

PAPER • OPEN ACCESS

Comparative assessment of supervised machine learning algorithms for predicting geometric characteristics of laser cladded inconel 718

To cite this article: Hao Yang *et al* 2024 *Mater. Res. Express* 11 046516View the [article online](#) for updates and enhancements.

You may also like

- [Out-of-equilibrium gene expression fluctuations in the presence of extrinsic noise](#)
Marta Biondo, Abhyudai Singh, Michele Caselle *et al.*
- [Microstructure and Mechanical Properties of Cermet Composite Coating on TC4 Surface](#)
Lijuan Zheng, Yinkai Xie, Kuo Zhang *et al.*
- [Study on microstructure and mechanical properties of Ni60 + WC/Ni35/AISI1040 functional surface gradient structure of remanufacturing chute plate for the mining scraper by a low cost high power CO₂ laser cladding technique](#)
J Luo, J J Gao, S W Gou *et al.*

EDINBURGH
INSTRUMENTS

SCAN ME

FS5 SPECTROFLUOROMETER WITH MICRO PL UPGRADE

- + High Sensitivity (Single Photon Counting)
- + 200 nm - 1700 nm Spectral Range
- + Fluorescence Lifetimes (TCSPC) from < 25 ps
- + Phosphorescence Lifetime (MCS) 10 ns to seconds



VISIT OUR WEBSITE FOR MORE DETAILS

[edinst.com](https://www.edinst.com)



PAPER

Comparative assessment of supervised machine learning algorithms for predicting geometric characteristics of laser clad Inconel 718

OPEN ACCESS

RECEIVED

24 November 2023

REVISED

8 April 2024

ACCEPTED FOR PUBLICATION

17 April 2024

PUBLISHED

29 April 2024

Hao Yang¹, Heran Geng¹, Marco Alfano^{2,3}  and Junfeng Yuan^{1,2} ¹ School of Mechatronic Engineering, China University of Mining and Technology, Xuzhou 221116, People's Republic of China² Department of Mechanical and Mechatronics Engineering, University of Waterloo, 200 University Avenue West, Waterloo, ON N2L 3G1, Canada³ Dipartimento di Scienze e Metodi dell'Ingegneria, Università di Modena e Reggio Emilia, Via Amendola 2, Padiglione Morselli, 42122 Reggio Emilia, ItalyE-mail: yuanjf@cumt.edu.cn and yuanjfacademia@outlook.com

Original content from this work may be used under the terms of the [Creative Commons Attribution 4.0 licence](https://creativecommons.org/licenses/by/4.0/).

Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.

**Keywords:** laser cladding, IN718, machine learning, performance prediction, processing variables**Abstract**

Laser cladding, an innovative surface modification and coating preparation process, has emerged as a research hotspot in material surface modification and green remanufacturing domains. In the laser cladding process, the interaction between laser light, powder particles, and the substrate results in a complicated mapping connection between process parameters and clad layer quality. This work aims to shed light on this mapping using fast evolving machine learning algorithms. A full factorial experimental design was employed to clad Inconel 718 powder on an A286 substrate comprising 64 groups. Analysis of variance, contour plots, and surface plots were used to explore the effects of laser power, powder feeding rate, and scanning speed on the width, height, and dilution rate of the cladding. The performance of the predictive models was evaluated using the index of merit (IM), which includes mean square error (MSE), mean absolute error (MAE), and coefficient of determination (R^2). By comparing the performance of the models, it was found that the Extra Trees, Random forest regression, Decision tree regression, and XGBoost algorithms exhibited the highest predictive accuracy. Specifically, the Extra Trees algorithm outperformed other machine learning models in predicting the cladding width, while the RFR algorithm excelled in predicting the associated height. The DTR algorithm demonstrated the best performance in predicting the cladding dilution rate. The R^2 values for width, height, and dilution rate were found to be 0.949, 0.954, and 0.912, respectively, for these three models.

1. Introduction

Laser cladding is an advanced surface modification and coating preparation technology, integrating material preparation and surface configuration, and is a crucial support technology for green remanufacturing technology. Laser cladding exploits the high energy density of a laser beam to melt and spray metal powder (or wire) onto the surface of the target substrate, forming coatings with high hardness, high abrasion resistance, high corrosion resistance, etc. The cladding process enables surface modification of critical parts as well as the repair of surface damage, and it has been widely used in the fields of aviation, aerospace, automotive, machinery and so on [1, 2].

The process parameters of the laser cladding process have a profound effect on the quality characteristics of the resulting clad layer (i.e., the geometrical and mechanical properties). The coupling of light, powder, and substrate properties leads to a complex mapping relationship between process parameters and cladding layer [3, 4]. In order to optimize laser cladding, many scholars are exploring the influence of the different process parameters on the coating morphology and mechanical [5, 6]. In the study by Liu *et al* 27 sets of experiments were selected using a full factorial design (FFD) to investigate the interplay between laser power (P), powder feed rate (F), and scanning speed (S) on the cross-sectional dimensions (i.e., width and height) of a clad deposit

consisting of ferrous self-melting alloy powder. A nonlinear fitting model was used to fit the relationship between geometrical parameters and process parameters [7]. Li *et al* used a single-track orthogonal experiment to explore the influence of process parameters (substrate tilt angle, P , F , and S) on the cladding layer width and height during the laser cladding deposition of Ni60A powder under the tilted substrate. The results showed that an increase in tilt angle would increase the cladding layer width and decrease the height, with the width being more significantly affected by laser power, and the height being more influenced by S and F [8]. Xu *et al* used Taguchi's method to design L_{16} orthogonal experiments to melt cladding of In718 on a A286 substrate [9]. They analyzed the effects of P , F , and S , on micro-hardness, load bearing capacity, yield strength, ultimate tensile strength, and elongation and failure. In addition, they have shown that using optimized process parameters, mechanical properties could be further improved. Despite the differences in the selection of process parameters by researchers in many scientific papers, it is still evident that P , F , and S generally have an important influence in different experimental studies.

The optimization of laser cladding process parameters is often divided into three stages: experimental design, predictive modeling, and parameter optimization. The research methodology adopted in each stage affects the final results. Currently, robust experimental designs for laser cladding often employ factorial experiments, Taguchi analysis methods, and response surface methodology [5–7].

To realize the complex mapping between laser cladding process parameters and cladding layer quality, three methods are commonly used: (i) statistical analysis method, to establish the regression model between the process parameters and the response [10–12]; (ii) using finite element analysis methods, the established three-dimensional model controls each parameter variable, simulates the experimental process of laser cladding, and predicts the desired experimental results [13–16]; (iii) application of machine learning (ML) algorithms such as Random forest regression (RFR), Support Vector Machine (SVM), Artificial Neural Network (ANN), and Deep Learning [16–21].

Alizadeh-Sh *et al* melted In718 alloy powder on the surface of A286 iron-based superalloy. They proposed an empirical statistical analysis based on the linear regression (LR) method to analyze the melting and cladding process. The critical geometrical characteristics (i.e., width, height, angle, dilution rate) required to avoid solidification cracking during the melting process were predicted [22]. Lian *et al* carried out laser cladding on curved surfaces and used the response surface method to establish mathematical models for width, height, and dilution rate. The authors obtained the relationship between the response and the process parameters, and experimentally verified the model's reliability. However, it is difficult to solely use regression equations to describe this correlation when dealing with nonlinear data and data characteristics involving complex polynomials [23]. Song *et al* established a three-dimensional finite element model of the laser cladding process of 7075 aluminum alloy powder on 2024 aluminum alloy substrate, and obtained the temperature field and residual stress field generated during the process, in order to analyze the effects of different laser power, scanning speed, cladding layer lengths, cladding layer patterns, and cladding angles on fatigue life of aircraft Fuselage [24]. Wolff *et al* predicted the effects of process parameters on temperature distribution, liquid metal flow, cladding layer geometry, and dilution rate in the melt pool during the laser cladding process by establishing a three-dimensional thermo fluid dynamics model and using a surface contour calculation method based on minimizing surface free energy [25]. Many assumptions need to be made, which often do not align well with the actual cladding process. This misalignment results in simulation outcomes that struggle to offer meaningful guidance for the actual process. Additionally, the finite element method often consumes a significant amount of time during the solving process.

To circumvent the above issues, many scholars have recently turned their attention to rapidly developing research methods such as ML to achieve better process predictions. Omar *et al* studied the applicability of common ML algorithms such as Gaussian Process Regression, Decision Tree Regression (DTR), Random Forest Regression (RFR), Support Vector Regression (SVR), Gradient Boosting Regressor, and Multi-layer Perceptron (MLP) in predicting friction welding process parameters. The results showed that Gradient Boosting Regressor (GBR), Support Vector Regressor (SVR), and Gaussian Process Regressor had the highest accuracy with a percentage error of less than 3% [17]. According to literature research, commonly used ML algorithms in establishing prediction models for laser cladding include Back-propagation Neural Network (BP), SVR, DTR, RFR, GBR, etc. [3, 20, 26–31].

Although numerous scholars have conducted detailed studies on predicting laser cladding process parameters, fewer have undertaken in-depth comparative analyses of these ML algorithms to determine their accuracy and applicability. Therefore, the primary focus of this study is to compare the predictive accuracy of commonly used ML algorithms in determining the geometric characteristics of laser-clad layers of Inconel 718. Specifically, our study centers on K-Nearest Neighbors (KNN), Back-propagation Neural Network (BP), Support Vector Regression (SVR), Decision Tree Regression (DTR), Random Forest Regression (RFR), Extra Trees, and XGBoost. Our objective is to establish which ML algorithms exhibit higher precision in predicting process parameters for laser cladding. We conducted 64 FFD experiments on Inconel 718 clad on an A286

