Qassim University جامعة القصيم

JEIT

Journal of Emerging and Innovative Technologies

Volume 1, Issue 1
July–December 2026

Received
03 September 2026

Accepted
20 October 2026

Published
05 November 2026

DOI
https://doi.org/10.71422/jeit.0007

License
CC BY-NC-SA 4.0

Cite this article

Omar HM, Irfan OM. Machine Learning–Based Flow Stress Prediction and Processing Map Optimization of Al/SiCp Metal Matrix Composites. J Emerg Innov Technol. 2026;1(1):79–105.

Journal website

Research article

# Machine Learning–Based Flow Stress Prediction and Processing Map Optimization of Al/SiCp Metal Matrix Composites

Hanafy M. Omar[1], Osama M. Irfan[1,*]

[1] Department of Mechanical Engineering, College of Engineering, Qassim University, Buraydah 51452, Saudi Arabia

*Corresponding author: o.ahmed@qu.edu.sa

## Abstract

This study benchmarks machine learning (ML) models for predicting the hot-deformation flow stress of a powder-metallurgy 2124 Al/10 vol.% $SiC_p$ composite. The dataset comprises 100 flow-stress values compiled from published hot-compression data [37] at 300–500 °C, 0.001–1.0 $s^{-1}$ and true strains of 0.1–0.5, forming a complete 5 × 4 × 5 grid. Eight algorithms (linear regression, random forest, gradient boosting, XGBoost, a neural network, LightGBM, CatBoost and a stacked ensemble) were trained on 15 features derived from temperature, strain rate, strain and the Zener–Hollomon parameter (activation energy 162 kJ $mol^{-1}$, obtained from a hyperbolic-sine fit). Each model was evaluated over 20 random, temperature-stratified 70/15/15 splits and by leave-one-temperature-out and leave-one-strain-rate-out tests. Gradient boosting and the stacked ensemble gave the lowest interpolation error (test MAPE = 5.2 ± 2.8% and 5.3 ± 2.5%; $R^2 \approx 0.997$), against 14.3% for the hyperbolic-sine equation fitted to the same data and 17.0 ± 6.9% for linear regression; differences among the boosted models were smaller than the split-to-split scatter. Errors were concentrated at the low stresses measured at 500 °C, and they increased markedly when a whole temperature or strain-rate level was withheld (mean MAPE of 25–33% and 15–27%, respectively, for the non-linear models). Grouped permutation importance attributed 83 ± 3% of the predictive importance to the Zener–Hollomon features. Because tree ensembles are piecewise constant, processing maps were derived from a Gaussian-process surrogate that provides smooth strain-rate sensitivities and their uncertainty. The highest efficiency of power dissipation ($\eta \approx 53$–$68\%$) occurs at 475–500 °C and about 0.002–0.01 $s^{-1}$, whereas the instability boundary is poorly constrained by four strain rates. The results apply to interpolation within the tested envelope of a single composite.

**Keywords:** Metal Matrix Composites; Flow Stress Prediction; Machine Learning; Ensemble Methods; Processing Maps; Gaussian Process.

## 1. Introduction

In the metalworking industry, accurate prediction of flow stress behavior during hot deformation is essential for manufacturing process optimization [1].

© 2026 The Author(s). Published by Qassim University under CC BY-NC-SA 4.0.

Flow stress, which represents the resistance of materials to plastic deformation at elevated temperatures, governs the design of forming equipment, determination of processing parameters, and prediction of final product quality [2], [3]. Understanding and predicting flow stress behavior is particularly essential for developing processing maps that delineate safe and unsafe hot working conditions [4]. The use of conventional aluminum alloys in automotive, aerospace and defense components is limited mainly by their loss of strength and stiffness at elevated temperatures; the addition of ceramic reinforcing particles is an established way of mitigating this limitation [5]. Metal matrix composites (MMCs) represent an important class of advanced materials that combine the properties of metallic matrices with ceramic reinforcements, offering superior specific strength, stiffness, and elevated temperature performance compared to conventional alloys [6], [7]. Particulate-reinforced aluminum composites offer enhanced stiffness, wear resistance, and elevated-temperature strength, making them suitable for aerospace, automotive, and defense applications [8], [9]. Hot workability is the ability of a material to deform plastically at elevated temperatures without flow localization, cracking or fracture [15]. Accurate characterization of hot deformation behavior is essential for designing thermomechanical processing routes and ensuring defect-free manufacturing [10]. For Al/SiC❑$_p$ composites, conventional trial-and-error approaches to process optimization are inefficient, costly, and often unreliable due to the strong nonlinearity between processing parameters and material response [11].

Composites combine constituents so as to obtain combinations of properties that no single constituent provides [12]. Among various MMC systems, aluminum-based composites reinforced with silicon carbide particles have gained significant attention for aerospace and automotive applications due to their excellent strength-to-weight ratio and thermal stability [13], [14].

### 1.1 Hot Deformation Behavior and Processing Maps

Understanding the hot deformation behavior of MMCs is essential for designing efficient thermomechanical processing routes [16], [17]. Work hardening, dynamic recovery, and dynamic recrystallization mechanisms interact intricately to control the flow stress during hot working [3], [18]. Processing maps, developed based on the Dynamic Materials Model (DMM), provide a valuable framework for identifying optimal processing windows and unstable deformation regimes [19], [4]. For 2124 Al/SiC$_p$ composites, processing maps have been used to examine the effect of the reinforcement volume fraction on hot workability [15]. The development of accurate processing maps requires precise flow stress data across a wide range of temperatures and strain rates. Traditional approaches rely on constitutive equations, such as the hyperbolic sine law proposed by Sellars and McTegart [20], which relates flow stress to the Zener-Hollomon parameter [18], [20]:

$$Z = \dot{\varepsilon} \exp\left(\frac{Q}{RT}\right) = A¿ \tag{1}$$

where $Z$ is the Zener–Hollomon parameter, $\dot{\varepsilon}$ is the strain rate, $Q$ is the activation energy, $R$ is the gas constant, $T$ is absolute temperature, and $A$, $\alpha$, and $n$ are material constants. The Zener-Hollomon parameter serves as a master variable that captures the temperature-compensated strain rate, providing a unified descriptor for thermally activated plastic deformation processes.

### 1.2 Limitations of Conventional Modeling Approaches

In order to describe the flow stress of metals during hot deformation, many mathematical models have been proposed. The conventional method is to carry out regression with experimental data based on the

© 2026 The Author(s). Published by Qassim University under CC BY-NC-SA 4.0.

models to obtain the constants in the models. However, most of the factors affecting the flow stress are nonlinear, which makes the accuracy of the flow stress predicted using the regression method relatively low. Conventional regression-based methods for flow stress prediction suffer from several limitations [21], [2]. First, they demonstrate poor accuracy in capturing nonlinear relationships, with traditional constitutive models typically achieving mean absolute percentage error (MAPE) values of 10–15% [22]. For example, regression analysis applied to 2124 Al/10% SiCp composites yielded mean errors of up to 13.5% [37], indicating that the accuracy of flow stress predictions using regression methods remains unsatisfactory for many engineering applications. Second, there is difficulty in modeling complex interactions, as the interplay between temperature, strain rate, strain, and microstructural evolution is highly nonlinear [23]. Third, regression models exhibit limited extrapolation capability and often fail outside the experimental data range [24]. Finally, determining optimal material constants requires iterative optimization procedures involving manual parameter tuning [25].

### 1.3 Soft Computing and Machine Learning Approaches

To address the limitations of conventional approaches, soft processing techniques including fuzzy logic, artificial neural networks (ANNs), and genetic algorithms have been explored for flow stress prediction [26], [27], [28]. Fuzzy logic systems have been applied to flow stress prediction by modeling the approximate and imprecise reasoning processes common in human problem-solving. A fuzzy logic unit is made up of an inference engine, a fuzzy rule base, membership functions, a defuzzifier, and a fuzzifier. The concept of fuzzy reasoning enables input-output mapping of flow stress data through if-then rules that handle uncertainty and vague information effectively. Previous studies have demonstrated that fuzzy logic systems can effectively model the nonlinear relationships in flow stress data [27]. Recent applications of fuzzy logic to metal forming include work on 2124 Al/10% SiC$❑_p$ composites, where fuzzy logic achieved a mean deviation of 6.1% compared to 13.5% for regression methods [37]. However, fuzzy systems require expert knowledge for membership function design and rule base construction, introducing subjectivity and limiting their generalizability across different material systems. Artificial neural networks have demonstrated improved performance over traditional models in several recent studies. Mandal et al. [22] developed ANN models for predicting high-temperature flow stress in AISI 304L stainless steel during hot torsion, with a reported error of approximately 6.0%. Quan et al. [29] compared ANN predictions with Arrhenius-type models for AA7050 aluminum alloy, reporting MAPE of 5.2% and a lower error than that of the Arrhenius-type model. Ji et al. [30] applied artificial neural networks to Alloy 617 flow stress prediction, achieving MAPE of 7.1%. Haghdadi et al. [21] applied ANNs to predict microstructural evolution during thermomechanical processing, demonstrating the potential of data-driven approaches in materials engineering. However, ANNs often suffer from overfitting when training data is limited and provide limited interpretability regarding which input features drive predictions, raising concerns for safety-critical applications. Ensemble tree algorithms such as random forests [31], gradient boosting [32] and XGBoost [33] have been applied increasingly in materials science [34], [35]. Ensemble methods combine multiple base learners to produce superior predictions through variance reduction and complementary learning. The fundamental principle is that aggregating diverse models with different inductive biases can capture complex patterns while maintaining robustness against individual model weaknesses. Ensemble methods have been applied to predict mechanical properties of metal matrix composites, with reported improvements over single-model approaches. Li et al. [36] applied XGBoost for predicting flow stress in Al-Zn-Mg-Cu aluminum alloys, achieving MAPE of 3.9%. The sequential boosting

© 2026 The Author(s). Published by Qassim University under CC BY-NC-SA 4.0.

mechanism with regularization controls the bias–variance trade-off, which often gives good accuracy on small tabular datasets. However, systematic comparative studies of these advanced ML techniques for flow stress prediction in MMCs remain limited.

### 1.4 Model Interpretability

A critical advancement in modern machine learning applications is model interpretability. The "black box" nature of complex models has been a persistent concern in safety-critical applications such as aerospace manufacturing. SHAP (SHapley Additive exPlanations) values provide a unified framework for explaining individual predictions [38]. Accumulated local effects plots have been developed to visualize feature effects when the inputs are correlated, a situation in which classical partial dependence plots can mislead [39]. Impurity-based importance measures built into tree ensembles are biased, and permutation importance is generally preferred [40].

### 1.5 Research Objectives

Despite these advances, several critical gaps remain in the literature. First, comprehensive comparative evaluations of multiple machine learning algorithms for flow stress prediction in aluminum composites are scarce, with most studies focusing on single algorithms without systematic comparison. Second, rigorous hyperparameter optimization using modern techniques such as extensive grid search or Bayesian optimization is often neglected, leading to suboptimal model performance. Third, ensemble methods specifically designed for constitutive modeling through stacked generalization remain underexplored, despite their proven effectiveness in other domains. Fourth, most studies report accuracy from a single random train–test split and do not test predictions at temperatures or strain rates that were withheld from training, so the reported errors mix interpolation with chance and say little about extrapolation. Finally, physics-based interpretability analysis to validate that models learn actual deformation mechanisms rather than spurious correlations is rarely performed, limiting confidence in model extrapolation beyond training data.

This study addresses these gaps through a systematic benchmarking study for 2124 Al/10 vol.% $SiC_p$ composites. The specific objectives are to: (1) compare eight machine learning algorithms (linear regression, random forest, gradient boosting, XGBoost, neural network, LightGBM, CatBoost and a stacked ensemble) under a repeated-split protocol that reports the split-to-split variability; (2) optimize the hyperparameters of every model by grid search with 5-fold cross-validation on the training data only; (3) build a stacked-generalization ensemble whose ridge meta-learner is fitted on out-of-fold predictions; (4) distinguish interpolation from extrapolation with leave-one-temperature-out and leave-one-strain-rate-out tests; (5) analyze interpretability with grouped permutation importance and partial dependence, and compare the learned response quantitatively with the hyperbolic-sine constitutive equation fitted to the same data; (6) quantify the difference between the ML models and the regression and fuzzy-logic approaches previously applied to this composite; and (7) construct processing maps from a smooth Gaussian-process surrogate, together with their uncertainty.

## 2. Data Source and Experimental Background

### 2.1 Material and Data Source

The flow-stress data analyzed in this study were not generated by the authors. They were compiled from a published hot-compression study of a 2124 Al/10 vol.% $SiC_p$ composite [37], which also reports the

© 2026 The Author(s). Published by Qassim University under CC BY-NC-SA 4.0.

regression (13.5%) and fuzzy-logic (6.1%) errors used as benchmarks in Sections 1 and 4.7. As described in [37], the composite was produced by the powder-metallurgy route: 2124 aluminum alloy powder was blended with 10 vol.% silicon carbide particles (average size 25 μm) in a tubular mixer for 4 hours.

The nominal composition limits of the 2124 matrix alloy [41] are listed in Table 1. 2124 is a high-strength Al–Cu–Mg alloy used in aerospace applications because of its high strength-to-weight ratio.

