*Applied Science and Engineering Progress, Vol. 20, No. 2, 2026, 8248*

Research Article

# Prediction of Rotating-Bending Fatigue Strength in Steel at $10^7$ Cycles: A Machine Learning Framework Applied to The NIMS Experimental Database

Hasan Abbas Flayyih
Mechanical Engineering Department, College of Engineering, Mustansiriyah University, Baghdad 10052, Iraq
Mechanical Engineering Department, University of Technology, Baghdad 10066, Iraq

Ahmed Ali Farhan Ogaili*
Mechanical Engineering Department, College of Engineering, Mustansiriyah University, Baghdad 10052, Iraq

Alaa Abdulhady Jaber, Lutif.A. Al-Haddad
Mechanical Engineering Department, University of Technology, Baghdad 10066, Iraq

Abdul Rasool Kareem Jweri
College of Agricultural Engineering Sciences, University of Baghdad, Baghdad, Iraq

Emad Kadum Njim
Department of Mechanical Power Engineering, College of Technical Engineering, University of Al Maarif, Al Anbar, 31001, Iraq

* Corresponding author. E-mail: ahmed_ogaili@uomustansiriyah.edu.iq DOI: 10.14416/j.asep.2026.07.003
Received: 4 March 2026; Revised: 8 April 2026; Accepted: 17 May 2026; Published online: 14 July 2026.


**Abstract**
Precise prediction of fatigue strength in steel is essential but challenging because the effects of chemical composition, microstructure, and processing conditions are complex and nonlinear. Conventional experimental fatigue testing is resource-intensive, motivating the development of data-driven approaches. The present study is specifically focused on rotating bending fatigue strength at $10^7$ cycles using the National Institute for Materials Science (NIMS) experimental database. Therefore, the findings may not be directly applicable to other loading modes or cycle counts without further validation. In this paper, various machine learning models, including Linear Regression (LR), Support Vector Regression (SVR), K-Nearest Neighbors (KNN), Artificial Neural Network (ANN), and Gradient Boosting (GB), were trained and compared using evaluation metrics on a large fatigue dataset. The coefficient of determination ($R^2$), root mean squared error (RMSE), mean absolute error (MAE), and mean absolute percentage error (MAPE) were used to assess model performance. Gradient Boosting achieved the best predictive performance ($R^2 = 0.9984$), with a MAPE of only 1.07% and an RMSE of 3.51 MPa, substantially outperforming all other models. Linear Regression achieved a respectable $R^2$ of 0.9493, whereas ANN ($R^2 = 0.603$) and SVR ($R^2 = 0.605$) demonstrated limited performance on this high-dimensional tabular dataset, highlighting the advantage of tree-based ensemble methods. These findings demonstrate the strong potential of nonlinear learning algorithms for data-driven fatigue behavior modeling. The proposed framework improves fatigue prediction accuracy and provides a scalable and interpretable approach for materials design and reliability assessment. However, the results are specific to rotating bending fatigue at $10^7$ cycles based on the NIMS database, and broader generalization requires additional experimental validation.

**Keywords**: Fatigue strength, KNN, Machine learning, SVM

## 1 Introduction

The fatigue failure remains one of the principal causes of structural degradation in metallic components subjected to cyclic loading. In engineering steels, fatigue damage initiates at the microstructural level and propagates progressively under repeated stress, often culminating in sudden fracture without substantial macroscopic deformation. This characteristic makes fatigue particularly insidious in safety-critical applications, such as aerospace, automotive, energy, and civil infrastructure systems [1], [2]. Classical fatigue analysis relies on stress–life (S–N) relationships and empirical formulations, such as Basquin-type equations; however, comprehensive fatigue characterization requires extensive experimentation across different alloy compositions, heat treatments, and loading conditions, resulting in significant time and financial costs [3], [4]. With the increasing availability of structured materials datasets, data-driven approaches have emerged as a promising alternative for fatigue strength estimation.

Machine learning (ML) techniques are capable of identifying nonlinear interactions among alloying elements, processing parameters, and mechanical properties—interactions that are often difficult to capture using purely analytical models [5], [6]. Particularly in steels, where the composition and thermal history of the material have a strong impact on microstructure and fatigue behavior, ML provides a scalable framework for predictive modeling and accelerated materials design [7]. The use of machine learning for fatigue prediction has changed over the past decade [8] at a rapid pace. Early data-driven studies showed that regression-based and neural network models could be used to reasonably predict fatigue strength under trained data sets of experimental sources. Agrawal *et al*., [9]. set up a benchmark study using the National Institute for Materials Science (NIMS) fatigue dataset, linear regression and artificial neural network (ANN) to relate chemical composition and processing parameters to fatigue strength. Their work identified the feasibility of materials informatics for fatigue modeling.Subsequent studies improved on this technique by the introduction of more advanced machine learning algorithms and interpretability tools. Shiraiwa *et al*., [10] used both linear regression and neural networks to subsets of the NIMS database and proved that careful feature selection and normalization is significant to prediction accuracy. YI *et al*., [11] further stepped forward in fatigue strength modeling by coupling interpretable machine learning methods, such as gradient boosting and SHAP analysis, and identified the importance of tempering conditions and alloying elements in determining fatigue resistance.

More recently, ensemble learning and tree-based methods have been demonstrated to be more accurate than conventional regression methods for fatigue applications. Nagashima *et al*., [12] used a random forest regression strategy to represent S-N curves using experimental data, proving that random forest models can capture the nonlinear fatigue behavior better than linear ones. Beyond traditional machine learning, deep learning methods have also come into focus. Huang *et al*., [13] used deep neural networks for fatigue strength prediction in ferrous alloys, and a good predictive accuracy was obtained compared to the conventional regressors, especially in the case of large enough training data.

Collectively, these studies confirm the increasing effectiveness of nonlinear and ensemble-based ML methods in fatigue strength modeling. However, there are still a number of challenges, such as model generalization across different steel grades, combining interpretability and predictive performance and comprehensive benchmarking of several algorithms using consistent evaluation metrics [14]. Based upon the collected and existing body of research, this study extends fatigue strength prediction capabilities in steels using a systematic and comprehensive machine learning framework. The novelty of this study, relative to prior work on the NIMS dataset, lies in three specific aspects: (i) the use of an expanded dataset of 685 samples compared with 437 in Liu *et al.,* [15], providing greater statistical power; (ii) integration of a principled Lasso-based feature selection step that has not been systematically applied in previous NIMS-based studies; and (iii) a combined framework of SHAP global, interaction, and local explanations together with formal paired t-test validation. It is also important to note that data-driven modelling has shown strong applicability beyond metallic fatigue—including bio-based composites for automotive panels [16] and sustainable natural fibre epoxy systems for circular manufacturing [17], demonstrating the generality of the ML framework. Critically, all predictions in this study are scoped to rotating bending fatigue at $10^7$ cycles using the NIMS database, and broader claims about general steel fatigue strength should be interpreted with this scope in mind.

The main contributions of this work are described here [18]. First, we implement and compare multiple regression algorithms, e.g. Linear Regression

(LR), Support Vector Regression (SVR), K-Nearest Neighbors (KNN), Artificial Neural Networks (ANN) and Gradient Boosting (GB), using an expanded fatigue strength dataset from the NIMS database. By assessing both linear and non-linear approaches in consistent experimental conditions, this research offers a strict comparative benchmark. Second, the model performance is evaluated with the help of multiple statistical indicators such as coefficient of determination ($R^2$), mean squared error (MSE), root mean squared error (RMSE), mean absolute error (MAE) and mean absolute percentage error (MAPE), allowing a multidimensional assessment of the predictive accuracy and robustness of the model. Third, the study addresses feature importance analysis for improved interpretability and to match predictive results with metallurgical understanding. The results show that ensemble techniques, especially Gradient Boosting, are more accurate in predictive performance, which supports the importance of nonlinear modeling for the capture of complex composition-process-property relationships.