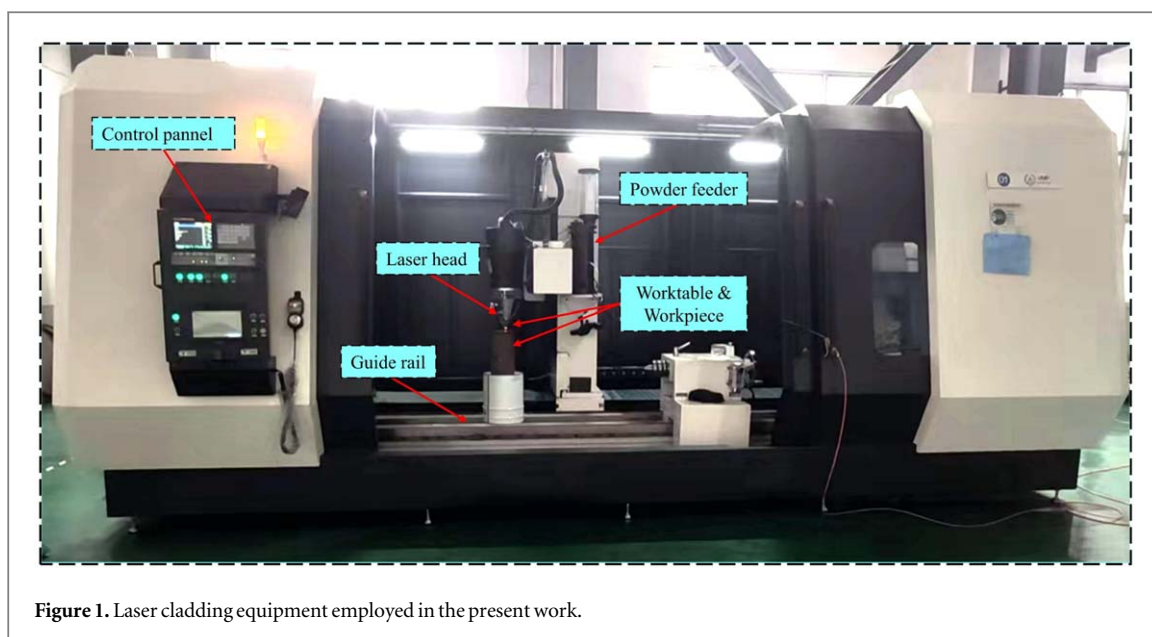


Figure 1. Laser cladding equipment employed in the present work.

substrate. First, we applied an analysis of variance to assess the contribution of each parameter to various responses and the statistical significance of the cladding process. Secondly, traditional analytical methods investigated the relationship between the three inputs—laser power (P), powder feed rate (F), and scanning speed (S)—and the responses—cladding width (W), cladding height (H), and dilution rate (D). Finally, we optimized the hyperparameters in the machine learning models using genetic algorithms (GA). Subsequently, we assessed the performance of these optimized ML algorithms in predicting the quality indices (IMs) of the laser cladding process parameters, and comparative evaluations were performed.

2. Materials and methods

2.1. Laser cladding

Figure 1 shows the ZKZM-Z06 laser cladding system employed in this work. The laser cladding system, which is installed in a 4-axis CNC machine, consists of 6KW fiber laser, a coaxial laser cladding head, and powder feeding and cooling systems. Disc-shaped samples ($\phi 150 \text{ mm} \times 10 \text{ mm}$) consisting of iron-based high-temperature alloy (A286) were used as the substrate material. Nickel-based high-temperature alloy In718 powder was selected as the cladding material. To ensure good flowability, In718 powder with spherical-shaped particles in the diameter of $45\text{--}150 \mu\text{m}$ was chosen, and its micrograph, obtained through a Scanning Electron Microscope (SEM, FEI Quanta 250), is shown in figure 2.

The chemical composition of the substrate was determined using carbon-sulfur analysis and atomic spectroscopy, as shown in table 1. Before the experiments, the powders were dried in a vacuum using an oven set at $100 \text{ }^\circ\text{C}$ for 2 h. Prior to laser cladding, the A286 substrates were ground using 180# SiC sandpaper to remove oxides and contaminants from the substrate surface and wiped clean using absolute ethanol (anhydrous ethanol). Single-track cladded layers were deposited on the pre-treated upper surface of the substrates, with representative dimensions of $5 \text{ mm} \times 8 \text{ mm} \times 10 \text{ mm}$. Subsequently, the samples were sequentially grounded using SiC sandpaper with grit size 400#, 800#, 1200#, 1500# and 2000# and then polished with W2.5 abrasive paste. The samples cross-section was subsequently etched using Kalling's reagent (i.e., a mixture of 100 ml water, 100 ml HCl, and 5 g CuCl_2). Finally, the etched cross-section was observed using an optical microscope (OM, LEICA DM4), and the geometric characteristics of the cladding layer were measured and evaluated using an image processing program (ImageJ, National Institute of Health, US).

2.2. Design of experiments

In this study, a full factorial experimental (FFD) approach was employed to design the experiments, considering the varying degrees of influence from different process parameters. The laser spot diameter was set constant at 5 mm, and argon gas was used as the shielding and carrier gas for powder delivery. The gas flow rate was maintained at 15 L/min. Based on literature research, three main process parameters, namely P , F , and S , were selected as independent variables. Each parameter was set at 4 levels, as shown in table 2. The 64 sets of parameter combinations obtained from the full factorial experimental design are shown in table 3.

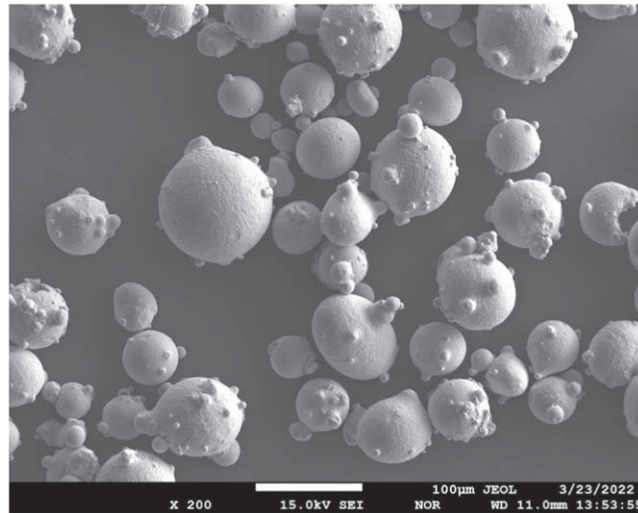


Figure 2. SEM micrograph of the In718 powder.

Table 1. Composition of A286 and In718.

Material	Nominal composition (wt%)								
A286	Ni	Cr	Ti	Mo	Mn	Si	Al	C	Fe
	24.64	14.81	2.05	1.26	1.2	0.6	0.24	0.06	Bal
In718	Ni	Cr	Nb	Mo	Ti	Al	Cu	C	Fe
	52.41	19.17	5.19	3.19	0.95	0.51	0.087	0.04	Bal

Table 2. Key parameters of laser cladding process in different levels.

Process parameters	Unit	Notations	Level 1	Level 2	Level 3	level 4
Laser power	W	P	800	1300	1800	2300
Powder feed speed	r/min	F	0.4	0.5	0.6	0.7
Scan speed	mm/s	S	3	4.5	6	7.5

Figure 3 presents the measured geometric characteristics of the cross-section of the single-track claddings. The geometric characteristics, including W, H, melt pool depth (h), cladding area (A_c), and fusion area (A_f), were measured using ImageJ software.

The dilution rate (D) is an important parameter which indicates the degree of bonding between the fusion cladding and the substrate. In the present work, it was calculated using equation (1) [32]:

$$D = \frac{A_f}{A_c + A_f} \quad (1)$$

where A_f is the area of the fusion zone, and A_c is the area of the cladding layer.

2.3. Machine learning methods and selection of hyperparameters

Drawing upon relevant literature and the authors' own investigations into mainstream machine learning techniques, this paper employs a selection of algorithms, including K-Nearest Neighbors (KNN), Back-propagation Neural Network (BP), Support Vector Regression (SVR), Decision Tree Regression (DTR), Random Forest Regression (RFR), Extra Trees, and XGBoost [3, 20, 26–31]. The objective is to compare their predictive performance in the context of this study.

To avoid issues such as overfitting and underfitting, this study employs 5-fold cross-validation. In this validation technique, k denotes the number of parts into which the data is divided; K-1 folds are used for training, and the remaining folds are used for testing the model. The evaluation metrics of the cross-validation set can be used to continuously adjust the hyper-parameters to obtain a reliable and stable model. Additionally, different ML algorithms have their characteristics, and the hyperparameters that significantly impact the prediction results may vary depending on the specific problem. Therefore, this study selects important

Table 3. Full factorial experimental data set.

No.	P (kW)	F (r/min)	S (mm/s)	W (μm)	H (μm)	D (%)	No.	P (kW)	F (r/min)	S (mm/s)	W (μm)	H (μm)	D (%)
1	0.8	0.4	3	1979	1529	0.73	33	1.8	0.4	3	4438	2388	10.28
2	0.8	0.4	4.5	1818	887	4.72	34	1.8	0.4	4.5	3897	1497	17.67
3	0.8	0.4	6	1568	638	2.23	35	1.8	0.4	6	3321	1158	18.9
4	0.8	0.4	7.5	1529	269	19.28	36	1.8	0.4	7.5	3422	986	11.84
5	0.8	0.5	3	2442	434	27.74	37	1.8	0.5	3	4242	2382	11.39
6	0.8	0.5	4.5	1489	1168	0	38	1.8	0.5	4.5	4052	1794	5.27
7	0.8	0.5	6	1557	832	0	39	1.8	0.5	6	3606	1378	9.04
8	0.8	0.5	7.5	1194	630	0	40	1.8	0.5	7.5	3279	1147	7.69
9	0.8	0.6	3	2424	2216	0	41	1.8	0.6	3	4687	3196	0.71
10	0.8	0.6	4.5	2263	1640	0	42	1.8	0.6	4.5	4349	2364	0.85
11	0.8	0.6	6	1636	998	0	43	1.8	0.6	6	3945	1687	1.01
12	0.8	0.6	7.5	—	—	0	44	1.8	0.6	7.5	3790	1556	1.04
13	0.8	0.7	3	—	—	0	45	1.8	0.7	3	4657	3493	0
14	0.8	0.7	4.5	—	—	0	46	1.8	0.7	4.5	4402	2780	0
15	0.8	0.7	6	—	—	0	47	1.8	0.7	6	3701	2020	0
16	0.8	0.7	7.5	—	—	0	48	1.8	0.7	7.5	3802	1325	0
17	1.3	0.4	3	3719	2156	5.16	49	2.3	0.4	3	5070	2293	26.58
18	1.3	0.4	4.5	3137	1212	15.91	50	2.3	0.4	4.5	4522	1574	40.39
19	1.3	0.4	6	2673	986	8.25	51	2.3	0.4	6	4213	1176	44.86
20	1.3	0.4	7.5	2465	790	6.74	52	2.3	0.4	7.5	3677	951	52.01
21	1.3	0.5	3	3719	2531	1.59	53	2.3	0.5	3	4764	2560	25.12
22	1.3	0.5	4.5	3107	1580	1.59	54	2.3	0.5	4.5	4337	1859	29.82
23	1.3	0.5	6	3351	1349	1.22	55	2.3	0.5	6	4088	1141	40.41
24	1.3	0.5	7.5	2816	939	3.85	56	2.3	0.5	7.5	3867	1028	37.76
25	1.3	0.6	3	3743	2941	0	57	2.3	0.6	3	4690	3217	12.2
26	1.3	0.6	4.5	3089	1990	0	58	2.3	0.6	4.5	4479	2186	20.23
27	1.3	0.6	6	2970	1366	0	59	2.3	0.6	6	4319	1764	19.93
28	1.3	0.6	7.5	2935	1194	0	60	2.3	0.6	7.5	4123	1521	16.57
29	1.3	0.7	3	4010	3244	0	61	2.3	0.7	3	4990	3541	5.51
30	1.3	0.7	4.5	3392	2180	0	62	2.3	0.7	4.5	4592	2566	5.4
31	1.3	0.7	6	2881	1675	0	63	2.3	0.7	6	4337	2151	3.78
32	1.3	0.7	7.5	2691	1319	0	64	2.3	0.7	7.5	4048	1751	5.15