**Table 1.** Nominal chemical composition limits of 2124 aluminum alloy (wt.%; single values are maxima) [41]

| Cu | Mg | Mn | Si | Fe | Cr | Zn | Ti | Others (total) | Al |
|---|---|---|---|---|---|---|---|---|---|
| 3.8–4.9 | 1.2–1.8 | 0.30–0.9 | 0.20 | 0.30 | 0.10 | 0.25 | 0.15 | 0.15 | Balance |

In the source study [37], the powder mixture was degassed at 450 °C for 1 hour under vacuum pressure of $10^{-4}$ torr to remove moisture and adsorbed gases. Vacuum hot pressing was performed at 450 °C under 75 MPa pressure to produce cylindrical compacts with dimensions of 10 mm diameter × 15 mm height. The green compacts were subsequently sintered at 500 °C for 2 hours in a vacuum atmosphere to achieve full densification and ensure strong interfacial bonding between the aluminum matrix and SiC reinforcement particles.

### 2.2 Hot Compression Test Conditions

In the source study [37], isothermal hot compression tests were performed on a servo-controlled universal testing machine of 100 kN capacity equipped with a three-zone resistance furnace. Specimens were heated to the test temperature (300, 350, 400, 450 and 500 °C) at 10 °C/min and held for 5 min; the temperature, monitored with K-type thermocouples, was maintained within ±2 °C.

The tests were performed at constant true strain rates of 0.001, 0.01, 0.1 and 1.0 $s^{-1}$, spanning three orders of magnitude, to a true strain of 0.5, with graphite foil as lubricant. Load–displacement data were converted to true stress–true strain curves, from which the flow stresses at true strains of 0.1–0.5 were taken.

### 2.3 Test Matrix

The dataset covers five strain levels (0.1, 0.2, 0.3, 0.4 and 0.5) for every combination of temperature and strain rate (Table 2), giving 100 flow-stress values (5 temperatures × 4 strain rates × 5 strain levels) that form a complete factorial grid over the processing space relevant to hot extrusion, forging and rolling. The grid structure has consequences for model validation, which are addressed in Section 3.2.

**Table 2.** Flow stress (MPa) of 2124 Al/10 vol.% SiCp at various temperatures, strain rates and strains, compiled from [37] (values corrected for deformation heating; see Section 2.4)

| Temperature (°C) | True strain | Strain rate ($s^{-1}$) | | | |
|---|---|---|---|---|---|
| | | 0.001 | 0.01 | 0.1 | 1.0 |
| 300 | 0.1 | 92.0 | 120.32 | 156.5 | 215.2 |
| | 0.2 | 87.78 | 118.62 | 152.8 | 212.5 |
| | 0.3 | 81.23 | 116.32 | 149.0 | 200.8 |
| | 0.4 | 77.82 | 115.37 | 145.4 | 195.62 |
| | 0.5 | 76.28 | 114.32 | 144.4 | 187.54 |

© 2026 The Author(s). Published by Qassim University under CC BY-NC-SA 4.0.

| Temperature (°C) | True strain | Strain rate ($s^{-1}$) | | | |
|---|---|---|---|---|---|
| | | 0.001 | 0.01 | 0.1 | 1.0 |
| 350 | 0.1 | 62.30 | 84.91 | 124.71 | 172.68 |
| | 0.2 | 59.91 | 82.87 | 123.35 | 169.32 |
| | 0.3 | 57.42 | 81.51 | 120.30 | 163.26 |
| | 0.4 | 55.91 | 80.32 | 118.10 | 157.85 |
| | 0.5 | 56.04 | 79.12 | 116.24 | 151.25 |
| 400 | 0.1 | 42.82 | 64.97 | 88.47 | 137.77 |
| | 0.2 | 41.62 | 64.83 | 87.21 | 133.62 |
| | 0.3 | 40.57 | 63.51 | 85.45 | 130.28 |
| | 0.4 | 41.30 | 61.89 | 83.92 | 125.28 |
| | 0.5 | 41.56 | 60.52 | 81.90 | 120.01 |
| 450 | 0.1 | 25.12 | 41.46 | 63.87 | 107.62 |
| | 0.2 | 27.28 | 41.74 | 63.03 | 105.03 |
| | 0.3 | 25.52 | 39.65 | 61.55 | 101.92 |
| | 0.4 | 23.89 | 38.85 | 59.12 | 97.75 |
| | 0.5 | 24.03 | 37.66 | 57.94 | 94.36 |
| 500 | 0.1 | 5.54 | 16.75 | 32.95 | 70.78 |
| | 0.2 | 8.10 | 17.82 | 32.13 | 67.67 |
| | 0.3 | 9.72 | 17.42 | 30.37 | 62.62 |
| | 0.4 | 7.95 | 16.72 | 28.75 | 54.28 |
| | 0.5 | 5.75 | 16.61 | 26.53 | 45.92 |

### 2.4 Deformation-Heating Correction

The flow stresses in Table 2 were corrected for deformation heating. The temperature rise at strain $\varepsilon$ was estimated as

$$\Delta T = \frac{\beta}{\rho C_p} \int_0^{\varepsilon} \sigma \, d\varepsilon \tag{2}$$

where $\beta$ = 0.9 is the fraction of plastic work converted into heat (Taylor–Quinney coefficient), $\rho$ = 2780 kg/m³ is the density and $C_p$ = 875 J/(kg·K) is the specific heat capacity of the composite. The corresponding stress correction is

$$\Delta\sigma = -\left(\frac{\partial \sigma}{\partial T}\right)_{\dot{\varepsilon},\varepsilon} \Delta T \tag{3}$$

where the temperature sensitivity $\partial\sigma/\partial T$ at constant strain rate and strain is evaluated by finite differences of the data in Table 2. Equation (2) assumes that all of the heat is retained in the specimen and is therefore an upper bound. At 0.001 $s^{-1}$ (test duration ≈ 500 s) heat is conducted to the dies and the deformation is nearly isothermal; adiabatic correction factors that increase from about zero at low strain rates to about one at 1–10 $s^{-1}$ describe this transition more realistically [42]. Table 3 lists the upper-bound values of $\Delta T$ and $\Delta\sigma$ at $\varepsilon$ = 0.5 for every test condition. The correction is 0.5–6.3 MPa (6–9% of the flow stress) at 0.001 $s^{-1}$ and 11–28 MPa (12–24%) at 1.0 $s^{-1}$, where the temperature rise reaches 38

© 2026 The Author(s). Published by Qassim University under CC BY-NC-SA 4.0.

K at 300 °C. Because the correction was applied with $\beta = 0.9$ at all strain rates, the flow stresses at the lowest strain rates may be slightly over-corrected; this is noted as a limitation in Section 4.10.

**Table 3.** Upper-bound temperature rise (ΔT) and stress correction (Δσ) at ε = 0.5 estimated from Eqs. (2) and (3), with Δσ as a percentage of the flow stress

| Temperature (°C) | Strain rate ($s^{-1}$): $\Delta T$ (K) / $\Delta\sigma$ (MPa) ($\Delta\sigma/\sigma$) | | | |
|---|---|---|---|---|
| | 0.001 | 0.01 | 0.1 | 1.0 |
| 300 | 15.6 / 6.3 (8%) | 21.8 / 15.3 (13%) | 27.9 / 15.7 (11%) | 37.9 / 27.5 (15%) |
| 350 | 10.9 / 3.8 (7%) | 15.2 / 8.2 (10%) | 22.5 / 14.0 (12%) | 30.5 / 20.6 (14%) |
| 400 | 7.7 / 2.5 (6%) | 11.8 / 4.9 (8%) | 15.9 / 9.3 (11%) | 24.3 / 13.8 (12%) |
| 450 | 4.7 / 1.7 (7%) | 7.4 / 3.3 (9%) | 11.4 / 6.3 (11%) | 19.0 / 14.1 (15%) |
| 500 | 1.4 / 0.5 (9%) | 3.2 / 1.3 (8%) | 5.7 / 3.6 (13%) | 11.6 / 11.2 (24%) |

### 2.5 Data Characteristics

The flow stress data exhibit strong dependencies on thermomechanical conditions that are consistent with established hot deformation behavior of aluminum alloys. The flow stress range spans 5.54–215.20 MPa, representing a 39-fold variation across the processing window. Temperature exerts a strong softening effect, with increasing temperature from 300 °C to 500 °C reducing flow stress by 67–94% due to enhanced thermal activation of dislocation motion, dynamic recovery, and potential dynamic recrystallization. Conversely, the strain rate effect shows that increasing strain rate from 0.001 to 1.0 $s^{-1}$ increases flow stress by 134% (300 °C, $\varepsilon = 0.1$) to more than ten-fold (500 °C, $\varepsilon = 0.1$) due to the rate sensitivity of thermally-activated deformation processes, where higher deformation rates outpace recovery mechanisms. The strain dependence generally shows flow stress decreasing with increasing strain, particularly at higher temperatures, indicating dynamic softening processes dominate over work hardening. These trends are consistent with thermally activated processes such as dislocation climb, cross-slip, and dynamic recovery/recrystallization that govern flow stress evolution in aluminum alloys and their composites at elevated temperatures.

## 3. Methodology

### 3.1 Feature Engineering

Physics-informed features were derived from the three primary variables, temperature ($T$, K), strain rate ($\dot{\varepsilon}$, $s^{-1}$) and strain ($\varepsilon$). The central derived variable is the Zener–Hollomon parameter, $Z = \dot{\varepsilon}\exp(Q/RT)$ (Eq. 1), which combines temperature and strain rate into a single temperature-compensated strain rate.

The activation energy $Q$ was determined from the data in Table 2 following Sellars and McTegart [20]. The constants $n_1$ and $\beta'$ were obtained as the mean inverse slopes of $\ln\sigma$ and $\sigma$, respectively, against $\ln\dot{\varepsilon}$ over the five temperatures, giving $\alpha = \beta'/n_1$; $n$ is the mean inverse slope of $\ln[\sinh(\alpha\sigma)]$ against $\ln\dot{\varepsilon}$, and $Q = R\,n\,S$, where $S$ is the mean slope of $\ln[\sinh(\alpha\sigma)]$ against $1/T$. With all strains pooled, this gives $\alpha = 0.0148$ $MPa^{-1}$, $n = 3.86$ and $Q = 161.7$ kJ $mol^{-1}$; a linear regression of $\ln Z$ on $\ln[\sinh(\alpha\sigma)]$ then gives a slope of 3.66 and $\ln A = 24.88$ ($A$ in $s^{-1}$) with $R^2 = 0.954$. Fitting each strain separately gives $Q = 156.8$–$173.5$ kJ $mol^{-1}$. The rounded value $Q = 162$ kJ $mol^{-1}$ was used for all Z-based features, and $R = 8.314$ J/(mol·K). The fitted value is higher than the activation energy for lattice self-diffusion in aluminum (about 142 kJ $mol^{-1}$), as is common for particle-reinforced aluminum composites. The strain-dependent hyperbolic-sine

© 2026 The Author(s). Published by Qassim University under CC BY-NC-SA 4.0.

equation (constants fitted separately at each strain) reproduces Table 2 with MAPE = 14.3% (15.0% with a single set of constants); it serves as the physics-based baseline in Section 4.

Additional derived features are the temperature-compensation term $\exp(Q/RT)$; the logarithmic transformations $\ln \dot{\varepsilon}$, $\ln T$ and $\ln Z$; the interaction terms $T \times \dot{\varepsilon}$, $T \times \varepsilon$, $\dot{\varepsilon} \times \varepsilon$ and $T \times \dot{\varepsilon} \times \varepsilon$; and the quadratic terms $T^2$, $\dot{\varepsilon}^2$ and $\varepsilon^2$. Together with $T$, $\dot{\varepsilon}$, $\varepsilon$ and $Z$, the feature matrix contains 15 variables. All of them are deterministic functions of $T$, $\dot{\varepsilon}$ and $\varepsilon$, so several are strongly correlated (Section 4.3); this affects the importance analysis (Section 3.7).

### 3.2 Data Splitting, Repeated Evaluation and Extrapolation Tests

For the interpolation tests, the 100 data points were divided into training (70 points), validation (15) and test (15) sets by random sampling stratified on temperature, so that each of the five temperatures contributes 14, 3 and 3 points to the three sets. The complete pipeline (splitting, scaling, hyperparameter search, training and evaluation) was repeated for 20 random seeds, and results are reported as the mean ± standard deviation (SD) over the 20 splits. Because the data form a complete grid, a random split leaves almost every test point with training points at the same temperature and strain rate, so random-split scores measure interpolation, mainly along the strain axis. Two grouped cross-validation schemes were therefore added: leave-one-temperature-out (LOTO; five folds, each withholding all 20 points at one temperature) and leave-one-strain-rate-out (LOSRO; four folds, each withholding all 25 points at one strain rate). Withholding 300 or 500 °C, or 0.001 or 1.0 $s^{-1}$, tests extrapolation; withholding an interior level tests interpolation across a gap of 50 °C or one decade of strain rate. Features were standardized with z-scores:

$$x_{scaled} = \frac{x - \mu}{\sigma} \tag{4}$$

where $\mu$ and $\sigma$ are mean and standard deviation computed from the training set only to prevent data leakage. This standardization is critical for gradient-based optimization algorithms and distance-based methods to converge efficiently.

