MACHINE LEARNING FOR PARAMETRICAL ANALYSIS OF FRICTION STIR WELDED ALUMINUM METAL MATRIX COMPOSITES

K. Saravanan, A. Giridharan

Department of Mechanical Engineering, Vellore Institute of Technology, Chennai Campus, India
*Corresponding author's e-mail address: giridharan.abimannan@vit.ac.in

ABSTRACT

The research focuses on the behaviour and process parametric influence on friction stir welded Al metal matrix composites reinforced with varied percentages of SiC, B_4C, and Mg. The experimentation involves fabrication of Al metal matrix composites followed by friction stir welding and, subsequently, evaluation of the joint properties in terms of mechanical strength, microstructural integrity, and quality. In comparison to other joints with varied base material compositions, the weld exhibits refined grains and uniform distribution of hybrid particles in the joint region, resulting in increased strength. Higher SiC composition adds to greater strength, better wear characteristics, and harness, whereas B_4C percentage is linked to hardness. The maximum ultimate tensile stress for a particular sample was determined to be around 160MPa, while the maximum percentage elongation was found to be around 165 for 10% SiC and 3% B_4C. As the amount of SiC declines and that of B_4C rises, the percentage elongation decreases. In samples with a B_4C weight percentage of 10%, the greatest hardness measured was around 103Hv. For a load of 30N, the wear rate was as high as 12gm/s with a SiC weight percentage of 10. For lower load values and a higher percentage of B_4C, the wear rate often decreased. Chemical properties are barely changed. Therefore, the materials keep their original qualities after welding. During the non-destructive testing process, no large cracks, pores, or clusters of pores are found, indicating that the weld is of good quality. To achieve a satisfactory weld, optimal ranges based on analysis using machine learning of rotary tool speed, tool linear velocity, transverse speed are maintained. Linear Regression algorithm, Random Forest algorithm and Lasso Regression algorithms are being used and the results are also compared. This work covers a wide range of topics, and the results are found to have improved significantly in most cases and is in good agreement with data previously presented in the literatures.

KEYWORDS: Metal matrix composites, linear regression, Machine Learning, tensile test, Random Forest Algorithm, optimization, microscopy.

1. INTRODUCTION

Aluminum alloys are reinforced with hard ceramic particles in numerous industrial applications to improve the mechanical properties of Al-MMC (Aluminum Metal Matrix Composites) [1] - [3]. Al-MMC is a lightweight material with high stiffness, hardness, strength, melting point, and wear resistance [4] - [6]. Aluminum metal matrix composites are only used for long-term applications such as military weaponry and aerospace because of their greater production costs. Al-MMC is also used in car components such pistons, engines, disk brakes, cylinder liners, and drum brakes [7], [8]. Within the weld region's matrix and reinforcing particles, fusion welding of aluminum MMC produces brittle intermetallic components. The stress produced by the weld reduces joint efficiency and exposes porosity and voids at the joint [9], [10].

A non-consumable spinning tool with higher toughness than the base material is pitched into the faying/butt ends of the plates to be welded in the friction stir welding (FSW) process. They are subjected to an axial force that is created along the joint line. The frictional heat generated by the tool during the spinning operation softens the material that is subjected to the rotational movement of the tool pin. The revolving tool moves the plastically warped material from the tool's

Date of submission/acceptance: 28.06.2022/15.12.2022 https://doi.org/10.35219/awet.2022.05

front to the tool's backside. The weld is easier to achieve with subsequent forging.

Ali et al. [11] studied and linked the hardness and tensile parameters of weld specimens in AA6061/SiC/B_4C composites to microstructural variation. When the time of exposure of the FSW tool is long, the nugget zone and grain growth are influenced, according to Palanivelet al. [12]. Storjohann, D et al [13] examined the degradations and microstructure evolutions of the reinforced Al MMC with Al_2O_3 during fusion welding and correlated it with the thermodynamic calculations. Decomposition of Al_2O_3 was predominant during the joining process. In SiC whisker-reinforced composites, during the friction joining process, the decomposition of SiC to Al_4C_3+Si occurred owing to the reaction with molten aluminium. Reorientation of SiC whiskers near the boundaries of the dynamically recrystallized and thermo-mechanically affected zone (TMAZ) was recorded [14]. Threaded cylinder profiled tool made of H13 steel with a D/d ratio of 2 yielded weld joints of highest strength. Besides, the weld zone microstructure of the Al-10% ZrB_2 MMC and Al-10% SiC welded using H13 threaded cylinder tool demonstrated considerable grain refinement with the precipitates homogeneously distributed [15].