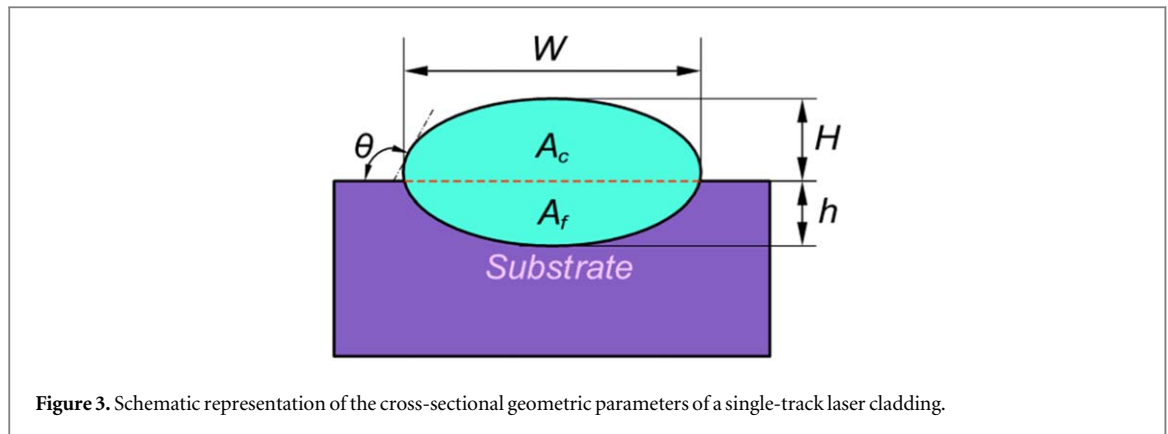


Figure 3. Schematic representation of the cross-sectional geometric parameters of a single-track laser cladding.

Table 4. Hyperparameter selection for algorithm implementation.

ML algorithm	Hyperparameters
KNN	Neighbor tree = 5, Vector distance algorithm = Euclidean distance
BP	Hidden layers = 3, Hide layer node tree = 100, Learning rate = 0.25, Activation function = Identity
SVR	Kernel function = Radial basis kernel function, Kernel function coefficients = Scale
DTR	Maximum depth = 30,
RFR	Number of trees = 100, maximum depth = 20
Extra Trees	Number of trees = 100, maximum depth = 20
XGBoost	Base learner = gblinear, Number of base learners = 100, Maximum depth = 10

hyperparameters based on previous research to ensure that the ML algorithms can achieve good predictive performance and generalization ability. For ML algorithm in table 4, heuristic algorithms are used to optimize the hyperparameters. The schematic is shown in figure 4.

2.3.1. K-nearest neighbor

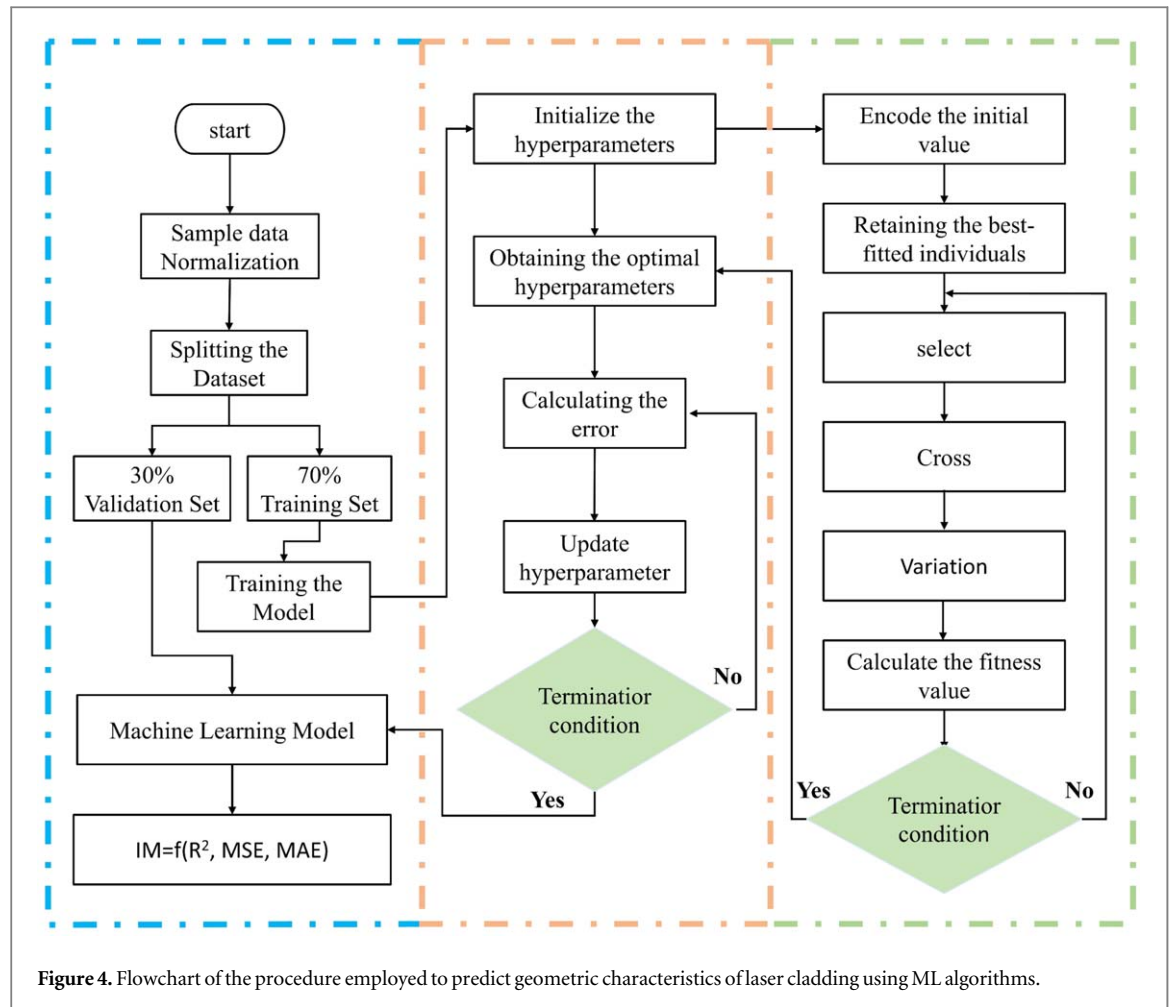
The K-nearest neighbor (KNN) regression algorithm operates on the principle of distance similarity. It selects the K closest neighbors to the sample to be predicted by computing the distance between the target sample and each sample in the training set. Prediction is then made based on the labels of these neighbors. The algorithm's workflow involves storing samples and their labels during the training phase, creating a sample space. In the prediction phase, the distance between a test sample and each training set sample is calculated, and the K nearest neighbors are chosen based on these distances. For regression problems, the average value of the K neighbors is used as the prediction result. By selecting an appropriate K value and distance metric, the KNN algorithm can accurately predict regression problems. Using Euclidean distance as the vector distance algorithm, and after optimizing the K value of KNN through the genetic algorithm (GA), the value of 5 was selected [26].

2.3.2. BP neural network

BPNN (BP) is a neural network regression algorithm based on the back-propagation algorithm, which is based on the principle of training the network weights through two phases: forward propagation and back-propagation, in order to enable the network to learn the mapping relationship between inputs and outputs. The loss function is minimized by continuously adjusting the weights and biases and is suitable for dealing with nonlinear problems. The complexity and learning capability of the network are determined by parameters such as the number of hidden layers and the number of neurons in each hidden layer. The learning rate determines the step size of each weight update. A learning rate that is too large may cause oscillations, while a learning rate that is too small may result in slow convergence. The number of iterations determines the number of training cycles. Too many iterations may lead to overfitting, while too few may result in underfitting [4, 5, 20, 27].

2.3.3. Support vector regression

Support Vector Regression (SVR) is a regression algorithm based on Support Vector Machines that fits the data by finding the optimal hyperplane and minimizing the prediction error as much as possible. It is suitable for dealing with high-dimensional data and linear problems. One of its most important hyperparameters is the kernel function. Through the comparison of different kernel functions, it was found that the model based on the radial basis is more suitable for predicting the geometric characteristics of the cladding [3].



2.3.4. Decision tree regression

Decision tree regression (DTR) is a regression algorithm based on a tree structure based on the principle of modeling the input data in segments by constructing a tree to predict continuous values of the output variables. In DTR, each tree node represents a feature, and the input data is partitioned based on that feature until a predetermined stopping condition is met. At each node, DTR uses some criterion to select the best feature and segmentation point, commonly used criteria include mean square error (MSE) and mean absolute error (MAE). During the prediction phase, the input data follows the branches of the tree and eventually reaches a leaf node, which stores the average or median value of the samples in that subset. This value serves as the prediction for that subset [17, 19]. There are important parameters that need to be set in DTR. These include the maximum depth of the tree, the minimum number of samples in a leaf node, and the minimum number of samples required for a split. These parameters determine the complexity and generalization ability of the tree. A larger maximum depth and minimum number of samples may lead to overfitting, while a smaller minimum number of samples may lead to underfitting. By adjusting the maximum depth through GA to optimize the Decision Tree Regression (DTR) and improve the model's performance and generalization ability, more accurate prediction results can be obtained.

2.3.5. Random forest regression

The Random Forest regression (RFR) algorithm is an ensemble learning method that constructs multiple decision trees by randomly selecting data and features. The final result is obtained by voting or averaging the predictions of these trees. It has high accuracy and generalization ability [17, 19]. In the Random Forest model, important hyperparameters include the number of decision trees and the maximum depth [27].