### 3.3 Machine Learning Algorithms

Eight regression models were compared. Ordinary least squares linear regression on the engineered features serves as the baseline. Random forest [31] averages the predictions of deep regression trees grown on bootstrap samples with random feature subsets. Gradient boosting [32], XGBoost [33], LightGBM and CatBoost build additive ensembles of shallow trees, each new tree being fitted to the negative gradient of the loss with respect to the current prediction; LightGBM grows trees leaf-wise with histogram-based splits, and CatBoost uses symmetric trees and ordered boosting. The neural network is a multilayer perceptron with two hidden layers of 50 tanh units, trained with the Adam optimizer on standardized inputs and a standardized target. XGBoost minimizes the regularized objective

$$L(\phi) = \sum_i l(\hat{y}_i, y_i) + \sum_k \Omega(f_k) \tag{5}$$

where $l$ is the squared-error loss, $\hat{y}_i$ is the prediction for sample $i$, $f_k$ is the $k$-th tree, and the complexity penalty of each tree is

$$\Omega(f) = \gamma T + \frac{1}{2}\lambda \sum_{j=1}^{T} w_j^2 \tag{6}$$

© 2026 The Author(s). Published by Qassim University under CC BY-NC-SA 4.0.

where $T$ here denotes the number of leaves, $w_j$ are the leaf weights, and $\gamma$ and $\lambda$ are regularization parameters. Regularization matters in the present problem because, with only 70 training points, unpenalized trees can reproduce the training data almost exactly (Section 4.1). The fixed settings, the search grids and the most frequently selected values for all models, including LightGBM and CatBoost, are listed in Table 4.

**Table 4.** Model settings, hyperparameter grids and the values most frequently selected over the 20 splits (5-fold cross-validation on the training set)

| Model | Fixed settings | Grid searched | Most frequently selected |
|---|---|---|---|
| Linear regression | Ordinary least squares; standardized features | — | — |
| Random forest | 200 trees | Max. depth (8, unlimited); min. samples per leaf (1, 2) | Depth 8 (11/20); leaf 1 (20/20) |
| Gradient boosting | 300 trees | Learning rate (0.05, 0.1); max. depth (2, 3) | 0.1 (20/20); depth 2 (19/20) |
| XGBoost | 300 trees; column subsample 0.8; $\lambda = 1$ | Learning rate (0.05, 0.1); max. depth (2, 3) | 0.1 (17/20); depth 2 (15/20) |
| Neural network | 2 × 50 tanh units; Adam; ≤ 5000 iterations; standardized target | L2 penalty ($10^{-4}$, $10^{-2}$); initial learning rate ($10^{-3}$, $10^{-2}$) | $10^{-2}$ (16/20); $10^{-2}$ (19/20) |
| LightGBM | 300 trees; min. 3 samples per leaf | Learning rate (0.05, 0.1); leaves (4, 8) | 0.1 (19/20); 4 leaves (14/20) |
| CatBoost | 300 iterations | Learning rate (0.05, 0.1); depth (3, 4) | 0.1 (20/20); depth 3 (19/20) |
| Enhanced ensemble | Base learners: tuned XGBoost, gradient boosting, random forest, neural network; ridge meta-learner | Ridge penalty: 13 values, $10^{-3}$–$10^{3}$ (internal cross-validation) | — |

### 3.4 Enhanced Ensemble Model

The enhanced ensemble uses stacked generalization [43]. The four base learners (XGBoost, gradient boosting, random forest and the neural network), each with the hyperparameters selected in Section 3.5, generate out-of-fold predictions for the 70 training points by 5-fold cross-validation. These predictions form the meta-feature matrix $\boldsymbol{H} = [\hat{y}_{XGB}, \hat{y}_{GB}, \hat{y}_{RF}, \hat{y}_{NN}]$, to which a ridge-regression meta-learner is fitted:

$$\min_{\alpha} \| y - H\alpha \|_2^2 + \lambda \| \alpha \|_2^2 \tag{7}$$

where $\alpha$ are the ensemble weights and the penalty $\lambda$ is selected by internal cross-validation. The base learners are then refitted on the whole training set, and the final prediction is

$$\hat{y}_{ensemble} = \sum_{k=1}^{4} \alpha_k \hat{y}_k \tag{8}$$

Fitting the meta-learner on out-of-fold predictions, rather than on the validation set, keeps the validation and test sets independent of every fitting step. Averaged over the 20 splits, the ridge weights were 0.72 ± 0.33 (gradient boosting), 0.29 ± 0.32 (XGBoost), 0.13 ± 0.10 (neural network) and −0.13 ± 0.15 (random forest); the large scatter and the negative random-forest weight indicate that the base predictions are highly collinear.

© 2026 The Author(s). Published by Qassim University under CC BY-NC-SA 4.0.

### 3.5 Hyperparameter Optimization

For every split, the hyperparameters were selected by grid search with 5-fold cross-validation on the 70 training points only, using the mean root-mean-square error over the folds as the criterion, and the selected model was refitted on the whole training set. The validation set was used only to report validation scores; it was not used for early stopping, model selection or fitting the meta-learner. In the LOTO and LOSRO tests the same search was repeated inside each training fold. The grids (Table 4) are deliberately small, because 70 training points cannot support fine tuning.

### 3.6 Performance Metrics

Model performance was quantified using multiple complementary metrics to provide comprehensive evaluation. The coefficient of determination ($R^2$) measures the proportion of variance explained:

$$R^2 = 1 - \frac{\sum_{i=1}^{n} (y_i - \hat{y}_i)^2}{\sum_{i=1}^{n} (y_i - \acute{y})^2} \tag{9}$$

Root mean square error (RMSE) quantifies average prediction error magnitude in the original units (MPa):

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n} (y_i - \hat{y}_i)^2} \tag{10}$$

Mean absolute error (MAE) provides a robust measure less sensitive to outliers:

$$MAE = \frac{1}{n}\sum_{i=1}^{n} ¿\, y_i - \hat{y}_i \vee ¿ \tag{11}$$

Mean absolute percentage error (MAPE) is reported because it is widely used in the literature, but it weights errors at low stresses heavily: an error of 4 MPa is 72% of the 5.54 MPa flow stress at 500 °C and 0.001 $s^{-1}$, but only 2% of the 215.2 MPa flow stress at 300 °C and 1.0 $s^{-1}$. It is therefore always reported together with RMSE and MAE:

$$MAPE = \frac{100\,\%}{n}\sum_{i=1}^{n} \left| \frac{y_i - \hat{y}_i}{y_i} \right| \tag{12}$$

### 3.7 Model Interpretability Analysis

Feature importance was evaluated for XGBoost by permutation importance on the test set of each of the 20 splits, with 30 permutations per split and the increase in RMSE as the error measure:

$$Importance(X_j) = \frac{1}{K}\sum_{k=1}^{K} [¿\, Error(perm(X_j)_k) - Error(X)] ¿ \tag{13}$$

where $K$ is the number of permutations. Because $Z$ and $\ln Z$ are monotonic transformations of each other (Spearman $\rho$ = 1.00) and several other features are strongly correlated (Section 4.3), permuting one feature at a time splits the credit among correlated features arbitrarily. The features were therefore permuted jointly in five physically defined groups: the Zener–Hollomon group ($Z$, $\ln Z$); the temperature group ($T$, $\ln T$, $T^2$, $\exp(Q/RT)$); the strain-rate group ($\dot{\varepsilon}$, $\ln \dot{\varepsilon}$, $\dot{\varepsilon}^2$); the strain group ($\varepsilon$, $\varepsilon^2$); and the interaction group ($T \times \dot{\varepsilon}$, $T \times \varepsilon$, $\dot{\varepsilon} \times \varepsilon$, $T \times \dot{\varepsilon} \times \varepsilon$). Importances are reported as the increase in test RMSE (MPa) and as a share of the total, with the SD across the 20 splits. Impurity-based importances were not used because

© 2026 The Author(s). Published by Qassim University under CC BY-NC-SA 4.0.

they are biased and are not on the same scale as permutation importances [40]. Partial dependence describes the marginal effect of a variable on the prediction, averaged over the other inputs [32]:

$$PD(x_j)=\frac{1}{n}\sum_{i=1}^{n} f(x_j, x_{i,-j}) \tag{14}$$

where $x_{i,-j}$ denotes the values of all inputs except $x_j$. Because the engineered features are functions of $T$, $\dot{\varepsilon}$ and $\varepsilon$, partial dependence was computed for the three primary variables, with all derived features recomputed consistently for each substituted value.

### 3.8 Gaussian-Process Surrogate for Processing Maps

The processing-map parameters (Section 4.9) require the strain-rate sensitivity $m = \partial \ln \sigma / \partial \ln \dot{\varepsilon}$ and its derivative. Tree ensembles are piecewise constant in each input: on a 41 × 61 grid of temperature and ln $\dot{\varepsilon}$ at $\varepsilon = 0.5$, the XGBoost prediction gave $m = 0$ at 47% of the grid steps and $m < 0$ at 23% of them, so it cannot provide a usable derivative. A Gaussian-process (GP) regression model [44] was therefore fitted to ln $\sigma$ with the standardized inputs $1000/T$, ln $\dot{\varepsilon}$ and $\varepsilon$, a squared-exponential kernel with a separate length scale for each input, and a white-noise term; the kernel hyperparameters were obtained by maximizing the marginal likelihood. The posterior mean is smooth, so $m$ and $\xi$ follow from finite differences on a fine grid, and 500 samples of the noise-free posterior surface give the uncertainty of $\eta$ and the probability that $\xi < 0$ at each grid point. Over the same 20 splits, the GP gave a test MAPE of 6.2 ± 2.2% ($R^2$ = 0.9945 ± 0.0023), comparable to that of the tree ensembles.

## 4. Results and Discussion

### 4.1 Model Performance over Repeated Splits

Table 5 gives the mean ± SD of $R^2$, RMSE, MAE and MAPE over the 20 splits for the training, validation and test sets, and Fig. 1 shows the test MAPE of every split.

**Table 5.** Performance of the eight models over 20 random temperature-stratified splits (mean ± SD; training n = 70, validation n = 15 and test n = 15 per split)

| Model | Training | | Validation | | Test | | | |
|---|---|---|---|---|---|---|---|---|
| | $R^2$ | MAPE (%) | $R^2$ | MAPE (%) | $R^2$ | RMSE (MPa) | MAE (MPa) | MAPE (%) |
| Linear regression | 0.9885 ± 0.0023 | 12.65 ± 1.33 | 0.9751 ± 0.0188 | 19.19 ± 10.09 | 0.9813 ± 0.0106 | 6.60 ± 1.12 | 5.04 ± 0.91 | 17.01 ± 6.87 |
| Random forest | 0.9985 ± 0.0005 | 2.24 ± 0.18 | 0.9870 ± 0.0113 | 7.26 ± 3.27 | 0.9895 ± 0.0059 | 5.07 ± 1.65 | 3.48 ± 0.98 | 6.98 ± 3.26 |
| Gradient boosting | 1.0000 ± 0.0000 | 0.47 ± 0.14 | 0.9954 ± 0.0055 | 5.40 ± 2.75 | 0.9970 ± 0.0024 | 2.60 ± 0.88 | 1.85 ± 0.55 | 5.16 ± 2.76 |
| XGBoost | 0.9999 ± 0.0000 | 0.73 ± 0.25 | 0.9945 ± 0.0061 | 5.86 ± 3.45 | 0.9962 ± 0.0031 | 2.91 ± 1.03 | 1.95 ± 0.53 | 5.82 ± 3.45 |
| Neural network | 0.9969 ± 0.0018 | 4.82 ± 1.71 | 0.9862 ± 0.0145 | 9.54 ± 4.44 | 0.9910 ± 0.0051 | 4.57 ± 1.24 | 3.54 ± 0.90 | 7.66 ± 3.13 |
| LightGBM | 0.9999 ± 0.0001 | 0.84 ± 0.43 | 0.9888 ± 0.0199 | 6.66 ± 3.42 | 0.9905 ± 0.0155 | 4.12 ± 2.89 | 2.76 ± 1.57 | 6.39 ± 3.98 |
| CatBoost | 0.9999 ± 0.0000 | 1.06 ± 0.21 | 0.9928 ± 0.0086 | 6.02 ± 3.61 | 0.9942 ± 0.0047 | 3.64 ± 1.54 | 2.35 ± 0.65 | 6.15 ± 3.94 |

© 2026 The Author(s). Published by Qassim University under CC BY-NC-SA 4.0.

| Model | Training | | Validation | | Test | | | |
|---|---|---|---|---|---|---|---|---|
| | $R^2$ | MAPE (%) | $R^2$ | MAPE (%) | $R^2$ | RMSE (MPa) | MAE (MPa) | MAPE (%) |
| Enhanced ensemble | 0.9997 ± 0.0002 | 1.48 ± 0.48 | 0.9955 ± 0.0049 | 5.61 ± 2.62 | 0.9968 ± 0.0023 | 2.70 ± 0.86 | 1.97 ± 0.56 | 5.33 ± 2.50 |

Four observations follow from Table 5 and Fig. 1. First, the tree-based models, the neural network and the ensemble reach mean test MAPE values of 5.2–7.7% and $R^2$ of 0.990–0.997, and the split-to-split SD (2.5–4.0 percentage points of MAPE) is as large as the differences between them. Gradient boosting had the lowest mean test MAPE (5.2 ± 2.8%), followed by the enhanced ensemble (5.3 ± 2.5%) and XGBoost (5.8 ± 3.5%), but the ranking changes from split to split; XGBoost, for example, ranged from 1.4% to 15.7%. Paired Wilcoxon signed-rank tests over the 20 splits (not corrected for multiple comparisons) show that gradient boosting and the ensemble are more accurate than random forest, the neural network and LightGBM ($p \leq 0.008$), whereas the differences among gradient boosting, XGBoost, CatBoost and the ensemble are not significant ($p \geq 0.07$). The stacked ensemble therefore brings no measurable gain over its best base learner.

