A generalized machine learning framework to estimate fatigue life across materials with minimal data

Srinivasan, Dharun Vadugappatty; Moradi, Morteza; Komninos, Panagiotis; Zarouchas, Dimitrios; Vassilopoulos, Anastasios P.

DOI
10.1016/j.matdes.2024.113355

Publication date
2024

Document Version
Final published version

Published in
Materials and Design

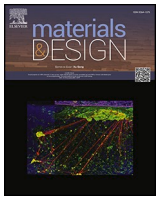
Citation (APA)
Srinivasan, D. V., Moradi, M., Komninos, P., Zarouchas, D., & Vassilopoulos, A. P. (2024). A generalized machine learning framework to estimate fatigue life across materials with minimal data. *Materials and Design*, *246*, Article 113355. https://doi.org/10.1016/j.matdes.2024.113355

Materials & Design 246 (2024) 113355

Contents lists available at ScienceDirect

Materials & Design

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



A generalized machine learning framework to estimate fatigue life across materials with minimal data

Dharun Vadugappatty Srinivasan [a,*], Morteza Moradi [b], Panagiotis Komninos [b], Dimitrios Zarouchas [b], Anastasios P. Vassilopoulos [a]

[a] Composite Mechanics Group (GR-MeC), Ecole Polytechnique Fédérale de Lausanne (EPFL), Station 16, CH-1015 Lausanne, Switzerland
[b] Center of Excellence in Artificial Intelligence for Structures, Prognostics & Health Management, Aerospace Engineering Faculty, Delft University of Technology, Kluyverweg 1, Delft 2629 HS, the Netherlands

ARTICLE INFO

Keywords:
Composites
Fatigue
Metal alloys
Machine learning
Void
Minimal data

ABSTRACT

In this research, a generalized machine learning (ML) framework is proposed to estimate the fatigue life of epoxy polymers and additively manufactured AlSi10Mg alloy materials, leveraging their failure surface void characteristics. An extreme gradient boosting algorithm-based ML framework encompassing Synthetic Minority Oversampling TEchnique (SMOTE), categorical data encoding, and external loop cross-validation is developed to evaluate the fatigue life across materials. The influence of different training strategies based on materials, input features, encoding method, and data standardization on the model performance is explored. Additionally, the importance of anti-data-leakage and anti-overfitting measures over the ML model performance is addressed. The result shows that the data-leakage-free, external loop cross-validated model can estimate the fatigue life of selective epoxy polymers and metal alloys with an average R^2 of 0.71 ± 0.06 using a mere 12 to 27 experimental data points per material category. Whereas the model trained with data-leakage and overfitting results in high R^2 of 0.9.

1. Introduction

Material defects are primarily introduced during manufacturing and assembly processes such as casting, additive manufacturing, sintering, welding, and adhesive bonding. For instance, manual manufacturing of epoxy materials results in more and bigger voids as compared with machine-assisted manufacturing techniques [1]. In metals, lack of fusion, porosities, shrinkage, and inclusions are the common intrinsic defects that are detrimental to fatigue life [2]. These defects can act as sources of fatigue crack initiation leading to the development of large cracks over the component's lifetime, ultimately leading to catastrophic failure. By understanding the influence of voids and their unique characteristics, the structural integrity of components can be continuously monitored and maintained to prevent complete failure.

To evaluate the fatigue performance of materials, specimens of different sizes and geometries such as dog bone specimens are typically subjected to specified stress levels, load ratios, frequencies, environments, and geometries [3–7]. The resulting stress level (S) and the number of cycles to failure (N) data are then used to plot the S-N curves. Materials tested under the same geometry and loading conditions provide scattered fatigue life that can be primarily attributed to material defects, including voids [6]. When the size of the defect in the crack initiation point is correlated with the life, the scattering is greatly reduced [8]. Several studies have focused on investigating the effects of defects and notches on fatigue life, e.g., [9,10] showing that the defects or notches decrease the effective load-bearing cross-section and increase the stress concentration factor, resulting in higher local stress in the material than the applied stress. In epoxy polymers, the fatigue crack initiation phase which accounts originally for 60 % to 90 % of the total lifetime is shortened by the notches [10].

In addition to the experimental investigations, analytical models such as Murukami's two variables (defect size and applied stress) and Z's three variable parameters (defect size, location, and applied stress) have been used to correlate the defect features with fatigue life. The fatigue limit of a material is also influenced by the defect type and morphology [11]. The 3-dimensional (3D) and 2-dimensional (2D) characteristics of the inherent defects or voids can be obtained by computed microtomography [1] and microscopies, respectively. Following the fatigue

* Corresponding author.
E-mail address: dharun.srinivasan@epfl.ch (D.V. Srinivasan).

https://doi.org/10.1016/j.matdes.2024.113355
Received 10 May 2024; Received in revised form 22 September 2024; Accepted 29 September 2024
Available online 2 October 2024

failure, the scanning electron microscopic (SEM) images of the fracture surface are taken and the 2D void data [6] is retrieved by using an image analysis software (Fiji) [12] or interactive machine learning programs (ilastik) [13].

Fatigue is a complex phenomenon that is influenced by multiple factors, making it challenging to comprehend the interplay between these factors using either analytical models or experimental results[14] Moreover, conducting fatigue experiments is a time-consuming, laborious, and expensive task. The use of data-driven approaches for estimating the fatigue life of materials, particularly in metal alloys [9,15–33] and fiber-reinforced polymeric composites [34–39] has gained significant attention in the last two decades due to its potential to overcome the limitations of traditional analytical and experimental methods. Machine learning (ML) models including artificial neural networks (ANN) [15–17,21,26,31,34,35], deep neural networks (DNN) [29], support vector regression (SVR) [18,21,29–31,40–43], multi-graph attention networks (Multi-GAT) [28], random-forest regression (RFR) [21,40,43], kernal ridge regression [27], gaussian process regression [43], physics informed neural networks (PINNS) [9] and gradient boost regression (XGBoost) [23,37], have been utilized in fatigue life estimation. These models aim to estimate the correlation between different factors such as microstructure-fatigue crack driving force [44] and defect features to fatigue life [9,18,20,23,25].

For training and validating the ML models, the data points obtained through fatigue experiments are small due to the limited number of specimens, and in some cases, additional data points were created with virtual sample generators such as Monte Carlo simulations, linear interpolation (LI), nearest neighbor interpolation (NNI), and linear interpolation with the Gaussian mixture model (LIGMM). For example, Monte Carlo simulations were used to increase the original experimental data points from 22 to an additional 32,800 points for training an ANN model [26]. In another study [32], the experimental data points of 45 were augmented to 615 using LIGMM for training ANN, RFR, and SVR models, with RFR providing the best R^2. Typically, ANN is more suitable for a large dataset where extrapolation of the available data points is possible and for small datasets, the XGBoost model estimates the fatigue life very well [23]. XGBoost model is also used in evaluating the influence of various chemical components and processing parameters on the tensile strength and plasticity of the steel [45] and its accuracy is better than the linear regression models, K- nearest neighbors (KNN), decision trees, and gradient boosting regression models. Despite the success of these approaches, there are still several research gaps that need to be addressed.

Recent studies [21–23,37] reporting ML models with high R^2 ($>$0.80) values could be potentially misleading because either these models were not cross-validated or prone to data-leakage [46]. This is a serious challenge since the model could be overfitted for a selective test dataset and often fails to generalize. Therefore, the common pitfalls [47] need to be addressed while developing a robust and reliable ML framework. While existing studies have used ML models to correlate fatigue life with various factors such as microstructure, defect features, and crack driving force, most of these models are limited to specific material systems and cannot be easily extended to other materials with different properties. Furthermore, the fatigue life of additively manufactured metallic materials is influenced by factors such as printing direction and post-processing techniques, resulting in distinct S-N curves. Despite these variations, these materials have been treated as a single class in the development of ML models, without a framework for evaluating fatigue life across different S-N curves or materials. However, recent studies [48,49] show the potential for generalizing ANN models to extend predictions for various materials and fatigue scenarios. To the best of the authors' knowledge, there is no framework developed for estimating the wide range of fatigue life (10^2 to 10^6 cycles) (i) with minimal dataset, (ii) across different materials and (iii) with anti-data-leakage and anti-overfitting strategies. In ML-based fatigue life prediction, the minimal dataset can comprise 10 to 30 fatigue data points.

To tackle these challenges, an ML approach based on the extreme gradient boosting algorithm is introduced in this research, incorporating the Synthetic Minority Over-sampling Technique (SMOTE), categorical data encoding, and external loop cross-validation techniques. Its purpose is to estimate the fatigue life of distinct epoxy polymers based on the fatigue failure surface void features, using relatively limited experimental data points while maintaining its generalizability.

Expanding its applicability, the developed ML framework is subsequently applied to a combined dataset encompassing epoxy polymer and additively manufactured AlSi10Mg alloy data from literature [23]. The model's performance is evaluated using established metrics like error bands, coefficient of determination (R^2) and mean squared error (*MSE*). Notably, this framework exhibits remarkable versatility, enabling its utilization across an array of materials, even when only small datasets are available. Moreover, it showcases the ability to fairly estimate fatigue life over a wide spectrum of cycles, spanning from 10^2 to 10^6 cycles. The data-leakage and overfitting sources are identified and their effect on model performance is investigated with simplistic models. The ML model can be employed to construct reliable S-N curves without excluding experimental data from specimens containing voids.

2. Methodology

The proposed ML framework for estimating the fatigue life of materials is depicted in Fig. 1. The first step was to collect the raw data consisting of the void characteristics of the epoxy materials [6,7] and metal alloy specimens [23] after fatigue failure. Subsequently, outliers within the raw dataset were removed and the dataset was partitioned into training and test datasets. The training dataset undergoes four preprocessing steps (i) synthetic sampling, (ii) data standardization, (iii) data encoding and (iv) feature selection through RFR, before training the XGBoost regression model. These preprocessing steps assist the framework in handling datasets comprising multiple materials and categorical data types, while concurrently reducing dimensionality for improved fatigue life estimation. Ultimately, the model's hyperparameters were tuned, and cross-validation procedures were conducted to ensure its generalizability. The pre-processing techniques and ML training procedures used in this framework are summarized in the following sections.