There are adequate literatures on FSW of SiC and B4C reinforcement Al MMC, for which industrially acceptable mechanical property ranges and microstructural integrities for dependability have been determined. Reports on weldability research on these MMCsin particular, are limited and lack an interdisciplinary approach. This opens up more possibilities for research into temperature effects on microstructures, process parametric effects on weld efficiency, and mechanical behaviour analysis. As a result, this research aims to fabricate Al-MMC reinforced with various SiC, B4C, and Mg particles, subject it to FSW, followed by mechanical tests and microstructural investigations of the welded samples. Subsequently, process parametric influences on the weld performance are established. Furthermore, process parameters are generally distributed over a large range and the complexity of the thermomechanical behaviour keeps the individual effects enveloped. This limitation stresses on the requirement of research on mechanisms to arrive at parameters of importance and thus help in identifying the initial process parameters for optimum builds. To avoid the issue of overfitting and under fitting of models, techniques of regularization and dropout is implemented.

The ML techniques widely adopted by various researchers for the industrial application processes give results with high accuracy in various domains of applications excepting the latest manufacturing techniques. This is because of a lack of a wide range of training data due to the expense of collecting experimental values. Results of ML show higher accuracy and reproducibility for any number of analysis/trials and thus, proves significant in optimizing the performance of industrial processes. Use of AI methods would serve the requirement of selection and prediction of appropriate input parameters for an optimum weld for any of the welding techniques.

2. EXPERIMENTS

Metal matrix composites with AA6061 matrix and B_4C and SiC reinforcements in various percentage compositions respectively is fabricated using the stir cast technique. The friction stir welding method is applied to each of these combinations independently during each of the trials which are carried out by altering the parameters of the FSW procedure. The welds are next examined for strength, microstructural integrity, and quality.

2.1. Preparation of Composite Materials (Phase-I)

The metal matrix composites of Al that are strengthened with B_4C and SiC are prepared in a crucible furnace. Degassing tablets containing hexachloroethane are used to remove dross. B_4C and SiC are pre-oxidized for 2 hours at 650°C before being placed into the liquid matrix (AA6061) and swirled at a constant rate. Heat treatment of B_4C is used to build a layer on SiC to increase molten metal bonding. Following the addition of B_4C and SiC, the melt is stirred at a consistent rate at a defined ideal speed and time, 500 rpm for 10 minutes. The molten metal is then poured into an iron die mold. There are no signs of macro casting flaws. Magnesium was added to the alloy throughout the stirring and melting process to improve particle wettability.

2.2. Setup for the Experiment

The outside shell of the electrical furnace has dimensions of 500x500x500 mm, the crucible volume is 180 mm diameter, 250 mm height, and the conical bottom is 60 mm, and the furnace height is 0.75 m.

Fig. 1. Stir casting set up

The furnace is powered by a 230V, 6kW motor. The working temperature of molten metal is 9900.1°C, and the maximum temperature that may be sustained is 1200°C. The stir casting unit was used to fabricate the Al-MMC's (Fig. 1). Spindle speed ranging from 250-500rpm, stirring time of 8 minutes, melt temperature of 900ºC, SiC&B_4C particles preheat temperature of 600ºC and powder feed rate ranging between 0.75 to 1.0g/s were during the stirring process.

2.3. Al-FSW MMC's Experimental Approach (Phase-II)

An average 325 mesh size of SiC and 30 nm of B4C particles are reinforced on Al6061. Mg is added to increase the wettability of SiC and B_4C powders and their behaviour in Al melts. Table 1 shows the composition percentages of the samples that will be used in the experiment.

Table 1. Composition of various samples in wt. %.

Sample no	Al [%]	SiC	B_4C [%]	Mg [%]
1	82	12	4	2
2	82	10	6	2
3	82	8	8	2
4	82	6	10	2

Fig. 2. Friction stir welded sample

The work piece samples of 150 x 150 x 6 mm are prepared for a friction stir weld. The FSW tool is composed of H-13 tool steel and has a 25 mm shoulder diameter. A cylindrical tool with a diameter of 6 mm and a height of 5.7 mm is used for the study. The FSW process parameters approximate ranges were found after a few trial-and-error trials utilizing FSW 3T 300NC machine. Following that, 81 trials were carried out based on the design of experiments (DOE), with different levels of process parameters. However, because the scope of the paper is focused on the properties and behaviours of the weld, the detailed information from the DOE-based trials is not discussed here. Nonetheless, a detailed study of DOE-based trials for process parametric optimization is conducted, but it is not presented here because it is outside the scope of this paper's objectives. For the trials, the FSW tool's rotating speed, travel speed, and plunge force were 1000 rpm, 75 mm/min, and 10 KN, respectively. For tensile test specimens, the welded specimens are cut along the cross-sections according to ASTM (E8M) standards. The welded specimen is shown in figure 2.

