EPJ Web of Conferences **345**, 01054 (2026) https://doi.org/10.1051/epjconf/202634501054
ICE3MT2025

Prediction of tensile strength in AA7075/SiC composites using random forest and support vector machine models

Veera Venkata Nagaraju[1,*], *Radhamanohar* Aepuru[1], *Ruben Fernandez* Urritia[1], *D. V. Sudarsan* Reddy[2], *Anmol* Singh[3], and *Kadapa* Vijaya Bhaskar Reddy[4]

[1]Departamento de Ingeniería Mecánica, Facultad de Ciencias Físicas y Matemáticas Universidad de Chile, Santiago-8370451, Chile
[2]Department of Electrical and Electronics Engineering, Mohan Babu University, Tirupati-517102, India
[3]Department of Chemistry, School of Physical Sciences, DIT University, Dehradun, 248009, India
[4]Department of Mechanical Engineering, Marri Laxman Reddy Institute of Technology and Management, Dundigal, Hyderabad – 500043, India

Abstract. In the present study, regression approaches of machine learning were implemented to predict the tensile strength of AA7075/SiC composites fabricated under different process parameters. SiC composition, compaction pressure, sintering temperature, and sintering time were considered as independent variables, while tensile strength was considered as a dependent response. This study presents an exploratory data analysis through parametric distribution using pairplots and boxplots, showing how process parameters affect the tensile strength of AMCs. GridSearchCV was used for hyperparameter tuning in the following two optimized supervised regression models: Random Forest and Support Vector Machine. The coefficient of determination (R^2), mean absolute error (MAE), and root mean square error (RMSE) were used to evaluate model performance. The Random Forest model demonstrated better predictive performance than the SVM model, as evident from $R^2 = 0.965$, MAE = 7.81, and RMSE = 9.50 versus $R^2 = 0.838$, MAE = 16.51, and RMSE = 20.43, respectively. This implies that the Random Forest algorithm is robust enough to provide a better generalization on nonlinear relationships between processing parameters and the tensile strength of composites and thus is reliable for optimization of composite manufacturing processes.

1 Introduction

The continuous development of materials science has been motivated by a need for structural materials that possess an optimum combination of high strength, low weight, and superlative performance under diversified service conditions. Among the various classes of engineering materials, MMCs have received considerable attention in aerospace, automotive, marine, and defense industries based on their prospects for attaining enhanced toughness and ductility by

[*] Corresponding author: naga.damu125@gmail.com

using metals as matrices in combination with hardness and wear resistance given by ceramic reinforcements [1]. Amongst the various classes of MMCs, the aluminum-based composites represent one of the most extensively studied categories due to their advantageous balance between mechanical properties, corrosion resistance, good machinability, and relatively low cost of production compared to other advanced composites. By varying the reinforcement type, particle size, and volume fraction, aluminum MMCs can be engineered to fit the specifications imposed by high-performance applications [2].

Among the aluminum alloy family, the AA7075 series has been one of the important candidate materials for reinforcement because it inherently possesses a high strength-to-weight ratio, good fatigue resistance, and excellent mechanical performance after proper heat treatment. Generally, AA7075 is strengthened by precipitation hardening and is a zinc-rich aluminum alloy widely used in aerospace structural components, automotive suspension, and defense equipment. Despite its attractive strength levels, poor wear resistance and moderate toughness are some of the limitations of AA7075, which make it unsuitable for applications under severe loading, high friction, or abrasive wear conditions. These demerits have encouraged the researchers to explore different composite fabrication routes in which AA7075 is used as a matrix, reinforced with ceramic particles to enhance the load-bearing capacity, stiffness, and wear resistance of the matrix material [3].

SiC has been identified among the most effective reinforcements for aluminum alloys due to its high hardness, excellent thermal stability, low density, and good interfacial compatibility with aluminum. Addition of SiC particles to AA7075 has been reported to result in significant improvements in properties such as tensile strength, hardness, and wear resistance as well as dimensional stability at high temperatures [4]. Their reinforcement in general contributes to the load transfer from the ductile matrix to the stiff ceramic phase, reduces dislocation motion, and refines the grain structures, which are all beneficial to mechanical performance. However, the extent of these improvements depends strongly on processing parameters like reinforcement composition, compaction pressure, sintering temperature, and sintering time. Even slight variations in these factors have been known to be associated with marked changes in microstructural features responsible for final mechanical properties, such as porosity, particle distribution, and interfacial bonding.