2.1. Data collection

Five different structural epoxy materials (BB, BT, TB, TT and SD) used in wind turbine blade and construction applications are considered in this study. BB epoxy (SPABOND™ 820HTA) and TT (SPABOND™ 840HTA) epoxy materials were provided by Gurit (UK) Ltd. SD epoxy (Sikadur-330) was supplied by SIKA Schweiz AG, Switzerland.

BB epoxy mainly contains short-glass fiber fillers, while TT is formulated with core–shell rubber particles. Both materials contain sag-resistance fillers to maintain a high-thickness bond line in wind turbine rotor blades. Additionally, two hybrid epoxy materials such as BT and TB were developed by combining BB and TT epoxy materials at weight ratios of 75:25 and 50:50, respectively. The epoxy base and hardener materials (100:34) were manually mixed with a wooden spatula and degassed under vacuum for 5 min before applying them inside the mold cavity. The mixed epoxy material was cured inside a convection oven at 70 °C for 2 hrs.

SD epoxy is also a thixotropic two-component epoxy adhesive, mainly used in construction applications. The base and hardener of SD epoxy were mixed at a weight ratio of 4:1, degassed and applied inside the mold. After 7 days of room temperature curing (20 ± 1°C), the epoxy panel was de-molded and post-cured at 90 °C for 30 min. All five different epoxy material panels were cut into ASTM D638-22 Type II specimens using a water jet cutting technique. Aluminum end tabs having a thickness of 2 mm were bonded at the specimen gripping areas

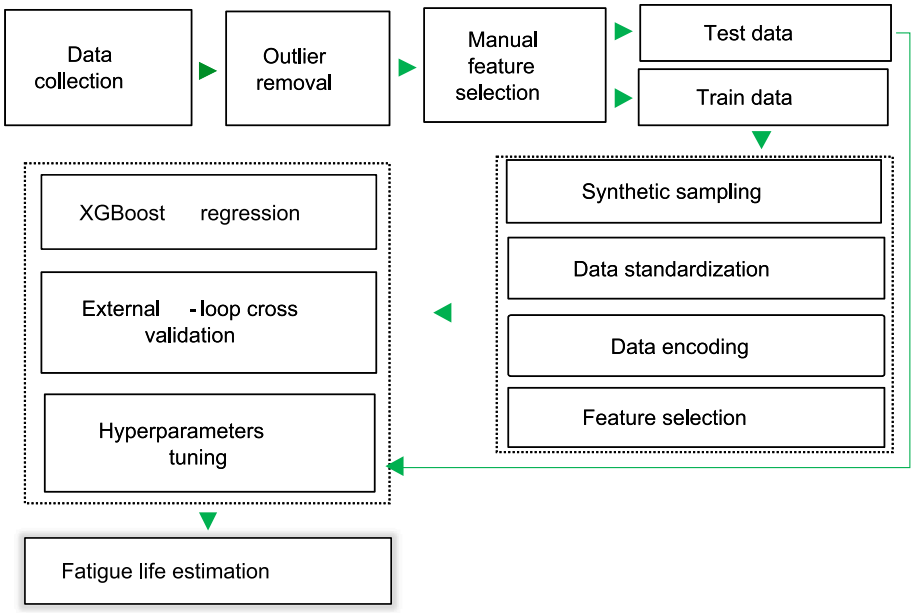


Fig. 1. Proposed ML framework based on XGBoost model.

to avoid any slippage during fatigue experiments. Table 1 shows the tensile properties of the investigated epoxy materials.

MTS® 810 Landmark servo-hydraulic (5 kN load cell) and MTS® Acumen electro-mechanical machines (3 kN load cell) were used to perform the tensile-tensile fatigue experiments of ASTM D638-22 Type II epoxy specimens. The fatigue experiments were load-controlled and conducted at four different stress levels to cover a wide range of fatigue lives spanning from 10^2 to 10^6 cycles. The fatigue loading frequency of 10 Hz and stress ratio of 0.1 was maintained for all specimens. The fatigue experiments were conducted at a controlled ambient temperature of 23 ± 2°C and relative humidity of 40 ± 10 %. The quasi-static and fatigue performance of the epoxy materials were comprehensively discussed in [6,7]. The S-N curve slopes of these materials were similar (0.075), there was no notable increase in the hysteresis energy and the stiffness degradation was between 1 % to 6 % depending on the applied load level.

In the previous works [6,7], the fatigue life of specimens with premature failure due to manufacturing defects was discarded to construct reliable S-N curves. In contrast, this research considers the fatigue life of all epoxy materials, especially with large defects on the failure surface (refer to Fig. 2). There is a total of 81 experimental data points in the five-epoxy materials category and the number of data points in each material category is also mentioned in Fig. 2.

Following the fatigue failure, a scanning electron microscope (SEM), ZEISS Gemini SEM 300 was used to capture images of the fractured surfaces of 81 specimens (Fig. 3a). As voids formed due to air entrapment during the manufacturing process, the void shapes were predominantly elliptical with smooth contours. To determine the void characteristics, SEM images were analyzed using the open-source software ilastik [13] (Fig. 3b). The "pixel classification" workflow was employed to label SEM image pixels based on pixel features such as

Table 1
Averaged tensile properties of epoxies.

Property	BB [6]	BT [6]	TB [6]	TT [6]	SD [7]
Modulus (GPa)	5.9	4.9	4.3	2.9	4.1
Ultimate strength (MPa)	71.4	62.4	55.1	45.7	48.2
Failure strain (mm/mm)	0.0146	0.0159	0.0181	0.0422	0.0146

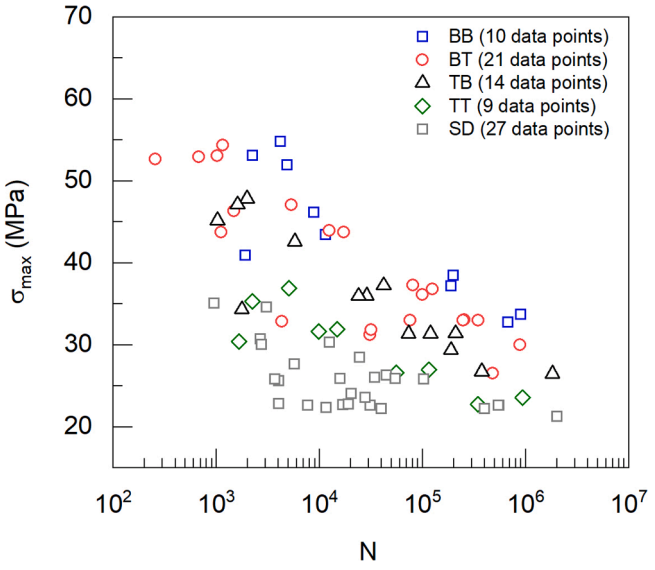


Fig. 2. Applied maximum stress versus cycles to failure of epoxy polymers.

smoothed pixel intensity, edge filters, and texture descriptors. The software has an in-built Random Forest Classification algorithm that was trained at least 16 images with selected pixel features and user annotations. The algorithm was able to classify the annotations such as epoxy material (black), void (grey), and specimen edges (white). The exterior areas were manually removed using Fiji software [12] to collect void data only from the specimen cross-section. 'Fill holes', 'erode', 'dilate', and 'analyze particles functions' were used to identify voids, remove noise, and automatically calculate void parameters such as area, perimeter, axes of the fitted ellipse, and angle of the ellipse based on the pixel scale of the images (Fig. 3c). Additionally, various features such as angle, size, aspect ratio, jaggedness, and location were calculated for the critical void (Fig. 3d).

Void size (S): This parameter is obtained by taking the square root of the critical void area. The minimum detectable void area was on the order of $10^{-4}\ mm^2$.

Aspect ratio (AR): This parameter is provided by the ratio of the major

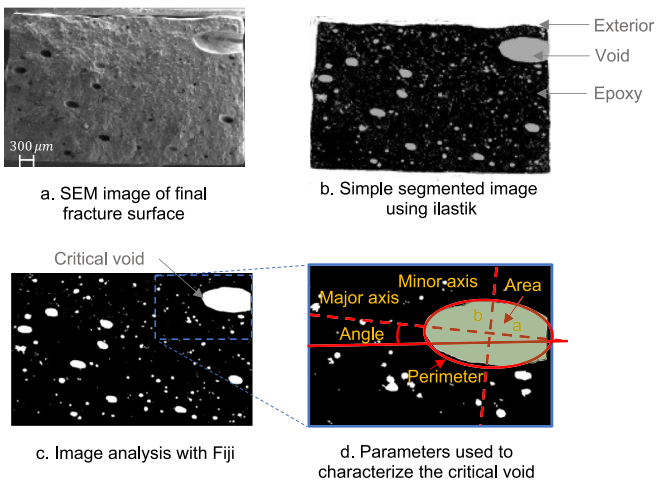


Fig. 3. Digital image analysis and void data collection.

(a) and minor axis (b) lengths of the fitted ellipse.

Jaggedness (J): This parameter also relates the void shape to an ellipse and is calculated using the formula $\frac{\sqrt{2}Perimeter}{2\pi\sqrt{a^2+b^2}}$.

Location (L): This parameter is obtained by manually selecting the void contour, measuring the distance between the void centroid and the specimen edge, and further dividing it by the void size.

Angle (A): This parameter refers to the angle between the major axis of the ellipse and the horizontal axis of the image and is provided in radians.

While previous studies only consider critical void features [23], other voids across the cross-section can interact with the fatigue crack and accelerate the final brittle failure [6]. Due to the limited dataset, secondary void features are not considered in this study. In this work, eight raw input features are considered (i) Material (M), (ii) Young's modulus (E), (iii) Maximum applied stress (σ_{max}), (iv) Aspect ratio (AR), (v) Size (S), (vi) Location (L), (vii) Jaggedness (J) and (viii) Angle (A). 'Material' is the only categorical data and the output feature is the fatigue life, also referred as cycles to failure. The distribution of the input and output features are shown in Fig. 4.