2.3.6. Extra-trees

Extra trees, also known as Extremely Randomized Trees, further increase the randomness on top of the Random Forest. Extra trees randomly select features and split points when constructing each decision tree instead of optimizing criteria. This additional randomness can enhance the diversity of the model and reduce the risk of

overfitting, making it suitable for handling high-dimensional data and nonlinear problems. The important hyperparameters of Extra-trees are the same as those in the Random Forest algorithm.

2.3.7. XGBoost

XGBoost trains each weak classifier by optimizing the gradient of the loss function and obtains the final prediction by weighted summation. The gradient boosting algorithm adjusts the sample and classifier weights in each iteration to minimize the loss function. The accuracy and efficiency of the model is improved by using regularization and parallel processing, which is suitable for dealing with high-dimensional data and nonlinear problems [17, 20]. After optimization through GA, the learning rate is set to 0.1.

Random Forest and Extra Trees belong to the Bagging model, while XGBoost belongs to the Boosting model. The Bagging model is a parallel ensemble learning method that constructs multiple base classifiers by randomly sampling the training set with replacement. The final prediction is made through voting or averaging. The Boosting model is a sequential ensemble learning method that iteratively trains multiple weak classifiers. In each iteration, the sample weights are adjusted based on the predictions from the previous round to improve the model's performance.

2.4. Index of merit for the evaluation of the precision of ML algorithms

Using the 59 samples obtained from the experiments, the dataset was randomly divided into a training set (70%) and a validation set (30%). To eliminate the dimensional impact among response values, a min-max scaling preprocessing was performed. In particular, the data was scaled linearly between [0,1] using the following equation (2):

$$X' = \frac{X - X_{\min}}{X_{\max} - X_{\min}} \quad (2)$$

To mitigate the randomness of the ML algorithm, five consecutive calculations were performed using the hyperparameters shown in table 4. The best-performing result in predictive performance was selected as the accuracy evaluation metric for the chosen algorithm. Notice that, in this study, we adopted an index of merit (IM) introduced by Barrionuevo *et al* [17, 19] to assess the predictive accuracy, and which is defined as:

$$IM = \sqrt{(1 - R^2)^2 + MSE + (MAE)^2} \quad (3)$$

Where:

$$R^2 = 1 - \frac{\sum_i^N (y_i - \hat{y}_i)^2}{\sum_i^N (y_i - \bar{y})^2} \quad (4)$$

$$MSE = \frac{1}{N} \sum_i^N (y_i - \hat{y}_i)^2 \quad (5)$$

$$MAE = \frac{1}{N} \sum_i^N |y_i - \hat{y}_i| \quad (6)$$

This index combines the above metrics of to provide a comprehensive measure of algorithm accuracy. Notice that a value closer to 0 indicates better overall performance.

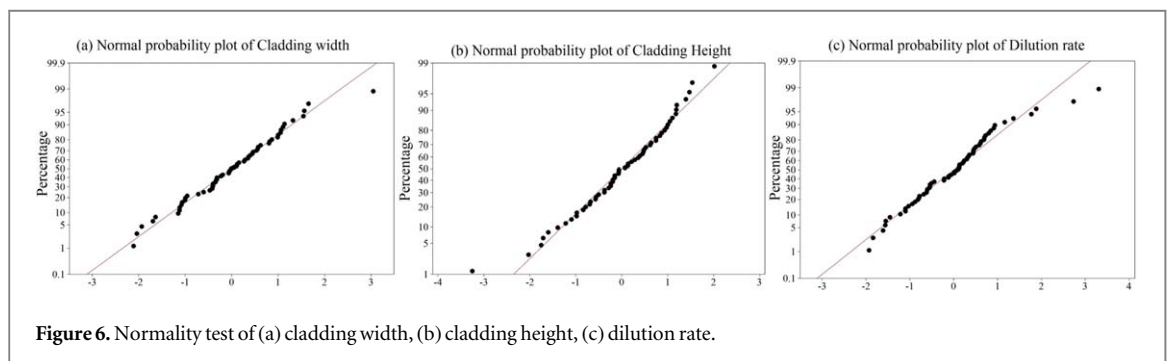
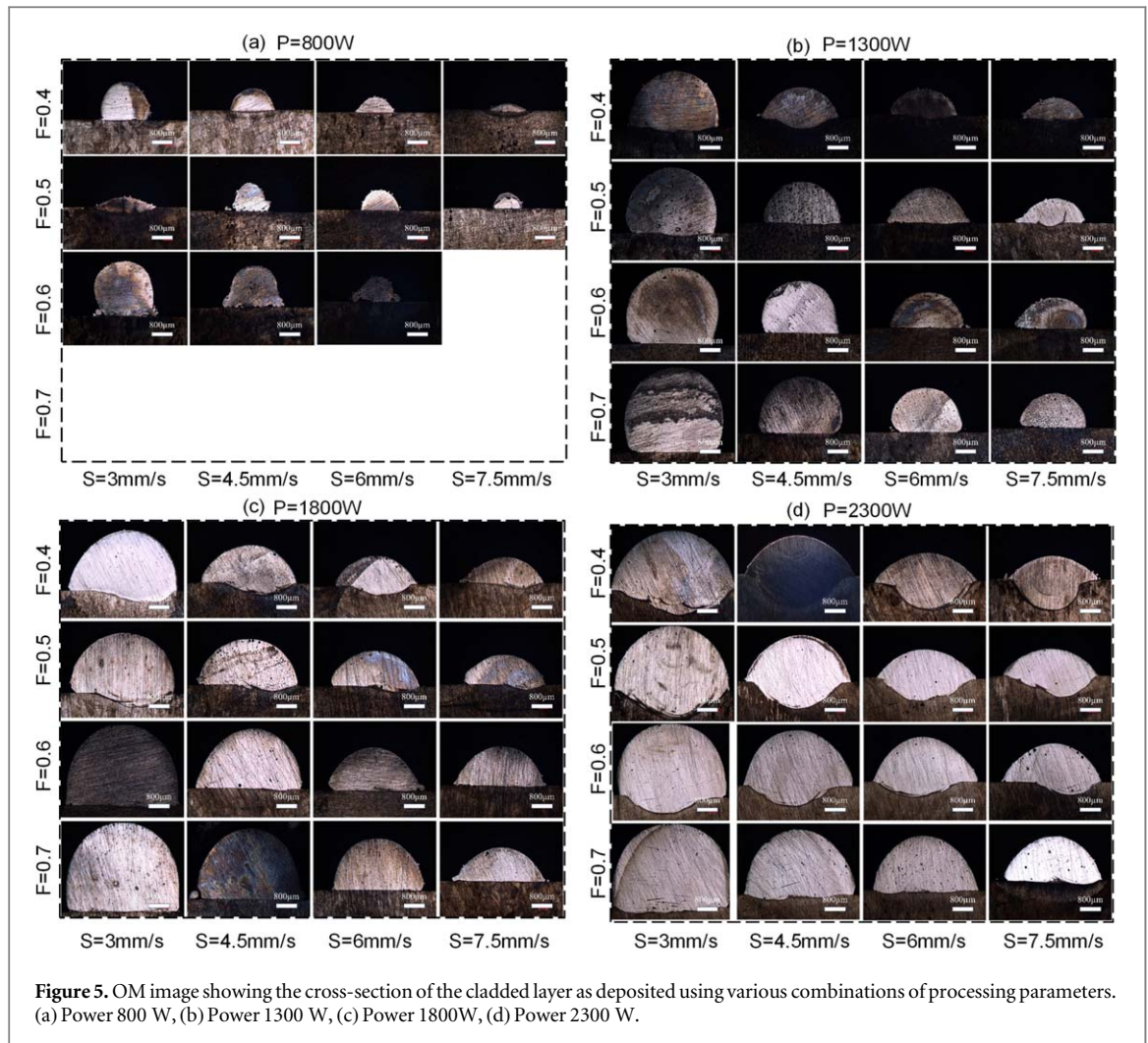
3. Results and discussion

Five samples were dislodged after the experiment and 59 samples were obtained. The single track cross-section maps obtained from optical microscopy (OM) are shown in figure 5, representing the claddings prepared at 800 W, 1300 W, 1800 W, and 2300 W laser power (P), respectively.

3.1. Statistical analysis

Analysis of variance (ANOVA) can be used to investigate whether the process parameters significantly affect the results of the laser cladding process. ANOVA is usually based on the assumption of normal distribution, so the results of the obtained responses need to be tested for normality. Figure 6 shows the residual normal probability plots for the measured W, H, and D. It can be observed that the data points are distributed along a straight line, indicating that W, H, and D all follow a normal distribution.

The importance of p-values lies in their ability to help us perform hypothesis testing, i.e., to determine whether the sample data supports the null hypothesis. Suppose the p-value is smaller than the predetermined significance level (typically 0.05 or 0.01). In that case, we can reject the null hypothesis and conclude that there is a significant difference between the sample data and the hypothesized population mean.



The ANOVA results are shown in table 5, except for F and S , which have insignificant effects on W and D , respectively, the p -values of P , F and S on W , H , and D are all less than 0.01, indicating that all have extremely significant effects. Among them, the contribution rates of P , F , and S to W are 90.39%, 1.16%, and 8.45%, respectively, indicating that P has the greatest influence on W . For H , the contribution rates of P , F , and S are 43.85%, 9.30%, and 46.85%, respectively, indicating that P and F have the greatest impact on H . For D , the contribution rates of P and S are 63.80%, 0.48%, and 35.58%, respectively, indicating that P and S significantly influence the dilution rate.

