International Journal of Fatigue 190 (2025) 108588

Contents lists available at ScienceDirect

International Journal of Fatigue

journal homepage: www.elsevier.com/locate/ijfatigue

Explainable machine learning-based fatigue assessment of 316L stainless steel fabricated by laser-powder bed fusion

Xiru Wang [a], Moritz Braun [a,b,*]

[a] Institute of Ship Structural Design and Analysis, Hamburg University of Technology, Hamburg, Germany
[b] German Aerospace Center (DLR), Institute of Maritime Energy Systems, Geesthacht, Germany

ARTICLE INFO

Keywords:
Additive manufacturing
Fatigue life prediction
Fatigue strength assessment
Machine learning approaches
Gradient boosted trees
SHAP

ABSTRACT

Additive manufacturing (AM) and in particular laser-powder bed fusion has become a popular manufacturing techniques in recent years due to its significant advantages; however, the mechanical behavior of AM components often varies from components fabricated using conventional processes. For example, the fatigue behavior of components made by AM processes is heavily influenced by process-related defects and residual stresses in addition to applied stress amplitudes, stress ratio and surface conditions. Accounting for the interaction of these effects in fatigue design is difficult by means of traditional fatigue assessment concepts. Machine learning algorithms offer a possibility to account for such interactions and are easily applied once trained and validated. In this study, machine learning algorithms based on *gradient boosted trees* with the *SHapley Additive exPlanation* framework are used to predict defect location and fatigue life of additive manufactured AISI 316L specimens in as-built and post-treated manufacturing states, while also facilitating the understanding of the importance and interactions of various influencing factors.

1. Introduction

In recent years, additive manufacturing (AM) technology has experienced a rapid development and widespread application across various industries, including engineering, medicine, architecture, and education [1]. In comparison to conventional processing techniques, AM offers the capability to produce intricate geometries, reduce production time and costs, and enable customization and human interactions [2]

Among the various AM technologies, laser-based powder bed fusion (L-PBF) stands out as one of the most prevalent processes for 316L [3]. L-PBF utilizes a high-energy source to melt the metallic powder within a powder bed system [4]. 316L stainless steel is extensively used as a metal in AM applications due to its high corrosion resistance [5,6], high thermal conductivity [7], high melting point [8], and strong infrared absorptivity [9]; however, AM-produced components may generally exhibit significant characteristic imperfections when compared to other manufacturing methods. These imperfections include higher porosity and surface roughness [10–12], resulting from gaps left by powder shape and flow, which may impact the fatigue properties of the AM components [9,13]. In addition, the L-PBF process often leads to tensile residual stresses in AM components [14–16]. Thus, post-processing is frequently applied to mitigate these effects. Several studies have shown that the fatigue strength of L-PBF 316L components can be enhanced through post-processing, i.e., heat, machining or polishing [17–20].

Under external loading, force- and stress-related parameters, such as cyclic stress amplitude, stress ratio also influence the fatigue strength and number of cycles to failure. Traditional assessment methods, such as stress-life [21,22], strain-life [23,24], or fracture mechanics approaches [25–27], are commonly employed for the analysis of fatigue behavior and prediction of fatigue life; however, additively manufactured parts involves numerous influencing factors, various and uneven distributed defects [28] and complex interactions, which increases the complexity and challenges in predicting fatigue behavior using traditional assessment methods. Often, conventional assessment methods are unable to elucidate the interactions among influencing factors and their impact on the predicted results.

Due to the limitations of traditional assessment methods, machine learning (ML) approaches have provide an alternative for handling multivariate data, analyzing complex relationships between variables and performing prediction. Extensive literature supports the efficacy of ML models, showcasing their excellent performance and strong prediction accuracy [29,30]. Studies have investigated the advantages of ML models in predicting fatigue life and optimizing fatigue performance for AM-built parts and steel spring manufacturing [31–34]. In addition, ML models performed also well with relatively small experimental datasets in predicting fatigue life for L-PBF AlSi10Mg alloy [35,36]. Furthermore, ML methods were found effective in investigating the

* Correspondence to: German Aerospace Center (DLR), Institute of Maritime Energy Systems, Düneberger Str. 108, D-21502 Geesthacht, Germany.
E-mail address: moritz.braun@dlr.de (M. Braun).

https://doi.org/10.1016/j.ijfatigue.2024.108588
Received 29 January 2024; Received in revised form 31 August 2024; Accepted 3 September 2024
Available online 5 September 2024

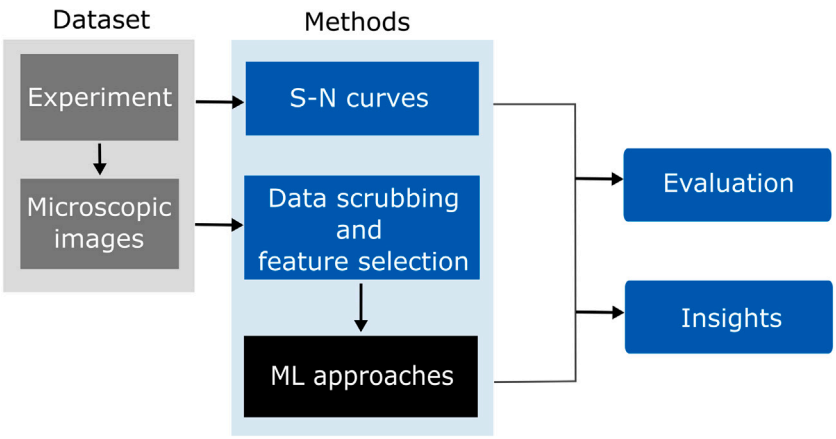

Fig. 1. The flowchart of fatigue strength assessment of AM specimens in this study.

influence of post-treatments on AM parts [37]. Moreover, ML models can be used to explore the relationship between defect features and the fatigue life of L-PBF specimens and different steels [38,39]. On the other hand, the interpretation of results can become difficult as model complexity increases, and the cooperative game theory-based *SHapley Additive exPlanation* (SHAP) framework technique provides a way to measure the importance of each input attribute for the model output [40,41]. Braun et al. [42,43] employed ML models based on *gradient boosted trees* to analyze fatigue behavior and used the SHAP framework to investigate feature importance and their interactions on predictions.

Therefore, building upon the analysis of fatigue strength using conventional fatigue assessment methods, this study further uses advanced ML techniques to predict both fatigue location and lifetime, subsequently conducting a comparative analysis, a practice seldom observed in previous studies. In addition, there is limited research that utilizes ML to analyze the impact of certain parameters on the prediction of AM specimens as well as the interactions between these influencing parameters. Thence, this work delves into an investigation of the mutual influence between various parameters and their importance on the outcomes of predictions by combining XGBoost with SHAP value.

The aim of this study is to integrate machine learning models with fatigue data analysis to enhance predictive ability using limited experimental data. By combining ML models with detailed characterization data, such as defect size, defect location, and post-treatment conditions, this approach provides deeper insights beyond traditional stress-life (S-N) curve analysis. Additionally, the use of SHAP values facilitates an intuitive understanding of the importance and interaction of various factors on fatigue strength.

This study is structured as follows. First, stress-life (S-N) curves are fitted to analyze the fatigue strength of all 80 specimens under different manufacturing and test conditions (including as-built (AB), heat-treated (HT), and turned and heat-treated (THT) specimens). Second, Pearson correlation coefficients (PCC) are used to select the features, reducing redundant input variables that are dependent on each other. Then, ML models are built to predict fatigue failure location and fatigue life of the AM specimens. In addition, the SHAP value theory is used to rank the main potentially important influencing parameters of fatigue life prediction and analyze their interactions. Meanwhile, the effective mean stress approach is introduced to understand the effect of residual stresses and mean stresses on fatigue life prediction.

2. Methodology

Fig. 1 shows the general flowchart on fatigue strength assessment of additive manufactured AISI 316L samples based on experimental fatigue tests and machine learning.

2.1. Data collection

As described in the study from Braun et al. [17], all AISI 316L specimens were prepared using laser-based powder bed fusion technique in a Renishaw AM250 L-PBF machine. The selection of standard specimen geometry follows ASTM E466-15 [1]. To investigate the effect of residual stress as well as the benefit of surface treatment, some as-built samples were machined and then heat-treated at 650°C for 2 h, followed by furnace cooling. The specimen geometry with its image is shown in Fig. 2. Since the surface condition cannot be changed with heat treatment [17], the surface of AB specimens was presumed to exhibit similarities to that of HT specimens. The measured maximum roughness R_t and arithmetic mean roughness R_a for HT (same as AB) and THT specimens are listed in Table 1. In addition, the impact of heat treatment on microstructure in dry air can be considered negligible, since heat treatment at 650°C primarily aims to reduce residual stresses without significantly altering the grain boundary microstructure [44]. Although grain boundary sensitization occurs within 10 nm according to the study by Tekin et al. [45], this effect is not significant enough to influence fatigue performance in dry air.

2.2. Basic relations of fatigue strength assessment

The most commonly used statistical method for estimating the fatigue strength of tested specimens is linear regression, represented by an S-N curve. The S-N curve typically delineates the relationship between stress range $\Delta\sigma$ or stress amplitude σ_a and fatigue life N_f, using the material constant C.

$$\log N_f = \log C + m \log \Delta\sigma \tag{1}$$

Stress amplitude σ_a is half the stress range $\Delta\sigma$, $\sigma_a = \Delta\sigma/2$ and stress ratio R is defined as:

$$R = \frac{\sigma_{min}}{\sigma_{max}} = \frac{\sigma_m - \sigma_a}{\sigma_m + \sigma_a} \tag{2}$$

where σ_{min} and σ_{max} are the minimum and maximum applied stresses, respectively. σ_m represents the mean stress.