To demonstrate the generalizability of the developed framework for estimating the fatigue life of metallic alloys, a dataset of 27 points was considered from [23]. This dataset also focuses on the critical defect characteristics, along with the fatigue life of additively manufactured AlSi10Mg alloys. These alloys were fabricated using the laser powder bed fusion technique with two different printing directions: x and z. Notably, the z-direction printed alloy specimens (Alz) exhibited a higher amount of defects compared to those printed in the x-direction (Alx), consequently resulting in reduced tensile-tensile fatigue performance (refer to Fig. 5). Therefore, these printed metallic alloys were considered as two different materials.

As the epoxy polymers and the metal alloy datasets have the same raw input features, they can be consolidated to increase the data points as well as to evaluate the proposed framework's ability to predict fatigue life across different material categories.

2.2. Outlier removal and manual feature selection

In this exploratory data analysis, the Z-score technique was applied to remove extreme data points based on fatigue life, setting the score to ± 1.96, which corresponds to a 95 % confidence level. Following this data trimming, a manual feature selection process was implemented to segregate the required input and output features.

2.3. Synthetic minority over-sampling TEchnique (SMOTE) and data standardization

The combined dataset comprises five types of epoxy polymers (BB, BT, TB, TT, and SD) and two AlSi10Mg alloys (Alx and Alz). Among these, SD epoxy (designated as the 'majority class') encompasses 27 data points, whereas the TT epoxy (designated as the 'minority class') has only 9 data points, resulting in data imbalance. Data imbalance can negatively impact ML algorithms' estimation accuracy, i.e., algorithms tend to perform better on the majority class than on the minority class [48]. This bias was overcome by employing a data balancing technique called SMOTE. SMOTE involves generating synthetic samples of the minority material by selecting the five nearest neighbors to each minority material sample and introducing synthetic examples along the line segments connecting them. The SMOTE technique's pseudocode is available in [48]. For example, a randomly selected training dataset (orange bars in Fig. 6a) can be balanced with SMOTE data (green bars in Fig. 6a), thus the total number of training data points was also reasonably increased. As shown in Fig. 6b and 6c, the newly generated dataset almost has the original input data distribution.

Considering the broad spectrum of fatigue lives, which can range from a few hundred cycles to several tens of thousands of cycles, the fatigue life was transformed from the integer domain to the logarithmic domain using a log transformation. Further, the input features can be standardized (Z) as follows:

$$Z = \frac{X - \mu}{\sigma} \tag{1}$$

here, X, μ and σ denote the input feature value, mean and the corresponding standard deviation. Notably, SMOTE and data standardization were applied only to the training dataset to avoid any data-leakage. Subsequently, the calculated mean and standard deviation were considered to normalize the testing data.

2.4. Data encoding

The main objective is designing an ML framework for different materials (including epoxy polymers and metallic alloys) hence, the materials need to be regarded as categorical data that can be transformed into numerical values. For this purpose, one-hot encoding, ordinal encoding, target encoding, catboost encoding and count encoding are commonly used in ML [49]. In the one-hot encoding, each material category is represented as a binary vector with a length equal to the number of categories. The vector has only one element as 1 (also known as "hot" or "on"), which represents the corresponding category, and all other elements are 0 (also known as "cold" or "off"). The categorical variable "Material" with five categories/cardinality: BB, BT, TB, TT and SD can be represented as a binary vector using one-hot encoding, as shown in Fig. 7. With these additional encoded features, the total raw input features were increased from 8 to 12.

An anticipated limitation associated with employing one-hot encoding for features with a higher cardinality (>4) is the potential emergence of a high-dimensional feature space. This can subsequently lead to suboptimal model performance and high computational costs, as previously highlighted in [49]. To address this concern, an alternative approach, namely target encoding or mean encoding, was employed for both polymer and combined polymer-metallic alloy datasets. This method entails substituting each material category with the average of the target variable (log of cycles to failure) specific to that material category, as illustrated in Fig. 8. Notably, target encoding presents a more informative data representation compared to conventional one-hot encoding or ordinal encoding techniques [49]. Both encoding techniques were applied only to the training data portion to circumvent data-leakage.

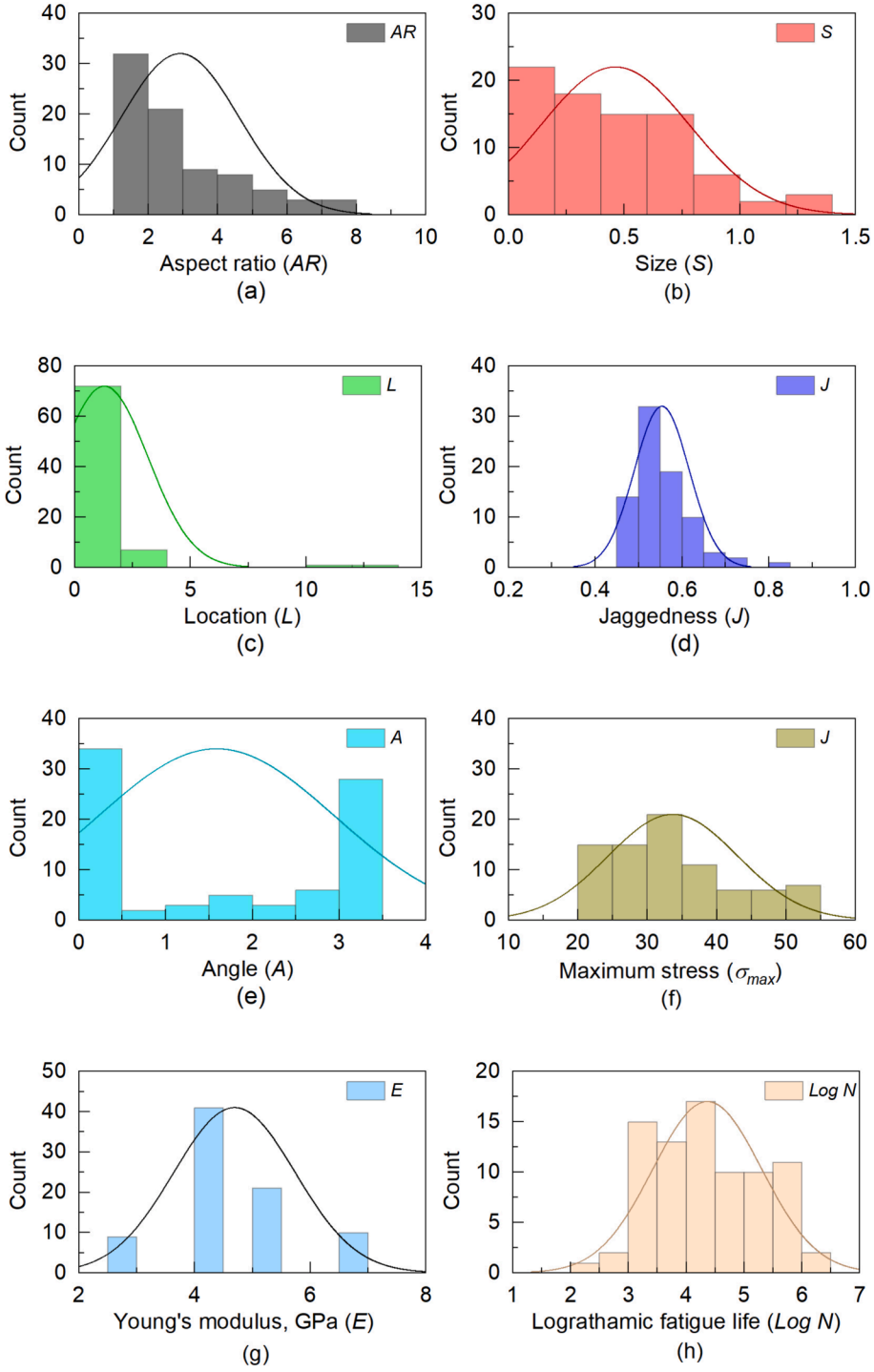


Fig. 4. Input and output features distribution of five epoxy polymers: (a) *AR*, (b) *S,* (c) *L*, (d), *J*, (e) *A*, (f) σ_{max} (g) *E*, and (h) *Log N*.

2.5. Pearson correlation coefficient (PCC)

Among the considered input features, some of them are highly correlated with each other which negatively impacts the performance of the model. Multicollinearity can cause problems when interpreting the regression coefficients because it becomes difficult to determine the independent effect of some input features on the output. The model may assign too much importance to the correlated input features, which may not be relevant to the fatigue life, or omit the importance of uncorrelated features to the fatigue life, resulting in poor performance. Correlated input features can also lead to increased computational complexity and training time due to the big values for the relevant coefficients. When any one of the features from the correlated pairs is removed, the model becomes simpler and easier to interpret, and the computational resources needed for training are reduced. In this regard, the correlation coefficient between any two features (*X1, X2*) can be identified utilizing their covariance (*cov* (*X1, X2*)) and standard deviations (σ_{X1}, σ_{X2}), in the form of the Pearson correlation coefficient as follows:

$$\rho(X1, X2) = \frac{cov(X1, X2)}{\sigma_{X1}\sigma_{X2}} \tag{2}$$

To understand the importance of the feature selection process,

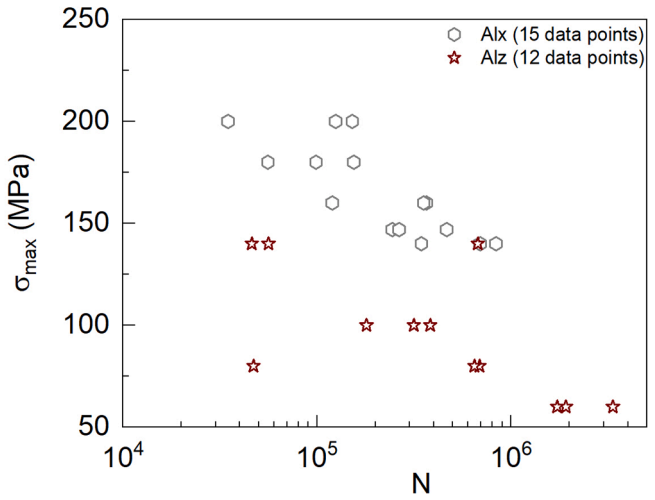


Fig. 5. Applied maximum stress versus cycles to failure of additively manufactured AlSi10Mg alloys. Replotted from [23].

features that have a PCC higher than 0.7 were removed for developing the ML model with "selective" features.