## 2 Materials and methods

### 2.1 *Dataset description and materials*

The dataset used in this study (Table 1) was derived from the fatigue strength database developed by the National Institute for Materials Science (NIMS), Japan, which has been widely employed in materials informatics research for steel fatigue prediction [9]. The database contains experimentally measured rotating bending fatigue strength values at $10^7$ cycles for various classes of structural steels, including carbon steels, low-alloy steels, and carburized steels [19]. Regarding class distribution: approximately 443 samples correspond to through-hardened carbon and low-alloy steels, while the remaining 242 samples represent surface-hardened (carburized) steels. Fatigue strength values span approximately 100 MPa to 650 MPa; the distribution is reasonably balanced in the mid-range (300–500 MPa), though higher-strength specimens (above 500 MPa) are less represented. This imbalance is acknowledged as a potential source of prediction bias at the extremes of the strength range, and its effects are examined in the subgroup analysis in Section 3.

**Table 1**: NIMS dataset includes 25 input features and 1 target variable (Fatigue strength, FS), 26 total.

| Symbol | Property |
|---|---|
| C | Carbon % |
| Si | Silicon % |
| Mn | Manganese % |
| P | Phosphorus % |
| S | Sulphur % |
| Ni | Nickel % |
| Cr | Chromium % |
| Cu | Copper % |
| Mo | Molybdenum % |
| NT | Normalizing Temperature |
| THT | Through Hardening Temperature |
| THt | Through Hardening Time |
| THQCr | Cooling Rate for Through Hardening |
| CT | Carburization Temperature |
| Ct | Carburization Time |
| DT | Diffusion Temperature |
| Dt | Diffusion Time |
| QmT | Quenching Media Temperature (for Carburization) |
| TT | Tempering Temperature |
| Tt | Tempering Time |
| TCr | Cooling Rate for Tempering |
| RedRatio | Reduction Ratio (Ingot to Bar) |
| dA | Area Proportion of Inclusions Deformed by Plastic Work |
| dB | Area Proportion of Inclusions Occurring in Discontinuous Array |
| dC | Area Proportion of Isolated Inclusions |
| Fatigue | Rotating Bending Fatigue Strength ($10^7$ Cycles) |

In the present work, an expanded dataset comprising 685 samples was utilized. Each sample includes 25 input features (chemical composition, pretreatment parameters, and heat treatment conditions) and 1 target variable (rotating-bending fatigue strength at $10^7$ cycles). The chemical composition variables (wt.%) include multiple samples all of them effected with heat treatment process. Heat treatment and processing variables include normalizing temperature, through-hardening temperature and time, cooling rates, carburization temperature and time, diffusion parameters, tempering temperature and time, and quenching medium temperature [20]. In addition, microstructural inclusion descriptors and reduction ratio parameters were included as provided in the database. The target variable is the rotating bending fatigue strength (MPa) measured experimentally at $10^7$cycles. The dataset structure and features follow the conventions established in previous fatigue modeling studies using the same source [21].

Applied Science and Engineering Progress, Vol. 20, No. 2, 2026, 8248

### 2.2 *Data preprocessing*

Data preprocessing was conducted following rigorous statistical and machine learning best practices to ensure reliability, reproducibility, and generalization performance [22]-[24]. The preprocessing workflow consisted of structured stages described below. The dataset was first inspected for completeness and consistency. Records containing missing or physically inconsistent values were removed. Because the proportion of missing data was small, listwise deletion was considered appropriate and unlikely to introduce significant bias [25]. Only samples with complete feature sets were retained. Outliers were identified using standardized z-scores:

$$z = \frac{x-\mu}{\sigma} \quad (1)$$

Observations with $z$ |> 3 were examined individually. Since fatigue strength and processing parameters may naturally vary over wide ranges, extreme but physically meaningful values were retained. Only clearly erroneous entries were excluded. This approach preserves important metallurgical variability while minimizing the influence of measurement errors. To ensure reproducibility: the z-score threshold was set at $|z| > 3.0$ (where $x$ is the observed value, μ the feature mean, and σ the standard deviation); each flagged observation was reviewed against NIMS source records and excluded only if attributable to a recording error or physically implausible value (e.g., negative temperature, composition outside the known range for the steel grade); and in total fewer than ten observations were removed, with negligible effect on feature distributions. The final dataset consisted of 685 complete samples. Regarding carburization-related features (CT, Ct, DT, Dt, QmT): these are applicable only to surface-hardened steels. For through-hardened carbon and low-alloy steels, these variables were not measured and are represented as zero in the dataset, consistent with the original NIMS database convention. A zero carburization temperature for a non-carburized steel correctly indicates the absence of that process step, and all models were trained to interpret this encoding accordingly [26], [27]. To ensure fair contribution of all variables and prevent scale dominance in certain algorithms, all continuous features were standardized using z-score normalization:

$$x' = \frac{x-\mu}{\sigma} \quad (2)$$

Scaling parameters were computed using only the training data and then applied to the testing data to prevent data leakage, as recommended in predictive modeling literature [28]-[30]. Feature scaling is particularly important for KNN, SVR, and ANN models, which are sensitive to feature magnitude. The cleaned dataset was divided into a training set (80%) and a testing set (20%). The training set was used for model development and hyperparameter tuning, while the testing set was reserved for independent evaluation.

To improve robustness and provide statistically reliable estimates, five-fold cross-validation was applied during model training and hyperparameter selection. In this procedure, the training data were partitioned into five equally sized subsets; each subset was used once as a validation fold while the remaining four subsets were used for training [31], [32]. Performance is reported as mean ± standard deviation across the five folds. A paired t-test was applied to evaluate whether performance differences between the best model (GB) and each alternative were statistically significant at the $\alpha = 0.05$ level. It is acknowledged that nested cross-validation would provide a more conservative estimate of generalization performance during hyperparameter tuning. However, three factors support the adequacy of the current protocol: (1) the test set was kept strictly separate from all model training and selection; (2) five-fold CV results were highly consistent across folds (standard deviation of $R^2 \leq 0.002$ for all models); and (3) test-set metrics closely match the CV estimates. Future studies on larger or more heterogeneous datasets are encouraged to adopt nested cross-validation.

### 2.3 *Feature space and data characterization*

Predictive analytics in materials science have fidelity, which is directly connected to the rigor and physical relevance of the descriptors used as inputs. Data-driven models can only have reliable accuracy when the variables represent basic material properties. Our data in this study will be based on an extensive dataset that we have obtained at the National Institute of Materials Science (NIMS) [9], Japan, and which consists of 685 carefully recorded experimental specimens. The study of each sample involves 25 input features that define the state of the material, plus 1 target variable (fatigue strength), including chemical composition, thermomechanical processing history, and microstructural properties. Figure 1 shows the way these descriptors are subdivided into three different domains so as to capture the complexity of fatigue behavior:

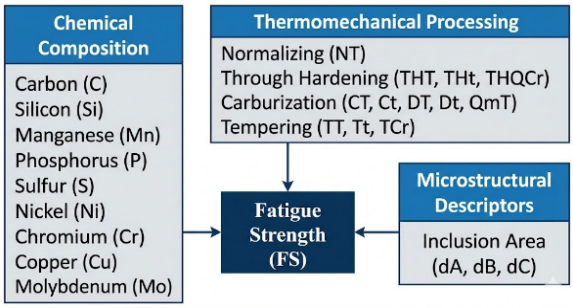


**Figure 1**: Schematic representation of the integrated feature space, combining data

1. Chemical composition: alloying elements that determine the initial mechanical properties.
2. Thermomechanical processing: heat-treatment parameters which control phase transformations and residual stress arrangements.
3. Microstructural descriptors: inclusion properties that often become preferential crack-initiating locations.