Traditional optimization of mechanical performance in composite systems relies strongly on experimental methodologies involving trial and error. While such experiments are obviously useful, they tend to be very time-consuming, resource intensive, and poorly suited for the exploration of large multidimensional parameter spaces. For example, optimization of SiC composition, compaction pressure, and sintering conditions would require a huge number of experiments, each involving careful material preparation, processing, and mechanical testing. This makes it difficult to establish accurate predictive relationships between input parameters and output responses, especially in nonlinear interactions of complex variables. Therefore, in recent years, there has been an increased focus on the incorporation of computational modeling and machine learning into materials development to accelerate the design and optimization process while reducing experimental costs.

Machine learning, a branch of artificial intelligence, has great potential and has been successfully applied in recent years to uncover hidden patterns in data for dependable predictions in nonlinear complex systems. Unlike conventional regression or analytical modeling, which requires explicit assumptions about functional relationships, ML algorithms are data-driven and can automatically learn from historical datasets in order to build predictive models. Applications of machine learning vary widely in materials science, from alloy design and phase diagram prediction to microstructure-property correlation and estimation of mechanical properties, to name a few. The ML models can make fast predictions of the material behaviors under new conditions by leveraging existing experimental or simulated datasets and hence guide experiments and accelerate the discovery of advanced materials [5]. Among the broad categories of ML algorithms, ensemble methods and kernel-based methods have received focused attention with regard to regression issues. Random Forest is a decision tree-based ensemble learning technique with manifold applications found in the prediction of material properties because of its efficacy in handling nonlinear relationships, robustness against overfitting, and providing feature importance rankings. In RF, several decision trees are constructed on bootstrapped samples of the dataset, and their predictions are aggregated to yield a final output, hence reducing variance and improving generalization. It also makes RF appropriate for problems where property prediction depends on complex interactions between multiple input variables, which is the case with MMCs [6].

On the other hand, the Support Vector Machine is a kernel-based learning method aimed at constructing the optimal hyperplanes in high-dimensional feature spaces to separate or fit data points. In its use for regression tasks-known as Support Vector Regression, or SVR-the aim is to perform minimization of prediction errors within a predefined margin, hence becoming very effective in nonlinear function approximation. The robustness of SVMs comes with the ability to utilize kernel functions (linear, polynomial, or RBF) that turn the input data into higher dimensions where linear regression becomes feasible. Although SVM models are powerful, they are sensitive to the choice of such hyperparameters as the regularization parameter (C) and kernel coefficient (gamma) that have to be laboriously tuned for optimal performance [7].

The focus of this study is the prediction of the tensile strength for AA7075/SiC composites using Random Forest and Support Vector Machine regression models. The input parameters that will be taken into account include SiC composition, compaction pressure, sintering temperature, and sintering time. Tensile strength is the target response variable. Both models use GridSearchCV for hyperparameters optimization so as to get the optimal predictive capability. Models will be assessed statistically using coefficient of determination (R^2), mean absolute error, and root mean square error, amongst others. In addition, pairplots and boxplots will be used to illustrate the variation of process parameters against tensile strength, and regression plots with confidence intervals will be used in assessing model accuracy [8].

2 Materials and methods

The base matrix material used in the present study is AA7075 aluminum alloy due to its high strength-to-weight ratio, while reinforcement was done by silicon carbide (SiC) particles of size 20–40 µm in order to improve the mechanical performance of composites. Composites with variable SiC composition (0–9 wt.%) were fabricated by the powder metallurgy route, which enables homogeneous dispersion of reinforcement particles. Powders were mechanically mixed to attain homogeneity, and then uniaxial cold compaction at pressures ranging from 400 to 800 MPa was carried out using a hardened steel die. The green compacts were then sintered in an argon atmosphere within a temperature range of 350–580 °C for 90–145 minutes to ensure adequate metallurgical bonding between matrix and reinforcement with minimum porosity. After sintering, the specimens were furnace-cooled to room temperature and subsequently machined into tensile test specimens according to ASTM E8/E8M standards [9]. Tensile testing at room temperature was carried out using a UTM, and the average tensile strength of three replicates for each processing condition was recorded to ensure the accuracy and repeatability of experiments. Thus, the obtained dataset comprises four independent variables, namely SiC composition, compaction pressure, sintering temperature, and sintering time, while tensile strength is considered the dependent response. EDA was carried out by employing Python libraries wherein pairplots were used to visualize inter-variable correlation as depicted in Figure 1, and boxplots were generated to investigate the effect of an individual parameter on tensile strength as depicted in Figure 2.