2.6. Feature importance by random forest regression

Feature engineering helps to identify the key input features that contribute the most to the accuracy of the model. Feature importance scores of all input features were determined by random forest regression and the working flow of the algorithm is shown in Fig. 9. Random forest algorithm constructs multiple decision trees (Tree 1, Tree 2, …, Tree *i*) using randomly selected subsets of the data and features to reduce the correlation between the decision trees. During the construction of each tree, the algorithm determines the importance of each input feature by measuring how much the tree nodes that use this feature decrease the *MSE* in the fatigue life. After constructing all the trees, the importance scores of each feature are averaged across all trees to obtain the final feature importance rankings.

2.7. Extreme gradient boosting regression

The main backbone of the proposed framework is the XGBoost (eXtreme Gradient Boosting) regression algorithm which is trained with different strategies for predicting fatigue life. It is a popular and powerful learning algorithm that utilizes a distributed gradient-boosted decision tree framework for predictive modeling [50]. As illustrated in Fig. 10, it is an ensemble learning algorithm that works by iteratively building decision trees that correct the errors of the previous tree and

Categorical feature

Material
BB
BT
TB
TT
SD

One-hot encoding →

Numerical data

BB	BT	TB	TT	SD
1	0	0	0	0
0	1	0	0	0
0	0	1	0	0
0	0	0	1	0
0	0	0	0	1

Fig. 7. One-hot encoding of epoxy polymer "Material" feature.

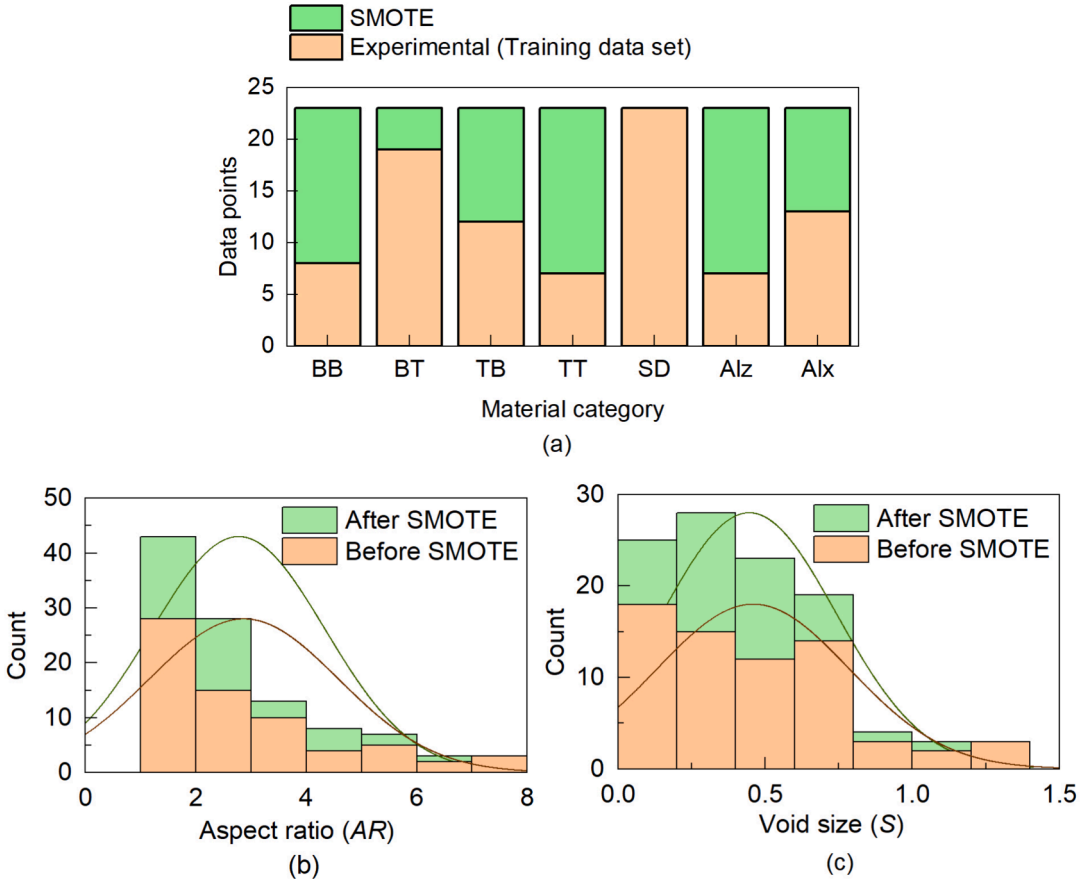


Fig. 6. SMOTE: (a) balancing the training data points among the materials, (b) distribution of *AR* and (c) distribution of *S*.

Categorical feature

S.No	Material	Log N
1	BB	1
2	BB	3
3	BB	5
4	BT	3
5	BT	4
6	BT	6
7	TB	4
8	TB	5
9	TB	6
10	TT	4
11	TT	5
12	SD	2
13	SD	3

→ Target encoding

Material	Log N mean	Encoding
BB	3	3
BT	4.33	4.33
TB	5	5
TT	4.5	4.5
SD	2.5	2.5

→ Numerical data

S.No	Material	Material
1	BB	3
2	BB	3
3	BB	3
4	BT	4.33
5	BT	4.33
6	BT	4.33
7	TB	5
8	TB	5
9	TB	5
10	TT	4.5
11	TT	4.5
12	SD	2.5
13	SD	2.5

Fig. 8. Working principle of a mean target encoding method (only for illustration purposes, not the actual data).

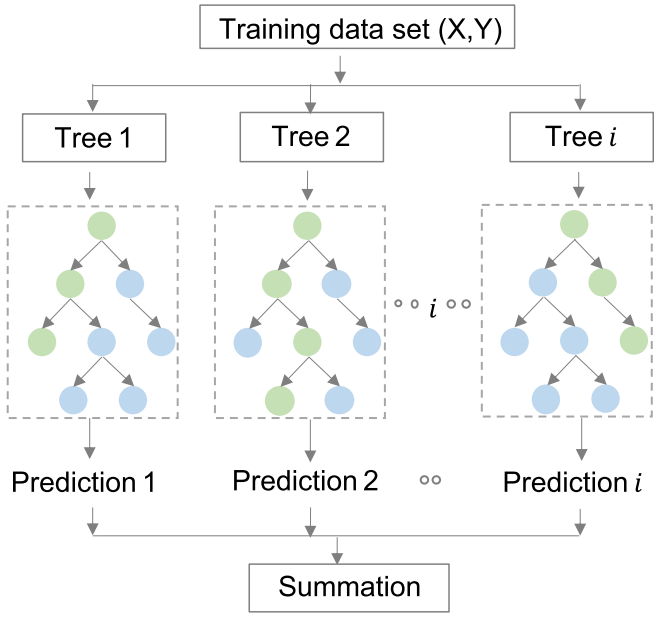


Fig. 9. General working flow of random forest algorithm.

can be regularized (α_i) to avoid overfitting. The process involves dividing the dataset into multiple subsets and training simple decision trees on each subset, with each tree providing a prediction for a specific subset of the data. The predictions of these weak learners are then aggregated to produce a final strong learner with better-predicting performance and less error.

Additionally, to address the problem of "over-specialization" where the decision trees added in subsequent iterations have a limited contribution to the performance of the model, particularly in the case of unseen data, the "Dropouts meet Multiple Additive Regression Trees (DART)" method was utilized [51].

2.8. External loop cross-validation and hyperparameter tuning

Overfitting poses a common challenge in ML models, prompting the application of cross-validation (CV) methods to validate the model's effectiveness [52]. As shown in Fig. 11, this research adopts the external loop, also referred to as nested cross-validation to overcome this challenge, as a technique to precisely anticipate model performance by synergizing the 5-fold outer and 3-fold inner loops. In the outer loop, model performance assessment unfolds through the partitioning of the dataset into a training set (80 %) and a distinct test set (20 %) via a stratified K-fold technique. The test dataset is only reserved for model performance evaluation without any preprocessing, while the training set (80 %) undergoes further division into diverse folds utilizing 3-fold cross-validation.

Within the inner loops, the hyperparameters such as maximum depth, learning rate, subsample, and the number of estimators were tuned with the "BayesSearchCV" algorithm [53]. The search space of these parameters along with the step size is provided in Table 2. This Bayesian optimization algorithm works by iteratively selecting hyperparameter configurations to evaluate based on a probabilistic model of the objective function's behavior across the hyperparameter space. In this case, the objective function is to minimize the negative root mean squared error. This approach efficiently explores the hyperparameter space, often requiring fewer evaluations compared to exhaustive search methods like grid search or random search.

Subsequently, leveraging these refined hyperparameters, the model is trained on the entire training set and subsequently appraised against the initially reserved test set. Aggregation of performance metrics from these 5 different outer loop evaluations yields a more precise and robust estimation of the model's ability to generalize. The average *MSE* and R^2 values of these five external loops are also reported to counter overfitting and establish a reliable evaluation framework for the ML models.