Considering that mean stress will also be used in this study, the relationship between mean stress, stress amplitude, and stress ratio is described as follows:

$$\sigma_m = -\frac{R+1}{R-1} \cdot \sigma_a \tag{3}$$

2.3. Data preprocessing

To enhance the accuracy of the subsequent ML models, preprocessing of the test dataset was conducted before training the models. Outliers (runouts in fatigue tests) were removed from the dataset,

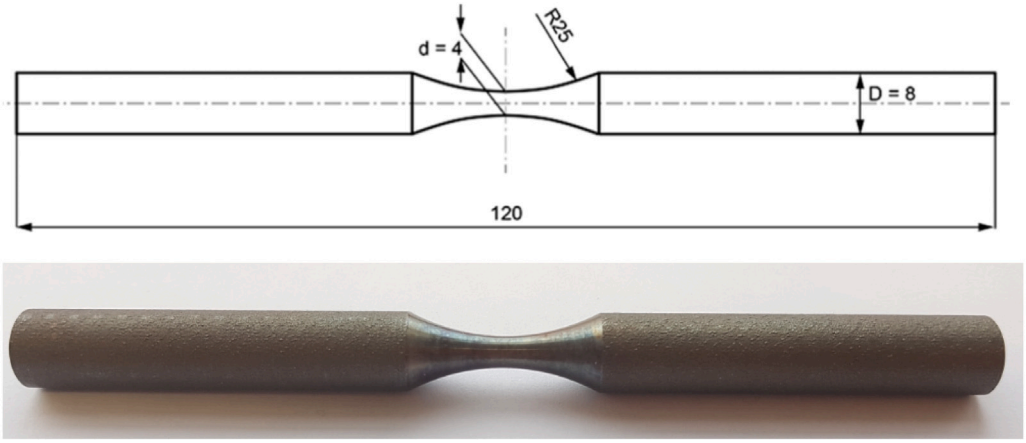

Fig. 2. Specimen geometry and image after turning and heat treatment.
Source: Taken from Braun et al. [17]

Table 1
Measured roughness for HT and THT samples, taken from [17].

Test series	Maximum roughness R_t [µm]	Arithmetic mean roughness R_a [µm]
Heat-treated	41.93	6.30
Turned and Heat-treated	5.06	1.02

and force-related features were excluded due to their correlation with stress-related features. Categorical features, such as the state of specimens and defect location, were encoded into numerical features. Furthermore, Murakami [46] emphasized the relationship between prediction of maximum defect size and the prediction of fatigue strength. Thus, $\sqrt{area}$, the square root of defect projected area to the maximum stress direction is used as the defect size parameter, because the material defects cannot be described as a simple circle [46]. As an example, Fig. 3 illustrates a pore defect on the fracture surface of an AB specimen and the measurement of its defect size.

2.4. Feature extraction

Due to the potential impact of causal dependency on the performance of ML prediction models [47], it is essential to ensure that the input features are independent of each other [31] to avoid over-complex model [39]. Thus, the Pearson correlation coefficient (PCC) is used to evaluate the correlation between features. The PCC is calculated as the ratio of the covariance to the product of the standard deviations of two features.

$$PCC = \frac{\sum_{i=1}^{n}(X_{1i} - \bar{X_1})(X_{2i} - \bar{X_2})}{\sqrt{\sum_{i=1}^{n}(X_{1i} - \bar{X_1})^2}\sqrt{\sum_{i=1}^{n}(X_{2i} - \bar{X_2})^2}} \tag{4}$$

Here n is the sample size, X_1 and X_2 are the features with their mean value $\bar{X_1}$ and $\bar{X_2}$

According to Eq. (4), PCC takes values between −1 and 1. A PCC value closer to zero indicates a lower correlation between features. A value of 0 signifies no linear dependency between these two features. PCC values of −1 or 1 indicate perfect negative or positive correlation between the chosen features, respectively.

Fig. 4 represents the heat map of PCC of all features that may influence the defect of location. All samples are classified into three states according to their post-processing: as-built, heat-treated, as well as turned and heat-treated. These categorical features are then encoded in alphabetical order and utilized as input for the machine learning model. It is well known that the state of the specimens and the temperature of heat treatment are closely related. For as-built specimens, the temperature of treatment is set to 20°C; however, for both HT and THT specimens, the temperature of treatment is 650°C. Therefore, state of specimens is selected for further assessment. As Eq. (3) shows, a mean stress of 0MPa yields a stress ratio $R = -1$, regardless of the value of stress amplitude. This could possibly explain the PCC value of 0.024 between σ_a and σ_m, and a PCC of 0.93 for σ_m and R. On the other hand, regardless of the individual magnitudes of σ_m and σ_a, as long as they are equal, the stress ratio R will always be 0. Hence, to further investigate the influence of σ_a, σ_m, and R on fatigue life prediction model, this study separately analyzes the following two cases due to the strong correlation between σ_m and R: (1) using σ_a and σ_m as input parameters; (2) taking σ_a and R as influencing parameters. In addition, the remaining features can all be used as input, because their PCCs are less than 0.5 [35].

Considering the effect of residual stress which occurs in the production processes, mean stress effect and their interaction, effective mean stress $\sigma_{m,eff}$ and effective stress ratio R_{eff} are introduced for ML fatigue life prediction models.

$$\sigma_{m,eff} = \sigma_m + \sigma_{RS,stab} \tag{5}$$

$$R_{eff} = \frac{\sigma_{min} + \sigma_{RS,stab}}{\sigma_{max} + \sigma_{RS,stab}} = \frac{\sigma_{m,eff} - \sigma_a}{\sigma_{m,eff} + \sigma_a} \tag{6}$$

in which $\sigma_{RS,stab}$ represents the cyclically stabilized residual stress at $N_f = 10^6$, as suggested by Hensel [48]; however, the current study used initial residual stresses σ_{RS} that were measured before testing. With the help of $\sigma_{m,eff}$ and R_{eff}, the effective mean stress approach in this study offers an alternative to evaluate the combined effect of nominal mean stresses (due to external loading) and residual stresses for each state of specimens.

According to Fig. 5, the state of specimens, applied stress amplitude σ_a, effective defect area $\sqrt{area}$ as well as the distance of defect center to surface L are uncorrelated features and should therefore be used for the ML models. In addition, the residual stress is perfectly correlated to state of specimens, rendering it unnecessary to be considered as an independent input parameter; however, given its significance and following the effective mean stress approach [48], residual stress is analyzed in conjunction with the state of specimens, mean stress and stress ratio. Therefore, R_{eff} and $\sigma_{m,eff}$ are chosen for the evaluation.

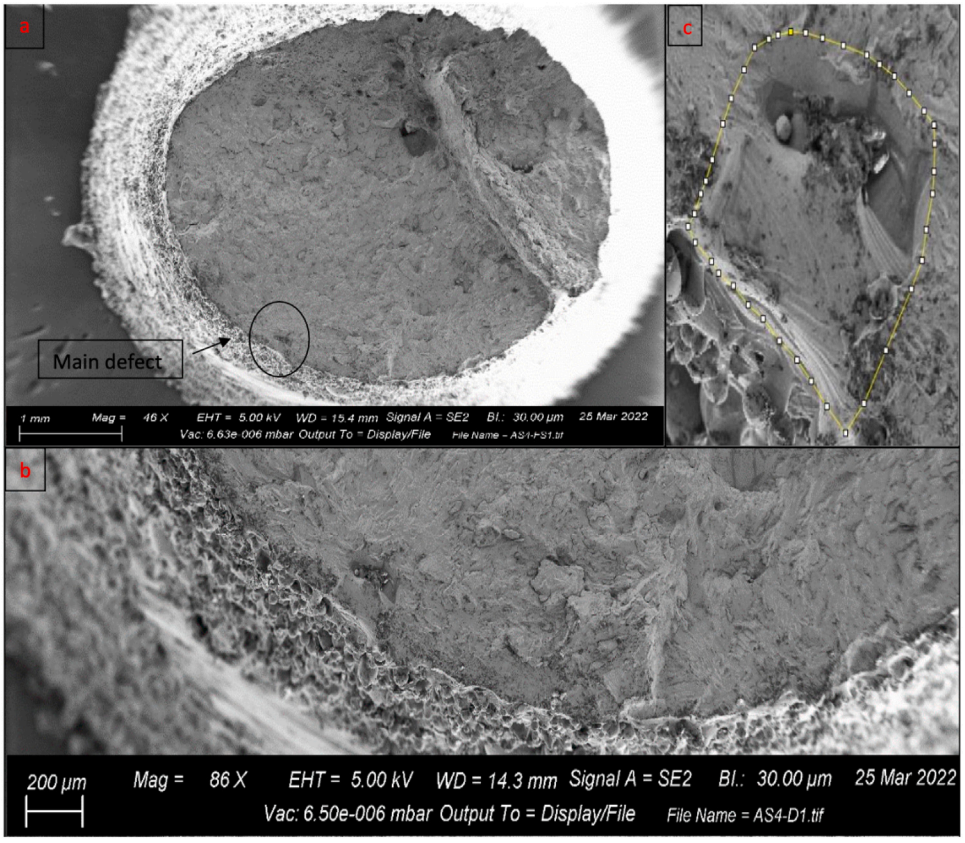

Fig. 3. (a) Defect surface of AB specimen no.4 with failure initiation from a subsurface lack of fusion defect. (b) Zoomed view. (c) Measurement of defect size.

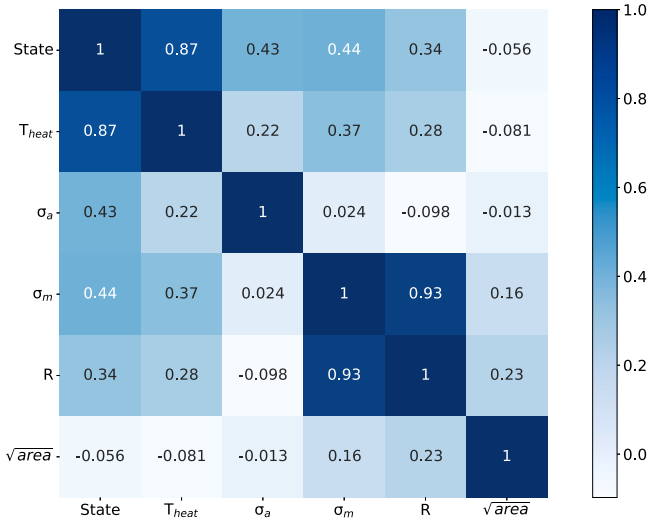

Fig. 4. PCC of all features for defect location prediction.