Second, the tree ensembles reproduce the training data almost exactly (training MAPE 0.5–2.2%) while their validation and test MAPE are 5–7%, a gap that reflects variance rather than bias at this sample size. On individual splits the test MAPE may be higher or lower than the validation MAPE; both sets contain only 15 points, and differences of one or two percentage points between them, or between models on a single split, are within the noise.

Third, linear regression on the engineered features explains most of the variance (test $R^2 = 0.981 \pm 0.011$) but has a test MAPE of 17.0 ± 6.9%. Its absolute error (RMSE = 6.6 ± 1.1 MPa) is only 2–2.5 times that of the boosted models (2.6–2.9 MPa), but it is spread evenly over the stress range and therefore becomes large in relative terms at the low stresses measured at 500 °C (Section 4.5). Training the linear model on $\ln \sigma$ instead of $\sigma$ reduced its test MAPE to 10.1 ± 3.0%, and the same change reduced the XGBoost test MAPE to 4.7 ± 1.9%, which indicates that the power-law character of the data should be reflected in the target as well as in the features. Fourth, out-of-fold predictions from ten repetitions of 5-fold cross-validation on all 100 points give the same picture: $R^2 = 0.9969 \pm 0.0009$ and MAPE = 4.6 ± 0.4% for XGBoost, against $R^2 = 0.983 \pm 0.001$ and MAPE = 16.3 ± 0.9% for linear regression. Because the $R^2$ values of all models lie between 0.98 and 1.00, Fig. 1 compares the models by MAPE.

© 2026 The Author(s). Published by Qassim University under CC BY-NC-SA 4.0.

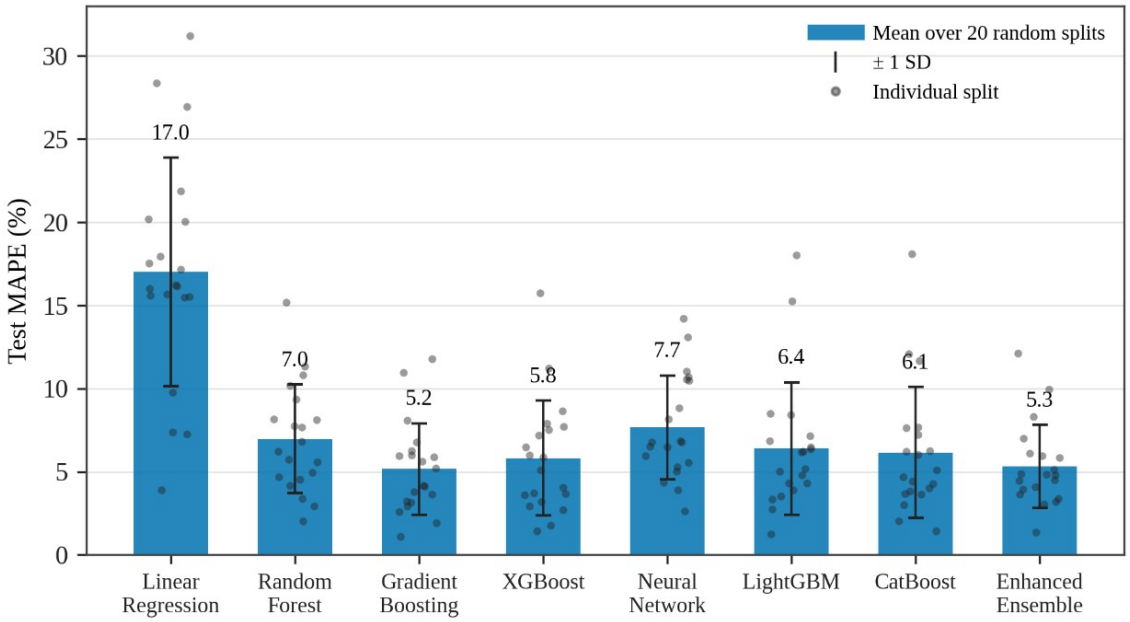



**Figure 1.** Test MAPE of the eight models over 20 random splits (bars: mean; error bars: ±1 SD; dots: individual splits; n = 15 test points per split)

### 4.2 Prediction at Withheld Temperatures and Strain Rates

Table 6 lists the MAPE obtained when all points at one temperature or one strain rate are withheld from training. When an interior temperature (350, 400 or 450 °C) is withheld, the error is 1.4–31.5%, depending on the model and the level, and no model is consistently best (e.g., XGBoost gives 31.5% at 400 °C and CatBoost 1.4% at 350 °C). Withholding an end temperature is much more severe. The tree-based models cannot predict outside the stress range seen in training and give errors of 11–13% at 300 °C and 88–131% at 500 °C, where the stresses are lowest; the linear model and the neural network extrapolate unstably (122% and 35% at 300 °C). The strain-rate tests show the same pattern for the non-linear models: 3.6–20% for the interior strain rates, 27–52% at 0.001 $s^{-1}$ and 14–46% at 1.0 $s^{-1}$. The linear model diverges when 1.0 $s^{-1}$ is withheld, because its quadratic strain-rate terms are then unconstrained. The GP surrogate of Section 3.8 behaves similarly (mean MAPE of 44.6% for LOTO and 27.0% for LOSRO).

The random-split errors in Table 5 therefore describe interpolation within the tested grid. For a process engineer who needs a prediction at a temperature or strain rate that has not been tested, errors of the order of 10–30% inside the envelope, and much larger errors outside it, are the realistic expectation.

**Table 6.** Test MAPE (%) when all points at one temperature (leave-one-temperature-out, 20 test points per fold) or at one strain rate (leave-one-strain-rate-out, 25 test points per fold) are withheld

| Model | Temperature withheld (°C) | | | | | | Strain rate withheld ($s^{-1}$) | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | 300 | 350 | 400 | 450 | 500 | Mean | 0.001 | 0.01 | 0.1 | 1.0 | Mean |
| Linear regression | 121.7 | 6.4 | 3.6 | 11.1 | 56.4 | 39.8 | 121.0 | 24.2 | 12.8 | >1000 | >1000 |
| Random forest | 13.0 | 4.5 | 3.5 | 20.7 | 87.9 | 25.9 | 39.8 | 19.9 | 11.1 | 14.4 | 21.3 |
| Gradient boosting | 11.1 | 4.0 | 4.7 | 13.0 | 99.2 | 26.4 | 32.4 | 11.3 | 5.5 | 15.8 | 16.3 |
| XGBoost | 11.0 | 7.0 | 31.5 | 11.4 | 100.8 | 32.3 | 41.1 | 10.2 | 10.6 | 24.2 | 21.5 |

© 2026 The Author(s). Published by Qassim University under CC BY-NC-SA 4.0.

| Model | Temperature withheld (°C) | | | | | | Strain rate withheld ($s^{-1}$) | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | 300 | 350 | 400 | 450 | 500 | Mean | 0.001 | 0.01 | 0.1 | 1.0 | Mean |
| Neural network | 35.3 | 9.6 | 6.7 | 9.3 | 90.9 | 30.3 | 30.0 | 6.0 | 8.0 | 45.9 | 22.5 |
| LightGBM | 10.9 | 6.8 | 5.0 | 6.8 | 97.1 | 25.3 | 32.7 | 3.6 | 9.5 | 15.8 | 15.4 |
| CatBoost | 11.2 | 1.4 | 12.8 | 7.0 | 130.6 | 32.6 | 52.4 | 19.9 | 9.4 | 24.4 | 26.5 |
| Enhanced ensemble | 10.5 | 5.9 | 15.2 | 12.1 | 93.0 | 27.4 | 27.2 | 8.0 | 5.4 | 24.2 | 16.2 |

### 4.3 Model Interpretability and Feature Importance

Fig. 2 shows the Spearman rank-correlation matrix of the 15 features. $Z$ and ln $Z$ are perfectly rank-correlated ($\rho$ = 1.00), and 25 of the 105 feature pairs have $|\rho| > 0.9$, so single-feature importances are unreliable for this feature set. The grouped permutation importances of XGBoost are shown in Fig. 3 and Table 7.

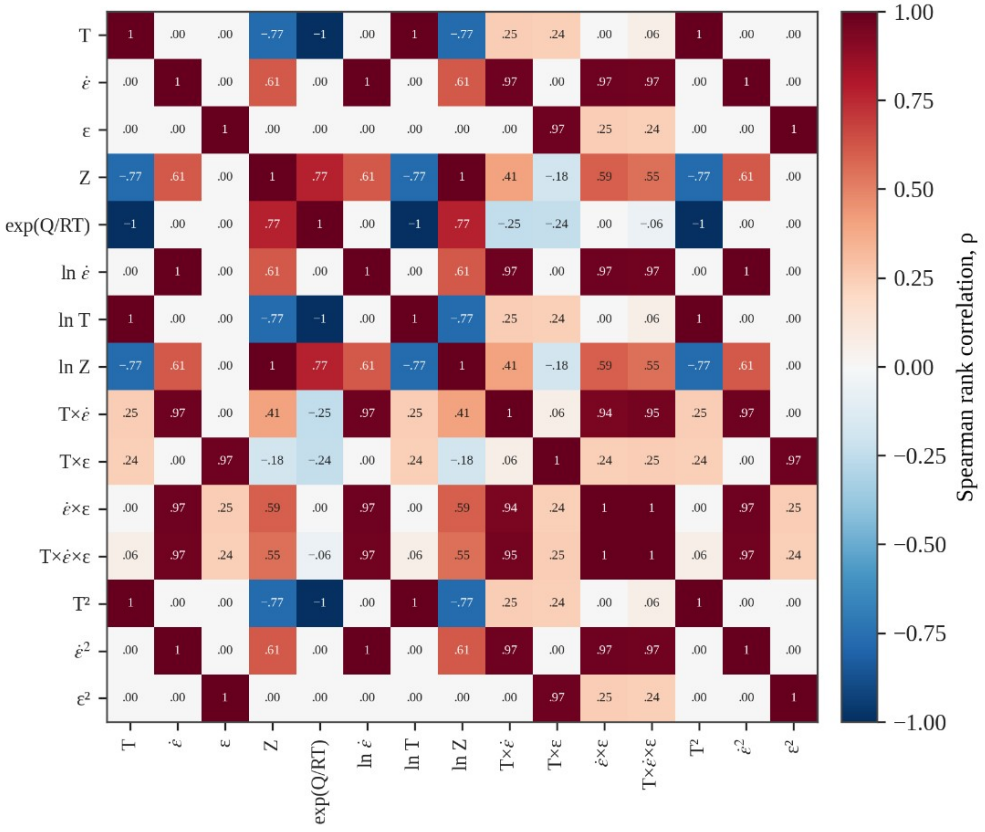



**Figure 2.** Spearman rank-correlation matrix of the 15 input features (all 100 data points)

© 2026 The Author(s). Published by Qassim University under CC BY-NC-SA 4.0.