3. MECHANICAL TESTING

FSW samples were subjected to strength, quality, and metallurgical integrity.

3.1. Tensile Testing

Tensile tests were performed at room temperature using an INSTRON 8801 universal tensile testing equipment with a capacity of 40 tonnes (High Wycombe, UK). The ASTM (E8M) standard was used to prepare the specimens. Figure 5 depicts the results for tested samples. It can be seen that the first sample, which contains less B_4C than the other samples, results in a higher final tensile stress for the first sample. The use of Al in increasing percentages enhanced the tensile strength [13]. This is due to the high plasticity and density of Al, as well as to the fact that at the right temperature, the tensile strength of Al in the region improves, contributing to the total strength. The ultimate tensile strength of SiC and B_4C improves by 10% and 3%, respectively.

Tensile strength falls when the percentage of SiC decreases and B_4C rises, suggesting proportionality between SiC percentage and inverse proportionality for B_4C percentage. The even dispersion of fine Al-SiC particles is to blame for this. Higher SiC percentages resulted in more Al-SiC particle precipitates during the solidification of the homogeneous solution, resulting in the highest composite strength. The maximum yield stress achieved is 142.58 MPa, which is well within most applications' acceptable range. With a decrease in SiC, and a rise in B_4C content, the percentage elongation dropped from 6.48 to 4.08. The huge number of dislocations in the aluminium metal matrix accounts for its great strength. Due to the resulting internal tensions, additional dislocations were produced by SiC and B_4C particles during the production of Al-composite. Because of the discrepancy between the thermal expansion coefficients, each of the reinforcement particles cools at a different pace from the weld process temperature. Furthermore, the FSW process is responsible for the homogenous dispersion of these particles at the joint due to friction and grinding action. The dislocation density of SiC and B_4C particles, plastic deformation, and interactions between dislocations are all elements that contribute to the strengthening of composite joints after the FSW process. At high temperatures, SiC has a high strength-to-weight ratio and a high strength retention. When this is added to a metal matrix composite, it improves chemical stability and reduces non-catastrophic failures, which are not discussed because they are beyond the scope of this study. It is crucial to note, however, that processing SiC at high temperatures is only recommended when the

reinforcement will be processed at high temperatures. The use of monolithic ceramics is also increased, resulting in composites that are more fragile. The thermal expansion coefficient between the reinforcement and the matrix contributes to the thermal cooling stress from the processing temperature during the FSW process.

Figure 3 shows SEM micrographs of worn surfaces under varying stresses of 10 N and 20 N. Because of the frictional heat produced during welding, the particles agglomerate in the weld zone. The sample used for the characterization is 4*3*0.5 cm^3. Figure 3 show SEM micrographs of the 20 N load. Along with the SEM micrographs, the matching EDAX analysis is also shown. Al and reinforcing particle peaks were found. Small peaks of Fe were seen, indicating that the reinforcing particles abraded the steel surface. Particles are evenly dispersed, and there is no wear on the worn surface at greater weights. The presence of delamination with severe plastic deformation in a few cases indicates adhesion wear. The existence of oxidative driven wear is indicated by the presence of O peaks in all EDAX studies. It can be deduced from figure 3 that the friction coefficient increases as the load increases, resulting in material loss from the surface.