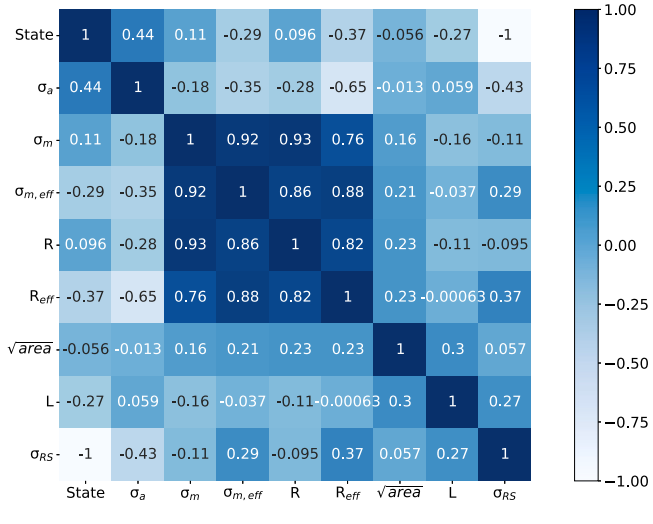

Fig. 5. PCC of all features for fatigue life prediction.

2.5. Model training and evaluation

ML models are increasingly used in the fatigue field for fatigue life estimation [49] and can be applied for prediction after training [43,50]. Furthermore, while in this study, the defect sizes used as input values are ascertained through examination of fracture surfaces using scanning electron microscope (SEM), the application of ML models in practical scenarios becomes feasible if fracture dimension data can be obtained through non-destructive computer tomography scans. Certainly, traditional fatigue assessment methods can also perform predictions based on the input data, but have limitations, e.g. they cannot provide insights into the interactions between different parameters or their influence on the outcome. In contrast, ML has the capability to analyze complex relationships between input variables and make predictions while also identifying which parameters have the most significant impact on the predicted outcome. Therefore, ML offers a more comprehensive and insightful approach to prediction than other methods. Thus, ML methods are used in this study to evaluate the fatigue behavior of AM parts and SHAP values are employed to investigate the interactions between the main influencing features, as illustrated in Fig. 6.

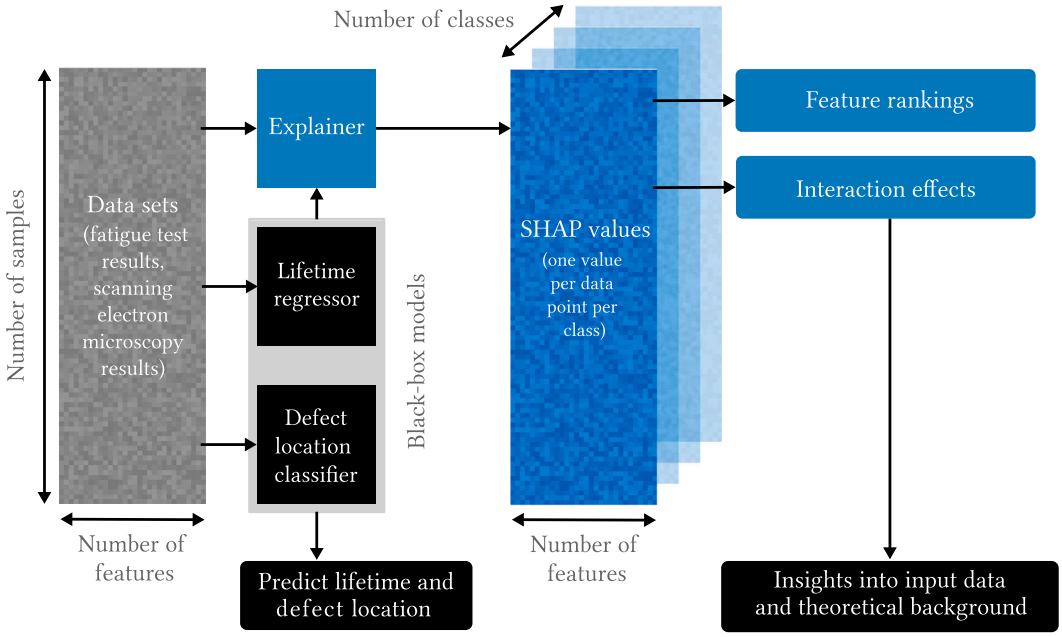

Fig. 6. Flowchart presenting the ML approach for defect location and fatigue life prediction via ML models. Both models are combined with an additional explainer model to generate SHapley Additive exPlanation (SHAP) values. These can be used for model agnostic interpretation, e.g., to determine feature rankings and interaction effects. Figure based on Braun and Kellner [42].

2.5.1. XGBoost models

Among all ML approaches, a decision tree (DT) is one of the most popular tools for classification and prediction problems. Although DT models perform well for small datasets [51] and are interpretable, there have certain limitations. Song et al. [52] highlighted that decision trees can result in overfitting and underfitting for small database, which can restrict generalizability and robustness of the resulting models. Therefore, the utilization of ensemble methods involving the construction of a collection of individual decision trees was employed to enhance the accuracy of classification outcomes [53]. In addition, tree-based gradient boosting models create a strong estimator during the averaging combination of weak estimators, using iterative learning with adjusted instance distribution [35,54]. Unlike traditional gradient boosting techniques, XGBoost utilizes a Taylor expansion to approximate the loss function, achieving a superior balance between bias and variance [55]. The loss function of XGBoost is defined in Eq. (7).

$$L(y_i, \hat{y}_i^{(m)}) = \sum_{i=0}^{n} L(y_i, \hat{y}_i^{(m-1)} + f_m(x_i)) + \Omega(f_k) \tag{7}$$

where n the numbers of samples, $\hat{y}_i^{(m-1)}$ is the current estimate of the true value y_i after the previous iteration, $f(x)$ is the tree that used for final predicted value after K trees: $\hat{y}_i = \sum_{k=0}^{K} f_k(x_i)$. $\Omega(f_k)$ is used to set the penalty of the complicated trees.

In addition, XGBoost exhibit superior performance compared to conventional deep models when applied to tabular-style datasets [56]. Due to their elevated precision and interpretability, tree-based models have emerged as prominent non-linear predictive models in contemporary research [57]. On the other hand, XGBoost is noted for its fast computation and excellent performance, outperforming other tree algorithms in both accuracy and time for both classification and regression problems [58,59]. Additionally, due to its capability to accept sparse matrices as input [60], XGBoost is well-suited for handling the fatigue data used in this study, which includes categorical variables such as the state of specimens. Moreover, among all gradient boosting models, the XGBoost model has been demonstrated to be applicable for assessing the fatigue behavior of AM components [35] and effective in estimating the fatigue life of AM parts [61]. Furthermore, in conjunction with SHAP framework, the XGBoost model performed also well for the fatigue behavior prediction of welded joints [43]. Thus, XGBoost was chosen in this study to investigate the fatigue behavior of AM parts.

The hyperparameters of the model were chosen as follows: 800 gradient boosted trees with a learning rate of 0.08, a subsample size of 60%. The total data was split into 85% training samples and 15% testing samples, in order to ensure sufficient test data because of limited given data. In addition, five-fold cross-validation was used in order to estimate the performance of the model, which means that the training set was again split into 5 subsets for training and testing. To avoid overfitting and improve the generalization, the ML training would be stopped if no improvement was detected over 3 training iterations.

In this study, a defect location classifier was built to predict where the defect initiated, i.e., center, subsurface and surface. The objective employs the softmax function as defined in Eq. (8), while multiclass logloss serves as the evaluation metric for early stopping calculated, see Eq. (9).

$$\hat{y}_{ij} = \frac{e^{z_{ij}}}{\sum_{k=1}^{c} e^{z_{ik}}} \tag{8}$$

$$\text{LogLoss} = -\frac{1}{n} \sum_{i=1}^{n} \sum_{j=1}^{c} y_{ij} \log(\hat{y}_{ij}) \tag{9}$$

with n the numbers of samples, c the number of classes, $\hat{y}_{ij}$ is the predicted probability of sample i belonging to class j, z_{ij} is the score for class j of instance i, y_{ij} is a binary indicator (0 or 1) if class label j is the correct classification for sample i.

In addition, a regression model was used for the fatigue life prediction. Squared error (SE) is used as the objective $L_{SE} = \frac{1}{2}(y_i - \hat{y}_i)^2$, root-mean-square error (RMSE) is used as the evaluation metric for early stopping, see Eq. (13).

The evaluating parameters for the defect location classifier include the accuracy (ACC) and the Matthew correlation coefficient (MCC), which are based on the number of true positive predictions (TP), true negatives (TN), false positives (FP), and false negatives (FN) [43]. The accuracy represents the percentage of correct predictions among all predictions, ranging from 0 to 1 and it can be obtained by:

$$ACC = \frac{TP + TN}{TP + TN + FP + FN} \tag{10}$$

On the other hand, MCC was used because it can handle unbalanced datasets, i.e. different size of samples for each class. For binary classifier, MCC yields a high value only if the model correctly predicts all positive or negative instances [62]. MCC rangs from −1 to 1, where −1 indicates complete misclassification and 1 indicates complete classification. Additionally, in this study, MCC also addresses the multiclass scenario. It can be viewed as a discrete multivariate version of the Pearson Correlation Coefficient (PCC) [63], as introduced in Section 2.4, where prediction and true label serve as two features.

$$MCC_{\text{binary}} = \frac{TP+TN}{\sqrt{(TP+FP)(TP+FN)(TN+FP)(TN+FN)}} \tag{11}$$

To assess the predictive performance of the XGBoost model, two metrics, R-squared (R^2) and root-mean-square error (RMSE) are utilized.

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

$$RMSE = \sqrt{\frac{1}{n}\sum(y_i - \hat{y}_i)^2} \tag{13}$$

where y_i is the true value, $\hat{y}_i$ is the predicted value, n is the number of samples and $\bar{y}$ is the average value of generated dataset. R^2 represents the similarity between predicted and true value, with a range from 0 to 1. The closer R^2 to 1 is, the more accurate the model is. RMSE indicates the average deviation between observed value and predicted results. A smaller RMSE indicates better model performance.