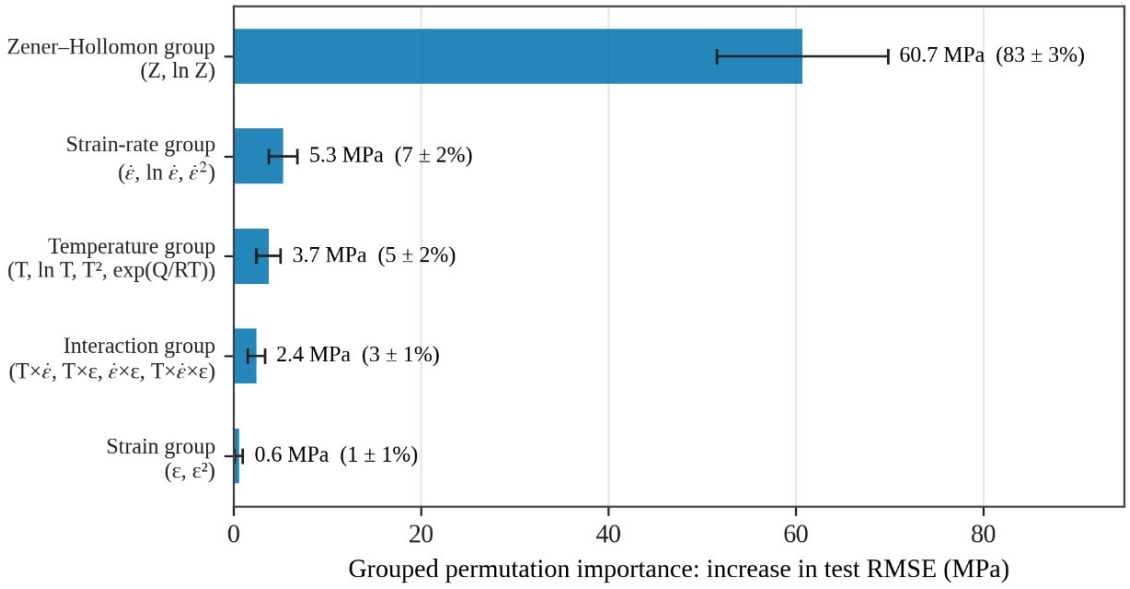



**Figure 3.** Grouped permutation importance of the XGBoost model, expressed as the increase in test RMSE when each feature group is permuted (bars: mean over 20 splits, 30 permutations per split; error bars: ±1 SD across splits); the share of the total importance is given in parentheses

**Table 7.** Grouped permutation importance of the XGBoost model on the test sets (mean ± SD over 20 splits; 30 permutations per split)

| Rank | Feature group | Increase in test RMSE (MPa) | Share of total (%) |
|---|---|---|---|
| 1 | Zener–Hollomon ($Z$, ln $Z$) | 60.7 ± 9.1 | 83.3 ± 3.5 |
| 2 | Strain rate ($\dot{\varepsilon}$, ln $\dot{\varepsilon}$, $\dot{\varepsilon}^2$) | 5.3 ± 1.6 | 7.2 ± 1.9 |
| 3 | Temperature ($T$, ln $T$, $T^2$, exp($Q/RT$)) | 3.7 ± 1.3 | 5.3 ± 2.0 |
| 4 | Interactions ($T \times \dot{\varepsilon}$, $T \times \varepsilon$, $\dot{\varepsilon} \times \varepsilon$, $T \times \dot{\varepsilon} \times \varepsilon$) | 2.4 ± 0.9 | 3.4 ± 1.2 |
| 5 | Strain ($\varepsilon$, $\varepsilon^2$) | 0.6 ± 0.4 | 0.8 ± 0.6 |

The Zener–Hollomon group accounts for 83 ± 3% of the total importance: permuting $Z$ and ln $Z$ together raises the test RMSE by about 61 MPa, compared with 5.3 MPa for the strain-rate group and 3.7 MPa for the temperature group. The strain terms contribute least (about 1%), consistent with the weak strain dependence of the data (the mean flow stress falls by only 11% from $\varepsilon$ = 0.1 to 0.5). When the features are permuted one at a time, the credit for $Z$ is split between $Z$ (47 ± 8 MPa) and ln $Z$ (12 ± 5 MPa), which shows why the grouped analysis is needed. The result is consistent with the Sellars–McTegart description [20], in which $Z$ controls thermally activated flow. It should not be over-interpreted, however: $Z$ was supplied as an input, and heavy use of it shows that $Z$ is an efficient summary of $T$ and $\dot{\varepsilon}$ for this dataset, not that the model has learned a deformation mechanism.

### 4.4 Partial Dependence Analysis

Partial dependence computed for the primary variables (Section 3.7) reproduces the monotonic trends of the data. For the XGBoost model trained on all 100 points, the average predicted flow stress decreases from 138.0 MPa at 300 °C to 28.7 MPa at 500 °C, increases from 43.1 MPa at 0.001 $s^{-1}$ to 131.2 MPa at 1.0 $s^{-1}$, and decreases slightly from 86.3 MPa at $\varepsilon$ = 0.1 to 76.9 MPa at $\varepsilon$ = 0.5; these values agree with the corresponding averages of Table 2 to within 0.1 MPa, as expected for a model evaluated on its own

© 2026 The Author(s). Published by Qassim University under CC BY-NC-SA 4.0.

training grid. The more demanding test is the behavior between the grid points. On a 41 × 61 grid of temperature and ln $\dot{\varepsilon}$ at $\varepsilon$ = 0.5, the XGBoost prediction increased with temperature at 12.6% of adjacent grid steps and decreased with strain rate at 22.7% of them, because of its piecewise-constant form, whereas the linear model and the GP surrogate violated these monotonicity conditions at no more than 2.6% of the steps for $\varepsilon$ = 0.1–0.5. Monotonic trends at the data points therefore do not guarantee a physically consistent surface between them.

### 4.5 Prediction Accuracy and Residual Analysis

Fig. 4 compares the XGBoost predictions with the measured values for one representative split (seed 0), with the training, validation and test points shown separately; the 15 test predictions are listed in Table A1 (Appendix A). For this split the test set gives $R^2$ = 0.998, RMSE = 2.31 MPa and MAPE = 7.2%. The MAPE is dominated by a single point: at 500 °C, 0.001 $s^{-1}$ and $\varepsilon$ = 0.1, the measured stress is 5.54 MPa and the prediction 9.70 MPa, an error of 75%, while the other 14 test points have errors between −5.1% and +6.0% (MAPE 2.3%). $R^2$, RMSE and MAPE must therefore be read together.

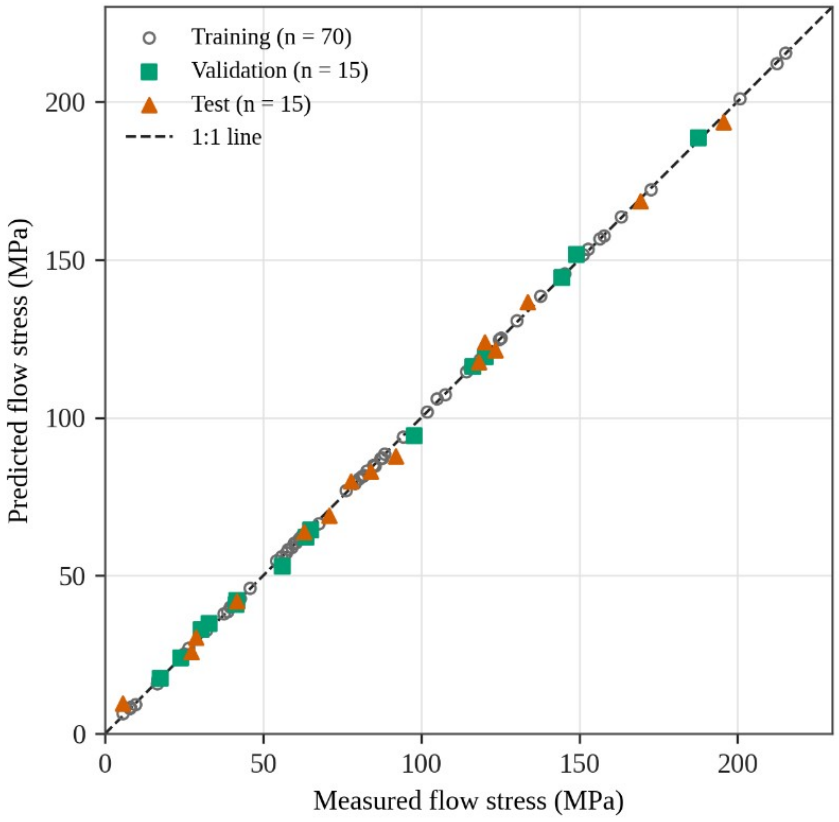



**Figure 4.** XGBoost predictions versus measured flow stress for one split (seed 0): training (n = 70), validation (n = 15) and test (n = 15) points with the 1:1 line; test set R2 = 0.998, RMSE = 2.31 MPa, MAPE = 7.2%

Fifteen test residuals are too few to assess the error distribution, so Fig. 5 uses the out-of-fold predictions for all 100 points from ten repetitions of 5-fold cross-validation (1000 predictions per model). The residuals are not homoscedastic in relative terms. For XGBoost, the SD of the absolute residual is similar over the whole range (2.0–3.4 MPa in the stress bands 0–25, 25–50, 50–100 and 100–250 MPa), so the relative error grows as the stress falls: the MAPE is 1.5% above 100 MPa, 2.5% for 50–100 MPa, 4.5% for 25–50 MPa and 19.3% below 25 MPa. By temperature, the out-of-fold MAPE is 1.6–

© 2026 The Author(s). Published by Qassim University under CC BY-NC-SA 4.0.

2.8% at 300–450 °C and 14.1% at 500 °C. Linear regression shows the same pattern with larger errors (residual SD 5.2–6.2 MPa; MAPE 3.5% above 100 MPa and 80% below 25 MPa).

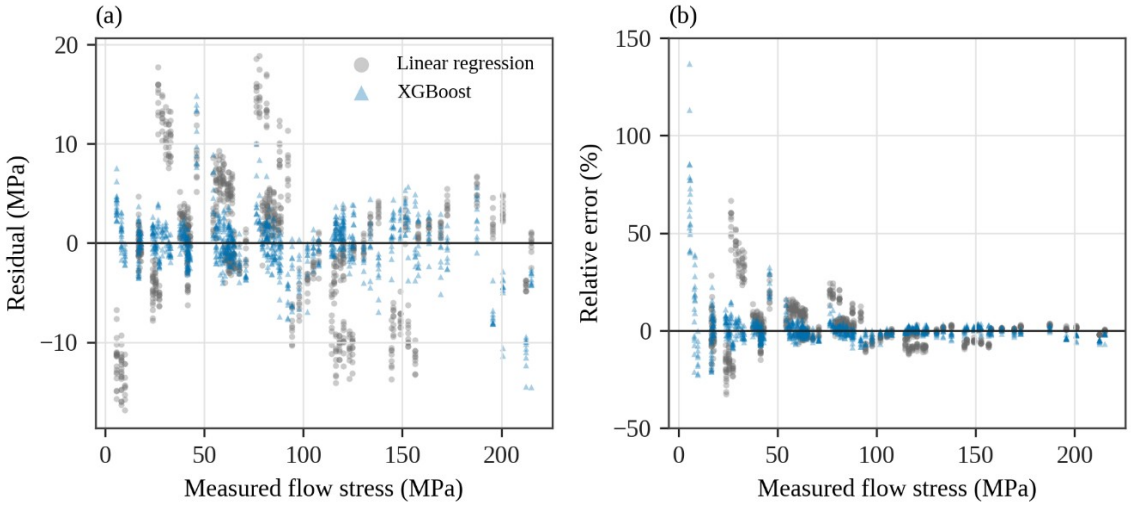



**Figure 5.** Out-of-fold residuals of XGBoost and linear regression for all 100 points (ten repetitions of 5-fold cross-validation): (a) absolute residual and (b) relative error versus measured flow stress

The pooled XGBoost residuals (Fig. 6) have a mean of −0.05 MPa and an SD of 2.86 MPa, but their distribution is more peaked and heavier-tailed than a normal distribution (excess kurtosis 5.4), so normality should not be assumed when prediction intervals are constructed.

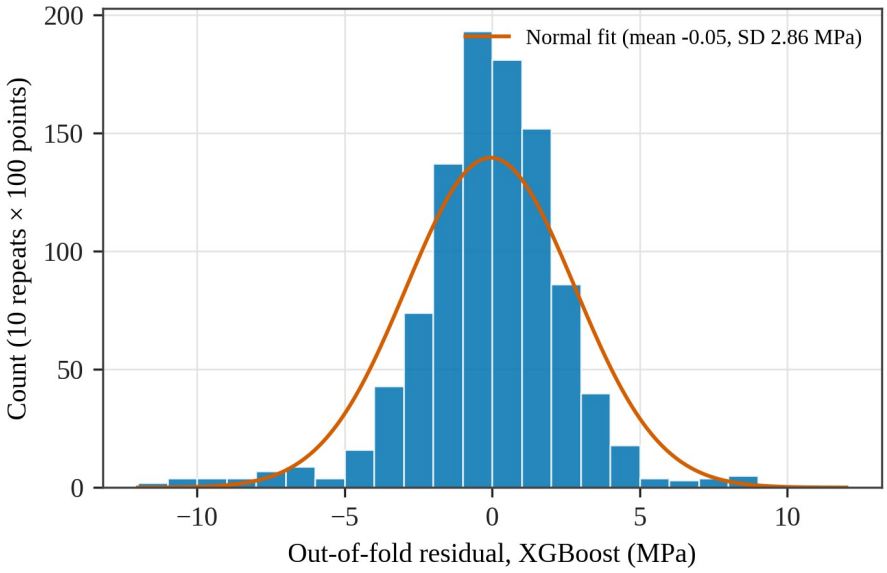



**Figure 6.** Histogram of the pooled out-of-fold XGBoost residuals (1000 predictions), with a normal distribution of the same mean and SD for reference

The models are therefore least reliable where the flow stress is lowest, i.e. at 500 °C and low strain rates. This region also shows the largest scatter in the data themselves: at 500 °C and 0.001 $s^{-1}$, the flow stress varies non-monotonically between 5.5 and 9.7 MPa with strain.

© 2026 The Author(s). Published by Qassim University under CC BY-NC-SA 4.0.