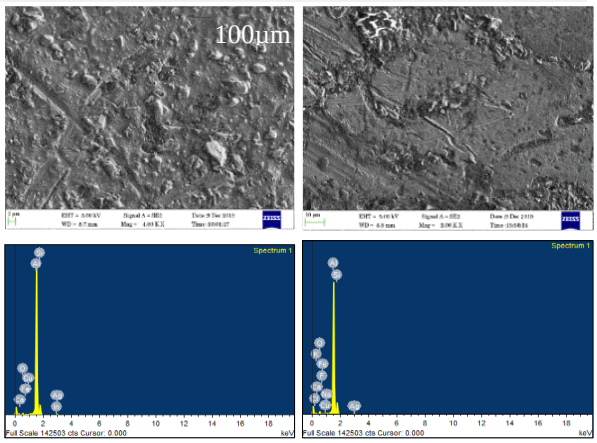

Fig. 3. SEM Images with EDAX at the FSW interface

The material loss from the surface is investigated via SEM analysis. Wear debris and ploughing marks can be seen, but they aren't significant enough to influence the material's behaviour. With increasing velocity, the friction coefficient rises. The rate of wear of composites increased as the load rose. This is because of direct metal-on-metal contact, which causes wear debris to accumulate. For 5 wt. percent to 10 wt. percent SiC, the wear rate was practically constant, and from 10 wt. percent to 15 wt. percent, the wear rate increased linearly. By restricting dislocation mobility, the reinforcing particles reduced plastic deformation. When Al is subjected to counter steel abrasion, it quickly combines with air oxygen and generates aluminium oxide on the surface. The energy dispersive spectroscopy (EDAX) examination reveals all the ingredients that occurred from the steel abrasion, including reinforcing particles, Fe, and O. The reinforcement particles eroded off the surface when the normal load rose, resulting in a higher wear rate. Elemental analysis of the macro regions in weld zone were carried out with SEM integrated with EDAX. This presented details on distribution of alloying elements in the stir zone of FSW welds. SEM image was analysed at different magnification of 50X, 500X, 1000X and 2000X.The carbon percentage in base metal, HAZ, TMZ and Nugget zone is 31.2, 26.4, 33.5 and 16.2 respectively. The oxygen percentage varied between 6-8%, the silicon percentage varied between 4-10%, the Fe percentage varied between 2-3.7% and the Boron one varied between 3-4% as observed from the elemental analysis.

4. NON-DESTRUCTIVE WELD TESTING

A radiography test is performed to determine the quality of the welds.

4.1. Radiography Examination

Friction stir welded Al-MMC samples are placed between the radiation sources (Ir 192, Co 60, and Cs 137 (in rare situations) and the radioactive film pores, cracks, and discontinuities are identified using differences in picture intensity. The sample size used for the test is 10*4* 0.5 cm^3. The images revealed the presence of negligible discontinuities in the tested specimens. The welded specimens are free of flaws and discontinuities. There are no pores or pore clusters to be found.

The formation of tiny pore clusters in the matrix metal occurs in a small number of samples, mostly welded with high tool rotating speeds and low transverse speeds. Pores around the reinforcing particle clusters may degrade the region's strength marginally, but they have no effect on the weld qualities.

5. CHARACTERISATION OF WELD MICROSTRUCTURE

The tests are carried out on the cross-sections of Al-MMC cylindrical pellets reinforced with SiC and B_4C particles. For Al, Si, and B compositions, elemental mapping is used to study the crystalline phase particle distributions. The metallic infiltration segregation in the ceramic matrix channel in the composite cools to the point of detection, and SEM integrated EDAX analysis is used to identify it (Fig. 3). Si and Al phase grains of varying sizes are dispersed in the small gaps between B_4C particles. The Al phase controls the solid solubility of various reinforcing particles, and it's worth noting that there's very little Al in SiC pellets. The Al phase is circumscribed by composite with B_4C and Si particles where the high angle pitted samples intersect with the electron beam and the B phase is circumscribed by composite with B_4C and Si particles

where the high angle pitted samples interact with the electron beam. The Al-MMC consists primarily of finer domains with diameters ranging from 0.25 to 2 m.

This may also be linked to the materials' better hardness and bond stability. The quality of the joints has been assessed using microhardness measures and tensile testing. Reduced welding times have been observed that are due to an increase in pressure during FSW, which, in turn, lead to a rise in joint strength. When aluminium alloys are friction stir welded with austenitic steel, they become elastic. The microstructure of the weld specimen is shown in figure 4.

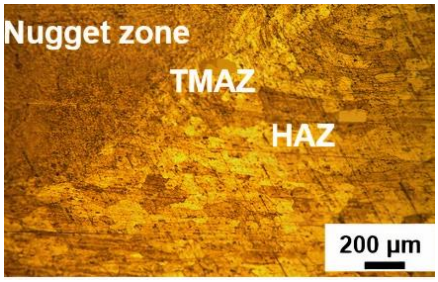

Fig. 4. Microstructure of the friction stir welded specimen

6. PARAMETRIC ANALYSIS USING MACHINE LEARNING

Various inference algorithms, which have the ability to simplify a mathematical function into a familiar form, are termed parametric machine learning algorithms. The key factors attributed to these studies on algorithm focus on the straightforwardness and speed of the data. Some of the key literatures attributed to this investigation on ML algorithms are discussed in brevity in this section.

The study involves application of Linear Regression, Lasso, and Random Forest on the experimental data of friction stir welding. These algorithms are trained using experimental data sets, based on which they predict different combinations of parameters, that are subsequently validated for determining its accuracy.