3.2. Contour plot and surface plot analysis

To visually examine the relationship between the process parameters and the geometrical characteristics of the cladding, surface and contour plots of the response results versus the cladding parameters are made from the 59 sets of experimental data obtained. The concept of powder distribution density (I) is introduced to synthesize the powder volume measure [33]:

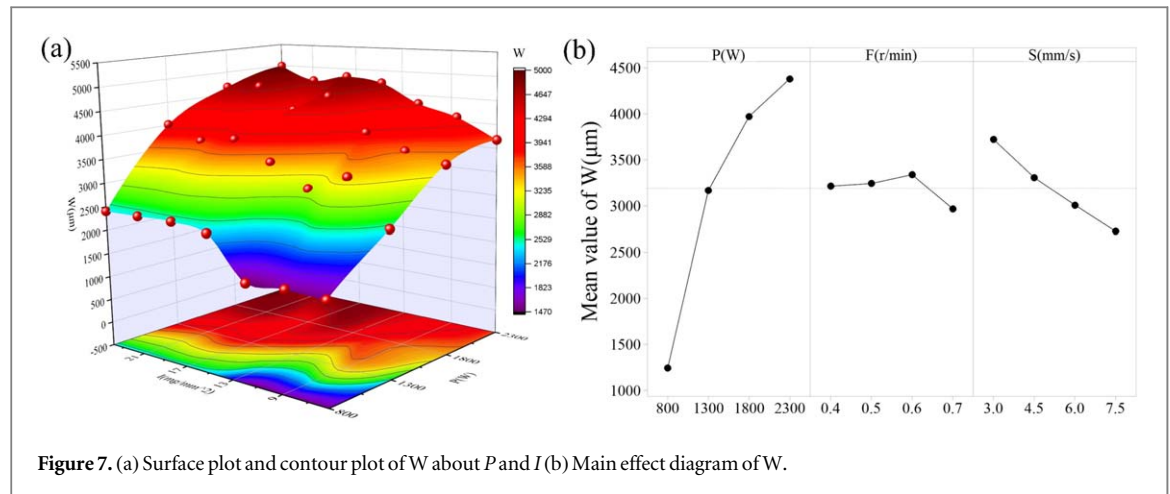


Table 5. ANOVA for cladding width, cladding height, dilution rate.

Source		Laser power(P)	Feeding rate(F)	Scanning speed(S)	Residual Error	Total
DF		3	3	3	54	63
Cladding width (W)	Adj SS	93201486	1199200	8713793	11373119	114487598
	Adj MS	31067162	399733	2904598	210613	
	F value	147.51	1.9	13.79		
	P value	0	0.141	0		
	Contribution	90.39%	1.16%	8.45%		
Cladding Height (H)	Adj SS	17019726	3607601	18188283	10246486	49062097
	Adj MS	5673242	1202534	6062761	189750	
	F value	29.9	6.34	31.95		
	P value	0	0.001	0		
	Contribution	43.85%	9.30%	46.85%		
Dilution rate (D)	Adj SS	0.49161	0.27485	0.00401	0.27956	1.05002
	Adj MS	0.163869	0.091616	0.001337	0.005177	
	F value	31.65	17.7	0.26		
	P value	0	0	0.855		
	Contribution	63.80%	35.68%	0.52%		

$$I = \frac{F}{S * d} \quad (7)$$

where d is the laser spot diameter. Surface plots are drawn with I and P as independent variables and each response as dependent variable, which enables the surface and contour plots of the same response to be displayed in a single plot.

3.2.1. Cladding width

As illustrated in figure 7(a), the contour plot shows the highest gradient variation along the direction of increasing power, while the gradient variation along the direction of I is relatively smaller. This suggests that P has a more significant impact on W compared to I . Additionally, as shown in figure 7(b), a positive correlation between P and W is observed. Initially, increasing F enhances the influence on W , which then decreases. Conversely, increasing S results in a decrease in W . This ultimately leads to a non-linear relationship between I and W . It follows that when I is constant, an increase in P within the melting power range can melt more powder per unit time, thereby improving powder utilization and increasing W . When P is constant and maintained within the melting power range, increasing I leads to more powder per unit time, resulting in a larger W . However, if I becomes too large, it can cause shielding effects, leading to a decrease in W . This explanation effectively accounts for the overall increasing trend of W .

3.2.2. Cladding height

Similarly, as shown in figure 8(a), the contour plot exhibits the highest gradient variation along the direction of I , while the gradient variation along the direction of P is relatively smaller. This indicates that H is more influenced by I than by P . Further observation of figure 8 reveals that increasing P leads to a gradual decrease in the slope of

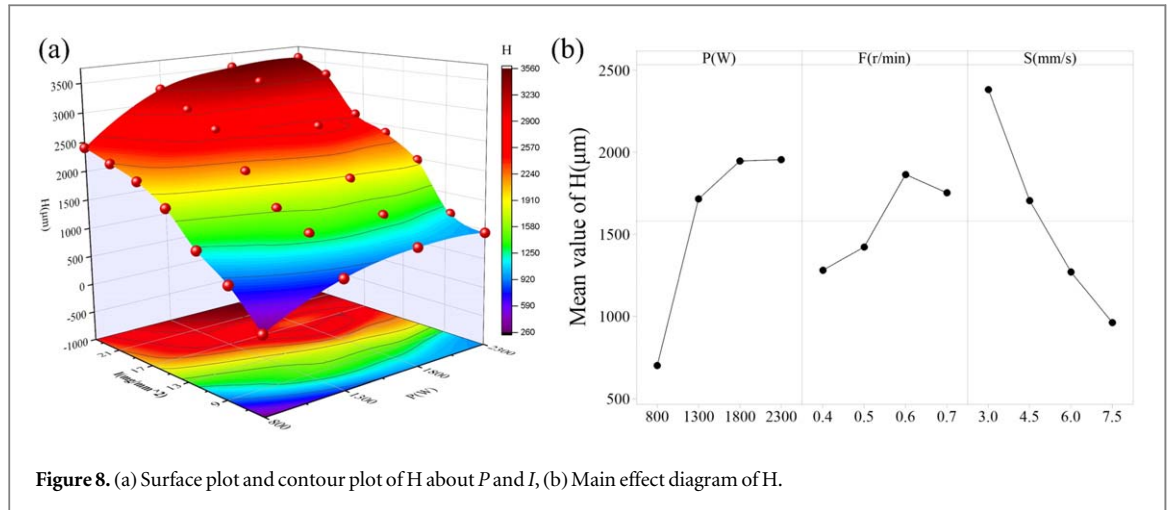


Figure 8. (a) Surface plot and contour plot of H about P and I, (b) Main effect diagram of H.

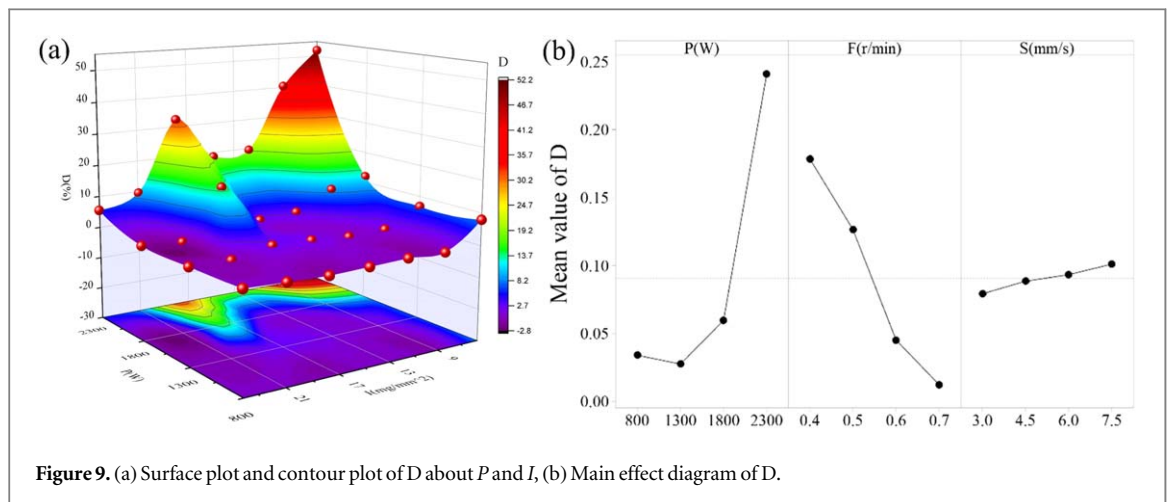


Figure 9. (a) Surface plot and contour plot of D about P and I, (b) Main effect diagram of D.

H, increasing F initially increases H and then decreases it, and increasing S leads to a decrease in H . Ultimately, this results in a non-linear relationship between I and H . Within the melting range, when P remains constant, a larger I leads to a greater amount of powder deposition per unit time, resulting in a larger H . When I is constant, increasing P leads to a larger amount of melted powder, thereby increasing H .

3.2.3. Dilution rate

In the laser cladding process, where high-energy lasers are used to melt cladding powder and form a melt pool on the substrate, the deposition of cladding material on the substrate surface is key to enhancing its surface properties. One crucial indicator is forming a sufficiently strong metallurgical bond between the deposited material and the substrate, or previously deposited layer [20]. The dilution rate (D) is typically used to gauge the bonding strength in laser cladding processes. The surface and contour plots shown in figures 9(a) and (b) show that D increases with increasing P . This is because it increases the melt pool area, which allows more powder material to be melted and bonded to the substrate material, inducing more material to fill the microscopic voids on the surface of the substrate, which increases the contact area between the materials, and in turn increases D . Moreover, the overall effect of I on the dilution rate exhibits an increasing trend.

3.3. Evaluation of model performance

To address the non-linear mapping between process parameters and the characteristics of the cladded layer, eight ML algorithms, namely linear regression (LR), KNN, BP, SVR, DTR, RFR, Extra Trees, and XGBoost, were employed to predict the Geometric Characteristics of laser cladding. The accuracy performance metric, IM, was obtained for each algorithm's prediction of the laser cladding response values, and the results are summarized in table 6. Cross-validation is a method for assessing the performance of a model by dividing the dataset into a training set and a validation set, which allows for an assessment of the model's ability to generalize over unseen data.

Table 6. Statistical evaluation of global accuracy performance indicators for various prediction algorithms in predicting geometric characteristics.