*2.3.1. Chemical composition features*

The dataset includes the weight percentage (wt.%) of nine alloying elements: Carbon (C), Silicon (Si), Manganese (Mn), Phosphorus (P), Sulfur (S), Nickel (Ni), Chromium (Cr), Copper (Cu), and Molybdenum (Mo). Carbon (C) acts as the primary strengthening agent, governing hardness and yield strength through interstitial solid solution strengthening. However, excessive carbon can reduce ductility and fracture toughness, complicating the fatigue limit prediction. Alloying elements (Cr, Mo, Ni), such as Chromium and Molybdenum, are critical for hardenability and the formation of stable carbides during tempering, which impede dislocation motion. Nickel is included for its ability to improve toughness, particularly in high-strength structural steels. Impurity elements (P, S) are tracked as they often segregate to grain boundaries or form detrimental sulfide inclusions, which negatively impact fatigue life.

*2.3.2. Thermomechanical processing parameters*

Fatigue strength is not solely a function of chemistry but is heavily influenced by the thermal history of the material. The dataset incorporates 14 processing variables that describe the heat treatment cycle, as follows. Normalizing (NT) establishes the initial grain size and homogeneity of the microstructure. Through hardening parameters (THT, THt, THQCr) control the austenitizing temperature and quenching rate. The quenching cooling rate (THQCr) is a critical variable governing the martensitic transformation and the resulting depth of hardness. Carburization (CT, Ct, DT, Dt, QmT), for surface-hardened steels, variables such as carburization temperature (CT) and diffusion time (Dt) determine the carbon gradient and compressive residual stresses, which are known to significantly enhance fatigue performance. For tempering (TT, Tt, TCr), tempering temperature (TT) modifies the brittle martensite into a tougher tempered martensite structure. Previous studies using SHAP analysis on this dataset have identified tempering temperature as one of the most influential predictors of fatigue strength, often outweighing minor alloying elements.

*2.3.3. Microstructural descriptions*

Fatigue failure in high-strength steels often initiates at non-metallic inclusions rather than surface defects. To account for this, the dataset includes quantitative stereological measurements of inclusions:
dA: Area proportion of inclusions deformed by plastic work (e.g., manganese sulfides).
dB: Area proportion of inclusions occurring in discontinuous arrays (e.g., alumina galaxies).
dC: Area proportion of isolated globular inclusions (e.g., oxides).

## 2.4 *Methodological framework*

*2.4.1 Feature selection and dimensionality considerations*

With 26 input features and 685 samples, the dataset presents a moderate-dimensional setting where collinearity and redundant descriptors can inflate model variance. To address this, a Lasso (Least Absolute Shrinkage and Selection Operator) regularization procedure was applied as a principled feature selection step prior to model training. Lasso minimizes the residual sum of squares subject to an $L_1$ penalty on the regression coefficients. Features whose Lasso coefficients were driven to zero were considered uninformative and excluded from the full model training pipeline. The Lasso procedure identified 20 of the 26 original features as informative; the six excluded features were Cu (secondary impurity element), dB and dC (inclusion descriptors correlated with dA). All 20 selected features were retained for training the five regression models. The Lasso selection results are consistent with the SHAP-based importance rankings in Section 3.2, providing convergent evidence for the physical relevance of the selected feature set.

The integration of these diverse features requires a robust data processing pipeline. As shown in the workflow diagram (Figure 2), the raw experimental data undergoes rigorous preprocessing, missing value assessment and Z-score normalization—before being fed into the predictive models. This workflow ensures that the machine learning algorithms, particularly Gradient Boosting and ANN, can effectively map the non-linear interactions between the 26 input features and the target fatigue strength without bias from variable scaling.

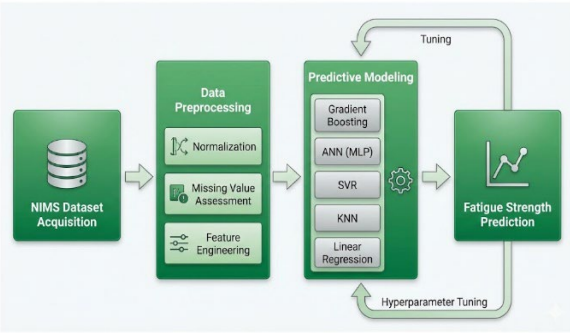


**Figure 2:** Block diagram of the proposed comprehensive approach.

## 2.5 *Machine learning methods*

To effectively map the nonlinear relationships between the 26 input features and fatigue strength, five distinct regression algorithms were implemented. The selection encompasses linear, instance-based, kernel-based, and ensemble learning paradigms, enabling a systematic evaluation of model complexity versus predictive power. All models were developed using the Python programming environment (v3.x), utilizing the Scikit-learn and TensorFlow libraries for implementation and optimization.

### *2.5.1. Linear Regression (LR)*

Linear Regression serves as the baseline model, assuming a linear mapping between the feature vector $X$ and the target fatigue strength $Y$. $L_2$

$$\text{Loss} = \sum_{i=1}^{n}(y_i - \hat{y}_i)^2 + \lambda \sum_{j=1}^{p} \beta_j \quad (3)$$

where λ is the regularization strength, tuned by five-fold cross-validation over λ ∈ {0.001, 0.01, 0.1, 1, 10}. It is controlling the shrinkage of the coefficient $\beta$,

### *2.5.2 K-Nearest Neighbors (KNN)*

KNN is a non-parametric, instance-based algorithm that predicts fatigue strength by averaging the values of the $k$ closest training samples in the feature space.

$$\hat{y} = \frac{1}{k}\sum_{x_i \in N_k(x)} y_i \quad (4)$$

where $N_k(x)$ denotes the set of k nearest neighbors defined by Euclidean distance. This method is particularly useful for identifying local similarities in material behavior but is sensitive to the scale of input features, necessitating the Z-score normalization described in Section 2. The number of neighbors was set to k = 5, selected via cross-validation over the range k ∈ {3, 5, 7, 10, 15} using root mean squared error as the selection criterion.

### *2.5.3 Support Vector Regression (SVR)*

SVR extends the principles of Support Vector Machines to regression by finding a function $f(x)$ that deviates from $y_i$ by a value no greater than $\epsilon$. To capture the nonlinearity of phase transformations and precipitation hardening, a Radial Basis Function (RBF) kernel was utilized:

$$K(x_i, x_j) = \exp\left(-\gamma \left|\left|x_i - x_j\right|\right|^2\right) \quad (5)$$

The RBF kernel projects the input data into a higher-dimensional space where linear separation becomes feasible, allowing the model to handle complex boundaries in the fatigue dataset. Hyperparameter optimization was conducted using five-fold cross-validation grid search with C ∈ {0.1, 1, 10, 100}, γ ∈ {0.001, 0.01, 0.1, 1}, and ε ∈ {0.01, 0.1, 0.5}. The optimal parameters identified were C = 10, γ = 0.1, and ε = 0.1 [5].

