energy consumption prediction model with those of nine other models. The results demonstrate the superiority of the BO-RF-based FDM energy consumption prediction model.

The proposed approach in this study is data-driven, which may have limitations in model interpretability. In future work, we will integrate data-driven and mechanistic knowledge-driven models to enhance model interpretability and improve predictive performance.

DECLARATIONS

Authors' contributions

Conceptualization, methodology, data analysis and interpretation, manuscript perfection: Li Y
Supervision, review and editing, formal analysis: Bao H
Original draft preparation, methodology: Yang J
Data curation, investigation: Ke Q, Zhou Y

Availability of data and materials

Not applicable.

Financial support and sponsorship

This research is supported by the National Key Research and Development Program (2020YFB1711604) and the National Natural Science Foundation of China (51505119).

Conflicts of interest

All authors declared that there are no conflicts of interest.

Ethical approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

REFERENCES

1. Huang SH, Liu P, Mokasdar A, Hou L. Additive manufacturing and its societal impact: a literature review. *Int J Adv Manuf Technol* 2013;67:1191-203. DOI
2. Suárez L, Domínguez M. Sustainability and environmental impact of fused deposition modelling (FDM) technologies. *Int J Adv Manuf Technol* 2020;106:1267-79. DOI
3. Griffiths C, Howarth J, De Almeida-rowbotham G, Rees A, Kerton R. A design of experiments approach for the optimisation of energy and waste during the production of parts manufactured by 3D printing. *J Clean Prod* 2016;139:74-85. DOI
4. Enemuoh EU, Duginski S, Feyen C, Menta VG. Effect of process parameters on energy consumption, physical, and mechanical properties of fused deposition modeling. *Polymers* 2021;13:2406. DOI PubMed PMC
5. Galetto M, Verna E, Genta G. Effect of process parameters on parts quality and process efficiency of fused deposition modeling. *Comput Ind Eng* 2021;156:107238. DOI
6. Elkaseer A, Schneider S, Scholz SG. Experiment-based process modeling and optimization for high-quality and resource-efficient FFF 3D printing. *Appl Sci* 2020;10:2899. DOI
7. Camposeco-negrete C. Optimization of printing parameters in fused deposition modeling for improving part quality and process sustainability. *Int J Adv Manuf Technol* 2020;108:2131-47. DOI
8. Al-ghamdi KA. Sustainable FDM additive manufacturing of ABS components with emphasis on energy minimized and time efficient lightweight construction. *Int J Lightweight Mater Manuf* 2019;2:338-45. DOI
9. Hassan MR, Jeon HW, Kim G, Park K. The effects of infill patterns and infill percentages on energy consumption in fused filament fabrication using CFR-PEEK. *Rapid Prototyping J* 2021;27:1886-99. DOI
10. Feng M, Hua Z, Hon KKB. A qualitative model for predicting energy consumption of rapid prototyping processes - a case of fused deposition modeling processe. *IEEE Access* 2019;7:184825-31. DOI
11. Tian W, Ma J, Alizadeh M. Energy consumption optimization with geometric accuracy consideration for fused filament fabrication processes. *Int J Adv Manuf Technol* 2019;103:3223-33. DOI
12. Alizadeh M, Esfahani MN, Tian W, Ma J. Data-driven energy efficiency and part geometric accuracy modeling and optimization of green fused filament fabrication processes. *J Mech Des* 2020;142:041701. DOI
13. Lei Y, Yang B, Jiang X, Jia F, Li N, Nandi AK. Applications of machine learning to machine fault diagnosis: a review and roadmap. *Mech Syst Signal Pr* 2020;138:106587. DOI
14. Yang J, Li S, Wang Z, Yang G. Real-time tiny part defect detection system in manufacturing using deep learning. *IEEE Access* 2019;7:89278-91. DOI
15. Xiong Y, Guo L, Tian D, Zhang Y, Liu C. Intelligent optimization strategy based on statistical machine learning for spacecraft thermal design. *IEEE Access* 2020;8:204268-82. DOI
16. Xia C, Pan Z, Polden J, Li H, Xu Y, Chen S. Modelling and prediction of surface roughness in wire arc additive manufacturing using machine learning. *J Intell Manuf* 2022;33:1467-82. DOI
17. Li J, Cao L, Hu J, Sheng M, Zhou Q, Jin P. A prediction approach of SLM based on the ensemble of metamodels considering material efficiency, energy consumption, and tensile strength. *J Intell Manuf* 2022;33:687-702. DOI
18. Baturynska I, Martinsen K. Prediction of geometry deviations in additive manufactured parts: comparison of linear regression with machine learning algorithms. *J Intell Manuf* 2021;32:179-200. DOI
19. Cai R, Wang K, Wen W, Peng Y, Baniassadi M, Ahzi S. Application of machine learning methods on dynamic strength analysis for additive manufactured polypropylene-based composites. *Polym Test* 2022;110:107580. DOI
20. Gor M, Dobriyal A, Wankhede V, et al. Density prediction in powder bed fusion additive manufacturing: machine learning-based techniques. *Appl Sci* 2022;12:7271. DOI
21. Kumar P, Jain NK. Surface roughness prediction in micro-plasma transferred arc metal additive manufacturing process using K-nearest neighbors algorithm. *Int J Adv Manuf Technol* 2022;119:2985-97. DOI
22. Ranjan N, Kumar R, Kumar R, Kaur R, Singh S. Investigation of fused filament fabrication-based manufacturing of ABS-Al composite structures: prediction by machine learning and optimization. *J Mater Eng Perform* 2023;32:4555-74. DOI
23. Wu D, Wei Y, Terpenny J. Predictive modelling of surface roughness in fused deposition modelling using data fusion. *Int J Prod Res* 2019;57:3992-4006. DOI
24. Li Z, Zhang Z, Shi J, Wu D. Prediction of surface roughness in extrusion-based additive manufacturing with machine learning. *Robot Cim Int Manuf* 2019;57:488-95. DOI
25. Zhang L, Zhong YJ, Kan HY, Zhang BK. Optimization method for high energy efficiency process parameters in fused deposition manufacturing. *J Mech Des Manuf* 2021;3:149-52+6. (in Chinese). DOI
26. Zhang L, Zhang BK, Bao H, Zhang C, Zhang WW. Carbon emissions quantitative methodology of product fused deposition manufacturing. *J Mech Eng* 2017;5:50-9. (in Chinese). Available from: https://d.wanfangdata.com.cn/periodical/jxgcxb201705006. [Last accessed on 27 Aug 2024]
27. Li CB, Zhu YT, Li L, Chen XZ. Multi-objective CNC milling parameters optimization model for energy efficiency. *J Mech Eng* 2016;21:120-9. (in Chinese). Available from: https://d.wanfangdata.com.cn/periodical/jxgcxb201621015. [Last accessed on 27 Aug 2024]
28. He H, Garcia EA. Learning from imbalanced data. *IEEE Trans Knowl Data Eng* 2009;21:1263-84. DOI