### 4.6 Physical Validation

The measured and predicted flow stresses follow the expected trends of hot deformation. Fig. 7 shows the temperature dependence at $\varepsilon = 0.5$: at every strain rate the flow stress decreases monotonically with increasing temperature.

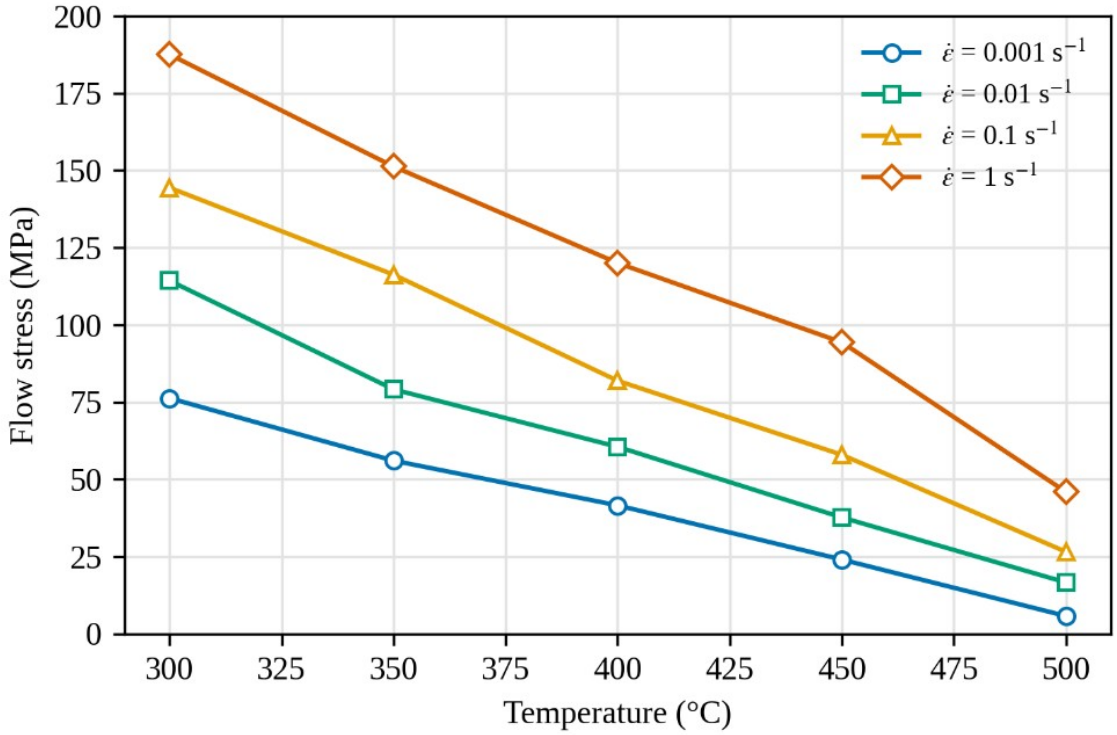



**Figure 7.** Flow stress versus temperature at ε = 0.5 for the four strain rates (data of Table 2)

At 0.01 s$^{-1}$ and $\varepsilon = 0.5$, the flow stress decreases from 114.3 MPa at 300 °C to 16.6 MPa at 500 °C, a reduction of 85%, reflecting the increasing contribution of thermally activated dislocation climb, dynamic recovery and, possibly, dynamic recrystallization at higher temperatures.

The flow stress increases with strain rate at every temperature (Fig. 8), confirming the rate-sensitive behavior characteristic of hot deformation.

© 2026 The Author(s). Published by Qassim University under CC BY-NC-SA 4.0.

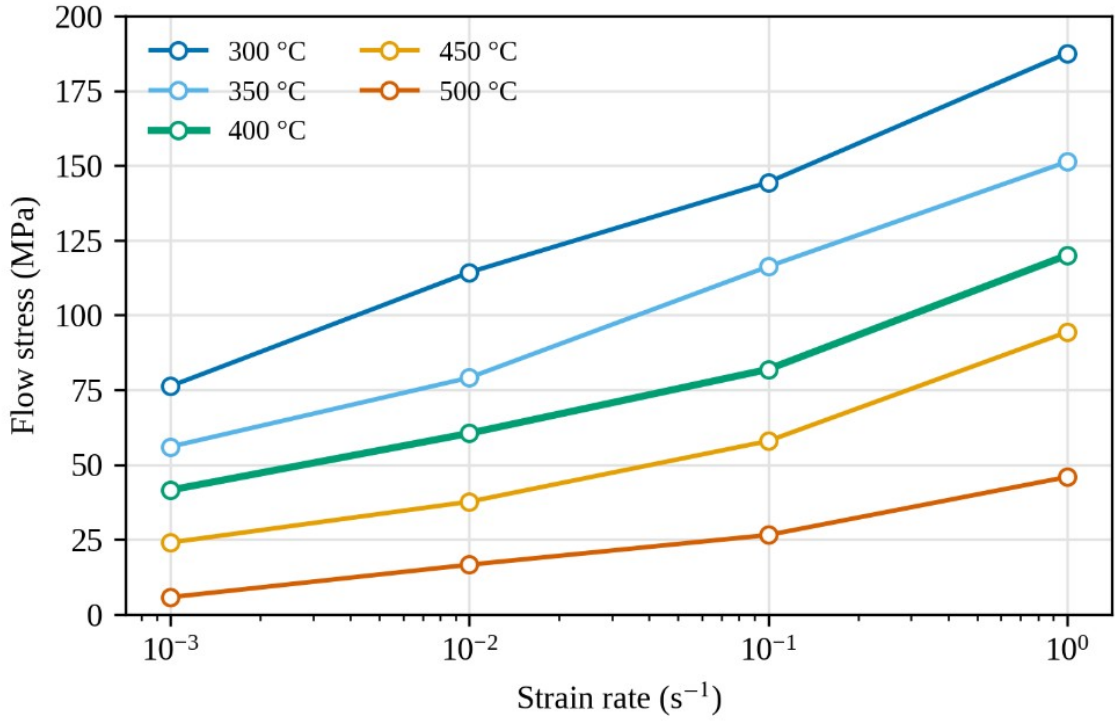



**Figure 8.** Flow stress versus strain rate (logarithmic scale) at ε = 0.5 for the five test temperatures (data of Table 2)

At 400 °C and $\varepsilon = 0.5$, the flow stress increases from 41.6 MPa at 0.001 $s^{-1}$ to 120.0 MPa at 1.0 $s^{-1}$. The local strain-rate sensitivity, estimated from quadratic fits of ln $\sigma$ against ln $\dot{\varepsilon}$, is 0.08–0.21 at 300–450 °C and reaches about 0.46 at 500 °C and 0.001 $s^{-1}$, where recovery processes have the most time to operate.

Fig. 9 plots the flow stress against ln $Z$. The data collapse onto a single band, and the hyperbolic-sine equation (Eq. 1) with the constants of Section 3.1 describes the trend, although points with similar values of $Z$ differ by 20–30 MPa because of the strain dependence and because a single value of $Q$ does not remove all of the temperature dependence.

© 2026 The Author(s). Published by Qassim University under CC BY-NC-SA 4.0.

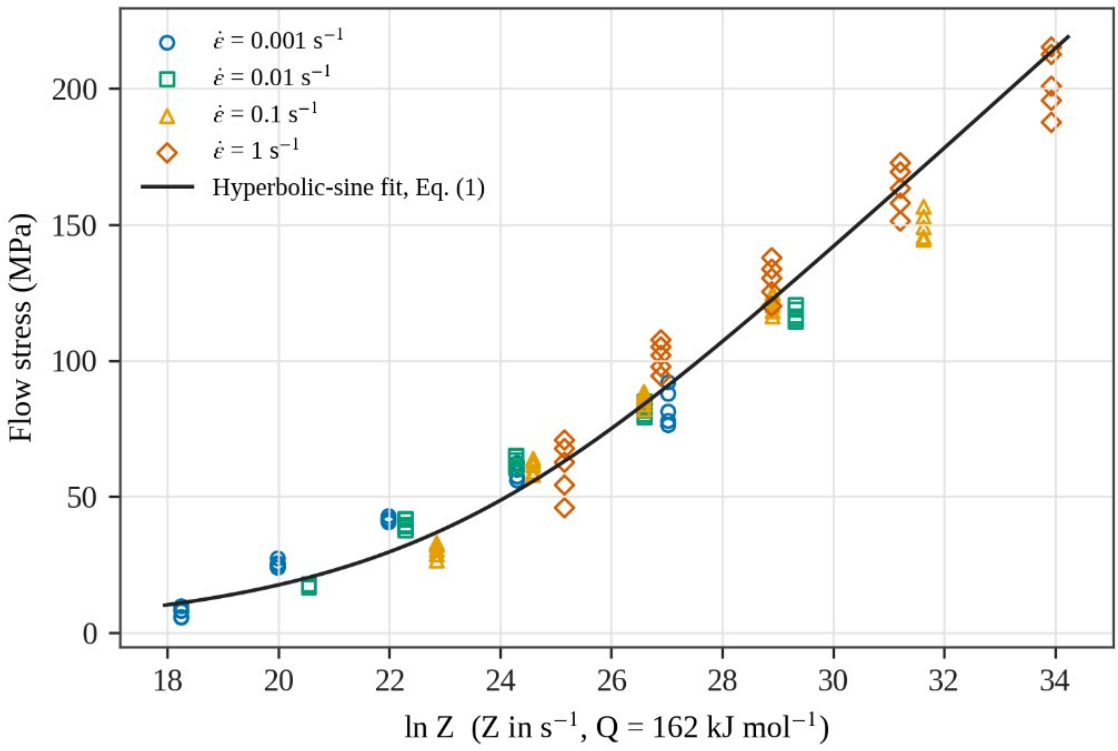



**Figure 9.** Flow stress versus ln Z (Q = 162 kJ mol−1) for all 100 data points, with the fitted hyperbolic-sine equation (Eq. 1; α = 0.0148 MPa−1, slope 3.66, ln A = 24.88)

The ML models reproduce this relationship because $Z$ is one of their inputs; the relevant question is whether they improve on it. Relative to the hyperbolic-sine equation fitted to the same data (MAPE = 14.3% with strain-dependent constants and 15.0% with a single set), the boosted models reduce the interpolation error to about 5% (Table 5), presumably by capturing the strain dependence and the departure from a single-$Q$ description at 500 °C.

### 4.7 Comparison with Literature

Table 8 lists errors reported for flow-stress models in the literature together with the present results. The comparison is given for context only and is not a ranking: the materials, dataset sizes, validation protocols and error definitions (MAPE, average absolute relative error or mean deviation) differ between the studies, and the present values refer to interpolation within a complete test grid.

**Table 8.** Reported errors of flow-stress models (context only; the error definitions, materials, dataset sizes and validation protocols differ and the values are not directly comparable)

| Reference | Material | Method | Error measure and data | Error (%) |
|---|---|---|---|---|
| [37] | 2124 Al/SiC$_p$ | Regression | Mean deviation, as reported | 13.5 |
| [37] | 2124 Al/SiC$_p$ | Fuzzy logic | Mean deviation, as reported | 6.1 |
| Mandal et al. [22] | AISI 304L steel | ANN | As reported | 6.0 |
| Quan et al. [29] | Al–Zn–Mg–Cu alloy | ANN | As reported | 5.2 |
| Ji et al. [30] | Alloy 617 | ANN | As reported | 7.1 |
| Li et al. [36] | Al–Zn–Mg–Cu alloy | XGBoost | As reported | 3.9 |

© 2026 The Author(s). Published by Qassim University under CC BY-NC-SA 4.0.

| Reference | Material | Method | Error measure and data | Error (%) |
|---|---|---|---|---|
| Present study | 2124 Al/$SiC_p$ | Hyperbolic-sine equation | MAPE, fit to all 100 points | 14.3 |
| Present study | 2124 Al/$SiC_p$ | Linear regression | Test MAPE, 20 splits ($n$ = 15 each) | 17.0 ± 6.9 |
| Present study | 2124 Al/$SiC_p$ | Gradient boosting | Test MAPE, 20 splits ($n$ = 15 each) | 5.2 ± 2.8 |
| Present study | 2124 Al/$SiC_p$ | XGBoost | Test MAPE, 20 splits ($n$ = 15 each) | 5.8 ± 3.5 |
| Present study | 2124 Al/$SiC_p$ | Enhanced ensemble | Test MAPE, 20 splits ($n$ = 15 each) | 5.3 ± 2.5 |
| Present study | 2124 Al/$SiC_p$ | XGBoost | MAPE, leave-one-temperature-out (mean) | 32.3 |

Against the regression (13.5%) and fuzzy-logic (6.1%) errors reported for the same composite [37], the mean test MAPE of gradient boosting and XGBoost (5.2–5.8%) is about 60% lower than the regression value and similar to the fuzzy-logic value. Against the linear model and the hyperbolic-sine equation fitted in this work (17.0% and 14.3%), the reduction is about two-thirds. Which comparison is appropriate depends on the error definition and validation protocol of [37], which are not identical to those used here. No improvement is claimed over the ANN and XGBoost studies of other alloys [22], [29], [30], [36], because those results were obtained on different materials and datasets.