Two regression algorithms, Random Forest and Support Vector Machine, were applied for predictive modeling using the scikit-learn library. Random Forest is an ensemble-based method whose main focus is reducing variance by aggregating multiple decision trees. With the help of GridSearchCV, its optimized hyperparameters included the number of estimators (100, 200, 300), maximum tree depth (None, 10, 20), and minimum samples per split (2, 5, 10). Support Vector Regression represents another kernel-based approach that can capture nonlinear dependencies. The optimized hyperparameters of Support Vector Regression were the penalty parameter C (0.1, 1, 10), kernel type (linear, rbf), and kernel coefficient gamma (scale, auto). To accomplish this, both models were trained on 80% of the dataset, then tested on 20%, where tuning of hyperparameters was done on the training set by 5-fold cross-validation using the coefficient of determination, R^2, as a metric for evaluation. Then, the predictive capability of the tuned models was assessed on the test dataset using several statistical measures such as R^2, MAE, and RMSE. Actual versus predicted regression plots of tensile strength were generated for both models, including a best-fit regression line with 95% confidence intervals and printed metrics in order to provide a clear visual assessment of model accuracy [10]. The overall methodology adopted a step-by-step workflow that comprises composite fabrication, tensile testing of the fabricated specimen, preparation of the dataset, exploratory analysis, development of machine learning models with optimized hyperparameters, evaluation using statistical metrics, and visualization performance of regression. Therefore, with such a structured approach, experimental data are combined systematically with machine learning methodologies for the prediction of tensile strength [11].

3 Results and discussion

This means that the mechanical performance of AA7075/SiC composites is very dependent on reinforcement composition, compaction pressure, and sintering conditions. Considering this, the current study seeks to integrate regression machine learning models with experimental datasets to predict tensile strength based on the following independent process variables: SiC composition, compaction pressure, sintering temperature, and sintering time. The results will be presented in two stages: exploratory data analysis of experimental results and predictive modeling through Random Forest (RF) and Support Vector Machine (SVM) regression.

Figure 1 presents the correlations of process variables with tensile strength as derived from the pairplot. In general, it was found that SiC composition increased tensile strength to a certain optimum beyond which clustering and porosity could offset the strengthening effect. Compaction pressure was found to be strongly positively correlated to TS, in agreement with its role of reducing porosity and improving particle bonding during sintering. Sintering temperature and time showed nonlinear effects, with moderate values favoring improvement in density and bonding, and with excessive values risking grain coarsening or interfacial degradation that could weaken tensile strength.