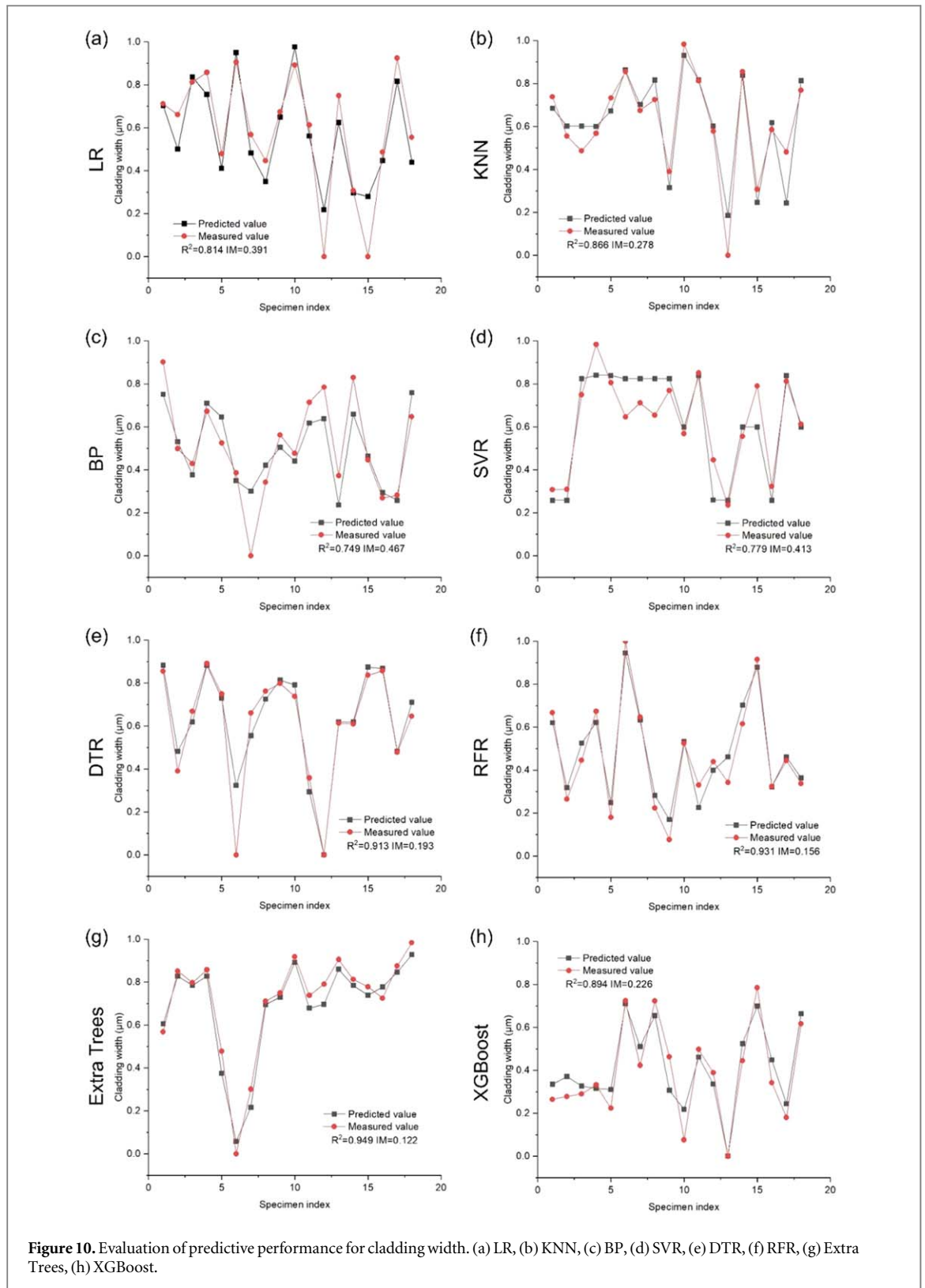
ML algorithm		5-fold cross-validation				Validation set			
		R ²	MSE	MAE	IM	R ²	MSE	MAE	IM
LR	W	0.672	0.017	0.092	0.571	0.814	0.03	0.093	0.391
	H	0.579	0.025	0.111	0.692	0.672	0.018	0.093	0.572
	D	0.206	0.033	0.146	0.986	0.569	0.027	0.133	0.709
KNN	W	0.738	0.015	0.089	0.480	0.866	0.007	0.089	0.278
	H	-0.125	0.038	0.0139	1.004	0.604	0.026	0.136	0.669
	D	-0.229	0.054	0.174	0.991	0.297	0.026	0.098	0.931
BP	W	0.119	0.021	0.108	1.002	0.749	0.019	0.081	0.467
	H	-0.078	0.032	0.178	1.025	0.718	0.012	0.089	0.505
	D	-0.651	0.072	0.198	0.666	0.414	0.013	0.098	0.842
SVR	W	0.694	0.018	0.106	0.546	0.779	0.01	0.076	0.413
	H	0.104	0.048	0.175	1.028	0.332	0.042	0.172	0.929
	D	-3.024	0.095	0.282	8.155	-0.394	0.133	0.338	0.980
DTR	W	0.881	0.004	0.05	0.238	0.913	0.007	0.05	0.193
	H	0.484	0.027	0.116	0.792	0.836	0.009	0.097	0.330
	D	0.342	0.028	0.1	0.904	0.912	0.005	0.037	0.186
RFR	W	0.705	0.016	0.086	0.526	0.931	0.004	0.05	0.156
	H	0.698	0.018	0.089	0.537	0.954	0.003	0.045	0.114
	D	0.423	0.01	0.065	0.830	0.743	0.02	0.085	0.477
Extra Trees	W	0.853	0.007	0.062	0.292	0.949	0.003	0.044	0.122
	H	0.662	0.018	0.085	0.584	0.872	0.008	0.067	0.264
	D	0.428	0.018	0.085	0.832	0.87	0.01	0.069	0.272
XGBoost	W	0.534	0.023	0.096	0.737	0.894	0.006	0.068	0.226
	H	0.681	0.014	0.08	0.555	0.831	0.015	0.077	0.342
	D	0.532	0.024	0.077	0.738	0.766	0.021	0.084	0.446

As shown in table 6, among the eight prediction methods used in this study, the performance of 5-fold cross-validation in terms of the IM metric is inferior to that of the validation set data. This is because it provides a more stringent evaluation of the model. In 5-fold cross-validation, the data is divided into five mutually exclusive subsets for model training and prediction, with each subset serving as the test set in turn. The cross-validation results reflect the model's performance on new, unseen data. This implies that the model is tested on a broader range of data, leading to a higher prediction error. Therefore, the predictive performance of cross-validation may exhibit larger variance and error [34, 35]. Furthermore, looking at the IM results of the 5-fold cross-validation among the eight prediction methods, the generalization capabilities of LR and tree-based models outperform KNN, BP, and SVR. This is due to LR being a parametric model that assumes a linear relationship between features and the target. This simple assumption makes LR less prone to overfitting during training, hence it has better generalization capabilities. The higher IM of tree-based models in CV could be attributed to their binary decision structure for prediction, which is insensitive to outliers. Moreover, models like Random Forest Regression (RFR), Extra Trees, and XGBoost are ensemble models that improve prediction performance by integrating multiple decision trees. The ensemble method can average out the noise of individual decision trees, enhancing the model's stability and robustness.

As discerned from figure 10(b)–(h) exhibit a high degree of fit for cladding width prediction. Table 6 reveals that the maximum R² value is for Extra Trees at 0.949, with KNN having the minimum at 0.866. Considering the predictive performance evaluation indicators of MSE, MAE, and R² as per equation (2), the IM values for DTR, RFR, Extra Trees, and XGBoost in predicting cladding width are 0.192, 0.156, 0.122, and 0.226 respectively, indicating that Extra Trees has the best predictive performance of cladding width.

Observing figure 11(e)–(g) demonstrate superior fitting results for cladding height prediction. Table 6 shows that the maximum R² value is for RFR at 0.954, with DTR having the minimum at 0.836. Considering the predictive performance evaluation indicators of MSE, MAE, and R² as per equation (3), the IM values for DTR, RFR, Extra Trees and XGBoost in predicting cladding height are 0.330, 0.114, 0.264 and 0.342 respectively, indicating that RFR has the best predictive performance of cladding height.

The observation of figures 12(e)–(h) shows that the algorithms DTR, RFR, Extra Trees and XGBoost have better predictive performance for the prediction of dilution rate, and through table 6 it is known that the largest R² value is DTR with 0.912 and the smallest is RFR with 0.743. According to equation (3), the IM evaluation metrics that combine the prediction performance of MSE, MAE and R², the IM values of DTR, RFR, Extra Trees and XGBoost for the prediction models of melting width are 0.186, 0.477, 0.272 and 0.446, respectively, which



indicates that RFR has the best prediction performance. All four algorithms have IM values greater than 0.3 for dilution rate, indicating that despite min-max Scaling data preprocessing, the small dilution rate and numerous zero values result in a decline in fitting results.

From figures 10, 11 and 12, it can be seen that tree-based algorithms DTR, RFR, Extra Trees, and XGBoost have the best predictive performance and accuracy. LR, KNN, BP, and SVR have poorer predictive performance due to the complex mapping relationship between laser cladding process parameters and cladding layer quality. Simple LR cannot accomplish such a complex prediction task. The KNN algorithm predicts based on the distance between data points, but when dealing with high-dimensional data, distance calculation becomes very