## 4.8 Discussion

### 4.8.1 Accuracy of Machine Learning Relative to Constitutive Equations

Within the tested envelope, the non-linear ML models reduce the interpolation error of the hyperbolic-sine description by about two-thirds (from 14.3% to about 5% MAPE). Three factors contribute. First, the tree ensembles and the neural network represent the strain dependence and the curvature of $\ln \sigma$ against $1/T$ without an assumed functional form. Second, the Zener–Hollomon features give every model, including the linear one, a physically appropriate coordinate: most of the variance is explained once $Z$ is available (test $R^2$ = 0.981 for linear regression). Third, stacking adds little: the enhanced ensemble and gradient boosting differ by 0.2 percentage points of mean MAPE, well within the split-to-split scatter. Outside the tested envelope, none of the models is reliable (Table 6).

### 4.8.2 Physical Consistency and Interpretability

The grouped permutation analysis attributes 83 ± 3% of the XGBoost importance to the Zener–Hollomon features, in line with the Sellars–McTegart framework [20]. Physical consistency was also checked quantitatively (Sections 4.4 and 4.6): partial dependence reproduces the monotonic trends of the data, but the tree-based surface is not monotonic between the grid points, whereas the linear model and the GP surrogate are nearly so. Interpretability analyses of this kind show which inputs a model relies on; they do not by themselves show that the model will extrapolate, and the LOTO and LOSRO results (Table 6) show that it does not.

### 4.8.3 Computational Efficiency and Practical Implementation

All trained models predict a single point in less than 1 ms on a standard computer, which is fast enough for use in process simulation or control. The models could be coupled with finite-element simulations through a user material subroutine, but for this purpose a smooth surrogate (the GP or the linear model

© 2026 The Author(s). Published by Qassim University under CC BY-NC-SA 4.0.

in the engineered feature space) is preferable to a tree ensemble, because implicit finite-element solvers require continuous derivatives of the flow stress.

The trained XGBoost model occupies about 0.3 MB in its native format, and all models were implemented with standard open-source Python libraries (scikit-learn, XGBoost, LightGBM and CatBoost), which supports reproducibility.

#### 4.8.4 Implications for Manufacturing

Within the tested envelope, a flow-stress model with an interpolation error of about 5% is adequate for estimating forming loads and press capacity and for comparing candidate temperature–strain-rate windows. At temperatures or strain rates that have not been tested, the errors in Table 6 apply, and additional tests are needed before a prediction is used for die or equipment design. Real-time adaptive control would also require prediction intervals, which the GP surrogate provides but the tree ensembles do not.

### 4.9 Processing Map Development

Processing maps were constructed with the dynamic materials model (DMM) [45], using the GP surrogate of Section 3.8 because the tree ensembles do not provide a usable strain-rate derivative. The efficiency of power dissipation $\eta$, the strain-rate sensitivity $m$ and the instability parameter $\xi$ were evaluated as follows.

Power dissipation efficiency:

$$\eta = \frac{2m}{m+1} \tag{15}$$

where $m$ represents the strain rate sensitivity parameter calculated from the GP posterior mean:

$$m = \left.\frac{\partial \ln \sigma}{\partial \ln \dot{\varepsilon}}\right|_{T,\varepsilon} \tag{16}$$

Instability parameter:

$$\xi = \frac{\partial}{\partial \ln \dot{\varepsilon}} \ln\left[\frac{m}{m+1}\right] + m \tag{17}$$

Regions with $\xi < 0$ are predicted to be prone to flow instability, and regions with $\xi > 0$ and high $\eta$ are candidate processing windows. The parameters were computed on a 41 × 61 grid spanning 300–500 °C and 0.001–1.0 $s^{-1}$ at $\varepsilon$ = 0.1, 0.3 and 0.5, with the derivatives in Eqs. (16) and (17) evaluated by central differences in ln $\dot{\varepsilon}$. The uncertainty was obtained by repeating the calculation for 500 posterior samples of the GP surface.

Fig. 10a shows the map at $\varepsilon$ = 0.5, with the experimental conditions marked so that the extent of interpolation is visible, and Fig. 10b shows the posterior SD of $\eta$ together with the contours at which the posterior probability of $\xi < 0$ reaches 0.5 and 0.9. The most robust feature is a domain of high efficiency at 475–500 °C and about 0.002–0.01 $s^{-1}$, where the posterior-mean $\eta$ is 53–68%, with a posterior SD of 3–6 percentage points at 500 °C. At 450 °C, $\eta$ is about 30–39% over most of the strain-rate range. At 400–450 °C and 0.01–0.1 $s^{-1}$, $\eta$ is 22–34%, clearly lower than in the high-temperature domain.

© 2026 The Author(s). Published by Qassim University under CC BY-NC-SA 4.0.

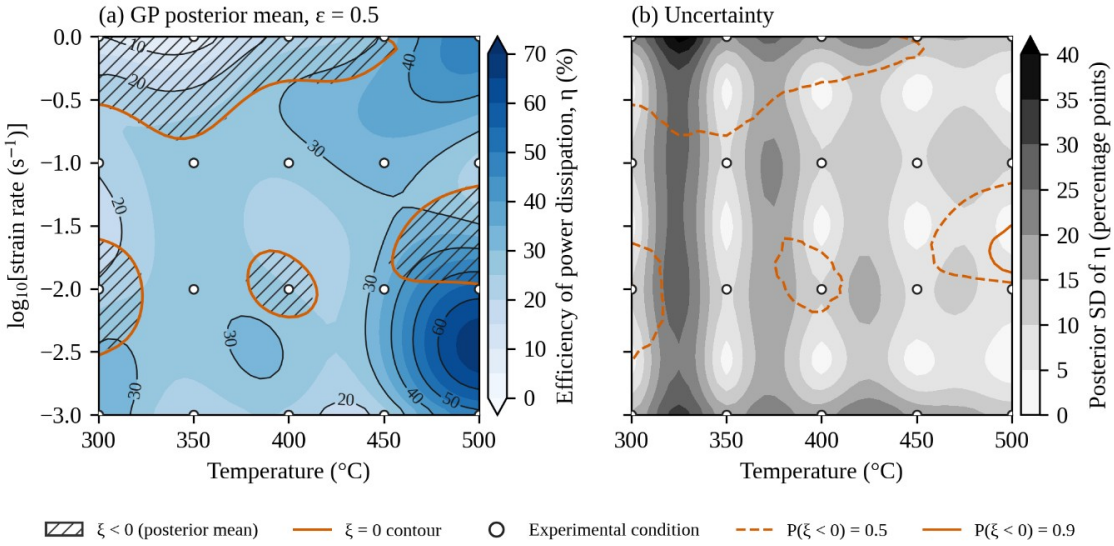



**Figure 10.** Processing map of 2124 Al/10 vol.% SiCp at ε = 0.5 derived from the Gaussian-process surrogate: (a) contours of η (%), region with ξ < 0 at the posterior mean (hatched) and experimental conditions (open circles); (b) posterior SD of η with the contours P(ξ < 0) = 0.5 and 0.9

The posterior-mean map shows $\xi < 0$ at strain rates above about 0.3 $s^{-1}$ at 300–450 °C, and in several smaller regions (Fig. 10a). However, the posterior probability of $\xi < 0$ exceeds 0.9 over less than 1% of the map and exceeds 0.5 over about 22% of it (Fig. 10b), so the instability boundary is poorly constrained. This is expected: $\xi$ depends on the derivative of $m$ with respect to ln $\dot{\varepsilon}$, i.e. on the second derivative of ln $\sigma$, and four strain rates spanning three decades carry little information about that curvature. The posterior SD of $\eta$ is also large (about 15–40 percentage points) midway between the tested temperatures, where the GP interpolates. The map should therefore be read as identifying a probable high-efficiency domain at the highest temperature and lowest strain rates and as indicating possible instability at high strain rates and lower temperatures. Confirming either requires metallographic examination of specimens deformed under these conditions, which was outside the scope of this data-based study. The high efficiency at 500 °C and 0.001–0.01 $s^{-1}$ is consistent with dynamic recrystallization or with grain-boundary sliding, but compression data alone cannot distinguish between these mechanisms.

Coupling a smooth probabilistic surrogate with the DMM makes the uncertainty of the processing map explicit. With the present data this uncertainty is large; additional strain rates (for example, 0.003, 0.03 and 0.3 $s^{-1}$) and intermediate temperatures would reduce it most efficiently.

### 4.10 Limitations

The study has the following limitations. (i) The dataset is small (100 points) and was compiled from a single published source; it covers one composite (2124 Al with 10 vol.% $SiC_p$) produced by one route, so the results cannot be transferred to other matrices, reinforcement fractions or particle sizes without new data. (ii) The accuracy figures in Table 5 refer to interpolation within a complete grid; the LOTO and LOSRO tests (Table 6) show that the errors are several times larger for withheld conditions and very large outside the tested envelope. (iii) The deformation-heating correction applied to the data uses $\beta = 0.9$ at all strain rates and is an upper bound at low strain rates (Table 3); the uncorrected data were not re-analyzed. (iv) Tree-based models do not provide usable derivatives, so the processing map relies on the GP surrogate, whose instability predictions are uncertain with only four strain rates. (v) No

© 2026 The Author(s). Published by Qassim University under CC BY-NC-SA 4.0.

metallographic validation of the processing map was performed. (vi) The tree ensembles provide point predictions only; prediction intervals are available only from the GP surrogate.

## 5. Conclusions

Eight machine learning algorithms were benchmarked for predicting the flow stress of a powder-metallurgy 2124 Al/10 vol.% $SiC_p$ composite from 100 published hot-compression data points (300–500 °C, 0.001–1.0 $s^{-1}$, $\varepsilon$ = 0.1–0.5), using Zener–Hollomon-based features ($Q$ = 162 kJ $mol^{-1}$ from a hyperbolic-sine fit), 20 repeated random splits, and leave-one-temperature-out and leave-one-strain-rate-out tests.

For interpolation within the test grid, gradient boosting, the stacked ensemble and XGBoost gave mean test MAPE values of 5.2–5.8% ($R^2 \approx 0.997$), about one-third of the error of the hyperbolic-sine equation fitted to the same data (14.3%) and of linear regression on the engineered features (17.0%). The differences among the boosted models were smaller than the split-to-split scatter, so no single algorithm can be declared best, and the errors were concentrated at the low stresses measured at 500 °C.

Grouped permutation importance attributed 83 ± 3% of the predictive importance to the Zener–Hollomon features. Partial dependence reproduced the monotonic trends of the data, but the tree-based response surfaces were not monotonic between the grid points.

Because boosted trees are piecewise constant, the processing maps were derived from a Gaussian-process surrogate. The map identifies a high-efficiency domain ($\eta \approx$ 53–68%) at 475–500 °C and about 0.002–0.01 $s^{-1}$; possible instability at high strain rates and lower temperatures is indicated but is not well constrained by four strain rates.

For the non-linear models, the errors rose to mean MAPE values of 25–33% when a whole temperature level was withheld and 15–27% when a whole strain-rate level was withheld, which is the realistic accuracy for untested conditions. The results apply to one composite and one dataset; data at additional strain rates and intermediate temperatures, together with metallographic validation, are needed before the models or maps are used for process design.

## Appendix A. Test-Set Predictions for a Representative Split

Table A1: Measured and predicted flow stresses for the 15 test points of split seed 0 (Fig. 4)

| $T$ (°C) | $\dot{\varepsilon}$ ($s^{-1}$) | $\varepsilon$ | Measured (MPa) | XGBoost (MPa) | Error (%) | Enhanced ensemble (MPa) | Error (%) |
|---|---|---|---|---|---|---|---|
| 300 | 0.001 | 0.1 | 92.00 | 87.90 | -4.5 | 88.29 | -4.0 |
| 300 | 0.001 | 0.4 | 77.82 | 79.94 | +2.7 | 79.73 | +2.5 |
| 300 | 1 | 0.4 | 195.62 | 193.57 | -1.0 | 194.50 | -0.6 |
| 350 | 0.1 | 0.2 | 123.35 | 121.42 | -1.6 | 122.04 | -1.1 |
| 350 | 0.1 | 0.4 | 118.10 | 117.62 | -0.4 | 118.06 | -0.0 |
| 350 | 1 | 0.2 | 169.32 | 168.56 | -0.4 | 169.05 | -0.2 |
| 400 | 0.1 | 0.4 | 83.92 | 83.01 | -1.1 | 83.69 | -0.3 |
| 400 | 1 | 0.2 | 133.62 | 136.65 | +2.3 | 136.78 | +2.4 |

© 2026 The Author(s). Published by Qassim University under CC BY-NC-SA 4.0.

| $T$ (°C) | $\dot{\varepsilon}$ ($s^{-1}$) | $\varepsilon$ | Measured (MPa) | XGBoost (MPa) | Error (%) | Enhanced ensemble (MPa) | Error (%) |
|---|---|---|---|---|---|---|---|
| 400 | 1 | 0.5 | 120.01 | 123.80 | +3.2 | 124.16 | +3.5 |
| 450 | 0.001 | 0.2 | 27.28 | 25.90 | -5.1 | 25.36 | -7.1 |
| 450 | 0.01 | 0.2 | 41.74 | 42.00 | +0.6 | 40.75 | -2.4 |
| 450 | 0.1 | 0.2 | 63.03 | 63.85 | +1.3 | 62.72 | -0.5 |
| 500 | 0.001 | 0.1 | 5.54 | 9.70 | +75.1 | 9.00 | +62.5 |
| 500 | 0.1 | 0.4 | 28.75 | 30.47 | +6.0 | 29.37 | +2.1 |
| 500 | 1 | 0.1 | 70.78 | 69.12 | -2.3 | 70.51 | -0.4 |

## References

[1] T. Pan, C. Song, Z. Gao, T. Xia, and T. Wang, “The prediction of flow stress in the hot compression of a Ni-Cr-Mo steel using machine learning algorithms,” Processes, vol. 12, p. 441, 2024, doi: 10.3390/pr12030441.