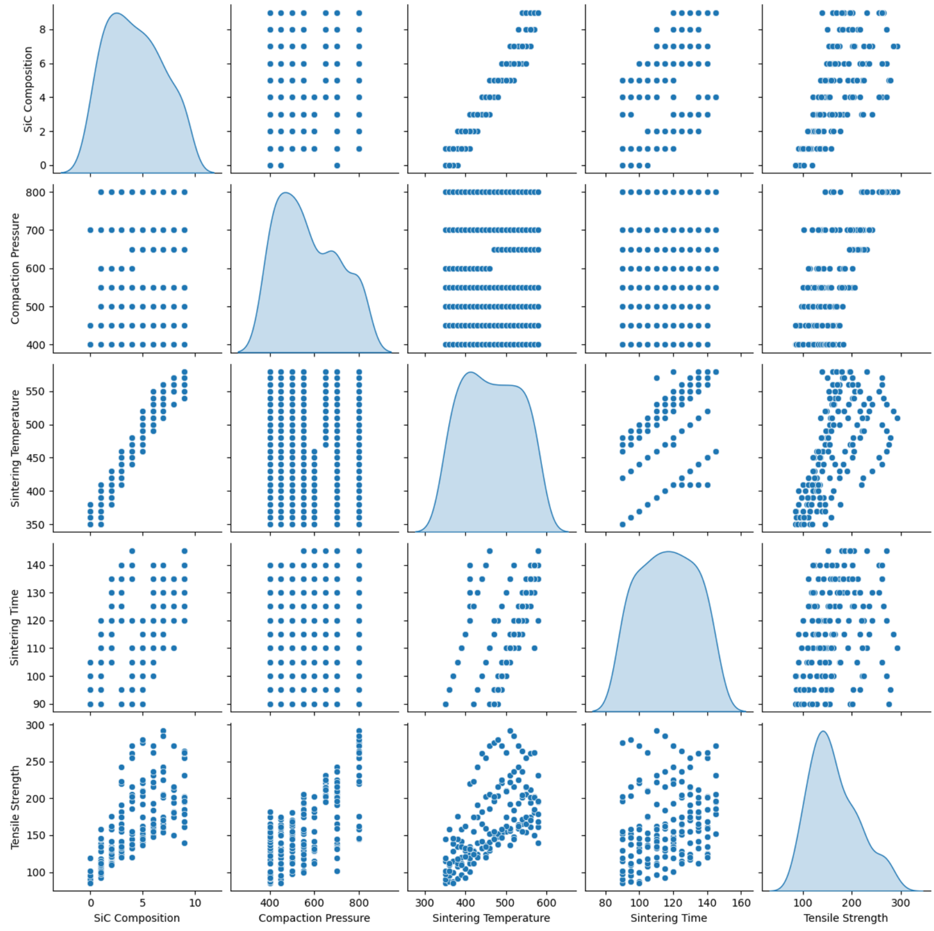

Fig. 1. Pairplot of process parameters (SiC composition, compaction pressure, sintering temperature, and sintering time) and tensile strength of AA7075/SiC composites.

Figure 2, boxplots of tensile strength for each of the processing parameters, provided further detail on the distribution of tensile strength with respect to each of the parameters. Trends were evident, with compaction pressure demonstrating higher tensile performance consistently with increased levels. SiC reinforcement increased the median tensile strengths, but data variability suggested sensitivities to the processing conditions. These exploratory analyses emphasize the complexity of process-property relationships and motivate further application of more sophisticated machine learning regression methods that are capable of modeling nonlinear dependencies.

The optimization of the Random Forest regression model using GridSearchCV yielded an optimum, featuring 200 estimators with a maximum depth of 20 and a minimum of 2 samples per split. The optimized RF model exhibited excellent predictive performance on the test dataset. Figure 3(a) presents the actual vs. predicted regression plot of the tensile strength, showing a strong clustering of data points around the regression line with very little scatter. The shaded area in the figure also represents the 95% confidence interval, which is narrow, hence reliable.