2.5.2. Explanation of SHAP value

While many machine learning models are designed for accurate predictions, it is equally important to understand why a model makes certain predictions and the quantitative relationships between the model's predictions and input features. Therefore, the need for model interpretability arises. Some machine learning models, like linear regression or decision trees, are inherently interpretable, but more complex models often act as black boxes, making their prediction methods hard to interpret. In such cases, as a personalized model-agnostic interpreter, SHAP can be used to explain the original model by accessing input data and prediction results. By leveraging concepts from game theory, the process used by the original model to make the prediction is mimicked. SHAP has been used in a number of studies for feature analysis and importance ranking [31,42,43,50,64]. The calculation of SHAP values can be found in Appendix.

3. Results

3.1. S-N curves

All test data for different states and manufacturing conditions with a stress ratio of $R = 0$ are presented in Fig. 7 and analyzed using designed S-N curve with best-fit slopes. In addition, the mean fatigue strength amplitude at $N_f = 2 \times 10^6$ cycles for a probability of survival of $P_s = 50\%$ is noted in the figure. Comparing to as-built specimens, the mean fatigue strength amplitude of the heat-treated specimens was improved from 89.2 MPa to 104 MPa, corresponding to an increase of 16.6% because of a reduction of residual stress. Furthermore, an 87.5% improvement in fatigue strength was achieved through the implementation of the additional turning treatment; however, it was proved by Braun et al. [17] that surface roughness changes would only result in an approximately 7.1% variance in fatigue strength, significantly lower than the impact of residual stresses. In addition, turning treatment induces compressive residual stresses on the surface, which contribute to improved fatigue strength [65]; however, this effect caused by turning is typically small, and the residual stresses are further reduced by subsequent heat treatment. Hence, process-related defects have more significant impact on the fatigue strength of L-PBF components because most defects are located in the region that was removed by turning treatment.

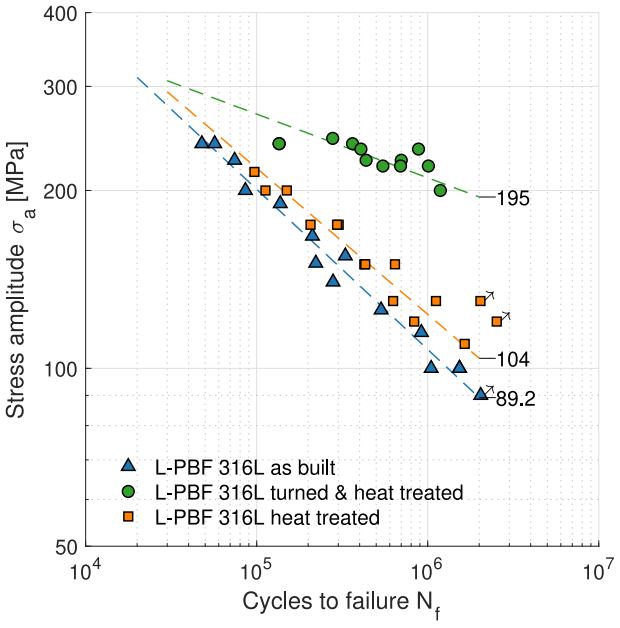

Fig. 7. S-N curves for different L-PBF production conditions with a stress ratio of $R = 0$.

3.2. Analysis of mean stress and residual stress effect

It is widely known that mean stresses have a significant impact on mean fatigue strength. Increased mean stress generally leads to a decrease in fatigue strength and cycles to failure [66]. Furthermore, in AM processes, residual stresses combined with external stress can result in high effective mean stresses, but as the external stress increases, residual stress relaxation occurs [17]. In this case, to analyze the effect of mean stress, residual stress as well as their interaction on fatigue strength, tests with alternating load $R = -1$ and $R = 0.5$ were conducted for both as-built and heat-treated specimens. The results of tests are presented in Fig. 8 and the results of mean fatigue strength at $N_f = 2 \times 10^6$ cycles as well as best-fit slopes of the S-N curve are summarized in Table 2. The residual stresses of AB and HT specimens were measured before testing. Unfortunately, the residual stresses of THT specimens could not be measured as all specimens were tested. Thus, the value of THT residual stress is assumed to be 0MPa. This assumption is based on the fact that the turning induces compressive residual stresses close to the surface, where the majority of cracks initiate. In addition, as indicated in Fig. 7, heat treatment contributes again a reduction of the residual stress state.

For as-built specimens, the mean fatigue strength for $R = -1$ is higher than for $R = 0$, which agrees with the general knowledge that higher mean stress results in a decrease of fatigue strength. Moreover, Fig. 8(a) shows that the best-fit slope exponent of mean curve for $R = -1$ is lower, indicating that in the absence of load-related mean stress, residual stress has a higher impact on low-stress levels than on high-stress levels. Comparing with the results for as-built and heat-treated specimens with $R = -1$, it can be concluded that heat treatment is relative effective in reducing residual stress and improving fatigue strength; however, there is no great difference between the mean fatigue strength and the best-fit slopes for both states with $R = 0.5$. One possible explanation for this observation is that at higher mean stresses, the maximum stress is also higher, which leads to a residual stress relaxation. As a result, the effect of residual stresses is reduced compared to that of mean stresses. This also explains the higher best-fit slope for both states for $R = 0.5$, compared to $R = -1$ or $R = 0$. Furthermore, this phenomenon also explains the reduction in fatigue strength for heat-treated specimens from $R = -1$ to $R = 0$.

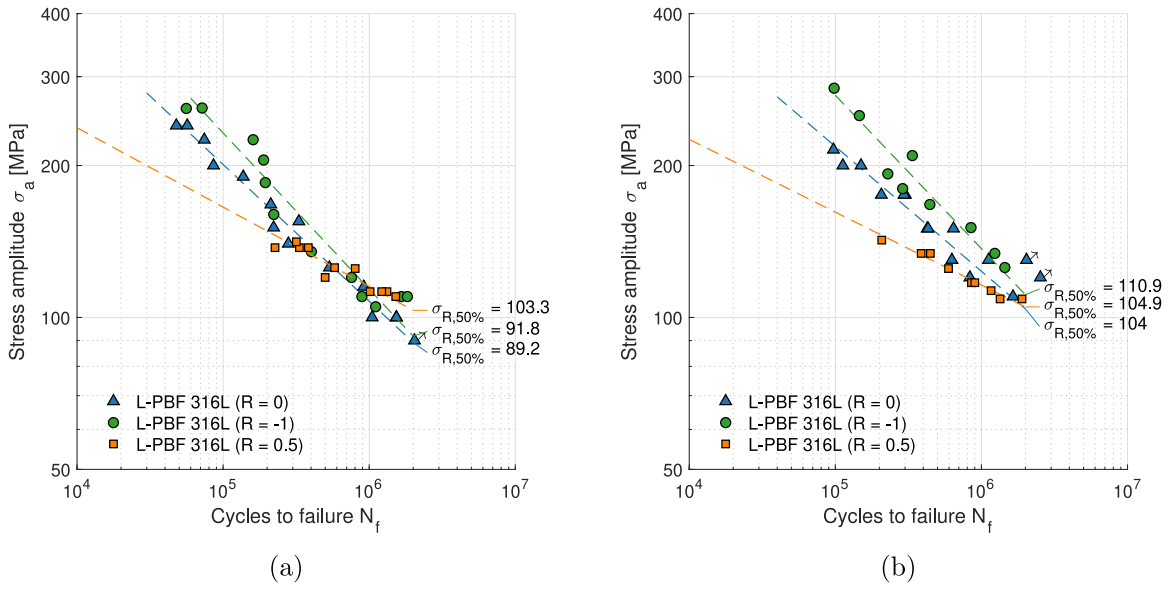

Fig. 8. Comparison of S-N curves with $R = 0$, $R = -1$ and $R = 0.5$ for L-PBF samples in as-built state (**a**) and in heat-treated state (**b**).

Table 2
Fatigue test data with calculated mean fatigue strength at $N_f = 2 \times 10^6$.

Test series	Stress ratio R	Best-fit slope	Residual stress level σ_{RS}[MPa]	Mean fatigue strength $\sigma_{R,50\%}$ at $N_f = 2 \times 10^6$[MPa]	Mean stress σ_m[MPa]
As-built	0	3.69	159.2	89.2	90–240
	−1	3.23		91.8	0
	0.5	6.36		103.3	330–423.8
Heat-treated	0	4.04	76.8	104	110–215
	−1	3.29		110.9	0
	0.5	6.93		104.9	326.5–426.9
Turned and heat-treated	0	9.27	0*	195	200–245

* Assumed value.

3.3. Performance of ML models

To judge the precision of an ML model, different metrics can be used, see Section 2.5.1. As explained in Section 2.4, the defect location classifier model was trained separately with (1) σ_a and σ_m; (2) σ_a and R. Table 3 shows the performance of defect location classifier with case (1), which was run 5 times with different random seeds. For case (1), the model showed the best performance and achieved an accuracy (ACC) score of 0.80 ± 0.06 and MCC of 0.36 ± 0.31, indicating a relatively high accuracy level. Meanwhile, the classifier after cross-validation (CV) had an ACC value of 0.70 ± 0.15 and MCC of only 0.17 ± 0.25. A lower MCC means that it is slightly better than random prediction. The likely reason for this outcome is data deficiency. Additionally, despite data cleaning and preprocessing, including outlier removal and feature selection to reduce dimensionality, overfitting still occurs due to the limited size of the training set and the complexity of the classifiers [67]. Furthermore, Subramanian and Simon [68] suggested that overfitting can be a significant issue when the relationship between the outcome and the predictor variables is weak. The results of prediction are reflected in the confusion matrix depicted in Fig. 9(a). As shown in Fig. 9(a), only eight data points were available for testing the built classifier model. Among these tested values, six of them were classified as defect location at surface although their input features varies greatly. Consequently, when using cross-validation or different random seeds, it is possible that all predicted values fall into the "surface" category. In such cases, the MCC value equals 0, as the confusion matrix becomes all zeros except for one column [69]. This explains why the mean value of MCC is relatively low. Hence, the defect location prediction cannot be assessed adequately, despite the prediction results aligning well with expectations. In such a case, further investigation of fracture location by XAI is not meaningful. To evaluate defect location and defect type more accurately, additional tests are required.