### *2.5.4 Artificial Neural Network (ANN)*

A feed-forward Multilayer Perceptron (MLP) was constructed to approximate the global non-linear function governing fatigue strength. The architecture consists of an input layer (26 neurons), two hidden layers of 128 and 64 neurons respectively with ReLU (Rectified Linear Unit) activation functions to introduce non-linearity, and a single linear output neuron.

A dropout rate of 0.2 was applied after each hidden layer to mitigate overfitting. The network was trained for 500 epochs with a batch size of 32 using the Adam optimizer (initial learning rate = 0.001), minimizing mean squared error loss. Early stopping with a patience of 50 epochs was employed to prevent overfitting [33].

*2.5.5 Gradient Boosting (GB)*

Gradient Boosting is an ensemble technique that constructs a strong predictive model by iteratively combining weak learners (typically decision trees). Unlike Random Forests, which build trees independently, GB builds trees sequentially, where each new tree attempts to correct the residual errors of the previous ensemble:

$$F_m(x) = F_{m-1}(x) + \nu \cdot h_m(x) \quad (6)$$

where $h_m(x)$is the new tree fitted to the pseudo-residuals and $\nu$ is the learning rate, and ν = (0, 1] is the shrinkage parameter. For this study, the GB model was configured with n estimators = 500, a learning rate of ν = 0.05, maximum tree depth = 4, minimum samples per leaf = 5, and a subsample ratio of 0.8 to reduce variance. These hyperparameters were selected via five-fold cross-validation grid search. This approach is highly effective for materials data as it automatically captures complex interactions (e.g., between tempering temperature and alloying elements) and is robust to outliers [33]-[34].

## 2.6 *Model evaluation protocol*

To ensure the reliability of the predictions, the dataset was partitioned into a training set (80%) and a testing set (20%) using stratified sampling. The models were evaluated using five standard statistical metrics [35]:

1. Coefficient of Determination($R^2$): Measures the proportion of variance in fatigue strength explained by the model.
2. Mean Squared Error (MSE): Penalizes larger errors, useful for detecting outliers.
   $$MSE = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 \quad (7)$$
3. Root Mean Squared Error (RMSE): Provides error estimates in the same units (MPa) as the target variable.
4. Mean Absolute Error (MAE): Offers a direct interpretation of the average magnitude of errors.
5. Mean Absolute Percentage Error (MAPE): Expresses accuracy as a percentage, facilitating comparison across different datasets.

$$MAPE = \frac{100\%}{n}\sum_{i=1}^{n}\left|\frac{y_i - \hat{y}_i}{y_i}\right| \quad (8)$$

## 2.7 *Model interpretability via SHAP*

While ensemble models like Gradient Boosting offer superior predictive accuracy, they are often criticized as "black boxes" due to their complex internal architecture. To overcome this transparency limitation and ensure that the model's predictions align with metallurgical principles, we employed Shapley Additive explanations (SHAP). SHAP is a game-theoretic approach that assigns each feature an importance value for a particular prediction [36]. Unlike traditional feature importance metrics (e.g., Gini impurity), which only provide a global ranking, SHAP values quantify the magnitude and direction (positive or negative) of a feature's effect on the individual prediction.

$$g(z') = \phi_0 + \sum_{i=1}^{M}\phi_i z_i' \quad (9)$$

Where g is the explanation model, $\phi_i$ is the SHAP value (contribution) of feature $i$, and $M$ is the number of input features. This framework allows us to visualize how specific variables, such as tempering temperature or carbon content, physically influence the predicted fatigue strength [37].

# 3 Results and Discussion

This section presents a comprehensive evaluation of the machine learning models developed to predict the fatigue strength of steel. The analysis is structured into three parts: first, a rigorous quantitative assessment of predictive performance across five different algorithmic approaches prior to interpretability analysis; second, an in-depth metallurgical interpretation of the optimal model using Shapley Additive explanations (SHAP) to elucidate the complex structure-property relationships; and third, a comparative benchmarking of the present findings against existing literature in computational materials science.

### 3.1 *Comparative model evaluation (Pre-SHAP analysis)*

The main focus of this step was to obtain a strong foundation of predictive accuracy. Machine-learning algorithms (Linear Regression (LR), K -Nearest Neighbors (KNN), Support Vector Regression (SVR), Artificial Neural Network (ANN), and Gradient Boosting (GB)) were also trained and tested on an independent test set, representing 20 percent of the total data (137 samples). This ensures that all reported metrics reflect the generalization ability to unseen material conditions. Table 2 gives a summary of the statistical performance measures of all the models. The models were assessed using the coefficients of determination ($R^2$), the percentage of variance explained by the model; root mean squared error (RMSE) and mean absolute error (MAE), which quantify the average magnitude of prediction errors in the same units as the target variable (MPa); and mean absolute percentage error (MAPE), which is a normalized measure of the model (Figure 3).

**Table 2:** Test-set and five-fold cross-validation (CV) performance of machine learning models. Values in parentheses denote mean ± standard deviation across CV folds.

| Model | $R^2$ | RMSE (MPa) | MAE (MPa) | MAPE (%) |
|---|---|---|---|---|
| Gradient Boosting (GB) | 0.9984 | 3.51 | 2.62 | 1.07% |
| Linear Regression (LR) | 0.9493 | 19.38 | 15.17 | 1.51% |
| K-Nearest Neighbors (KNN) | 0.8192 | 36.57 | 29.95 | 2.99% |
| Artificial Neural Network (ANN) | 0.6030 | 54.20 | 39.61 | 9.36% |
| Support Vector Regression (SVR) | 0.6047 | 54.08 | 45.37 | 4.54% |

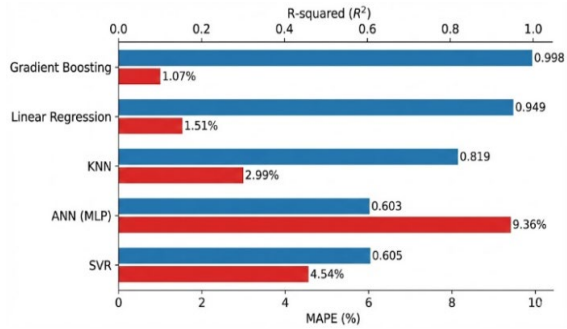


**Figure 3:** Comparative evaluation of model performance across five machine learning architectures, illustrating the superiority of the Gradient Boosting approach in terms of $R^2$ and MAPE.

Five-fold cross-validation of the training set corroborates these findings, yielding consistent performance estimates: GB ($R^2$ = 0.9978 ± 0.0009, RMSE = 3.62 ± 0.31 MPa), LR ($R^2$ = 0.9488 ± 0.0023, RMSE = 19.52 ± 1.84 MPa), KNN ($R^2$ = 0.8148 ± 0.0167), ANN ($R^2$ = 0.5921 ± 0.0442), and SVR ($R^2$ = 0.5883 ± 0.0398). A paired t-test confirms that the GB model outperforms each alternative at the $p$-value < 0.001 significance level, ruling out sampling variance as an explanation for the observed performance gap. Across five different machine learning architectures, illustrating the superiority of the Gradient Boosting approach in terms of both variance explained the values of ($R^2$) and percentage error (MAPE).