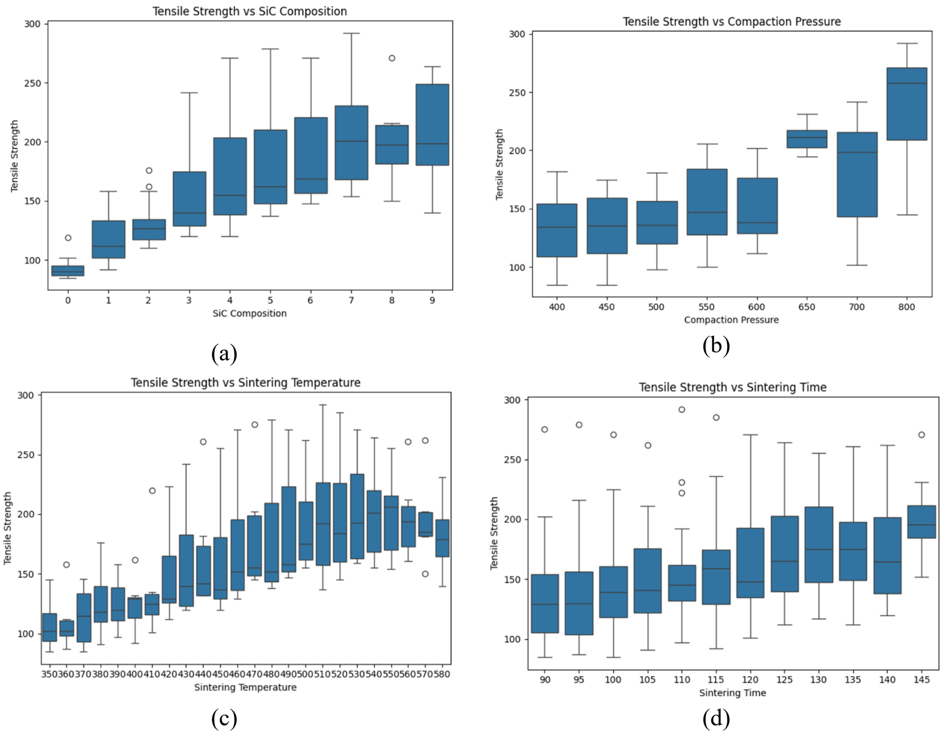

Fig. 2. Boxplots showing the variation of tensile strength with respect to (a) SiC composition, (b) Compaction pressure, (c) Sintering temperature, and (d) Sintering time.

This was further confirmed by the statistical performance evaluation. The RF model resulted in an R^2 of 0.965, which suggests that the explained variance of tensile strength in the model was 96.5%. The MAE was 7.81 MPa, which would imply that on average, the predictions fell short of the actual values by less than 8 MPa. The RMSE was 9.50 MPa, indicating a small standard deviation for the prediction errors. Various models used in present study are shown in Table 1. These metrics, therefore, establish RF as a robust model with which to predict tensile strength.

Table 1. Metrics of models used in the present study.

Model	R Square (R^2)	Mean Absolute Error (MAE)	Root Mean Squared Error (RMSE)
Random Forest (RF)	0.9649	7.8049	9.4946
Support vector Machine (SVM)	0.8380	16.5137	20.4257

From the perspective of material properties, nonlinear interactions can be handled by RF. The reinforcement strengthening effect by SiC is not a simple linear function; higher concentration leads to agglomeration and uniformity in dispersion. Similarly, the compaction pressure and sintering conditions interact in complex ways that affect densification and grain structure. These relationships were effectively captured while averaging multiple decision trees by Random Forest, giving stable and accurate predictions. Its high R-squared value demonstrates the suitability of ensemble learning approaches to materials property modelling [12].

Similarly, the hyperparameter tuning of the Support Vector Machine regression model was performed using GridSearchCV. The best parameters were identified as C = 10, kernel = rbf, and gamma = scale. Figure 3(b) provides the regression plot for the SVM predictions. In support of the previous discussion on RF, the SVM plot had a higher amount of scatter around the regression line, and there was a wider distribution of data points throughout the confidence interval. The model captured the general trends, but it was less precise in its ability to predict values over the range of tensile strength [13].

Quantitatively, the R^2 value was 0.838 for the SVM model; thus, about 83.8% of the variance in tensile strength was described by the model. The MAE for the model was about 16.51 MPa, over double that of RF, while the RMSE was 20.43 MPa, reflecting larger spreads in the prediction errors. These results thus suggest that although SVM is indeed capable of learning nonlinear relationships, it truly performed worse than that of RF for this dataset [14].

The limitations of SVM in this case may be attributed to the relatively small dataset size and high sensitivity of SVM modeling to hyperparameter tuning. While the rbf kernel is powerful for capturing nonlinear trends, it generalizes less well when noise or possibly overlapping effects of parameters are present in the dataset. In general, SVM models concentrate on boundary datapoints, which also reduces their robustness for property prediction in materials systems with complex internal interactions [15].