3. Results and discussion

3.1. Evaluation of ML framework for all epoxy polymers

To assess the influence of input feature selection, data standardiza-

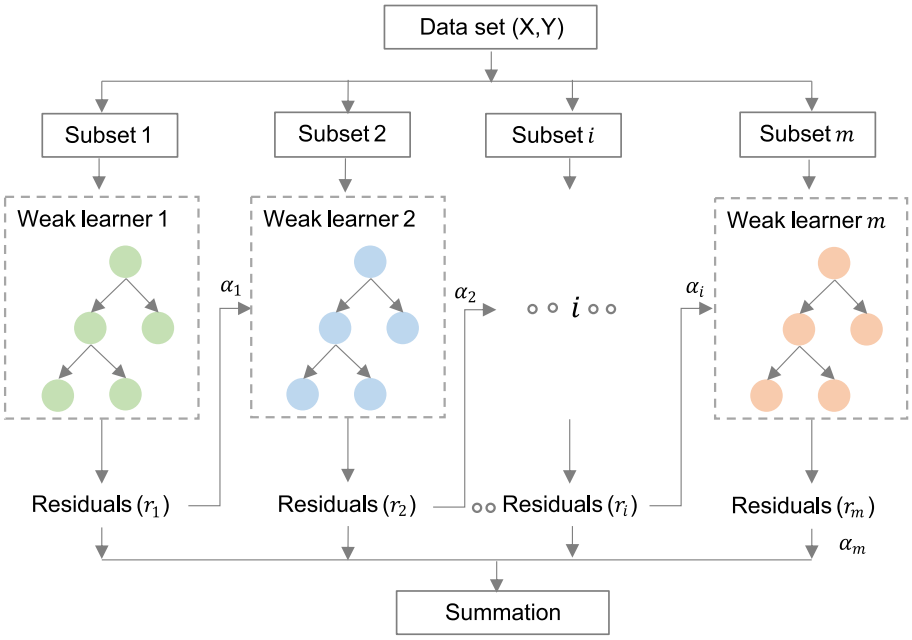


Fig. 10. General working flow of XGBoost algorithm.

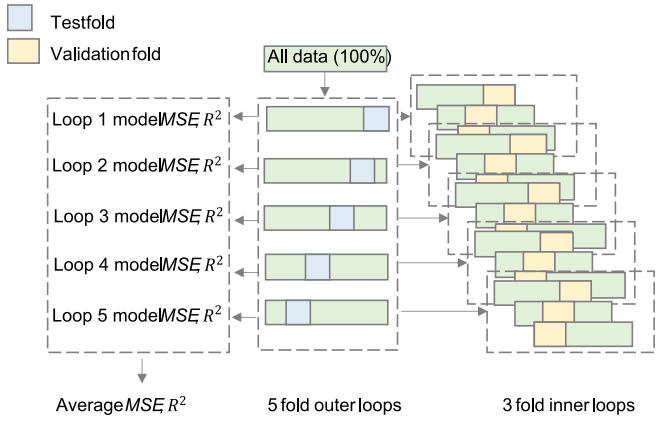


Fig. 11. External loop cross-validation.

Table 2
Search space for the hyperparameters.

Hyperparameters	Search space	Step size
Maximum depth	2–14	1
Learning rate	0.02–0.1	0.005
Subsample	0.5–0.7	0.05
Number of estimators	150–300	25

tion, and data encoding approaches on model performance, a set of eight distinct training strategies (E1, E2, E3, E4, E5, E6, E7 and E8) were constructed within this ML-based framework. 'E' refers to the input materials are all epoxy polymers (BB, BT, TB, TT and SD). The performance metrics of the external-loop cross-validated model are provided in Table 3. The average R^2 of the 5 outer loops with various training strategies varies between 0.38 and 0.44 with a high standard deviation. Although, one of the loops provides R^2 of 0.79, the other loops resulted in very low R^2 indicating the model is not robust. There is no clear evidence of the influence of different training strategies on the model's performance.

3.2. Evaluation of ML framework for metal alloys

Table 4 shows the performance metrics of the XGBoost model trained with a set of eight training strategies (M1, M2, M3, M4, M5, M6, M7 and M8) where 'M' refers to the input materials are two metal alloys (Alz and Alx). The average R^2 of the 5 outer loops with various training strategies varies between 0.04 and 0.06 with a high standard deviation. Although, the best R^2 is from 0.54 to 0.57, some of the loops resulted in negative R^2. This poor estimation of the fatigue life was due to insufficient representative data points in every external loop.

3.3. Evaluation of ML framework for all epoxy materials and metal alloys

Table 5 lists the performance metrics of the XGboost model trained with eight different strategies (EM1, EM2, EM3, EM4, EM5, EM6, EM7 and EM8) where 'EM' refers to the input materials consisting of all epoxy

Table 3
ML model based on epoxy materials data.

Training strategies	Input features	Standardization	Encoding	Average R^2	Average *MSE*	Best R^2	Best *MSE*
E1	"all"	✓	one-hot	0.44 ± 0.22	0.379 ± 0.168	0.76	0.162
E2	"all"	✓	target	0.38 ± 0.31	0.410 ± 0.203	0.79	0.140
E3	"all"	×	one-hot	0.44 ± 0.21	0.376 ± 0.152	0.76	0.162
E4	"all"	×	target	0.40 ± 0.31	0.401 ± 0.218	0.79	0.140
E5	"selective"	✓	one-hot	0.39 ± 0.24	0.410 ± 0.180	0.78	0.152
E6	"selective"	✓	target	0.38 ± 0.31	0.410 ± 0.203	0.79	0.140
E7	"selective"	×	one-hot	0.41 ± 0.25	0.396 ± 0.182	0.78	0.152
E8	"selective"	×	target	0.40 ± 0.31	0.401 ± 0.218	0.79	0.140

Table 4
ML model based on all metal alloy data.

Training strategies	Input features	Standardization	Encoding	Average R^2	Average *MSE*	Best R^2	Best *MSE*
M1	"all"	✓	one-hot	0.05 ± 0.39	0.173 ± 0.087	0.57	0.092
M2	"all"	✓	target	0.05 ± 0.39	0.173 ± 0.087	0.57	0.092
M3	"all"	×	one-hot	0.06 ± 0.39	0.168 ± 0.074	0.54	0.086
M4	"all"	×	target	0.06 ± 0.39	0.168 ± 0.074	0.54	0.086
M5	"selective"	✓	one-hot	0.04 ± 0.37	0.175 ± 0.084	0.56	0.106
M6	"selective"	✓	target	0.04 ± 0.37	0.175 ± 0.084	0.56	0.106
M7	"selective"	×	one-hot	0.06 ± 0.39	0.169 ± 0.075	0.54	0.086
M8	"selective"	×	target	0.06 ± 0.39	0.169 ± 0.075	0.54	0.086

Table 5
ML model based on all epoxy materials and metal alloys data.

Training strategies	Input features	Standardization	Encoding	Average R^2	Average *MSE*	Best R^2	Best *MSE*
EM1	"all"	✓	one-hot	0.54 ± 0.18	0.352 ± 0.089	0.74	0.263
EM2	"all"	✓	target	0.54 ± 0.14	0.352 ± 0.059	0.71	0.280
EM3	"all"	×	one-hot	0.54 ± 0.18	0.347 ± 0.089	0.74	0.259
EM4	"all"	×	target	0.55 ± 0.14	0.352 ± 0.060	0.71	0.272
EM5	"selective"	✓	one-hot	0.60 ± 0.17	0.301 ± 0.080	0.78	0.218
EM6	"selective"	✓	target	0.50 ± 0.20	0.382 ± 0.091	0.68	0.315
EM7	"selective"	×	one-hot	0.62 ± 0.17	0.290 ± 0.082	0.78	0.221
EM8	"selective"	×	target	0.49 ± 0.18	0.393 ± 0.085	0.68	0.329

and metal alloys. EM7 (selective, standardized and one-hot encoded) exhibits better performance with an average R^2 of 0.62 ± 0.17 than all other training strategies. The total number of data points was enhanced by combining the dataset of metals and polymers, leading to better prediction of the fatigue life than the model trained with only epoxy polymers or metal alloys. PINNs also handle minimal dataset [9], however it requires domain-specific knowledge and anti-overfitting measures due to the provided hard physical constraints.

3.4. Evaluation of ML framework for selective epoxy materials and all metal alloys

A set of eight different training strategies (SEM1, SEM2, SEM3, SEM4, SEM5, SEM6, SEM7 and SEM 8) were developed using a combined dataset encompassing selective epoxy polymers (BT, TB and SD) and additively manufactured AlSi10Mg alloys (Alz and Alx) and the importance of input feature selection, data standardization and data encoding approaches on model performance was assessed. The R^2 and *MSE* of the averaged loops and the best outer loop of the cross-validated model are provided in Table 6. Upon comprehensive comparison of the different training strategies, the XGBoost model, rooted in decision trees, exhibited less sensitivity to data standardization. The proposed ML framework demonstrates adeptness in handling input features with divergent scales, with or without resorting to standardization. In the context of data encoding techniques, predominantly one-hot encoding method yielded better performance than target encoding. When comparing the performance of models with all features and with selective features, the prior models have inferior performance. Overall, the SEM5 training strategy yielded a robust and best prediction (R^2 = 0.71 ± 0.06) and their tuned hyperparameters for each loop are listed in Table 7. As compared with EM7 strategy (all epoxy and metal data), SEM5 provides a substantial improvement (14.5 %) in R^2 due to the elimination of two epoxy materials with insufficient representative data points. Therefore, essential data points are very critical for the model to learn and evaluate. When SMOTE was not incorporated into the SEM5 training strategy, the model performance significantly decreased (R^2 = 0.54 ± 0.09) due to data imbalance.

Table 7
Tuned hyperparameters of SEM5 XGBoost model.

Tuned hyperparameters	Maximum depth	Learning rate	Subsample	Number of estimators
Loop 1	8	0.045	0.65	200
Loop 2	3	0.07	0.65	200
Loop 3	12	0.05	0.65	200
Loop 4	8	0.095	0.5	275
Loop 5	4	0.06	0.6	275

Fig. 12a shows the feature importance score of the training data in SEM 5 using random forest regression. While applying PCC, certain input features (Young's modulus, void jaggedness, void angle and Alx) are permanently removed due to high correlation whereas Alz was only incorporated in the second and fourth external cross-validation loops of the model. The loops having a higher score in applied maximum stress (σ_{max}) evaluated the fatigue life better than others. Similarly, the importance of void features for the best model is ranked as size (S) > location (L) > aspect ratio (AR). Due to one-hot encoding, most of the

Table 6
ML model based on selective epoxy materials and metal alloys data.