The performance evaluation of the fatigue life prediction model is summarized in Table 4, presenting the results of R-squared (R^2) and root-mean-square error (RMSE) for both cross-validation and test data. For the model considering effective mean stress $\sigma_{m,eff}$ and stress ratio R, the R^2 was found to be 0.83 and root-mean-square error (RMSE) was 0.17. These values were consistent with the average results obtained after cross-validation, which showed $R^2 = 0.78 \pm 0.03$, RMSE $= 0.17 \pm 0.03$. These outcomes indicate that the machine learning model performed well for fatigue life prediction despite the small dataset. Moreover, the higher R^2 value and lower RMSE value for the model using the effective mean stress approach demonstrate its superior performance compared to the model utilizing the nominal stress ratio R. Twelve data points were used for testing. The results of predictions based on $\sigma_{m,eff}$ with R_{eff} and R are listed in Table 5 and compared to the actual number of cycles to failure in Fig. 9(b). Fig. 9(b) illustrates that the deviation between true and predicted number of cycles to failure is mostly below a factor of two. Notably, the model using effective mean stresses which considers the effect of residual stress shows an

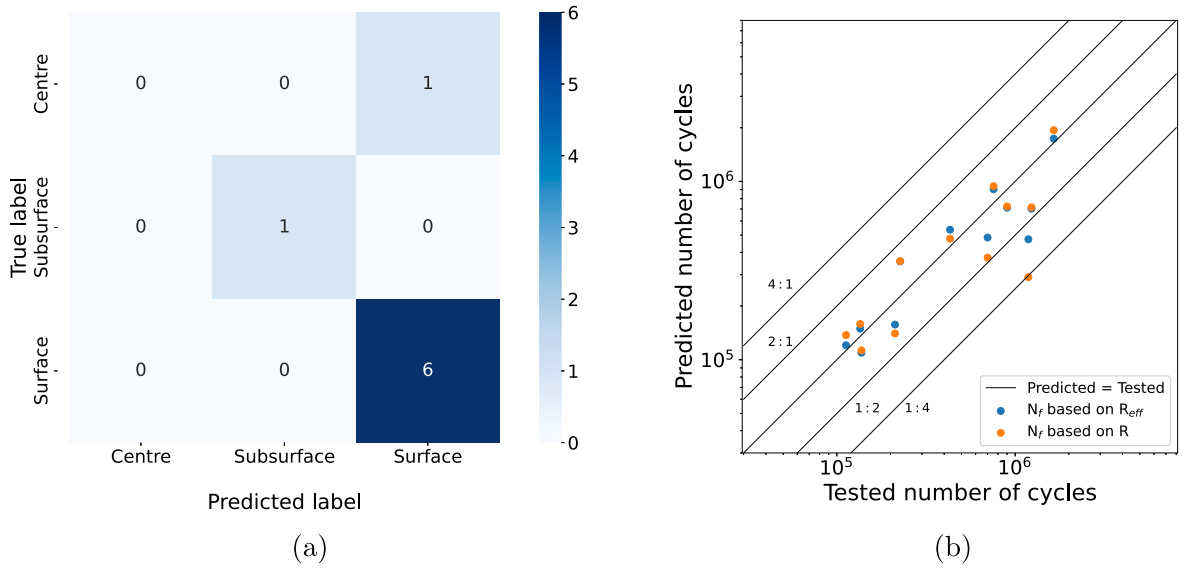

Fig. 9. (**a**) Confusion matrix for true versus predicted defect location with decision tree classifier. (**b**) Comparison of tested versus predicted number of cycles to failure based on R_{eff} and R.

Table 3
Test and cross-validation performance of the defect location prediction model considering σ_a and σ_m as input parameters.

	ACC after CV	MCC after CV	ACC tested	MCC tested
Defect location classifier	0.70 ± 0.15	0.17 ± 0.25	0.80 ± 0.06	0.36 ± 0.31

The results specify mean value ± standard deviation for $k = 5$.

Table 4
Test and cross-validation performance of the fatigue life prediction model with R_{eff} and R.

	R^2 after CV	RMSE after CV	R^2 tested	RMSE tested
Fatigue life prediction model based on R_{eff}	0.78 ± 0.03	0.17 ± 0.03	0.83	0.17
Fatigue life prediction model based on R	0.76 ± 0.03	0.20	0.70 ± 0.02	0.22 ± 0.01

The results specify mean value ± standard deviation for $k = 5$.

even better accuracy. This supports the conclusion that the ML model performed well, although there is a slight trend of underpredict for the number of cycles to failure around $N_f = 10^6$. Hence, effective mean stress $\sigma_{m,\text{eff}}$ and effective stress ratio R_{eff} are applied for further investigation, instead of nominal mean stress σ_m and nominal stress ratio R.

In addition, Table 5 shows percentage change between tested and predicted value based on R_{eff}; however, this value is not "symmetrical", which means that the change from tested value to predicted value is not same as that from predicted to tested value. Thus, log difference are used as an alternative to percentage change. Unlike percentage change, the log difference is symmetrical, regardless of which parameter we use as the base for calculation. The results further highlight the high accuracy of the fatigue life prediction model.

$$\text{Log difference} = \log_{10}(N_{\text{f,predicted}}) - \log_{10}(N_{\text{f,tested}}) \tag{14}$$

where $N_{\text{f,predicted}}$ is the predicted fatigue life by XGBoost and $N_{\text{f,tested}}$ is the experimental fatigue life for testing.

3.4. Explanation of fatigue life regressor

3.4.1. Importance ranking of all features

In addition to predicting the fatigue life of the additively manufactured specimens, *explainable AI* (XAI) is employed to gain insights into the influencing factors and their interactions. This approach proves valuable for practical applications and the advancement of AM techniques in industrial production.

The analysis of the relationship between the predictors and the output variable is based on the mean value of absolute SHAP values. In the case of the fatigue life regressor, the output is represented by the logarithmized number of cycles to failure (N_f). Therefore, a higher value of $\log(N_f)$ corresponds to a larger N_f value. Similarly, a higher SHAP value indicates a stronger effect of the corresponding input feature on achieving a higher number of cycles to failure [17]. By analyzing these SHAP values, we can understand how each predictor impacts the fatigue life prediction and the relative importance of different input features in determining the number of cycles to failure.

Fig. 10(a) displays the ranking of feature importance concerning fatigue life prediction. A higher rank indicates that the corresponding feature has a strong influence on the prediction; however, using average absolute SHAP values alone does not allow us to differentiate between the magnitude and prevalence of each feature's impact on the prediction [17]. In other words, while we can identify which features are relatively more important, we cannot ascertain whether their impact on the fatigue life prediction is larger or more widespread compared to other features. For a comprehensive understanding of the predictor's significance, Fig. 10(b) provides the SHAP values corresponding to each sample with respect to different features. This explain the impact of each feature more in detail. Each dot represents a single sample. Pile dots indicate that there is more than one sample with the same or similar value.

Table 5
List of tested versus predicted number of cycles to failure based on R_{eff} and R.

Number and state of specimens		Tested N_f	Predicted N_f based on R	Predicted N_f based on R_{eff}	Percentage difference	Log difference
53	THT	135 184	158 711	149 396	11%	0.04
78	HT	1 645 125	1 938 435	1 738 302	6%	0.02
75	HT	431 371	477 265	536 136	24%	0.09
37	AB	226 544	357 638	356 058	57%	0.20
6	AB	211 900	140 442	157 471	−26%	−0.13
21	AB	756 673	940 869	904 082	19%	0.08
9	AB	137 388	112 778	109 484	−20%	−0.10
57	HT	1 234 319	715 499	703 256	−43%	−0.24
70	HT	901 056	723 767	712 673	−21%	−0.10
81	HT	112 598	137 500	120 510	7%	0.03
43	THT	700 283	373 947	484 641	−31%	−0.16
42	THT	1 185 390	290 650	473 765	−60%	−0.40

Percentage change is difference between tested and predicted value based on R_{eff}, divided by tested value.

Stress amplitude is ranked highest and is followed by the specimen state, i.e, as-built, heat-treated, or turned and heat-treated. A high stress amplitude has a substantial negative contribution to lifetime prediction, while low values have a significant positive impact, which aligns with general expectations. The beeswarm plots for different states demonstrate that turned and heat-treated specimens have positive impact on higher number of cycles to failure (red dots). Heat-treated specimens perform also well (purple dots), whereas as-built specimens show a negative influence on fatigue life prediction (blue dots). These findings are in line with the study from Braun et al. [17], which suggests that surface roughness and residual stress reduction, as well as removal of sub-surface defects during HT and THT positively influence the fatigue behavior of AM components.

Similarly to stress amplitude, a high effective mean stress has a negative effect on fatigue life prediction; however, some samples with moderate values (purple dots) exhibit a dispersion of both negative and positive SHAP values. The negative purple dots represent HT specimens subjected to high mean stress and stress amplitude, resulting in a detrimental effect on the predicted fatigue life. In addition, defects closer to the surface negatively impact the number of cycles to failure.

Interestingly, low R_{eff} values have almost no contribution to the prediction (blue dots), while a R_{eff} value around 0 shows a negative impact on fatigue life prediction (purple dots with negative SHAP value). Conversely, higher R_{eff} values exhibit a positive effect on fatigue life predictions. This is thought to be related to the relaxation of tensile residual stresses at high stress levels and inline with the S-N curves presented in Section 2.2.

Compared with other considered parameters, the effective area of main defect has a lower impact on the fatigue life of AM parts, which also demonstrates the significance of residual stresses and the locations of defects (parameter L). In addition, no clear relation of effective area of main defect on fatigue life is observed. The possible reason for this may be the small size of the dataset, so further investigation is not pursued.