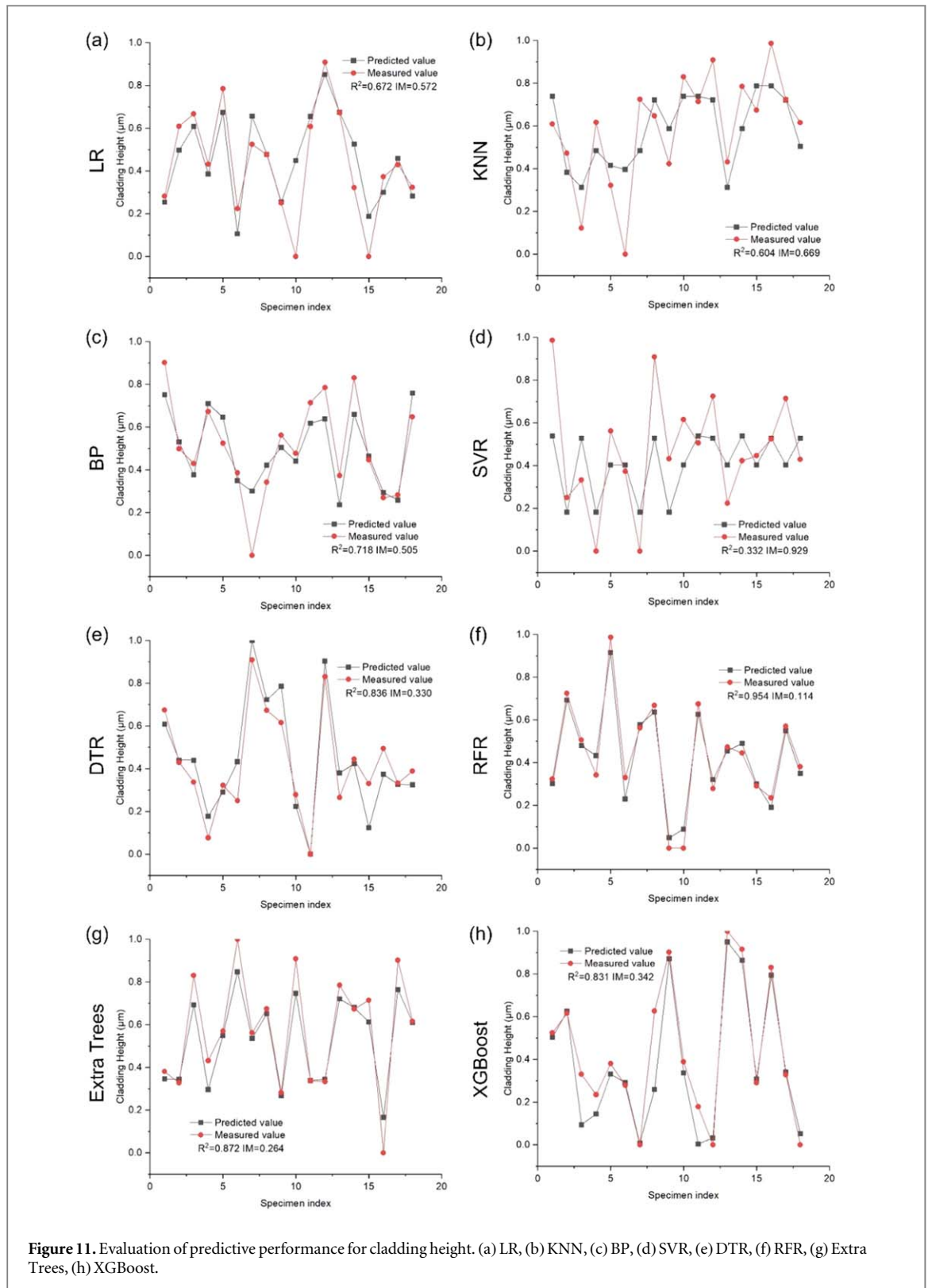


Figure 11. Evaluation of predictive performance for cladding height. (a) LR, (b) KNN, (c) BP, (d) SVR, (e) DTR, (f) RFR, (g) Extra Trees, (h) XGBoost.

difficult, leading to a decline in predictive performance. Due to the adoption of a single response, the cladding width, in determining the hyperparameters during the selection process, it may lead to poorer predictive performance when predicting the width and dilution rate. This is the reason why SVR and BP cannot fit the mapping relationship between height, dilution rate, and clad quality well [3–6].

The magnitude of the IM values calculated by the DTR, RFR, Extra Trees, and XGBoost prediction algorithms shows that Extra Trees is the optimal prediction model in the prediction process of cladding width, improving performance by 36.8%, 21.8%, and 46% compared to DTR, RFR, and XGBoost respectively. RFR has the best predictive performance in the prediction process of cladding height, improving performance by 56.4%,

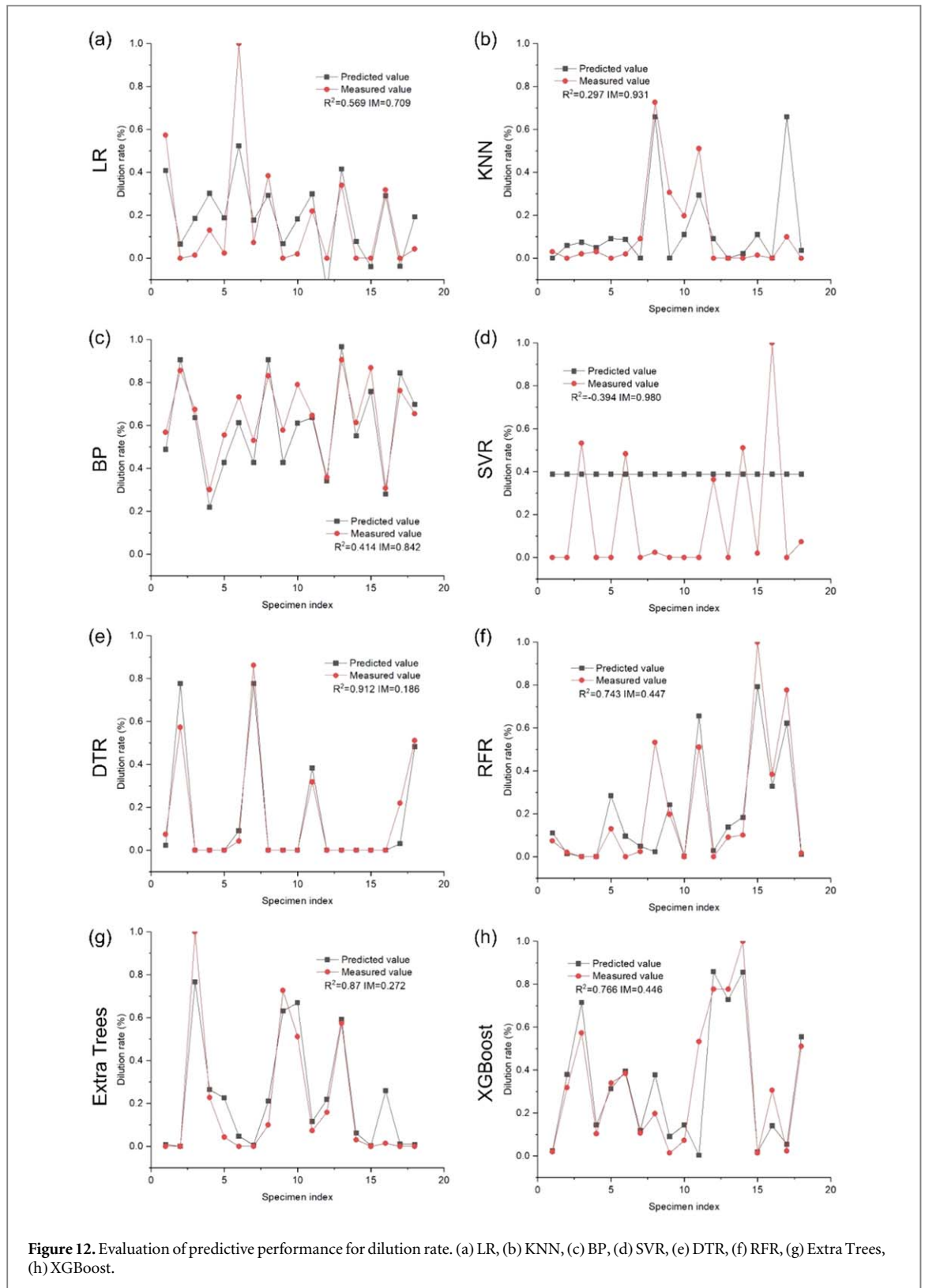


Figure 12. Evaluation of predictive performance for dilution rate. (a) LR, (b) KNN, (c) BP, (d) SVR, (e) DTR, (f) RFR, (g) Extra Trees, (h) XGBoost.

45.5%, and 57.9% compared to DTR, Extra Trees, and XGBoost respectively. DTR has superior predictive ability in the prediction process of dilution rate, improving performance by 58.4%, 31.6%, and 58.3% compared to RFR, Extra Trees, and XGBoost respectively. In general, the four tree-based ML algorithms DTR, RFR, Extra Trees, and XGBoost, by adopting tree structure, bagging, and boosting strategies, can effectively predict the results of the laser cladding process, providing powerful tools for the optimization and control of the laser cladding process.

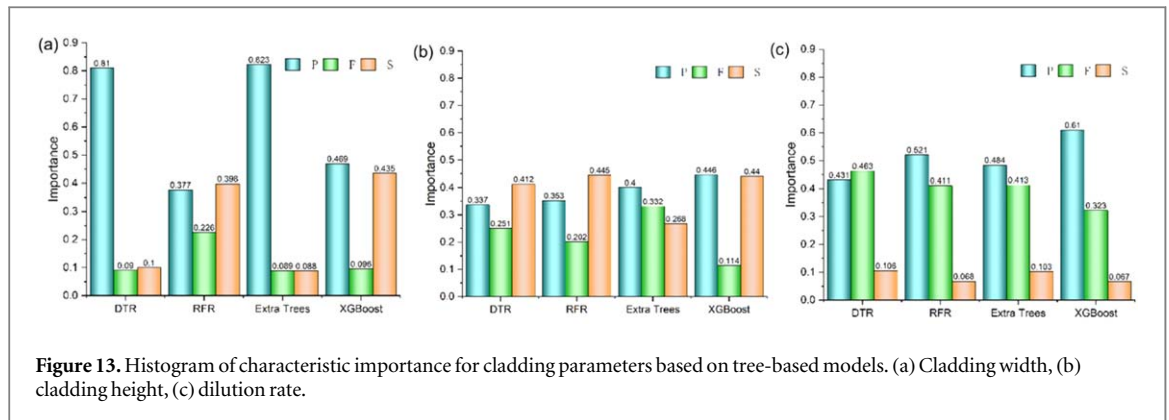


Figure 13. Histogram of characteristic importance for cladding parameters based on tree-based models. (a) Cladding width, (b) cladding height, (c) dilution rate.

3.4. Importance of different independent variables

Tree-based models possess a 'feature importance' attribute, which aids in analyzing the influence of different independent variables or features on the output of prediction results. Each independent variable is assigned a value from 0 to 1, with the total sum being 1. As seen in figure 13(a), among the factors influencing width in the laser cladding process, laser power has the highest score and the greatest contribution in the four tree-based model algorithms DTR, RFR, Extra Trees, and XGBoost, with respective values of 81%, 37.7%, 82.3%, and 46.9%. In the ANOVA of table 5, the contributions of each factor are 90.39%, 1.16%, and 8.45% respectively, indicating that the Extra Trees algorithm can predict width effectively. As seen in figure 13(b), in the laser cladding process, laser power and scanning speed contribute about 40% to height, which corresponds well with the ANOVA results in table 5 of 43.85%, 9.30%, and 46.85%. This suggests that the RFR algorithm, which has the smallest IM, can predict height and that the contribution rate of each factor to height is reliable. As seen in figure 13(c), in the laser cladding process, P and F are significant factors influencing D, both around 40%. Compared with the ANOVA results in table 5 of 63.80%, 35.68%, and 0.52%, DTR can be used as the prediction algorithm for dilution rate. The 'feature importance' attribute results are consistent with our ANOVA results, further validating the effectiveness of our chosen ML algorithms. From figure 13(c), it can be observed that P and F have a significant impact on D in the laser cladding process, both around 40%. This is consistent with the ANOVA results in table 5, which show percentages of 63.80%, 35.68%, and 0.52% respectively. Therefore, DTR can be considered as a suitable predictive algorithm for dilution rate. The 'feature importance' attribute results align with our ANOVA results, further validating the effectiveness of the ML algorithms we have chosen.