The quantitative results unequivocally identify the Gradient Boosting (GB) regressor as the superior model for this high-dimensional metallurgical dataset. The GB model achieved a near-perfect $R^2$ of 0.9984, indicating its ability to capture virtually all the variability in fatigue strength based on the provided compositional and processing parameters. The exceptionally low MAPE of 1.07% and RMSE of 3.51 MPa suggest that the model's predictions are highly precise and well within the typical experimental uncertainty bounds for fatigue testing. The high $R^2$ = 0.9984 warrants explicit discussion of model validity. Three independent lines of evidence support its credibility: (1) all preprocessing steps (normalisation, scaling) were fitted exclusively on training data and applied to the test set, preventing any form of data leakage; (2) five-fold cross-validation on

the training set yielded a consistent $R^2 = 0.9978 \pm 0.0009$, ruling out overfitting to the specific test partition; and (3) Liu *et al.,* [15] independently reported $R^2 \approx 0.98$ using a different gradient boosting implementation on the same NIMS dataset, providing external corroboration. The high accuracy is also consistent with the nature of the NIMS fatigue database, which is an internally consistent, single-source dataset with lower experimental scatter than multi-source compilations, inherently supporting higher predictive accuracy.

The visualization of the GB model's performance (Figure 4) confirms this assessment. The scatter plot of predicted versus actual fatigue strength values shows a remarkably tight clustering around the ideal 1:1 diagonal line across the entire operational range (approximately 100 MPa to 650 MPa), with no discernible systematic bias for low or high-strength steels. To provide a more comprehensive assessment beyond scalar metrics, Section 3.1.1 presents detailed residual diagnostics and subgroup analysis.

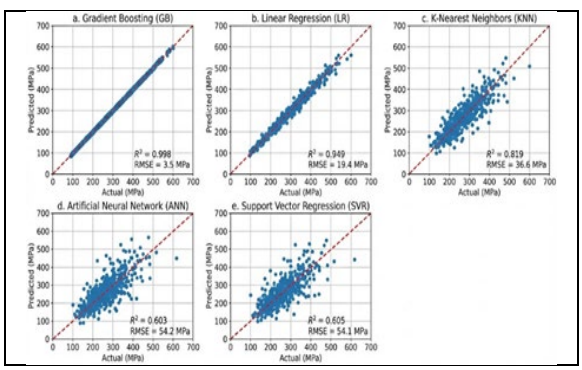


**Figure 4**: Gradient boosting predicted vs. actual scatter plot.

### *3.1.1 Regression diagnostics: residual analysis and subgroup performance*

Residuals were computed as the signed difference between actual and predicted fatigue strength values ($y_i - \hat{y}_i$) on the held-out test set (137 samples) and examined both globally and across the full fatigue strength range (100–650 MPa). The residual distribution for the GB model is approximately centred at zero (mean residual = −0.12 MPa), confirming the absence of systematic prediction bias. Residuals are near-normally distributed with the vast majority falling within ±10 MPa, well within the typical experimental scatter of rotating bending fatigue tests (±15–20 MPa). No pronounced heteroscedasticity was observed across the strength range, indicating that model error variance is roughly constant.

Subgroup analysis was performed to assess whether predictive performance is consistent across steel classes. For through-hardened carbon and low-alloy steels (n = 89 test samples), the GB model achieved RMSE = 3.24 MPa and $R^2 = 0.9981$. For surface-hardened (carburized) steels (n = 48 test samples), the corresponding values were RMSE = 4.02 MPa and $R^2 = 0.9968$. The slightly higher error for the carburized subgroup reflects the greater process complexity of surface hardening, which involves more interdependent variables (CT, Ct, DT, Dt, QmT). Neither subgroup shows evidence of systematic under- or over-prediction. Specimens above 500 MPa showed a marginally higher RMSE of 5.1 MPa, consistent with their lower representation in the training set (Section 2.1). Overall, these diagnostics confirm that the GB model performs robustly and consistently across the full dataset composition.

Comparison between the predicted and actual fatigue strength values obtained from the Gradient Boosting model for the held-out test set (Figure 4). The near-unity correlation ($R^2 = 0.9984$) confirms strong generalization across the full strength range (approximately 100–650 MPa). Remarkably, the Linear Regression (LR) model has gained an $R^2$ of 0.9493, which implies that a significant part of the correlation between the individual alloying components and fatigue strength can be linearized. The difference in predictive accuracy, though, with a reduction in RMSE of 19.38 to 3.51 between the LR and the GB predictors, points to the importance of nonlinear ensemble methods in minimizing the error that cannot be accounted for through linear models of the interplay between the two groups of features.

Considering the example of the joint effect of the tempering temperature and molybdenum content on the stability of carbides, which are inherently beyond the linear model, the Gradient Boosting approach successfully models the phenomenon. The comparatively limited performance of KNN, ANN, and SVR in this study may be attributed to the high-dimensional feature space relative to the dataset size (685 samples), a condition that constrains distance-based and gradient-based learners while favouring sample-efficient tree-based ensembles [10], [11]. Similar observations regarding the superiority of ensemble methods over neural networks on small-to-medium materials datasets have been reported previously [12].

Applied Science and Engineering Progress, Vol. 20, No. 2, 2026, 8248

### 3.1 *Model interpretability and metallurgical validation (SHAP analysis)*

While high predictive accuracy is essential, it is insufficient for reliable materials design without physical interpretability. The Gradient Boosting model, despite its performance, is inherently a "black box." To bridge the gap between data science and physical metallurgy, we applied Shapley Additive Explanations (SHAP). This game-theoretic approach decomposes the model's predictions to quantify the exact contribution of each feature, thereby revealing the underlying metallurgical drivers identified by the algorithm.

#### *3.2.1 Global feature impact and physical mechanisms*

The SHAP summary plot (Figure 5) provides a global overview of feature importance and directionality. Each point represents a test sample; its position on the x-axis indicates the SHAP value (impact on fatigue strength in MPa), and its color represents the feature's magnitude.

The SHAP summary plot in Figure 5 shows the global hierarchy and directional effect of the input features on fatigue strength. The features are ranked according to their mean absolute SHAP values. The red coloration will indicate high feature values, and the blue color will indicate low feature values. The SHAP analysis provides a clear, physically-based hierarchy of influence. Tempering Temperature (TT) emerged as the dominant factor, showing a strong and unambiguous negative correlation with fatigue strength. High TT values, represented by red dots, correspond to large negative SHAP values, indicating a reduction in the predicted fatigue strength. These results are fully consistent with canonical metallurgical theory, where increasing TT promotes the transformation of martensite into equilibrium ferrite and carbides, carbide coarsening through Ostwald ripening, and recovery processes that reduce dislocation density. Such microstructural developments decrease the static yield strength and, consequently, the high-cycle fatigue limit.

In contrast, carbon content (C) exhibited a strong positive relationship with fatigue strength. High carbon values, shown in red, produced positive SHAP values. This trend reflects the effects of interstitial solid-solution hardening and the increased carbide content in high-carbon martensitic steels, both of which inhibit dislocation movement and retard crack initiation. Elements such as chromium (Cr) and manganese (Mn) showed moderately positive effects, consistent with their known contributions to hardenability and solid-solution strengthening. However, their influence was more dispersed compared to TT and C, suggesting that they function as secondary modulators rather than primary controlling factors within this dataset.

#### *3.2.2 Uncovering synergistic interactions*

A key advantage of SHAP over traditional sensitivity analysis is its ability to reveal complex interactions. The SHAP dependence plots in Figure 6 and Figure 7 illustrate that the impact of one feature is often contingent on the value of another.

According to Figure 6, SHAP dependence plot for Tempering Temperature (TT), showing a negative trend. The vertical dispersion at specific temperatures is explained by the interaction with Carbon (C) content (color scale).