Training strategies	Input features	Standardization	Encoding	Average R^2	Average *MSE*	Best R^2	Best *MSE*
SEM1	"all"	✓	one-hot	0.49 ± 0.37	0.366 ± 0.185	0.76	0.190
SEM2	"all"	✓	target	0.49 ± 0.31	0.370 ± 0.145	0.72	0.227
SEM3	"all"	×	one-hot	0.52 ± 0.33	0.350 ± 0.164	0.76	0.192
SEM4	"all"	×	target	0.48 ± 0.38	0.373 ± 0.193	0.81	0.185
SEM5	"selective"	✓	one-hot	0.71 ± 0.06	0.227 ± 0.053	0.83	0.166
SEM6	"selective"	✓	target	0.61 ± 0.08	0.310 ± 0.062	0.74	0.227
SEM7	"selective"	×	one-hot	0.70 ± 0.07	0.240 ± 0.059	0.83	0.166
SEM8	"selective"	×	target	0.61 ± 0.08	0.310 ± 0.063	0.74	0.223

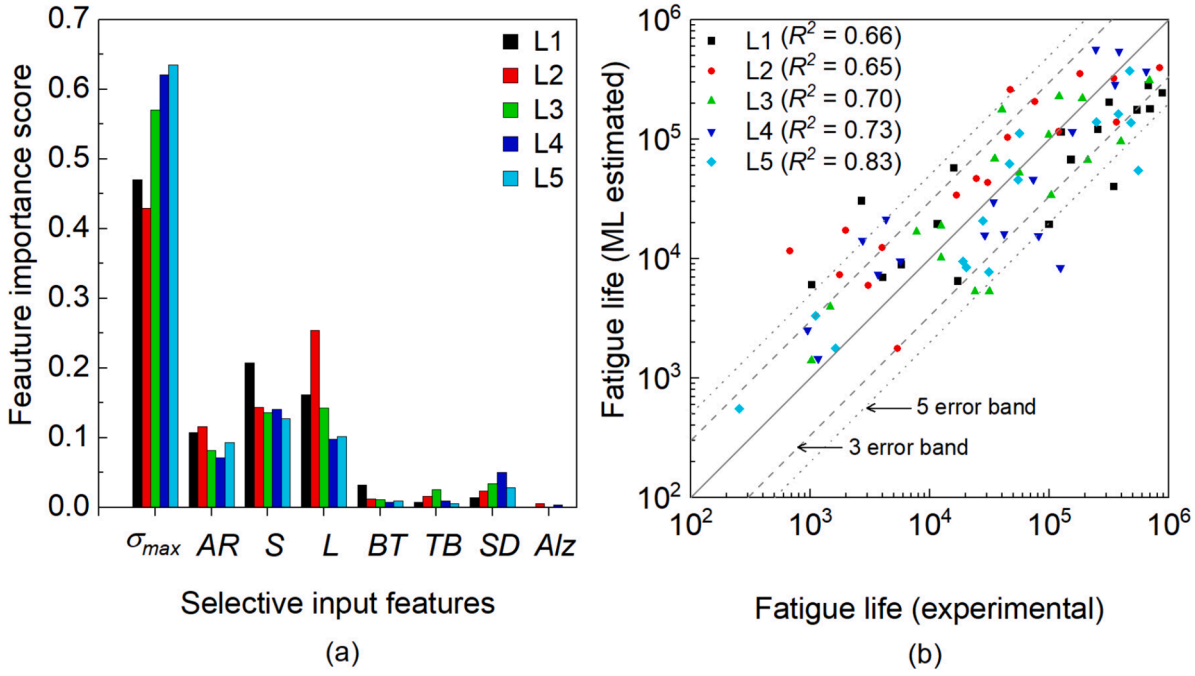


Fig. 12. SEM5: (a) feature importance score and (b) fatigue life estimation.

encoded data of materials (BT, TB, SD and Alz) was incorporated for better life estimation. Despite the relatively low importance score assigned to the material features, their inclusion in the model was strategic in improving the performance of the models trained with different classes of materials. Fig. 12b shows the experimental fatigue life (test data) versus the fatigue life estimated by the ML model for the five loops (L1, L2, L3, L4 and L5) in SEM5 strategy.

The void features were determined after fatigue failure, limiting the practical use of the developed model. One potential application is the development of machine learning (ML) aided S-N curves. As shown in Fig. 13, fatigue experimental data from specimens containing macro-voids are typically discarded when constructing the S-N curve. The developed model can estimate the fatigue life of these discarded specimens as if they contained no voids. To demonstrate this, the void characteristic values of the discarded specimens were artificially reduced by a factor of 100, simulating nearly defect-free specimens. The SEM5 model was then used to estimate their fatigue life (represented by blue solid circles). The ML model's estimated life almost coincides with the clean data. Although the specific ML model's application is limited, the proposed framework can be readily extended to other engineering applications, such as predicting fatigue life based on different inputs like manufacturing variables, loading parameters, and material compositions.

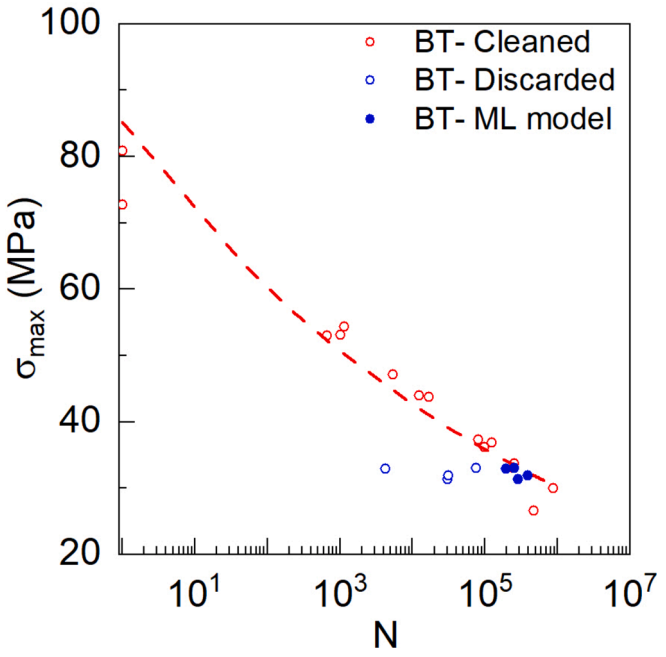


Fig. 13. ML-aided S-N curves.
Reproduced from [6].

The generalized ML framework refers to the proposed framework comprising different preprocessing steps and training strategies that can be applied to new materials and develop their corresponding models for making reliable predictions. While the developed model may require specific adjustments or re-training to accommodate the unique characteristics of each material system, the proposed framework remains the same and easily adaptable. Even with similar input and output features, differences in data distributions across materials (e.g., Gaussian vs. non-Gaussian) would necessitate re-training the model to ensure accuracy and performance. For example, the distribution of metal alloy input features such as the applied maximum stress, void angle, void aspect ratio are different from the epoxy polymers, each material system require unique training (Table 3 and Table 4). This also implies that training on one material type (e.g., polymers) and testing on another (e.g., metal alloys) would be challenging, highlighting the scope for future work.

3.5. Data-leakage and overfitting issues

Table 8 summarizes recent research results from the literature on ML-based fatigue life prediction of various additively manufactured metal alloys and compares them with this research. Most of the ML models [20–22,25–31,41] have predicted fatigue life with R^2 greater than 0.8 and some up to 0.98. Nevertheless, these over-optimistic findings could be related to data-leakage and overfitting. Data-leakage mainly occurs when the test data is reserved only after (1) data

Table 8
State of the art in ML-based fatigue life of materials.

Ref.	Year	Material	Data-leakage source	ML model	Total samples	Test portion	R^2(on the test portion) Generalizability			
							HV	*K-fold*	*LOOCV*	*NCV*
[27]	2021	Inconel 718	2	LR SVR KRR	28	?	0.71* 0.71* 0.64*	–	–	–
[42]	2021	Ti-6Al-4 V	1	KNN SVR	10	20 %	–	0.968 0.994	–	–
[26]	2022	Ti-6Al-4 V	1, 2	ANN	32,800	15 %	0.971	–	–	–
[23]	2022	AlSi10Mg	2	XGBoost	27	5 %	0.95*	–	–	–
[22]	2022	AlSi10Mg	2	RFR SVR	89	20 %	0.874 0.927	–	–	–
[9]	2022	AlSi10Mg	?	ANN PINN	12	20 %	–	0.322 0.591	–	–
[32]	2023	AlSi10Mg	1,2,4	ANN RFR SVR	1217	25 %	0.957 0.989 0.916	–	–	–
[21]	2023	Ti-6Al-4 V	2	ANN RFR SVR	29	31 %	–	–	0.848 0.827 0.807	–
[30]	2023	Ti alloy	3	Auto Glon RF SVR	2492	10 %	–	0.896 0.889 0.848	–	–
[31]	2023	GCr15 steel	1,2	ANN PINN SVR	110	10 %	0.848 0.943 –	– – 0.823	–	–
[29]	2024	Hastealloy X	?	DNN SVR	85 85	25 % 20 %	0.933 –	– 0.928	–	–
[28]	2024	Six Ti alloys	2,4	Multi GAT Auto Glon RF	571 1690 1690	9 % ? ?	–	0.904 0.919 0.855	–	–
Our work	2024	Three epoxies (BT, TB and SD) & AlSi10Mg	**No**	XGBoost	12 to 27 per material	20 %	–	–	–	0.71 ± 0.06

1: Data sampling, **2**: Data standardization/normalization, **3**: Data encoding and **4**: Feature selection.
?: No clear information has been provided.
*: R^2 on the training and test portion.
HV (Holdout-Validation): This method involves a single data division process where the dataset is split into test and training portions. The model is trained on the training data and evaluated on the test data.
K-fold Cross-Validation: Like holdout validation, but with an added step for optimizing hyperparameters. The training portion is further divided into different folds, and the model is trained and evaluated multiple times on different combinations of folds to find the optimal hyperparameters.
LOOCV (Leave-One-Out Cross-Validation): This technique splits a dataset into a training and a testing portion, using all but one observation as part of the training portion. It is repeated n (total number of observations) leaving out a different observation from the training portion each time.
NCV (Nested or External Loop Cross-Validation): This approach extends cross-validation by incorporating multiple data division processes to ensure thorough evaluation. The dataset is divided into test and training portions multiple times, covering all samples as the test set at least once. Within each iteration, the training portion is further divided into folds for hyperparameter optimization, resulting in a more robust model assessment.

sampling/augmentation, (2) data standardization/normalization, (3) data encoding and (4) feature selection. A lack of clean separation between training and test datasets before any of the above data preprocessing steps enables the model to access the test dataset, leading it to learn the relationships between input features and the output [46]. Following this learning procedure, the model can predict the concerned test dataset with a high R^2 but expected to fail when the specific dataset is changed. Additionally, there are studies [21,22,25,26,30,31] where no cross-validation was performed, which raises concerns about the overfitting of the model. Without cross-validation, the model may be overly tailored to the training data and may not perform well on unseen data. Although K-fold validation was performed in some cases, the generalizability of the model should be demonstrated with a more robust cross-validation method such as nested or external loop cross-validation. In a few studies, the test portion was either low (<20 %) [22,25,27,29,30] or not properly mentioned [27,28].