[2] Y. C. Lin, M.-S. Chen, and J. Zhong, “A critical review of experimental results and constitutive descriptions for metals and alloys in hot working,” Mater. Des., vol. 32, no. 4, pp. 1733–1759, 2011, doi: 10.1016/j.matdes.2010.11.048.

[3] H. J. McQueen and N. D. Ryan, “Constitutive analysis in hot working,” Mater. Sci. Eng. A, vol. 322, pp. 43–63, 2002, doi: 10.1016/S0921-5093(01)01117-0.

[4] S. V. S. N. Murty, B. N. Rao, and B. P. Kashyap, “Instability criteria for hot deformation of materials,” Int. Mater. Rev., vol. 45, no. 1, pp. 15–26, 2000, doi: 10.1179/095066000771048782.

[5] M. A. Ali, N. A. Mufti, M. Sana, M. Tlija, C. I. Pruncu, and L. Lamberti, “Comparison of machine learning algorithms for dynamic performance assessment in complex shapes manufacturing of hybrid particle-reinforced composite,” Expert Syst. Appl., vol. 275, art. 127022, 2025, doi: 10.1016/j.eswa.2025.127022.

[6] P. Pragathi and R. Elansezhian, “Mechanical and microstructure behaviour of aluminum nanocomposite fabricated by squeeze casting and ultrasonic aided squeeze casting: A comparative study,” J. Alloys Compd., vol. 956, art. 170203, 2023, doi: 10.1016/j.jallcom.2023.170203.

[7] F. Zhang, Z. He, K. Lu, Z. Zhan, Z. Li, and X. Wang, “Interfacial microstructure and mechanical properties of 2124 aluminum alloy reinforced by AlCoCrFeNi high entropy alloy,” J. Mater. Res. Technol., vol. 26, pp. 8846–8856, 2023, doi: 10.1016/j.jmrt.2023.09.166.

[8] S. Cai et al., “Hot deformation behavior and microstructural evolution of SiCp/Al-4.2Cu-4.1Mg composites,” J. Mater. Res. Technol., 2025, doi: 10.1016/j.jmrt.2025.11.245.

[9] S. Salifu, P. A. Olubambi, and L. Teffo, “Phase stability and microstructural properties of high entropy alloy reinforced aluminium matrix composites consolidated via spark plasma sintering,” Heliyon, vol. 10, no. 2, art. e24123, 2024, doi: 10.1016/j.heliyon.2024.e24123.

[10] Z. Seikh, M. Sekh, G. Mandal, B. Sengupta, and A. Sinha, “Metal matrix composites processed through powder metallurgy: A brief overview,” J. Inst. Eng. India Ser. D, pp. 1–4, 2024, doi: 10.1007/s40033-024-00789-4.

[11] K. Pooja, N. Tarannum, and P. Chaudhary, “Metal matrix composites: Revolutionary materials for shaping the future,” Discov. Mater., vol. 5, p. 35, 2025, doi: 10.1007/s43939-025-00226-6.

[12] A. Berghezan, “Composite materials,” in Proc. Int. Conf. Composite Materials, 1961, pp. 1–15.

[13] D. B. Miracle, “Metal matrix composites—from science to technological significance,” Compos. Sci. Technol., vol. 65, no. 15–16, pp. 2526–2540, 2005, doi: 10.1016/j.compscitech.2005.05.027.

[14] S. P. Rawal, “Metal-matrix composites for space applications,” JOM, vol. 53, no. 4, pp. 14–17, 2001, doi: 10.1007/s11837-001-0139-z.

[15] B. V. R. Bhat, Y. R. Mahajan, and Y. V. R. K. Prasad, “Effect of volume fraction of SiCp reinforcement on the processing maps for 2124 Al matrix composites,” Metall. Mater. Trans. A, vol. 31, pp. 629–639, 2000, doi: 10.1007/s11661-000-0006-5.

[16] A. A. Lakshmi, C. S. Rao, J. Gangadhar, C. Srinivasu, and S. K. Singh, “Review of processing maps and development of qualitative processing maps,” Mater. Today Proc., vol. 4, no. 2, pp. 946–956, 2017, doi: 10.1016/j.matpr.2017.01.106.

© 2026 The Author(s). Published by Qassim University under CC BY-NC-SA 4.0.

[17] T. Seshacharyulu, S. C. Medeiros, W. G. Frazier, and Y. V. R. K. Prasad, "Hot working of commercial Ti–6Al–4V with an equiaxed α–β microstructure: Materials modeling considerations," Mater. Sci. Eng. A, vol. 284, no. 1–2, pp. 184–194, 2000, doi: 10.1016/S0921-5093(00)00741-3.

[18] M. R. Rokni, A. Zarei-Hanzaki, A. A. Roostaei, and H. R. Abedi, "An investigation into the hot deformation characteristics of 7075 aluminum alloy," Mater. Des., vol. 32, no. 4, pp. 2339–2344, 2011, doi: 10.1016/j.matdes.2010.12.047.

[19] N. Jin, H. Zhang, Y. Han, W. Wu, and J. Chen, "Hot deformation behavior of 7150 aluminum alloy during compression at elevated temperature," Mater. Charact., vol. 60, no. 6, pp. 530–536, 2009, doi: 10.1016/j.matchar.2008.12.012.

[20] C. M. Sellars and W. J. McTegart, "On the mechanism of hot deformation," Acta Metall., vol. 14, no. 9, pp. 1136–1138, 1966, doi: 10.1016/0001-6160(66)90207-0.

[21] N. Haghdadi, A. Zarei-Hanzaki, A. R. Khalesian, and H. R. Abedi, "Artificial neural network modeling to predict the hot deformation behavior of an A356 aluminum alloy," Mater. Des., vol. 49, pp. 386–391, 2013, doi: 10.1016/j.matdes.2012.12.082.

[22] S. Mandal, P. V. Sivaprasad, S. Venugopal, and K. P. N. Murthy, "Artificial neural network modeling to evaluate and predict the deformation behavior of stainless steel type AISI 304L during hot torsion," Appl. Soft Comput., vol. 9, no. 1, pp. 237–244, 2009, doi: 10.1016/j.asoc.2008.03.016.

[23] Y. C. Lin, D.-X. Wen, J. Deng, G. Liu, and J. Chen, "A unified physically based constitutive model for describing strain hardening effect and dynamic recovery behavior of a Ni-based superalloy," J. Mater. Res., vol. 28, no. 14, pp. 1884–1895, 2013, doi: 10.1557/jmr.2013.17.

[24] Q. Wen, Z. Cao, S. Yang et al., "Research on predicting flow stress of 7075 aluminum alloy using machine learning models," Front. Mater., vol. 12, art. 1671753, 2025, doi: 10.3389/fmats.2025.1671753.

[25] H. R. R. Ashtiani, M. H. Parsa, and H. Bisadi, "Constitutive equations for elevated temperature flow behavior of commercial purity aluminum," Mater. Sci. Eng. A, vol. 545, pp. 61–67, 2012, doi: 10.1016/j.msea.2012.02.090.

[26] H. N. Han, J. K. Lee, H. J. Kim, and Y. S. Jin, "A model for deformation, temperature and phase transformation behavior of steels on run-out table in hot strip mill," J. Mater. Process. Technol., vol. 128, pp. 216–225, 2002, doi: 10.1016/S0924-0136(02)00454-5.

[27] P. S. Robi and U. S. Dixit, "Application of neural networks in generating processing map for hot working," J. Mater. Process. Technol., vol. 142, no. 1, pp. 289–294, 2003, doi: 10.1016/S0924-0136(03)00579-X.

[28] O. Sabokpa, A. Zarei-Hanzaki, H. R. Abedi, and N. Haghdadi, "Artificial neural network modeling to predict the high temperature flow behavior of an AZ81 magnesium alloy," Mater. Des., vol. 39, pp. 390–396, 2012, doi: 10.1016/j.matdes.2012.03.002.

[29] G.-Z. Quan, W.-Q. Lv, Y.-P. Mao, Y.-W. Zhang, and J. Zhou, "Prediction of flow stress in a wide strain rate range involving macro-micro mechanism for Al-Zn-Mg-Cu aluminum alloy," Mater. Des., vol. 79, pp. 94–114, 2015, doi: 10.1016/j.matdes.2015.04.032.

[30] G. Ji, F. Li, Q. Li, H. Li, and Z. Li, "A comparative study on Arrhenius-type constitutive model and artificial neural network model to predict high-temperature deformation behaviour in Alloy 617," Mater. Sci. Eng. A, vol. 528, no. 13–14, pp. 4774–4782, 2011, doi: 10.1016/j.msea.2011.03.017.

[31] L. Breiman, "Random forests," Mach. Learn., vol. 45, no. 1, pp. 5–32, 2001, doi: 10.1023/A:1010933404324.

[32] J. H. Friedman, "Greedy function approximation: A gradient boosting machine," Ann. Stat., vol. 29, no. 5, pp. 1189–1232, 2001, doi: 10.1214/aos/1013203451.

[33] T. Chen and C. Guestrin, "XGBoost: A scalable tree boosting system," in Proc. 22nd ACM SIGKDD Int. Conf. Knowledge Discovery and Data Mining, 2016, pp. 785–794, doi: 10.1145/2939672.2939785.

[34] Y. Liu, T. Zhao, W. Ju, and S. Shi, "Materials discovery and design using machine learning," J. Materiomics, vol. 3, no. 3, pp. 159–177, 2017, doi: 10.1016/j.jmat.2017.08.002.

[35] L. Ward, A. Agrawal, A. Choudhary, and C. Wolverton, "A general-purpose machine learning framework for predicting properties of inorganic materials," npj Comput. Mater., vol. 2, no. 1, pp. 1–7, 2016, doi: 10.1038/npjcompumats.2016.28.

[36] H.-Y. Li, X.-F. Wang, J.-Y. Wei, J.-T. Yan, and D.-D. Li, "Prediction of hot deformation behaviour of Al-Zn-Mg-Cu aluminum alloy using XGBoost machine learning method," Trans. Nonferrous Met. Soc. China, vol. 31, no. 4, pp. 1036–1051, 2021, doi: 10.1016/S1003-6326(21)65568-6.

[37] [Authors to insert: full reference of the published study that reports the hot-compression flow-stress data of Table 2 and the regression (13.5%) and fuzzy-logic (6.1%) errors for 2124 Al/10 vol.% SiCp.]

© 2026 The Author(s). Published by Qassim University under CC BY-NC-SA 4.0. 

[38] S. M. Lundberg and S.-I. Lee, "A unified approach to interpreting model predictions," in Adv. Neural Inf. Process. Syst., vol. 30, 2017, pp. 4765–4774.

[39] D. W. Apley and J. Zhu, "Visualizing the effects of predictor variables in black box supervised learning models," J. R. Stat. Soc. Ser. B Stat. Methodol., vol. 82, no. 4, pp. 1059–1086, 2020, doi: 10.1111/rssb.12377.

[40] C. Strobl, A.-L. Boulesteix, A. Zeileis, and T. Hothorn, "Bias in random forest variable importance measures: Illustrations, sources and a solution," BMC Bioinformatics, vol. 8, art. 25, 2007, doi: 10.1186/1471-2105-8-25.

[41] The Aluminum Association, International Alloy Designations and Chemical Composition Limits for Wrought Aluminum and Wrought Aluminum Alloys. Arlington, VA, USA: The Aluminum Association, 2018.

[42] R. L. Goetz and S. L. Semiatin, "The adiabatic correction factor for deformation heating during the uniaxial compression test," J. Mater. Eng. Perform., vol. 10, no. 6, pp. 710–717, 2001, doi: 10.1361/105994901770344593.

[43] D. H. Wolpert, "Stacked generalization," Neural Netw., vol. 5, no. 2, pp. 241–259, 1992, doi: 10.1016/S0893-6080(05)80023-1.

[44] C. E. Rasmussen and C. K. I. Williams, Gaussian Processes for Machine Learning. Cambridge, MA, USA: MIT Press, 2006.

[45] Y. V. R. K. Prasad, H. L. Gegel, S. M. Doraivelu, J. C. Malas, J. T. Morgan, K. A. Lark, and D. R. Barker, "Modeling of dynamic material behavior in hot deformation: Forging of Ti-6242," Metall. Trans. A, vol. 15, no. 10, pp. 1883–1892, 1984, doi: 10.1007/BF02664902.

© 2026 The Author(s). Published by Qassim University under CC BY-NC-SA 4.0.