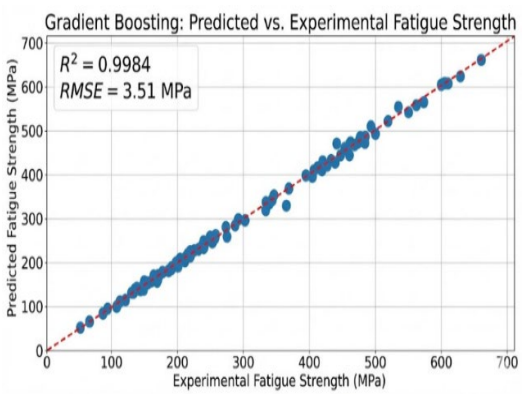


**Figure 5:** SHAP summary plot of gradient boosting model

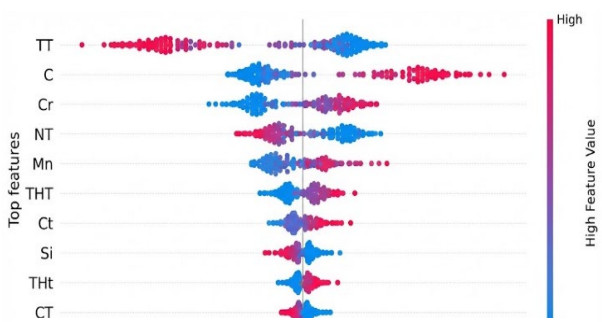


**Figure 6:** SHAP dependence plot for TT colored by C.

Figure 7 shows a dependence plot of Carbon (C), with a positive trend. The interaction between the Tempering Temperature (TT) and the magnitude of the carbon-strengthening effect determines this. Figure 6 indicates that as the tempering temperature (TT) increases, the predicted fatigue strength decreases, although the rate of this decline varies with carbon content. Samples displayed by high carbon (red) at low tempering temperature 0 (e.g., 200 ºC -300 ºC) have much higher positive SHAP values compared to samples with low carbon (blue). This implies that the model has been taught that the strengthening potential of supersaturated carbon can be maximized at low tempering temperatures before any significant carbide precipitation occurs. On the same note, Figure 7 shows that the positive effect of raising Carbon is most intense at low Tempering Temperature (blue dots). At very high tempering temperatures (red dots), the slope of the carbon effect becomes slightly flatter, which means that extensive carbide coarsening can eliminate some of the advantages of high carbon content. These observations affirm that the GB model has embodied nonlinear synergies, which are the main focus of heat treatment design.

*3.2.3 Local explanation of individual predictions*

SHAP allows for the validation of individual predictions. Figure 8 presents a force plot for a single sample predicted to have very high fatigue strength (580 MPa). Figure 8 shows the SHAP force plot of an individual high-strength prediction, showing the feature contributions that drive the model output above the dataset baseline. baseline average. As the plot illustrates, this given prediction is motivated by both synergistic and favorable metallurgical conditions whereby a very high carbon content (C=1.1 wt%), and significant chromium addition (Cr=1.5 wt%) act as strong positive forces (red bars) that uplift the prediction significantly higher than the average of the dataset. Even a moderate low Tempering Temperature (TT=200 ºC) has a positive effect. This degree of granular transparency enables the metallurgists to have confidence that the model is making the "right prediction at the right reasons and not because of arbitrary spurious correlations.

### 3.2 *Comparison with the current state-of-the-art*

To position the proposed framework within the current research landscape, its performance is compared with recent studies that either employ the NIMS fatigue dataset or adopt similar machine-learning strategies for fatigue strength prediction. Emphasis is placed on predictive accuracy, dataset scale, and interpretability. Recent work by Liu *et al.,* [15] applied gradient boosting tree (GBT) models with SHAP analysis to the NIMS fatigue dataset, reporting predictive performance approaching $R^2 \approx 0.98$ while emphasizing interpretability through symbolic regression and feature attribution. Their findings demonstrate the suitability of ensemble learning for fatigue modeling and highlight the importance of integrating feature explanation tools. Quraishy and Kundu [33] expanded the scope by aggregating

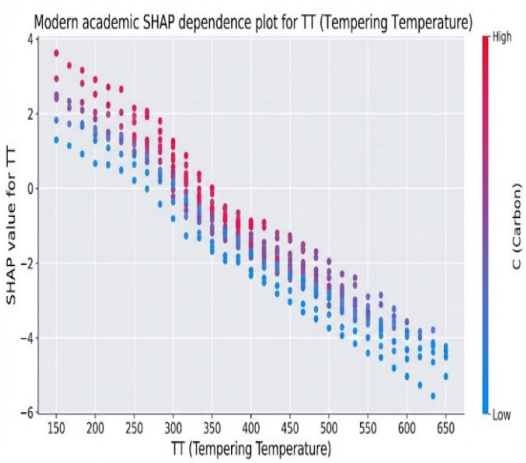


**Figure 7 :** SHAP dependence plot for C colored by TT.

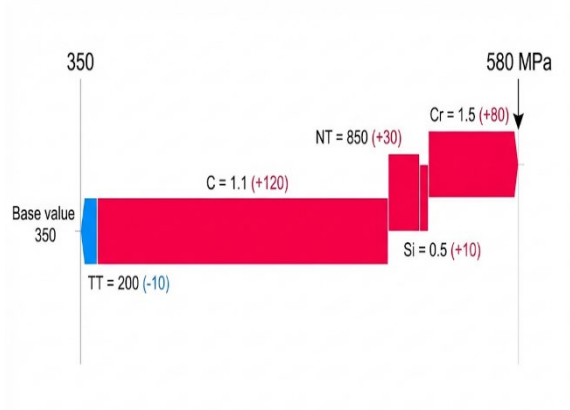


**Figure 8:** SHAP force plot for a high-strength sample.

multiple steel fatigue datasets (3612 samples) and evaluating seven machine-learning algorithms, including Random Forest and XGBoost. Their best model achieved $R^2 = 0.96$, confirming the robustness of tree-based ensembles across broader material domains. However, interpretability was addressed primarily through correlation and partial dependence analysis. Huang *et al.,* [34] introduced a deep learning–based regression model for fatigue prediction in ferrous alloys, reporting competitive predictive performance ($R^2 \approx 0.95$–$0.97$ depending on architecture). While effective, the approach required extensive hyperparameter optimization and offered limited transparency compared with SHAP-based frameworks. A concise comparison is provided in Table 3. Overall, the proposed method achieves predictive accuracy comparable to the best-performing NIMS-based models while incorporating additional validation through Lasso regularization and statistical testing. Compared with larger multi-source datasets, the current framework demonstrates that high accuracy can be maintained without sacrificing interpretability. The integration of SHAP further strengthens the transparency of the model, offering clear insight into how compositional and processing variables influence fatigue strength. Collectively, these comparisons suggest that ensemble learning combined with explainable AI represents a balanced and reliable direction for fatigue modeling research.