To further examine the data-leakage and overfitting issues, the ML model was intentionally trained with different data-leakages using SEM5 training strategy (3 epoxies + 2 metals alloys data, selective input features, standardized data and one-hot encoding) but with HV cross-validation. The R^2 values of the model are presented in Fig. 14. When data standardization (2), data encoding (3) and feature selection (4) were performed before the training and test data partitioning, there was no evidence of the data-leakage. The ML model trained with all leakages (1234) or only with data sampling leakage (1) shows similar R^2 values which is higher than the model trained with no data-leakage. This finding is only applicable for the XGBoost (tree-based ensemble) model trained with SMOTE augmented, minimal dataset. It is imperative to develop a robust machine learning framework for minimal datasets that addresses these issues and ensures reliable predictive performance across varied datasets.

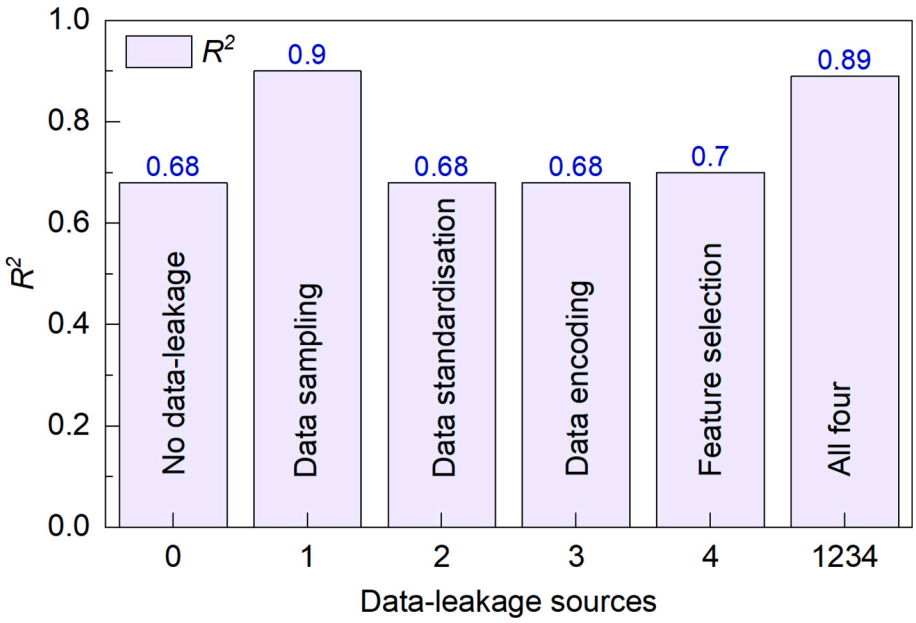


Fig. 14. Effect of data-leakage on the model performance.

4. Conclusions

In the present work, the raw data of void features from the fatigue failure surface of five distinct epoxy materials were collected and an ML framework-based extreme gradient boosting algorithm was developed to estimate fatigue life. Further, the framework was extended to evaluate the fatigue life of additively manufactured AlSi10Mg alloy. Several training strategies were developed to improve the model's accuracy and generalization and the following conclusions are derived:

1. The proposed ML framework can predict the fatigue life of selective epoxy polymers and metal alloys using a minimal dataset. Void size, location and aspect ratio are identified as the critical void parameters that reduce the fatigue life of the specimen as compared to void angle and jaggedness.
2. Data plays a crucial role in accurately predicting fatigue life. Especially, in a combined dataset of different materials, there should be enough representative data in each material category for the model to learn well.
3. While data standardization plays an insignificant role, data sampling through SMOTE, feature reduction and one-hot encoding emerge as important pre-processing steps for enhancing the model's overall performance.
4. Data-leakage and overfitting are serious issues in applied ML, and they are overlooked in recent studies; Data-leakage can occur via data sampling, data standardization, data encoding and feature selection that should be addressed for building a robust and generalized ML model. Data sampling (SMOTE) is identified as the main source of the data-leakage for the XGBoost model; therefore, data sampling should be performed only on the training dataset.

CRediT authorship contribution statement

Dharun Vadugappatty Srinivasan: Writing – original draft, Software, Methodology, Investigation, Formal analysis, Conceptualization. **Morteza Moradi:** Writing – review & editing, Software, Resources, Conceptualization. **Panagiotis Komninos:** Writing – review & editing, Software, Resources, Conceptualization. **Dimitrios Zarouchas:** Writing – review & editing, Resources, Conceptualization. **Anastasios P. Vassilopoulos:** Writing – review & editing, Validation, Supervision, Methodology, Funding acquisition, 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

Data will be made available on request.

Acknowledgments

The authors wish to acknowledge the support and funding of this research by the Swiss National Science Foundation, Switzerland (https://www.snf.ch/en) under the project "Bonded composite primary structures in engineering applications (BONDS, Grant No. IZCOZ0_189905)", the experimental assistance provided by the technical team of the structural engineering experimental platform (GIS-ENAC at EPFL, Switzerland). This article/publication is based upon work from COST Action CA18120 (CERTBOND - https://certbond.eu/), supported by COST (European Cooperation in Science and Technology).

References

[1] D.V. Srinivasan, A.P. Vassilopoulos, Manufacturing and toughening effects on the material properties of wind turbine blade adhesives, Polym. Test. 116 (2022) 107770, https://doi.org/10.1016/J.POLYMERTESTING.2022.107770.
[2] N. Sanaei, A. Fatemi, Defects in additive manufactured metals and their effect on fatigue performance: a state-of-the-art review, Prog. Mater Sci. 117 (2021) 100724, https://doi.org/10.1016/J.PMATSCI.2020.100724.
[3] M. Savvilotidou, T. Keller, A.P. Vassilopoulos, Fatigue performance of a cold-curing structural epoxy adhesive subjected to moist environments, Int. J. Fatigue 103 (2017) 405–414, https://doi.org/10.1016/J.IJFATIGUE.2017.06.022.
[4] M. Savvilotidou, A.P. Vassilopoulos, M. Frigione, T. Keller, Effects of aging in dry environment on physical and mechanical properties of a cold-curing structural epoxy adhesive for bridge construction, Constr. Build. Mater. 140 (2017) 552–561, https://doi.org/10.1016/J.CONBUILDMAT.2017.02.063.
[5] I.M.A. Foletti, J. Sena Cruz, A.P. Vassilopoulos, Fabrication and curing conditions effects on the fatigue behavior of a structural adhesive, Int. J. Fatigue 139 (105743) (2020), https://doi.org/10.1016/J.IJFATIGUE.2020.105743.
[6] D.V. Srinivasan, A.P. Vassilopoulos, Fatigue performance of wind turbine rotor blade epoxy adhesives, Polym. Test. 121 (2023) 107975, https://doi.org/10.1016/J.POLYMERTESTING.2023.107975.
[7] F. Mannino, D.V. Srinivasan, D. Fanteria, A.P. Vassilopoulos, Standard specimen geometries do not always lead to consistent fatigue results for epoxy adhesives, Int. J. Fatigue (2025) 108600, https://doi.org/10.1016/j.ijfatigue.2024.108600.

[8] S. Beretta, S. Romano, A comparison of fatigue strength sensitivity to defects for materials manufactured by AM or traditional processes, Int. J. Fatigue 94 (2017) 178–191, https://doi.org/10.1016/J.IJFATIGUE.2016.06.020.

[9] E. Salvati, A. Tognan, L. Laurenti, M. Pelegatti, F. De Bona, A defect-based physics-informed machine learning framework for fatigue finite life prediction in additive manufacturing, Mater. Des. 222 (2022) 111089, https://doi.org/10.1016/j.matdes.2022.111089.

[10] V.C. Beber, B. Schneider, Fatigue of structural adhesives under stress concentrations: Notch effect on fatigue strength, crack initiation and damage evolution, Int. J. Fatigue 140 (2020) 105824, https://doi.org/10.1016/J.IJFATIGUE.2020.105824.

[11] Y. Nadot, Fatigue from defect: influence of size, type, position, morphology and loading, Int. J. Fatigue 154 (2022) 106531, https://doi.org/10.1016/J.IJFATIGUE.2021.106531.

[12] J. Schindelin, I. Arganda-Carreras, E. Frise, V. Kaynig, M. Longair, T. Pietzsch, et al., Fiji: an open-source platform for biological-image analysis, Nat. Methods 9 (7) (2012) 676–682, https://doi.org/10.1038/nmeth.2019.

[13] S. Berg, D. Kutra, T. Kroeger, C.N. Straehle, B.X. Kausler, C. Haubold, et al., ilastik: interactive machine learning for (bio)image analysis, Nat. Methods 16 (12) (2019) 1226–1232, https://doi.org/10.1038/s41592-019-0582-9.

[14] A.P. Vassilopoulos, The history of fiber-reinforced polymer composite laminate fatigue, Int. J. Fatigue 134 (2020) 105512, https://doi.org/10.1016/J.IJFATIGUE.2020.105512.

[15] J. Chen, Y. Liu, Fatigue modeling using neural networks: a comprehensive review, Fatigue Fract. Eng. Mater. Struct. 45 (2022) 945–979, https://doi.org/10.1111/FFE.13640.

[16] P. Art, Y.M.I. Ak, I.L. Bu Kow Sk, J.F. El, R.B.S. Na, E.R. Hau, Determination of S-N curves with the application of artificial neural networks, Fatigue Fract. Eng. Mater. Struct. 22 (1999) 723–728, https://doi.org/10.1046/J.1460-2695.1999.T01-1-00198.X.