4. Conclusion

This paper uses the full-factor experimental design method to clad In718 coating on an A286 substrate, with laser power, powder feed rate, and scanning speed as inputs, and cladding width, cladding height, and dilution rate as responses. The effects of various factors on responses were studied using ANOVA and surface plots, and the laser cladding process was predicted using common ML algorithms, providing a reference for ML in laser cladding process prediction.

The conclusions of this study are as follows:

- (1) The ANOVA results show that laser power, scanning speed, and powder feed rate all have extremely significant effects on the laser cladding process, with only the powder feed rate having a significant effect on the dilution rate. Among them, laser power contributes the most to cladding width (90.39%) and dilution rate (63.80%), while powder feed rate contributes the most to cladding height (46.85%).
- (2) Through the analysis of contour and surface plots, an increase in laser power will increase cladding width, height, and dilution rate; an increase in powder distribution density will increase width and height and decrease dilution rate.
- (3) Through the application of eight machine learning algorithms, LR, KNN, BP, SVR, DTR, RFR, Extra Trees, and XGBoost, it was determined that DTR, RFR, Extra Trees, and XGBoost exhibited superior predictive performances. Specifically, in the prediction of cladding width, height, and dilution rate, Extra Trees, RFR, and DTR demonstrated the highest predictive accuracy, with Index of Merit (IM) values of 0.122, 0.114, and 0.186, respectively.
- (4) The tree-based models, with their 'feature importance' attribute, play a similar role to the contribution rate in ANOVA. They provide valuable insights into the influence of different variables on the prediction results,

aligning well with our ANOVA findings and further validating the effectiveness of our chosen ML algorithms.

Data availability statement

All data that support the findings of this study are included within the article (and any supplementary files).

Funding

The work was financially supported by the National Science Fund for Excellent Young Scholars (Oversea), China Postdoctoral Science Foundation Funded Project (Project No. 2021M693415), Jiangsu Provincial Postdoctoral Science Foundation Funded Project (Project No. 2020C340), Jiangsu Provincial Double-Innovation Doctor Program (Project No. 202031063), and China Postdoctoral International Exchange Program (Project No. PC2022061). M.A. gratefully acknowledges financial support received from University of Waterloo (Starter Grant).

ORCID iDs

Marco Alfano  <https://orcid.org/0000-0001-8000-4092>

Junfeng Yuan  <https://orcid.org/0000-0002-3968-0248>

References

- [1] Tamanna N, Crouch R and Naher S 2019 Progress in numerical simulation of the laser cladding process *Opt. Lasers Eng.* **122** 151–63
- [2] Fotovvati B, Dehghanhadikolaei A and Namdari N 2020 Laser-assisted coating techniques and surface modifications: a short review *Part. Sci. Technol.* **39** 738–47
- [3] Chen X, Xiao M, Kang D, Sang Y, Zhang Z and Jin X 2021 Prediction of geometric characteristics of melt track based on direct laser deposition using M-SVR algorithm *Materials (Basel)* **14** 7221
- [4] Gao J, Wang C, Hao Y, Wang X, Zhao K and Ding X 2022 Prediction of molten pool temperature and processing quality in laser metal deposition based on back propagation neural network algorithm *Opt. Lasers Eng.* **155** 108363
- [5] Deng Z, Chen T, Wang H, Li S and Liu D 2020 Process parameter optimization when preparing ti(c, n) ceramic coatings using laser cladding based on a neural network and quantum-behaved particle swarm optimization algorithm *Appl. Sci.* **10** 6331
- [6] Li T, Long H, Qiu C, Wang M, Li D, Dong Z and Gui Y 2022 Multi-objective optimization of process parameters of 45 steel laser cladding Ni60PTA alloy powder *Coatings* **12** 939
- [7] Liu H, Qin X, Huang S, Hu Z and Ni M 2018 Geometry modeling of single track cladding deposited by high power diode laser with rectangular beam spot *Opt. Laser Technol.* **100** 38–46
- [8] Wu Z, Li T, Li Q, Shi B, Li X, Wang X, Lu H and Zhang H-C 2019 Process optimization of laser cladding Ni60A alloy coating in remanufacturing *Opt Lasers Eng.* **120** 105718
- [9] Xu Z, Yuan J, Wu M, Arif A F M and Li D 2023 Effect of laser cladding parameters on Inconel 718 coating performance and multi-parameter optimization *Opt. Lasers Eng.* **158** 108850
- [10] Farayibi P K, Abioye T E and Clare A T 2016 A parametric study on laser cladding of Ti-6Al-4V wire and WC/W2C powder *Int. J. Adv. Manuf. Technol.* **87** 3349–58
- [11] Svetlizky D, Zheng B, Steinberg D M, Schoenung J M, Lavernia E J and Eliaz N 2022 The influence of laser directed energy deposition (DED) processing parameters for Al5083 studied by central composite design *J. Mater. Res. Technol.* **17** 3157–71
- [12] Erfanmanesh M, Abdollah-Pour H, Mohammadian-Semnani H and Shoja-Razavi R 2017 An empirical-statistical model for laser cladding of WC-12Co powder on AISI 321 stainless steel *Opt. Lasers Eng.* **97** 180–6
- [13] Bhatnagar S, Mullick S and Gopinath M 2021 A lumped parametric analytical model for predicting molten pool temperature and clad geometry in pre-placed powder laser cladding *Optik* **247** 168015
- [14] Chen B, Bian Y, Li Z, Dong B, Li S, Tian C, He X and Yu G 2023 Effect of laser beam profile on thermal transfer, fluid flow and solidification parameters during laser-based directed energy deposition of Inconel 718 *Materials (Basel)* **16** 4221
- [15] Jin Z, Kong X, Ma L, Dong J and Li X 2023 Prediction of primary dendrite arm spacing of the inconel 718 deposition layer by laser cladding based on a multi-scale simulation *Materials (Basel)* **16** 3479
- [16] Lv H, Liu Y, Chen H, Zhang W, Lv S and He D 2023 Temperature field simulation and microstructure evolution of Fe-based coating processed by extreme high-speed laser cladding for re-manufacturing locomotive axle *Surf. Coat. Technol.* **464** 129529
- [17] Barrionuevo G O, Mullo J L and Ramos-Grez J A 2021 Predicting the ultimate tensile strength of AISI 1045 steel and 2017-T4 aluminum alloy joints in a laser-assisted rotary friction welding process using machine learning: a comparison with response surface methodology *Int. J. Adv. Manuf. Technol.* **116** 1247–57
- [18] Xia C, Pan Z, Polden J, Li H, Xu Y and Chen S 2021 Modelling and prediction of surface roughness in wire arc additive manufacturing using machine learning *J. Intell. Manuf.* **33** 1467–82
- [19] Barrionuevo G O, Ramos-Grez J A, Walczak M and Betancourt C A 2021 Comparative evaluation of supervised machine learning algorithms in the prediction of the relative density of 316L stainless steel fabricated by selective laser melting *Int. J. Adv. Manuf. Technol.* **113** 419–33
- [20] Li Y, Wang K, Fu H, Zhi X, Guo X and Lin J 2021 Prediction for dilution rate of AlCoCrFeNi coatings by laser cladding based on a BP neural network *Coatings* **11** 1402
- [21] Zhang Y, Gong B, Tang Z and Cao W 2022 Application of a Bio-inspired algorithm in the process parameter optimization of laser cladding *Machines* **10** 263

- [22] Alizadeh-Sh M, Marashi S P H, Ranjbarnodeh E, Shoja-Razavi R and Oliveira J P 2020 Prediction of solidification cracking by an empirical-statistical analysis for laser cladding of Inconel 718 powder on a non-weldable substrate *Opt. Lasers Eng.* **128** 106244
- [23] Lian G, Zhang Y, Zhang H, Huang X, Chen C and Jiang J 2019 Investigation of geometric characteristics in curved surface laser cladding with curve path *Metals* **9** 947
- [24] Song M, Wu L, Liu J and Hu Y 2021 Effects of laser cladding on crack resistance improvement for aluminum alloy used in aircraft skin *Opt. Lasers Eng.* **133** 106531
- [25] Wolff S J, Gan Z, Lin S, Bennett J L, Yan W, Hyatt G, Ehmann K F, Wagner G J, Liu W K and Cao J 2019 Experimentally validated predictions of thermal history and microhardness in laser-deposited Inconel 718 on carbon steel *Addit. Manuf.* **27** 540–51
- [26] Kumar P and Jain N K 2022 Surface roughness prediction in micro-plasma transferred arc metal additive manufacturing process using K-nearest neighbors algorithm *Int. J. Adv. Manuf. Technol.* **119** 2985–97
- [27] Gao J, Wang C, Hao Y, Liang X and Zhao K 2022 Prediction of TC11 single-track geometry in laser metal deposition based on back propagation neural network and random forest *J. Mech. Sci. Technol.* **36** 1417–25
- [28] Czarnecki S, Sadowski Ł and Hoła J 2021 Evaluation of interlayer bonding in layered composites based on non-destructive measurements and machine learning: comparative analysis of selected learning algorithms *Autom. Constr.* **132** 103977
- [29] Giri J, Al-Lohedan H A, Mohammad F, Soleiman A A, Chadge R, Mahatme C, Sunheriya N, Giri P, Mutyarapwar D and Dhapke S 2023 A Comparative study on predication of appropriate mechanical ventilation mode through machine learning approach *Bioengineering (Basel)* **10** 418
- [30] Little K, KP B, Yang S, Noronha B, Campolo D and Accoto D 2021 Elbow motion trajectory prediction using a multi-modal wearable system: a comparative analysis of machine learning techniques *Sensors (Basel)* **21** 6501
- [31] Shanmugasundar G, Vanitha M, Čep R, Kumar V, Kalita K and Ramachandran M 2021 A comparative study of linear, random forest and adaboost regressions for modeling non-traditional machining *Processes* **9** 2015
- [32] Xi W, Song B, Zhao Y, Yu T and Wang J 2019 Geometry and dilution rate analysis and prediction of laser cladding *Int. J. Adv. Manuf. Technol.* **103** 4695–702
- [33] Yu T, Yang L, Zhao Y, Sun J and Li B 2018 Experimental research and multi-response multi-parameter optimization of laser cladding Fe313 *Opt. Lasers Eng.* **108** 321–32
- [34] Vakharia V, Shah M, Nair P, Borade H, Sahlot P and Wankhede V 2023 Estimation of lithium-ion battery discharge capacity by integrating optimized explainable-AI and Stacked LSTM model *Batteries*. **9** 125
- [35] Suthar V, Vakharia V, Patel V K and Shah M 2022 Detection of compound faults in ball bearings using multiscale-singan, heat transfer search optimization, and extreme learning machine *Machines*. **11** 29