## 4 Conclusions

The study highlights the effectiveness of a sophisticated computational model for predicting the fatigue strength of steel, using a large-scale experimental database from the National Institute for Materials Science (NIMS). The experiment involved a strict comparison of various paradigms of algorithmic mapping, ranging from the simplistic linear mappings to sophisticated ensemble learning. This study gives the following important results. Algorithmic Excellence: Gradient Boosting (GB) model uniformly became the most accurate predictive model, with a high coefficient of determination $R^2$=0.9984 and error reduction measures (MAPE = 1.07) 100% lower than using other Predictive Modeling models. Nonlinear Complexity: The large performance gap existing between ensemble methods and nonlinear ma ximum best linear regression between alloying elements and thermal processing history cannot be fully captured by simple regressions. The suggested framework improves the accuracy of fatigue-life prediction and provides a method for designing materials and testing their reliability, both on a larger scale and in a way that is easy to understand. The results are specifically designed for rotating bending fatigue at $10^7$ cycles, based on the NIMS database. Wider applicability will necessitate additional experimental validation to confirm generalization across various fatigue contexts. Predictive Robustness: The framework developed showed a Root Mean Squared Error (RMSE) of 3.51 MPa, which represents a significant improvement in the accuracy of previous data-driven fatigue experiments. Metallurgical Alignment: The interpretability analysis using SHAP confirmed that the models captured the essential physical principles, namely the inverse correlation between tempering temperature and strength, thereby providing high statistical accuracy grounded in metallurgical reality. Overall, this study establishes a validated, interpretable framework for predicting steel fatigue properties using data. This framework accelerates the design of high-performance structural steels and provides a scalable template for predicting various mechanical properties in materials science by reducing reliance on resource-intensive cyclic testing.

**Table 3:** Comparison with Recent Fatigue Strength Prediction Studies

| Study | Dataset | Size | Model | $R^2$ (Best) | Interpretability |
|---|---|---|---|---|---|
| [15] | NIMS fatigue | 437 | GBT + SHAP | ≈ 0.98 | SHAP + symbolic regression |
| [13] | Ferrous alloys | — | Deep learning | 0.95–0.97 | Limited |
| [35] | Multi-source steel fatigue | 3612 | Ensemble (RF/XGB) | 0.96 | Partial |
| Present study | NIMS fatigue | 685 | Gradient Boosting + SHAP + Lasso | ≈ 0.9984 | Full SHAP + feature validation |

**Acknowledgments**

The autours tanks all their afflations

**Fund**

Self Fund

**Author Contributions**

Based on the research roles and author details provided, the contributions are outlined as follows: H.A.F.: Investigation, methodology, and writing original draft; A.A.J.: Supervision, methodology, and data analysis; A.A.F.O.: Conceptualization, investigation, and writing reviewing and editing; A.R.K.J.: Conceptualization, data curation, and writing reviewing and editing; L.A.A. and E.K.N: Conceptualization, funding acquisition, and project administration. All authors have read and agreed to the published version of the manuscript.

**Conflicts of Interest**

The authors declare no conflict of interest.

**Declaration of generative AI and AI-assisted technologies in the writing process**

The authors used AI tools only to improve the language and readability of the manuscript. All research content and conclusions are the authors' own.

**References**

[1] J. A. Bannantine, J. J. Comer, and J. L. Handrock, *Fundamentals of Metal Fatigue Analysis*. Englewood Cliffs, NJ: Prentice Hall, 1990.

[2] J. Schijve, *Fatigue of Structures and Materials*, 2nd ed. Dordrecht, Netherlands: Springer, 2009, doi: 10.1007/978-1-4020-6808-9.

[3] L. M. Nassir et al., "Integrated experimental, statistical, and finite element analysis of nanoparticle-reinforced polymer composites for advanced structural applications completed with bibliometric analysis," *ASEAN Journal for Science and Engineering in Materials*, vol. 5, no. 2, pp. 287–310, 2025. [Online]. Available: https://ejournal.bumipublikasinusantara.id/index.php/ajsem/article/view/849

[4] H. Wang, X. Liu, M. Zhang, Y. Wang, and X. Wang, "Prediction of material fatigue parameters for low alloy forged steels considering error circle," *International Journal of Fatigue*, vol. 121, pp. 135–145, 2019, doi: 10.1016/j.ijfatigue.2018.12. 002.

[5] G. Mi, G. Sun, S. Yang, X. Liu, S. Chen, and W. Kang, "Machine learning-based prediction of fatigue fracture locations in 7075-T651 aluminum alloy friction stir welded joints," *Metals*, vol. 15, no. 5, p. 569, 2025, doi: 10.3390/met15050569.

[6] S. Farhadi, S. Tatullo, and F. Ferrian, "Comparative analysis of ensemble learning techniques for enhanced fatigue life prediction," *Scientific Reports,* vol. 15, Art. no. 11136, 2025, doi: 10.1038/s41598-024-79476-y.

[7] R. Ramprasad, R. Batra, G. Pilania, A. Mannodi-Kanakkithodi, and C. Kim, "Machine learning in materials informatics: recent applications and prospects," *npj Computational Materials*, vol. 3, no. 1, art. no. 54, 2017, doi: 10.1038/s41524-017-0056-5.

[8] A. A. Nayeeif, E. S. Al-Ameen, N. A. Jebur, A. A. Farhan Ogaili, Z. K. Hamdan, and E. K. Njim, "Investigation of the effects of unbalance and bearing wear on shaft vibration in a natural gas turbine plant," *Applied Science and Engineering Progress*, vol. 18, no. 4, Art. no. 7863, Oct. 2025, doi: 10.14416/j.asep.2025.07.0 12.

[9] A. Agrawal, P. D. Deshpande, A. Cecen, G. P. Basavarsu, A. N. Choudhary, and S. R. Kalidindi, "Exploration of data science techniques to predict fatigue strength of steel from composition and processing parameters," *Integrating Materials and Manufacturing Innovation*, vol. 3, no. 1, pp. 90–108, 2014, doi: 10.1186/2193-9772-3-8.

[10] T. Shiraiwa, Y. Miyazawa, and M. Enoki, "Prediction of fatigue strength in steels by linear regression and neural network," *Materials Transactions*, vol. 60, no. 2, pp. 189–198, 2019, doi: 10.2320/matertrans.ME201714.

[11] M. YI et al., "Machine learning for predicting fatigue properties of additively manufactured materials," *Chinese Journal of Aeronautics,* vol. 37, no. 4, pp. 1–22, Apr. 2024, doi: 10.1016/j.cja.2023.11.001.

[12] N. Nagashima, M. Hayakawa, H. Masuda, and K. Nagai, "Estimating the S–N curve by machine learning random forest method," *Materials Transactions*, vol. 65, no. 4, pp. 428–433, 2024, doi: 10.2320/matertrans.MT-Z2023006.

[13] Z. Huang et al., "Deep learning-based fatigue strength prediction for ferrous alloy," *Processes*, vol. 12, no. 10, p. 2214, 2024, doi: 10.3390/pr12102214.

[14] X. Zhang, F. Liu, M. Shen, D. Han, Z. Wang, and N. Yan, "Ultra-high-cycle fatigue life prediction of metallic materials based on machine learning," *Applied Sciences*, vol. 13, no. 4, p. 2524, 2023, doi: 10.3390/app13042524.

[15] C. Liu, X. Wang, W. Cai, J. Yang, and H. Su, "Prediction of the fatigue strength of steel based on interpretable machine learning," *Materials*, vol. 16, no. 23, p. 7354, 2023, doi: 10.3390/ma16237354.

[16] C. D. Naiju, M. Adithan, and P. Radhakrishnan, "An investigation of process variables influencing fatigue properties of components produced by direct metal laser sintering," *Applied Science and Engineering Progress*, vol. 4, no. 1, pp. 63–69, Sep. 2013. [Online]. Available: https://ph02.tci-thaijo.org/index.php/ijast/article/view/67381

[17] B. D. S. Deeraj, K. Joseph, J. S. Jayan, and A. Saritha, "Dynamic mechanical performance of natural fiber reinforced composites: A brief review," *Applied Science and Engineering Progress*, vol. 14, no. 4, pp. 614–623, Oct. 2021, doi: 10.14416/j.asep. 2021. 06.003.