[17] T.T. Pleunet, O.K. Chopra, Artificial neural networks and effects of loading conditions on fatigue life of carbon and low-alloy steels. In the Proceedings of ASME Pressure Vessels and Piping Conference, 1997.

[18] D.S. Leininger, F.C. Reissner, J. Baumgartner, New approaches for a reliable fatigue life prediction of powder metallurgy components using machine learning, Fatigue Fract. Eng. Mater. Struct. (2022), https://doi.org/10.1111/FFE.13921.

[19] J. Zhang, J. Zhu, W. Guo, W. Guo, A machine learning-based approach to predict the fatigue life of three-dimensional cracked specimens, Int. J. Fatigue 159 (2022) 106808, https://doi.org/10.1016/j.ijfatigue.2022.106808.

[20] J.A. Lee, M.J. Sagong, J. Jung, E.S. Kim, H.S. Kim, 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. 22 (2023) 413–423, https://doi.org/10.1016/J.JMRT.2022.11.137.

[21] J. Horňas, J. Běhal, P. Homola, S. Senck, M. Holzleitner, N. Godja, et al., Modelling fatigue life prediction of additively manufactured Ti-6Al-4V samples using machine learning approach, Int. J. Fatigue 169 (2023) 107483, https://doi.org/10.1016/j.ijfatigue.2022.107483.

[22] H. Wang, B. Li, F.Z. Xuan, Fatigue-life prediction of additively manufactured metals by continuous damage mechanics (CDM)-informed machine learning with sensitive features, Int. J. Fatigue 164 (2022) 107147, https://doi.org/10.1016/J.IJFATIGUE.2022.107147.

[23] X. Peng, S. Wu, W. Qian, J. Bao, Y. Hu, Z. Zhan, et al., The potency of defects on fatigue of additively manufactured metals, Int. J. Mech. Sci. 221 (2022) 107185, https://doi.org/10.1016/j.ijmecsci.2022.107185.

[24] J. Chen, Y. Liu, Probabilistic physics-guided machine learning for fatigue data analysis, Expert Syst. Appl. 168 (2021) 114316, https://doi.org/10.1016/J.ESWA.2020.114316.

[25] H. Li, Z. Tian, J. Zheng, K. Huang, B. Nie, W. Xu, et al., A defect-based fatigue life estimation method for laser additive manufactured Ti-6Al-4V alloy at elevated temperature in very high cycle regime, Int. J. Fatigue 167 (2023) 107375, https://doi.org/10.1016/J.IJFATIGUE.2022.107375.

[26] J. Li, Z. Yang, G. Qian, F. Berto, Machine learning based very-high-cycle fatigue life prediction of Ti-6Al-4V alloy fabricated by selective laser melting, Int. J. Fatigue 158 (2022) 106764, https://doi.org/10.1016/J.IJFATIGUE.2022.106764.

[27] Y.W. Luo, B. Zhang, X. Feng, Z.M. Song, X.B. Qi, C.P. Li, et al., Pore-affected fatigue life scattering and prediction of additively manufactured Inconel 718: An investigation based on miniature specimen testing and machine learning approach, Mater. Sci. Eng. A 802 (2021) 140693, https://doi.org/10.1016/J.MSEA.2020.140693.

[28] S. Zhu, Y. Zhang, B. Zhu, J. Zhang, Y. He, W. Xu, High cycle fatigue life prediction of titanium alloys based on a novel deep learning approach, Int. J. Fatigue 182 (2024) 108206, https://doi.org/10.1016/j.ijfatigue.2024.108206.

[29] L. Lei, B. Li, H. Wang, G. Huang, F. Xuan, High-temperature high-cycle fatigue performance and machine learning-based fatigue life prediction of additively manufactured Hastelloy X, Int. J. Fatigue 178 (2024) 108012, https://doi.org/10.1016/j.ijfatigue.2023.108012.

[30] S. Zhu, Y. Zhang, X. Chen, Y. He, W. Xu, A multi-algorithm integration machine learning approach for high cycle fatigue prediction of a titanium alloy in aero-engine, Eng. Fract. Mech. 289 (2023) 109485, https://doi.org/10.1016/J.ENGFRACMECH.2023.109485.

[31] Y.-K. Liu, J.-L. Fan, G. Zhu, M.-L. Zhu, F.-Z. Xuan, Data-driven approach to very high cycle fatigue life prediction, Eng. Fract. Mech. 292 (2023) 109630, https://doi.org/10.1016/j.engfracmech.2023.109630.

[32] T. Shi, J. Sun, J. Li, G. Qian, Y. Hong, Machine learning based very-high-cycle fatigue life prediction of AlSi10Mg alloy fabricated by selective laser melting, Int. J. Fatigue 171 (2023) 107585, https://doi.org/10.1016/J.IJFATIGUE.2023.107585.

[33] H. Wang, S.L. Gao, B.T. Wang, Y.T. Ma, Z.J. Guo, K. Zhang, et al., Recent advances in machine learning-assisted fatigue life prediction of additive manufactured metallic materials: A review, J. Mater. Sci. Technol. (2024), https://doi.org/10.1016/J.JMST.2024.01.086.

[34] A.P. Vassilopoulos, E.F. Georgopoulos, V. Dionysopoulos, Artificial neural networks in spectrum fatigue life prediction of composite materials, Int. J. Fatigue 29 (2007) 20–29, https://doi.org/10.1016/J.IJFATIGUE.2006.03.004.

[35] J.A. Lee, D.P. Almond, B. Harris, The use of neural networks for the prediction of fatigue lives of composite materials, Compos. A Appl. Sci. Manuf. 30 (1999) 1159–1169, https://doi.org/10.1016/S1359-835X(99)00027-5.

[36] G. Galanopoulos, E. Fytsilis, N. Yue, A. Broer, D. Milanoski, D. Zarouchas, et al., A data driven methodology for upscaling remaining useful life predictions: from single- to multi-stiffened composite panels, Composites Part C: Open Access 11 (2023) 100366, https://doi.org/10.1016/J.JCOMC.2023.100366.

[37] W. Sai, G.B. Chai, N. Srikanth, Fatigue life prediction of GLARE composites using regression tree ensemble-based machine learning model, Adv Theory Simul 3 (2020) 2000048, https://doi.org/10.1002/ADTS.202000048.

[38] A.P. Vassilopoulos, R. Bedi, Adaptive neuro-fuzzy inference system in modelling fatigue life of multidirectional composite laminates, Comput. Mater. Sci. 43 (2008) 1086–1093, https://doi.org/10.1016/J.COMMATSCI.2008.02.028.

[39] A.P. Vassilopoulos, E.F. Georgopoulos, T. Keller, Genetic programming in modelling of fatigue life of composite materials. Exp. Anal. Nano Eng. Mater. Struct. (2007) 201–202, https://doi.org/10.1007/978-1-4020-6239-1_99.

[40] Z. Zhan, H. Li, Machine learning based fatigue life prediction with effects of additive manufacturing process parameters for printed SS 316L, Int. J. Fatigue 142 (2021) 105941, https://doi.org/10.1016/j.ijfatigue.2020.105941.

[41] H. Chen, Y. Yang, S. Cao, K. Gao, S. Xu, Y. Li, et al., Fatigue life prediction of aluminum alloy 6061 based on defects analysis, Int. J. Fatigue 147 (2021) 106189, https://doi.org/10.1016/j.ijfatigue.2021.106189.

[42] H. Bao, S. Wu, Z. Wu, G. Kang, X. Peng, P.J. Withers, A machine-learning fatigue life prediction approach of additively manufactured metals, Eng. Fract. Mech. 242 (2021) 107508, https://doi.org/10.1016/J.ENGFRACMECH.2020.107508.

[43] L. Dang, X. He, D. Tang, Y. Li, T. Wang, A fatigue life prediction approach for laser-directed energy deposition titanium alloys by using support vector regression based on pore-induced failures, Int. J. Fatigue 159 (2022) 106748, https://doi.org/10.1016/J.IJFATIGUE.2022.106748.

[44] A. Rovinelli, M.D. Sangid, H. Proudhon, W. Ludwig, Using machine learning and a data-driven approach to identify the small fatigue crack driving force in polycrystalline materials, Npj Comput. Mater. 4 (2018) 1–10, https://doi.org/10.1038/s41524-018-0094-7.

[45] K. Song, F. Yan, T. Ding, L. Gao, S. Lu, A steel property optimization model based on the XGBoost algorithm and improved PSO, Comput. Mater. Sci. 174 (2020) 109472, https://doi.org/10.1016/J.COMMATSCI.2019.109472.

[46] S. Kapoor, A. Narayanan, Leakage and the reproducibility crisis in machine-learning-based science, Patterns 4 (2023) 100804, https://doi.org/10.1016/J.PATTER.2023.100804.

[47] M.A. Lones, How to avoid machine learning pitfalls: a guide for academic researchers 2021.

[48] N.V. Chawla, K.W. Bowyer, L.O. Hall, W.P. Kegelmeyer, SMOTE: synthetic minority over-sampling technique, J. Artif. Intell. Res. 16 (2011) 321–357, https://doi.org/10.1613/jair.953.

[49] A.-I. Udil, Encoding Methods for Categorical Data: A Comparative Analysis for Linear Models, Decision Trees, and Support Vector Machines, Delft University of Technology, The Netherlands, 2023.

[50] J.H. Friedman, Greedy Function Approximation: A Gradient Boosting Machine 2001;29:1189–232.

[51] K.V. Rashmi, R. Gilad-Bachrach, DART: dropouts meet multiple additive regression trees, J. Mach. Learn. Res. 38 (2015) 489–497.

[52] T.J. Bradshaw, Z. Huemann, J. Hu, A. Rahmim, A guide to cross-validation for artificial intelligence in medical imaging, Radiol. Artif. Intell. 5 (2023), https://doi.org/10.1148/RYAI.220232.

[53] skopt.BayesSearchCV — scikit-optimize 0.8.1 documentation n.d. https://scikit-optimize.github.io/stable/modules/generated/skopt.BayesSearchCV.html (accessed April 9, 2024).