3.4.2. Interactions between features on lifetime

Fig. 11(a) presents the SHAP dependence plot for the two most important features that have the highest effect on the fatigue life of the tested specimens. The plot reveals a clear trend: the number of cycles to failure decreases with increased stress amplitude. Furthermore, for the stress amplitude ranging from 200 to 250 MPa, the fatigue life can be improved by implementing additional turning and heat treatment. These results emphasize the significant influence of stress amplitude and the potential benefits of post-processing treatments in enhancing the fatigue performance of the tested components.

From Fig. 11(b), it is evident that high mean stresses in the range of 200 to 600 MPa have a negative effect on fatigue life, which is consistent with the observations in Fig. 10(b) (negative SHAP values). Moreover, high stress amplitude amplifies this negative effect. When combining Fig. 11(b) with Fig. 11(c), it becomes apparent that heat and machining treatment have stronger effects on increasing the fatigue life of samples that were loaded with mean stress less than 250 MPa. On the other hand, for effective mean stress larger than 300 MPa, AB and HT specimens perform similarly. As explained in Section 3.2, high loads result in a relaxation of residual stresses, similar to what is achieved through the HT process. Consequently, heat treatment does not result in a significant improvement in fatigue strength at high stress levels. In addition, compared to the stress amplitude in the Fig. 11(a), the range of SHAP value of effective mean stress is smaller. This observation corroborates the common assumption that mean stress has a lower impact on fatigue life compared to stress amplitude, regarding to all states of specimens. This finding is consistent with the results reported by Braun and Kellner [42] for welded joints using XAI.

Analyzing Fig. 11(d)–11(f) together is meaningful since stress amplitude, mean stress, and stress ratio are interrelated; however, it is important to note that the dots representing the SHAP values for each state are on the same vertical line and may overlap with each other. To enhance plot readability, random jitter was added to the feature values (see Fig. 11(d)). In the case of as-built additive manufacturing parts, a higher effective stress ratio (R_{eff}) leads to a reduced negative effect on fatigue life due to the relaxation of tensile residual stresses. Conversely, for heat-treated (HT) specimens, a higher R_{eff} indicates higher mean stress, resulting in a reduction of fatigue life.

In addition, it is apparent in Figs. 11(e) and 11(f) that for effective stress ratio less than 0.2, higher mean stress (up to about 250 MPa) induces higher negative influence on fatigue life. On the contrary, for points with effective stress ratio larger than 0.2, higher mean stress shows increasingly positive impact on fatigue life. For red points with mean stress over 400 MPa (see Fig. 11(f)), the maximum stress for these points exceeds the yield strength of the used material σ_{YS} = 510 MPa. This effect causes a plastic deformation and again a subsequent cyclic residual stress relaxation, leading to increased fatigue life.

Fig. 11(g) presents the dependence plot for the fatigue life regressor with the distance of the defect center to the surface (L) as the primary feature with $L < 200$ µm (surface and subsurface defects). The plots indicate that an increasing distance of the main defect L has a negative effect on fatigue life. In other words, the closer the defect is to the center of the specimens, the lower the number of cycles to failure; however, this effect is strongly related to the effective area of defects. The defects close to the surface are all comparably small. This observation can be supported by the findings of Solberg et al. [70]. They demonstrated the existence of a transition of crack initiation between low and high loading levels in their study of 316L specimens produced by L-PBF. For higher load levels, fatigue initiation is predominantly influenced by porosity and internal defects rather than surface defects [70]; nevertheless, it is important to note that further testing and investigation are necessary to support these findings.

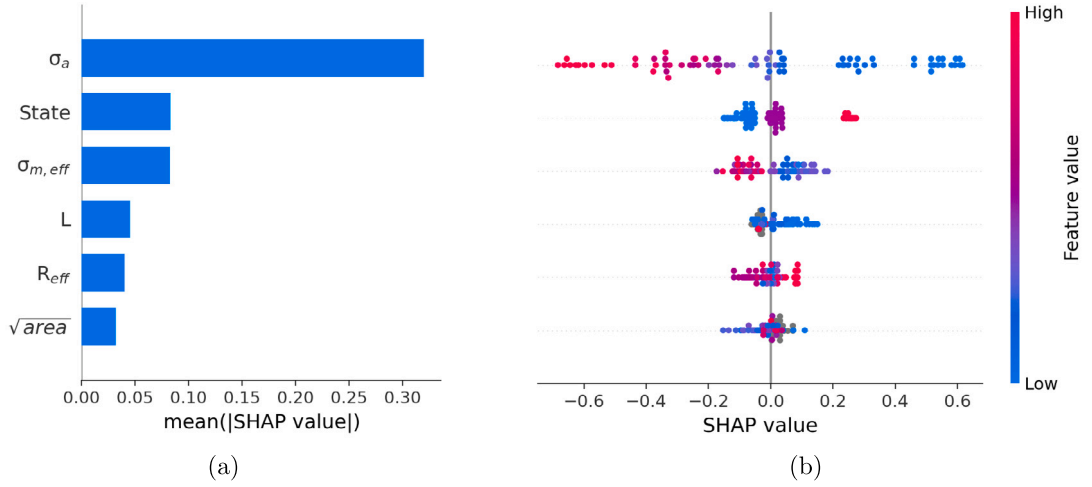

Fig. 10. SHAP summary plot for the lifetime regressor with (**a**) Feature importance ranking according to average SHAP value and (**b**) SHAP value of all selected features.

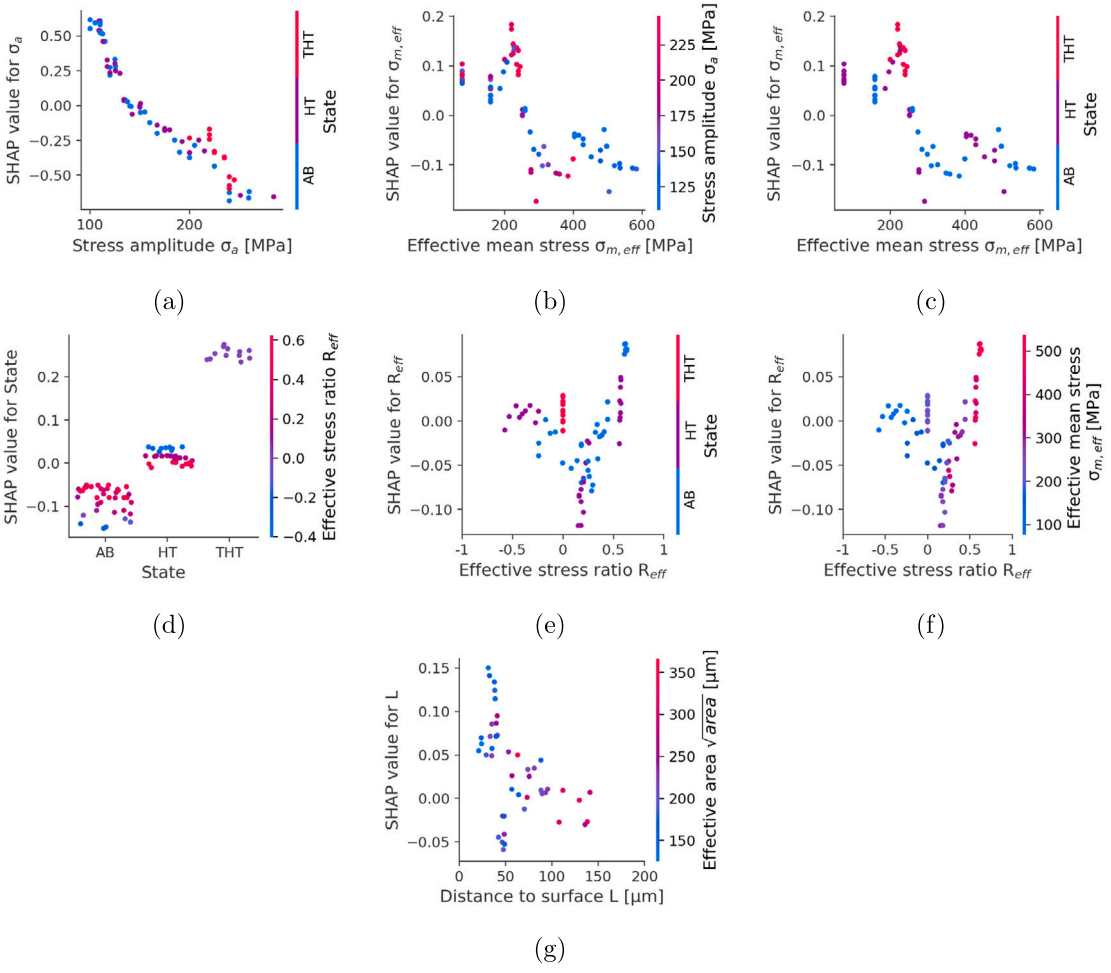

Fig. 11. Dependence plot for lifetime regressor with primary feature on horizontal axis and its SHAP value on vertical axis. The color of dots represents the value of secondary feature, showing the interaction between the primary and secondary feature. (**a**) Stress amplitude and state. (**b**) Effective mean stress and stress amplitude. (**c**) Effective mean stress and state. (**d**) State and effective stress ratio.(**e**) Effective stress ratio and state. (**f**) Effective stress ratio and effective mean stress. (**g**) Distance of the defect center to the surface (L) for $L < 200$ µm (surface and subsurface defects) and effective area.

4. Discussion

In general, machine learning models have shown a sound performance in predicting defect location and fatigue life of additive manufactured specimens, despite the relatively small databases. Furthermore, cross-validation has revealed the robustness of the models for further evaluation regarding input data and features, making them suitable for new data samples. In addition, Braun et al. [42,43] had

shown that XGBoost is suitable for large databases with a greater number of features. Therefore, the models are expected to exhibit even better performance with more extensive training datasets.

In addition, the outcomes obtained from the explanation models using SHAP values align well with theoretical expectations. Using the explanation model, all input features were ranked based on their influence on predicting the number of cycles to failure. For additive manufactured specimens, stress-related features (stress amplitude and mean stress) were found to have the most important impact on fatigue lifetime regressors, which agrees with the findings from Braun and Kellner [42] for butt-welded joints. The process state of the specimens (as-built, heat and turning treatment) also have a high impact. The combination of SHAP values and S-N curves has demonstrated that heat treatment and turning treatment can significantly improve the fatigue life of AM parts. In contrast, fatigue life appeared to be less influenced by defect-related parameters, namely the distance of defect center to the surface and effective area of the defect; nevertheless, more research is required on the interaction of defects with other properties such surface conditions, AM processes and size of components, see, e.g., [71,72].