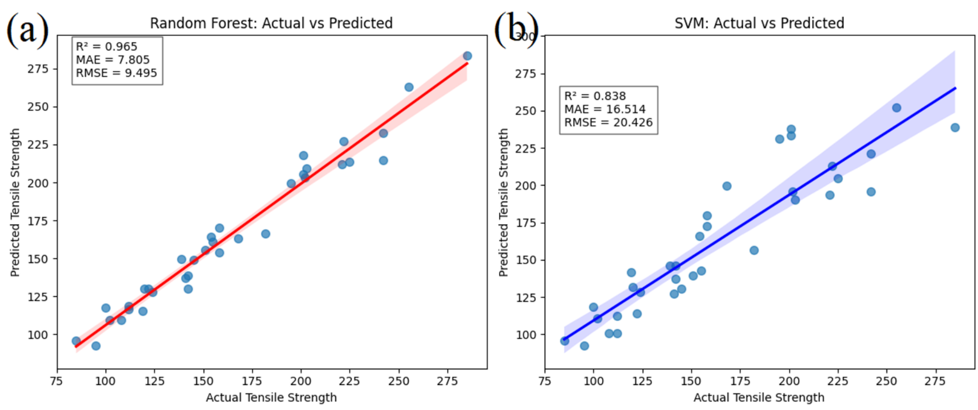

Fig. 3. Regression plots of predicted versus actual tensile strength using (a) Random Forest and (b) Support Vector Machine models.

A direct comparison between Figure 3(a) and Figure 3(b) indicates that the performance of Random Forest is much better than that of SVM in the prediction of tensile strength. In the RF regression plot, the predicted values were more closely aligned with the experimental results, while the SVM model had more scatter and deviation, especially for the largest and smallest values of tensile strength. This agrees with the numerical evaluation, where RF outperformed SVM significantly according to all three metrics: R^2, MAE, and RMSE.

From a materials engineering perspective, it can be noticed that the Random Forest model captured the nonlinear influence of compaction pressure, sintering temperature, and reinforcement composition on tensile strength better than the other methods. Its ensemble approach reduces overfitting, making the predictions robust in nature when interactions become complex and multidimensional. On the contrary, while being conceptually more powerful, SVM faced difficulty in generalizing across the dataset, making them sensitive to the tuning of parameters and data distribution.

The results highlight the importance of choice for machine learning models to better suit the nature of the data. In the case of composite materials with influences from multiple factors, ensemble-based methods such as Random Forest are more reliable than kernel-based models like SVM.

4 Conclusions

- The composites of AA7075/SiC were fabricated by using the powder metallurgy route. The main mechanical response was tensile strength. The independent parameters selected included SiC composition, compaction pressure, sintering temperature, and sintering time; tensile strength was the dependent variable.
- Exploratory data analysis, such as pairplots and boxplots, showed clear correlations between processing parameters and tensile strength. Some major contributors were identified to be compaction pressure and SiC content. Two regression models, including Random Forest and Support Vector Machine, were developed and optimized using GridSearchCV with 5-fold cross-validation.
- The Random Forest model had the best prediction performance with $R^2 = 0.965$, MAE = 7.81 MPa, and RMSE = 9.50 MPa, showing a close alignment between predicted and actual values (Figure 3a). Support Vector Machine captured the general nonlinear trends but showed lower accuracy with $R^2 = 0.838$, MAE = 16.51 MPa, RMSE = 20.43 MPa, and more scatter in the predictions (Figure 3b).
- Comparative analysis ascertained that the methods of ensemble learning, such as RF, are superior to kernel-based methods, such as SVM, in modeling nonlinear relationships in composite processing. The machine learning framework will reduce trial-and-error experimentation considerably in optimizing the processing parameters to achieve the mechanical performance of interest.
- The methodology can be extended to predict hardness, toughness, and wear resistance, and further improved by incorporating microstructural features that may be obtained from image analysis.
- Overall, this study showed that Random Forest regression is a reliable and robust method for the prediction of tensile strength in AA7075/SiC composites, thus opening up important pathways toward data-driven composite design and optimization.