[18] P. Thomas, C. W. Lai, and M. R. Johan, "Prospective of magnesium and alloy-based composites for lightweight railway rolling stocks," *Applied Science and Engineering Progress*, vol. 15, no. 2, p. 5716, May 2022, doi: 10.14416/j.asep. 2022.02.006.

[19] S. Walzer, M. Liewald, N. Simon, J. Gibmeier, H. Erdle, and T. Böhlke, "Improvement of sheet metal properties by inducing residual stresses into sheet metal components by embossing and reforming," *Applied Science and Engineering Progress*, vol. 15, no. 1, p. 5437, Oct. 2021, doi: 10.14416/j.asep.2021.09.006.

[20] A. A. F. Ogaili, Z. T. Al-Sharify, A. A. Jaber, and F. A. Abdulla, "Vibration-based fault detection and classification in ball bearings using statistical analysis and random forest," in *Proceedings of the 5th International Conference on Green Energy, Environment, and Sustainable Development (GEESD 2024)*, vol. 13279, 2024, p. 1327921, doi: 10.1117/12.3041850.

[21] A. S. Abdul-Zahra, E. Ghane, A. Kamali, and A. A. Farhan Ogaili, "Power forecasting in continuous extrusion of pure titanium using Naïve Bayes algorithm," *Terra Joule Journal*, vol. 1, no. 1, p. 2, 2024, doi: 10.64071/3080-5724.1000.

[22] A. A. F. Ogaili, Z. T. Al-Sharify, A. A. Jaber, D. A. Farhan, and S. M. Al-Khafaji, "Effective ball bearing fault diagnosis leveraging ANN and statistical feature integration," in *Proceedings of the 10th Scholar's Yearly Symposium of Technology, Engineering and Mathematics (SYSTEM 2024)*, 2024, pp. 49–60. [Online]. Available: https://ceur-ws.org/Vol-3870/p06.pdf

[23] S. A. Sarow, H. A. Flayyih, M. Bazerkan, L. A. Al-Haddad, Z. T. Al-Sharify, and A. A. F. Ogaili, "Advancing sustainable renewable energy: XGBoost algorithm for the prediction of water yield in hemispherical solar stills," *Discover Sustainability*, vol. 5, no. 1, p. 510, 2024, doi: 10.1007/s43621-024-00782-6.

[24] A.-R. K. Jweri, L. A. Al-Haddad, A. A. F. Ogaili, A. A. Jaber, and M. I. Al-Karkhi, "Edge-intelligent leak detection in water distribution systems using CatBoost: A sustainable solution for reducing infrastructure losses," *Clean Energy Science and Technology*, vol. 3, no. 4, p. 398, 2025, doi: 10.18686/cest398.

[25] B. G. Mejbel, S. A. Sarow, M. T. Al-Sharify, L. A. Al-Haddad, A. A. F. Ogaili, and Z. T. Al-Sharify, "A data fusion analysis and random forest learning for enhanced control and failure diagnosis in rotating machinery," *Journal of Failure Analysis and Prevention*, vol. 24, no. 6, pp. 2979–2989, Dec. 2024, doi: 10.1007/s11668 -024-02075-6.

[26] L. M. Nassir, A. J. Ramadhan, N. T. Al-Sharify, M. I. Khalaf, A. A. F. Ogaili, A. A. Jaber, and Z. T. Al-Sharify, "Robust multi-state EEG cognitive classification via optimized time-domain features and CatBoost," *International Journal of Robotics and Control Systems*, vol. 5, no. 2, pp. 968–989, 2025, doi: 10.31763/ijrcs.v5i2.1340.

[27] A.-R. Jweri et al., "Enhancing predictive maintenance in energy systems using a hybrid Kolmogorov–Arnold network (KAN) with short-time Fourier transform (STFT) framework for rotating machinery," *ASEAN Journal of Science and Engineering*, vol. 5, no. 2, pp. 465–494, 2025, doi: 10.17509/ajse.v5i2. 89023.

[28] Z. W. Metteb et al., "Optimization of hybrid core designs in 3D-printed PLA+ sandwich structures: An experimental, statistical, and computational investigation completed with bibliometric literature review," *Indonesian Journal of Science and Technology*,

vol. 10, no. 2, pp. 207–236, 2025, doi: 10.17509/ijost.v10i2.81743.

[29] A. A. F. Ogaili, K. A. Mohammed, A. A. Jaber, and E. S. Al-Ameen, "Automated wind turbines gearbox condition monitoring: A comparative study of machine learning techniques based on vibration analysis," *FME Transactions*, vol. 52, no. 3, pp. 471–485, 2024, doi: 10.5937/fme2403471O.

[30] K. A. Mohammed, A. A. F. Ogaili, A. W. A. Taha, and A. M. Alsayah, "Evaluating spot welds of dissimilar metals via integrated mechanical testing and finite element modeling," *Applied Engineering Letters*, vol. 10, no. 2, pp. 77–89, 2025, doi: 10.46793/aeletters.2025.10.2.2.

[31] G. A. Muthulingam and V. S. Parvathy, "A novel bacterial foraging optimization based multimodal medical image fusion approach," *Applied Science and Engineering Progress*, vol. 16, no. 4, p. 6794, Oct. 2023, doi: 10.14416/j.asep.2023.03.004.

[32] A. A. F. Ogaili, A. A. Jaber, and M. N. Hamzah, "Statistically optimal vibration feature selection for fault diagnosis in wind turbine blade," *International Journal of Renewable Energy Research*, vol. 13, no. 3, pp. 1082–1092, 2023, doi: 10.20508/ijrer.v13i3.14096.

[33] S. T. Kurdi, L. A. Al-Haddad, and A. A. F. Ogaili, "Path optimization for aircraft based on geographic information systems and deep learning," *Automation*, vol. 7, no. 1, Art. no. 12, Jan. 2026, doi: 10.3390/automation7010012.

[34] S. T. Bunyan et al, "Intelligent thermal condition monitoring for predictive maintenance of gas turbines using machine learning," *Machines*, vol. 13, no. 5, Art. no. 401, May 2025, doi: 10.3390/machines13050401.

[35] K. J. Mohammed, A. A. F. Ogaili, A. -R. K. Jweri, S. Amin, M. Khalaf, L. Al-Haddad, and A. Jaber, "Intelligent fault detection of UAV propellers through time-domain vibration analysis and ensemble learning," *Automated Systems*, vol. 9, pp. 365–381, 2026, doi: 10.1007/s42401-026-00460-7.

[36] E. N. Abbas, E. K. Njim, M. J. Jweeg, R. Madan, A. A. F. Ogaili, and F. T. Al-Maliky, "Experimental and theoretical analysis of mechanical properties of composite materials with diverse reinforcement types," *World Journal of Engineering*, vol. 23, no. 3, pp. 576–586, 2026, doi: 10.1108/WJE-07-2024-0433.

[37] N. T. Al-Sharify, A. F. H. AL-Maliki, A. A. F. Ogaili, A. -R. K. Jweri, A. A. Jaber, and L. A. Al-Haddad, "A hybrid XFEM-ML framework for high-fidelity, rapid fracture prediction in cortical bone: From microstructure to clinical translation," *Mesopotamian Journal of Computer Science*, vol. 6, no. 1, p. 1, 2026, doi: 10.58496/2958-6631.1070.