Moreover, a more detailed investigation of the interactions between primary and secondary features was conducted using SHAP dependence plots combined with the mean stress approach. The results further validate the consistency with theoretical expectations, laying the groundwork for future studies that can incorporate more advanced techniques, such as combining fatigue mechanisms/physics with ML tools. This could further improve prediction performance and provide deeper insights into the fatigue behavior of L-PBF metals. Moreover, leveraging data obtained from computer tomography scans, ML models can be applied to reduce the need for numerous fatigue tests in the future.

In summary, by using XAI, we can better comprehend the underlying factors that affect fatigue life and understand how they interact with each other. This knowledge can be leveraged to enhance the reliability and performance of AM components in practical applications, thereby facilitating the successful implementation and further development of AM technologies in industrial settings.

5. Conclusions

Additive manufacturing has demonstrated substantial advantages over conventional manufacturing processes and is extensively employed in diverse industries for production purposes. To predict defect location and fatigue life of additively manufactured metal components, an analysis was conducted using 80 fatigue test results on 316L stainless steel specimens in as-built and post-treated state. In addition, this study also helped to comprehend the influence of surface roughness, mean stress, and residual stress on fatigue life prediction. To this goal, explainable artificial intelligence based on SHAP values was utilized to investigate the primary influencing factors affecting fatigue strength and lifetime. The following conclusions are drawn from this study:

- The fatigue strength and lifetime of the additively manufactured specimens can be significantly improved by post-treatment, which causes residual stress relaxation, surface roughness reduction and also removes near-surface defects.
- The gradient boosted tree model, employed in this study, performed well for the prediction of number of cycles to failure, despite the small database; however, to improve fatigue failure locations prediction of additively manufactured 316L specimens and to evaluate its influencing features, more tests should be performed.
- Explainable machine learning based on the SHAP framework have proven to be a useful tool to rank influencing parameters and helped to investigate their mutual influence on the fatigue life predictions.
- Fatigue life predictions of additively manufactured AISI 316L samples are primarily influenced by stress-related features (stress amplitude), followed by process state (as-built, turning and/or heat treatment) . On the other hand, the impact of defect location and effective area size is lower than that of other evaluated variables.
- The efficacy of the effective mean stress approach using initial residual stresses has been demonstrated in its application to ML models and the SHAP framework, explaining the influence of mean stress and residual stress on the fatigue life of additively manufactured samples. As the mean stress increases, the fatigue life initially decreases until the maximum stress surpasses the yield strength of the materials. Subsequently, the fatigue life increases again with higher mean stress, due to cyclic residual stress relaxation.

CRediT authorship contribution statement

Xiru Wang: Writing – original draft, Visualization, Software, Investigation, Formal analysis, Data curation, Conceptualization. **Moritz Braun:** Writing – review & editing, Validation, Supervision, Software, Resources, Project administration, Methodology, Investigation, Data curation, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Declaration of Generative AI and AI-assisted technologies in the writing process

During the preparation of this work the authors used ChatGPT in order to check spelling and grammar. After using this tool, the authors reviewed and edited the content as needed and take full responsibility for the content of the publication.

Acknowledgment

The authors would like to acknowledge the support of Sebastian Springer, who performed the residual stress measurements.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Appendix

Calculation of SHAP value

SHAP value for a single feature F can be defined as the average of the marginal distribution of F:

$$\varphi(\upsilon) = \sum_{S \subseteq N \setminus \{i\}} \frac{|S|!(n-|S|-1)!}{n!} (\upsilon(S \cup \{i\}) - \upsilon(S)) \tag{15}$$

where n is the number of features with S the coalition subset of all input features. The characteristic function $\upsilon(S)$ describes the total expect value of coalition S.

The sum of SHAP values of all features equals the value of grand coalition:

$$\sum_{i \subseteq N} \varphi_i(\upsilon) = \upsilon(N) \tag{16}$$

Table A.6
Nomenclature.

Symbol	Description	Unit
$\Delta\sigma$	Stress range	MPa
$\sigma_a, \sigma_{max}, \sigma_{min}$	Stress amplitude, maximum stress, minimum stress	MPa
$\sigma_m, \sigma_{m,eff}$	Cyclic mean stress, effective mean stress	MPa
σ_{RS}, σ_{YS}	Residual stress, Yield strength	MPa
$\sigma_{RS,stab}$	Cyclically stabilized residual stress	MPa
$\sigma_{R,50\%}$	Reference mean fatigue strength at 2×10^6 cycles to failure	MPa
$\sqrt{area}$	Square root of effective area of main defect	µm
d	Diameter of specimens	mm
m	Slope of S-N curve	–
n	Number of specimens	–
$y_i, \hat{y}_i$	True value, predicted value	–
$\bar{y}$	Mean value of generated dataset	–
L	Distance of defect center to surface	µm
N_f	Number of cycles to failure	–
P_s	Survival probability	–
R, R_{eff}	Stress ratio, effective stress ratio	–
R_a, R_t	Arithmetic mean roughness, maximum roughness	µm
R^2	R-squared	–
T_{heat}	Heat treatment temperature	°C
X_1, X_2	First and second feature for PCC	–
$\bar{X}_1, \bar{X}_2$	Mean value of first and second feature for PCC	–
AB	As-built specimens	
ACC	Accuracy	
AM	Additive manufacturing	
CV	Cross-validation	
FN, FP	False negative and false positive predictions	
HT	Heat-treated specimens	
L-PBF	Laser-based powder bed fusion	
MCC	Matthews correlation coefficient	
ML	Machine learning	
PCC	Pearson correlation coefficient	
RMSE	Root-mean-square error	
SEM	Scanning electron microscope	
THT	Turned and heat-treated specimens	
TN, TP	True negative and true positive predictions	
XAI	Explainable artificial intelligence	
XGBoost	Extreme gradient boosting	

Nomenclature and abbreviations

See Table A.6 for the nomenclature and abbreviations.

References

[1] ISO/ASTM 52900:2021 Additive manufacturing—General principles—Fundamentals and vocabulary. ASTM International Geneva, Switzerland; 2021.

[2] Wong KV, Hernandez A. A review of additive manufacturing. Int sch res notices 2012;2012.

[3] Avanzini A. Fatigue behavior of additively manufactured stainless steel 316L. Mater 2023;16(1).

[4] Obeidi MA. Metal additive manufacturing by laser-powder bed fusion: Guidelines for process optimisation. Results Eng 2022;15:100473.

[5] Antony K, Arivazhagan N, Senthilkumaran K. Numerical and experimental investigations on laser melting of stainless steel 316L metal powders. J Manuf Process 2014;16(3):345–55.

[6] Shieu FS, Deng MJ, Lin SH. Microstructure and corrosion resistance of a type 316L stainless steel. Corros Sci 1998;40(8):1267–79.

[7] D'Andrea D. Additive manufacturing of AISI 316L stainless steel: A review. Metals 2023;13(8):1370.

[8] Romero A, García I, Arenas M, López V, Vázquez A. High melting point metals welding by concentrated solar energy. Sol Energy 2013;95:131–43.

[9] Bedmar J, Riquelme A, Rodrigo P, Torres B, Rams J. Comparison of different additive manufacturing methods for 316l stainless steel. Mater 2021;14(21):6504.

[10] Frazier WE. Metal additive manufacturing: a review. J Mater Eng Perform 2014;23:1917–28.

[11] Hauser T, Reisch RT, Breese PP, Lutz BS, Pantano M, Nalam Y, et al. Porosity in wire arc additive manufacturing of aluminium alloys. Addit Manuf 2021;41:101993.

[12] Cao L, Li J, Hu J, Liu H, Wu Y, Zhou Q. Optimization of surface roughness and dimensional accuracy in LPBF additive manufacturing. Opt Laser Technol 2021;142:107246.

[13] Poudel A, Yasin MS, Ye J, Liu J, Vinel A, Shao S, et al. Feature-based volumetric defect classification in metal additive manufacturing. Nature Commun 2022;13(1):6369.

[14] Mugwagwa L, Yadroitsava I, Makoana NW, Yadroitsev I. Residual stress in laser powder bed fusion. In: Fundamentals of laser powder bed fusion of metals. Elsevier; 2021, p. 245–76.

[15] Yang Y, Jamshidinia M, Boulware P, Kelly SM. Prediction of microstructure, residual stress, and deformation in laser powder bed fusion process. Comput Mech 2018;61:599–615.

[16] Sausto F, Carrion PE, Shamsaei N, Beretta S. Fatigue failure mechanisms for AlSi10Mg manufactured by L-PBF under axial and torsional loads: The role of defects and residual stresses. Int J Fatigue 2022;162:106903.

[17] Braun M, Mayer E, Kryukov I, Wolf C, Böhm S, Taghipour A, et al. Fatigue strength of PBF-LB/M and wrought 316L stainless steel: effect of post-treatment and cyclic mean stress. Fatigue Fract Eng Mater Struct 2021;44(11):3077–93.

[18] Chao Q, Thomas S, Birbilis N, Cizek P, Hodgson PD, Fabijanic D. The effect of post-processing heat treatment on the microstructure, residual stress and mechanical properties of selective laser melted 316L stainless steel. Mater Sci Eng A 2021;821:141611.

[19] Wang C, Loh YM, Cheung CF, Liang X, Zhang Z, Ho LT. Post processing of additively manufactured 316L stainless steel by multi-jet polishing method. J Mater Res Technol 2023;23:530–50.

[20] Elangeswaran C, Cutolo A, Gallas S, Dinh TD, Lammens N, Erdelyi H, et al. Predicting fatigue life of metal LPBF components by combining a large fatigue database for different sample conditions with novel simulation strategies. Addit Manuf 2022;50:102570.

[21] Wang Z, Wang Y, Gu H, Liao X. Study on fatigue behavior and crack propagation life of longitudinal fillet welded lap joint. Int J Fatigue 2023;167:107301.