References

1. N. Yilmaz, E. Baltaci, O. Baygin, T. Tüzüner, S. Ozkaya, A. Canakci, Effect of the usage of Er,Cr:YSGG laser with and without different remineralization agents on the enamel erosion of primary teeth, Lasers Med Sci 35 (2020) 1607–1620. https://doi.org/10.1007/s10103-020-03015-0.
2. K.V. Reddy, C.R. Reddy, Microstructural, mechanical and damping analysis of Al/Gr composites produced via friction stir processing: effect of rotational speed, Canadian Metallurgical Quarterly (2025) 1–10. https://doi.org/10.1080/00084433.2025.2508105.
3. V.S.S. Venkatesh, P.R. Vundavilli, M.M. Mahapatra, A. Goyal, A. Bhowmik, Extraction of kaolin and tribo informative analysis of the Al-kaolin composite through machine learning approaches, Sci Rep 15 (2025) 13370. https://doi.org/10.1038/s41598-025-97782-x.
4. V.S.S. Venkatesh, R.N. Rao, S.N. Singh, Tribo-informative Analysis of Spark Plasma Sintered Al-SiC-HAp Hybrid Composite Through Machine Learning Techniques, Arab J Sci Eng (2024). https://doi.org/10.1007/s13369-024-09835-4.

5. M.V.N.V. Satyanarayana, S.R. Maity, K.M. Pandey, V. Reddy, B. Pasha Mahammod, Effect of fabrication techniques on mechanical and microstructural behavior of AA7075/SiC/ZrC hybrid composite, Ceram Int 49 (2023) 37782–37792. https://doi.org/10.1016/j.ceramint.2023.09.106
6. V.K.S. Jain, K.K. Kumar, A. Kumar, S. Sundar, Influence of Cooling Media on the Electrochemical Behavior of Friction Stir Processed Al2014 Alloy, JOM 75 (2023) 526–536. https://doi.org/10.1007/s11837-022-05611-7.
7. A. Kumar, M.V.N.V. Satyanarayana, P.R. Mattli, A. Bhowmik, Enhanced Performance of AA7075/SiC/ZrC Hybrid Composite through Microwave Assisted Powder Metallurgy Techniques, J Mater Eng Perform 34 (2025) 4336–4349. https://doi.org/10.1007/s11665-024-09405-7.
8. K. Venkateswara Reddy, Dynamic mechanical and microstructural analysis of graphite-reinforced Al6061 alloy processed by friction stir processing, J Adhes Sci Technol (2025) 1–11. https://doi.org/10.1080/01694243.2025.2508752.
9. R. Kumar, P. Mastanaiah, A. Kumar, Microstructural Characterization and Mechanical Properties of Wire and Arc Additively Manufactured 18% Ni Maraging Steel 250, J Mater Eng Perform (2025). https://doi.org/10.1007/s11665-025-12339-3.
10. V. Reddy, Vaishali, From fabrication to prediction: unraveling the tensile strength of Al/SiN composites through machine learning, International Journal on Interactive Design and Manufacturing (IJIDeM) (2025). https://doi.org/10.1007/s12008-025-02274-x.
11. G. Manohar, A. Kumar, Exploring the Effects of Ball Milling and Microwave Sintering on Microstructural and Mechanical Behavior of AA7075/SiC/ZrO2 Hybrid Composites, Silicon (2024). https://doi.org/10.1007/s12633-024-02856-6.
12. P.R. Matli, R. Abdelatty, R.A. Shakoor, A. Azeem, S.S.R. Lingala, R.G. Kotalo, A.M.A. Mohamed, Characterization of microstructural and mechanical properties of hybrid Al/SiC/Al2O3 nanocomposites, Emergent Mater (2024). https://doi.org/10.1007/s42247-024-00636-x.
13. G. Manohar, Machine learning approaches for tensile strength prediction in Al/B4C metal matrix composites: a comparative analysis of neural networks and ensemble methods, International Journal on Interactive Design and Manufacturing (IJIDeM) (2025). https://doi.org/10.1007/s12008-025-02347-x
14. A. Bhowmik, A.J. Santhosh, Predictive Modelling of Tensile Strength and Hardness of Microwave Assisted AA7075/Sic Composite Through Supervised Machine Learning Algorithms, Results in Engineering (2025) 107793. https://doi.org/10.1016/j.rineng.2025.107793.
15. S. Thapliyal, R. Mamedipaka, Investigating the performance of aluminium-steel bimetallic structure fabricated through Wire and Arc Additive Manufacturing with interlayer approach, Structures 72 (2025) 108304. https://doi.org/10.1016/j.istruc.2025.108304.