[22] Allenov DG, Borisovna KD, Ghorbani S, Reza Kashyzadeh K. Simultaneous effects of cutting depth and tool overhang on the vibration behavior of cutting tool and high-cycle fatigue behavior of product: Experimental research on the turning machine. Int J Adv Manuf 2022;122(5–6):2361–78.

[23] Zamani P, Jaamialahmadi A, Da Silva LFM. Fatigue life evaluation of al-GFRP bonded lap joints under four-point bending using strain-life criteria. Int J Adhes 2023;103338.

[24] Pandian K, Neikter M, Bahbou F, Ganvir A, Hansson T, Pederson R. Fatigue behavior of low-temperature hot isostatic pressed electron beam powder bed fusion manufactured Ti-6Al-4V. J Alloys Compd 2023;171086.

[25] Steimbreger C, Gubeljak N, Vuherer T, Enzinger N, Ernst W, Chapetti M. Effect of welding processes on the fatigue behaviour of ultra-high strength steel butt-welded joints. Eng Fract Mech 2022;275:108845.

[26] Su H, Wang J, Du J. Experimental and numerical study of fatigue behavior of bridge weathering steel Q345qDNH. J Constr Steel Res 2019;161:86–97.

[27] Nourian-Avval A, Fatemi A. Fatigue life prediction of cast aluminum alloy based on porosity characteristics. Theor Appl Fract Mech 2020;109:102774.

[28] Wang H, Li B, Gong J, Xuan F-Z. Machine learning-based fatigue life prediction of metal materials: Perspectives of physics-informed and data-driven hybrid methods. Eng Fract Mech 2023;109242.

[29] Bao H, Wu S, Wu Z, Kang G, Peng X, Withers PJ. A machine-learning fatigue life prediction approach of additively manufactured metals. Eng Fract Mech 2021;242:107508.

[30] Zhan Z, Li H. Machine learning based fatigue life prediction with effects of additive manufacturing process parameters for printed SS 316L. Int J Fatigue 2021;142:105941.

[31] Wang H, Li B, Xuan F-Z. Fatigue-life prediction of additively manufactured metals by continuous damage mechanics (CDM)-informed machine learning with sensitive features. Int J Fatigue 2022;164:107147.

[32] Li J, Yang Z, Qian G, Berto F. Machine learning based very-high-cycle fatigue life prediction of Ti-6Al-4V alloy fabricated by selective laser melting. Int J Fatigue 2022;158:106764.

[33] Zhang M, Sun C-N, Zhang X, Goh PC, Wei J, Hardacre D, et al. High cycle fatigue life prediction of laser additive manufactured stainless steel: A machine learning approach. Int J Fatigue 2019;128:105194.

[34] Ruiz E, Ferreño D, Cuartas M, Arroyo B, Carrascal IA, Rivas I, et al. Application of machine learning algorithms for the optimization of the fabrication process of steel springs to improve their fatigue performance. Int J Fatigue 2022;159:106785.

[35] Peng X, Wu S, Qian W, Bao J, Hu Y, Zhan Z, et al. The potency of defects on fatigue of additively manufactured metals. Int J Mech Sci 2022;221:107185.

[36] Maleki E, Bagherifard S, Guagliano M. Correlation of residual stress, hardness and surface roughness with crack initiation and fatigue strength of surface treated additive manufactured AlSi10Mg: Experimental and machine learning approaches. J Mater Res Technol 2023;24:3265–83.

[37] Maleki E, Bagherifard S, Razavi S, Bandini M, du Plessis A, Berto F, et al. On the efficiency of machine learning for fatigue assessment of post-processed additively manufactured AlSi10Mg. Int J Fatigue 2022;160:106841.

[38] Li A, Baig S, Liu J, Shao S, Shamsaei N. Defect criticality analysis on fatigue life of L-PBF 17-4 PH stainless steel via machine learning. Int J Fatigue 2022;163:107018.

[39] He L, Wang Z, Ogawa Y, Akebono H, Sugeta A, Hayashi Y. Machine-learning-based investigation into the effect of defect/inclusion on fatigue behavior in steels. Int J Fatigue 2022;155:106597.
[40] Lubo-Robles D, Devegowda D, Jayaram V, Bedle H, Marfurt KJ, Pranter MJ. Machine learning model interpretability using SHAP values: Application to a seismic facies classification task. In: SEG international exposition and annual meeting. SEG; 2020, D021S008R006.
[41] Azadi M, Matin M. Shapley additive explanation on machine learning predictions of fatigue lifetimes in piston aluminum alloys under different manufacturing and loading conditions. Frattura ed Integrità Strutturale 2024;18(68):357–70.
[42] Braun M, Kellner L. Comparison of machine learning and stress concentration factors-based fatigue failure prediction in small-scale butt-welded joints. Fatigue Fract Eng Mater Struct 2022;45(11):3403–17.
[43] Braun M, Kellner L, Schreiber S, Ehlers S. Prediction of fatigue failure in small-scale butt-welded joints with explainable machine learning. Procedia Struct Integr 2022;38:182–91.
[44] Blinn B, Krebs F, Ley M, Teutsch R, Beck T. Determination of the influence of a stress-relief heat treatment and additively manufactured surface on the fatigue behavior of selectively laser melted AISI 316L by using efficient short-time procedures. Int J Fatigue 2020;131:105301.
[45] Tekin A, Martin J, Senior B. Grain boundary sensitization and desensitization during the ageing of 316L (N) austenitic stainless steels. J Mater Sci 1991;26:2458–66.
[46] Murakami Y. Material defects as the basis of fatigue design. Int J Fatigue 2012;41:2–10.
[47] Cui P, Athey S. Stable learning establishes some common ground between causal inference and machine learning. Nat. Mach. Intell. 2022;4(2):110–5.
[48] Hensel J. Mean stress correction in fatigue design under consideration of welding residual stress. WELD WORLD 2020;64(3):535–44.
[49] Chen J, Liu Y. Fatigue modeling using neural networks: A comprehensive review. Fatigue Fract Eng Mater Struct 2022;45(4):945–79.
[50] Lian Z, Li M, Lu W. Fatigue life prediction of aluminum alloy via knowledge-based machine learning. Int J Fatigue 2022;157:106716.
[51] Priyam A, Abhijeeta GR, Rathee A, Srivastava S. Comparative analysis of decision tree classification algorithms. Int J Curr Eng Technol 2013;3(2):334–7.
[52] Song Y-Y, Ying L. Decision tree methods: applications for classification and prediction. Shanghai Arch Psychiatry 2015;27(2):130.
[53] Dietterich TG. An experimental comparison of three methods for constructing ensembles of decision trees: Bagging, boosting, and randomization. Mach Learn 2000;40:139–57.
[54] Friedman J, Hastie T, Tibshirani R. Additive logistic regression: a statistical view of boosting (with discussion and a rejoinder by the authors). Ann Statist 2000;28(2):337–407.
[55] Qin C, Zhang Y, Bao F, Zhang C, Liu P, Liu P. XGBoost optimized by adaptive particle swarm optimization for credit scoring. Math Probl Eng 2021;2021:1–18.
[56] Xie P, Li Y, Deng B, Du C, Rui S, Deng W, et al. An explainable machine learning model for predicting in-hospital amputation rate of patients with diabetic foot ulcer. Int Wound J 2022;19(4):910–8.
[57] Lundberg SM, Erion G, Chen H, DeGrave A, Prutkin JM, Nair B, et al. From local explanations to global understanding with explainable AI for trees. Nat Mach Intell 2020;2(1):56–67.
[58] Zhang R, Li B, Jiao B. Application of XGboost algorithm in bearing fault diagnosis. In: IOP conference series: materials science and engineering. 490, IOP Publishing; 2019, 072062.
[59] Sharma N, Anju, Juneja A. Extreme gradient boosting with squared logistic loss function. In: Machine intelligence and signal analysis. Springer; 2019, p. 313–22.
[60] Chen T, Guestrin C. Xgboost: A scalable tree boosting system. In: Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining. 2016, p. 785–94.
[61] Konda N, Verma R, Jayaganthan R. Estimation of high cycle fatigue life of additively manufactured Ti6Al4V using data analytics. Procedia Struct Integr 2023;46:87–93.
[62] Chicco D, Jurman G. The advantages of the matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation. BMC Genom 2020;21(1):1–13.
[63] Stoica P, Babu P. Pearson-matthews correlation coefficients for binary and multinary classification and hypothesis testing. 2023, arXiv preprint arXiv:2305.05974.
[64] Lee JA, Sagong MJ, Jung J, Kim ES, Kim HS. Explainable machine learning for understanding and predicting geometry and defect types in Fe-Ni alloys fabricated by laser metal deposition additive manufacturing. J Mater Res Technol 2023;22:413–23.
[65] Lai W-J, Ojha A, Li Z, Engler-Pinto C, Su X. Effect of residual stress on fatigue strength of 316L stainless steel produced by laser powder bed fusion process. Prog Addit Manuf 2021;1–9.
[66] Pedersen MM. Introduction to metal fatigue, and no.:. Tech Rep ME-TR-11 2018;91.
[67] Ying X. An overview of overfitting and its solutions. In: Journal of physics: conference series. 1168, IOP Publishing; 2019, 022022.
[68] Subramanian J, Simon R. Overfitting in prediction models–is it a problem only in high dimensions? Contemporary clinical trials 2013;36(2):636–41.
[69] Jurman G, Riccadonna S, Furlanello C. A comparison of MCC and CEN error measures in multi-class prediction. PLOS ONE 2012;7(8):e41882.
[70] Solberg K, Guan S, Razavi SMJ, Welo T, Chan KC, Berto F. Fatigue of additively manufactured 316L stainless steel: The influence of porosity and surface roughness. Fatigue Fract Eng Mater Struct 2019;42(9):2043–52.
[71] Niu X, Zhu S-P, He J-C, Liao D, Correia JA, Berto F, et al. Defect tolerant fatigue assessment of AM materials: Size effect and probabilistic prospects. Int J Fatigue 2022;160:106884.
[72] Awd MMM. Machine Learning Algorithm for Fatigue Fields in Additive Manufacturing. Springer Nature; 2023.