In [16], the authors used the Linear Regression Analysis with L4 orthogonal array for deriving Optimum Parametric window for Weld Current (I amps), Gas flow rate (FR L/min) and Root Gap (G mm) for analysing their effect on Distortion (D) and Tensile Strength (TS). Correlation of current and gas flow rate affect the tensile strength, while that of voltage and gas flow rate affect the bead width. Prediction error with the regression analysis was found within 10% for the tensile strength and for the bead width.

In [17], the authors proposed a technique for joint hardness prediction based upon the thermogram sequence. Correlation analysis helped develop the interactions of temperature, welding linear energy and hardness. The linear regression model resulted in a prediction error in the joint area as low as 1.25%, while for HAZ it exceeded 15%. The regression model was trained separately for each seam hardness- temperature measurement data and this variation for each sample accounts for lower accuracy. The linear relationship developed for the system may be a simplified model may not represent the actual relationship. For a more accurate model, an enhanced method for assessing the temperature in HAZ may be required to overcome the existing drawbacks connected to the variable width of seam and HAZ for different process parameters.

Random Forest (RF) is a tree-based MLA, where both classification and regression could be conducted. Amperage and overall heat input guide the weld quality. There has been limited implementation of RF with high energy joining techniques. Statistical features representing good weld, weld with burn through it and lack of fusion, which are extracted from the sound signals, were given as an input to J48 and Random Forest algorithms. RF then classifies the welds as a good weld, or having lack of fusion or burn through, in the form of confusion matrix. This helps arrive at optimum range of weld current and heat input. Classification efficiency of Random Forest algorithm is found to be 88.69% [18].

In [19], the authors proposed an online penetration monitoring methodology for GTAW using a pattern-based feature extraction of weld pool images based on Active Appearance Model (AAM) to train a Random Forest (RF) supervised machine learning method. The RF model is then used to predict the penetration states and backside weld seam width, thus enabling online control of penetration in welding. Visual features were obtained from the segmented-out weld pool edge to be given as training and testing data set for the RF algorithm. RF algorithm is a high-speed process and is also found to have high prediction accuracy. Error rate of predicted penetrations, using root mean square error in this work was restricted to 2.11%.

Friction Stir Welding (FSW) is a solid-state weld process and is widely used for joining similar and dissimilar metals, especially lightweight non-ferrous materials like aluminium, copper, and magnesium alloys. Random Forest algorithm yielded highest coefficient of determination value of 0.926. Optimum process parameters of FSW for enhancing the tensile strength of the target material using the Machine Learning approach were estimated [20].

In [21], was proposed the use of Random Forest regression model, based on particle swarm optimization (PSO-RFR), to predict the welding parameters [22]. To improve the weld quality and to enhance the automation capability, an automatic reading and writing of the weld process parameters for the PLCs is designed. According to the requirements of this model of a cold rolling mill in a steel factory, the PSO-RFR algorithm is used to train the dynamic statistics in a parameter table from the PLC welder register. The training data included steel material code, thickness, welding current, torch speed, pressure

between steel plates, swap pressure, lap length and compensation. The model predicted the welding parameters based on RFR and the predicted values can be used in the actual weld process. Prediction efficiency of the model is significantly improved in comparison to the traditional RFR algorithm.

Weld quality prediction for the Resistance Spot Weld and for the correlation model was estimated using a polynomial regression model and a logistic regression model based on the features extracted from material information and welding process signals of electrode displacement and dynamic resistance information. Prediction models were used for expulsion occurrence, failure mode, indentation depth, and tensile shear strength. Backward elimination technique was used for model estimation using these regression models and prediction accuracy was above 85% for the target parameters. Multiple regression analysis is used to determine the relationship between the dependent variables of SMAW, namely bead width and weld bead hardness with welding current, arc voltage, welding speed, and electrode stick out. Repeated data was transformed and consolidated in terms of SNR which reflects the amount of variation in the data. SNR gives an idea about the control factors that may reduce the variation and improve the weld quality. The regression analysis done by the Minitab 15 version is within 95% confidence level and optimal parameter setting of weld bead width was predicted [23].

In [24], the authors studied how to obtain the desired geometry of the back bead in CO_2 arc butt welding. The regression model equation is obtained from welding process parameters through the correlation of the parameters to the back-bead. An inverse transformation is performed to this model to obtain the predicting equations for the process parameters to acquire desired back bead. The mean error rate proves the accuracy of the regression model at the analysis and verification level. The SPSS (Statistical Package for Social Science) was used in the regression analysis. The multiple regression analysis is modelled into a linear equation to obtain the geometry of the back-bead using the welding process parameters. The error rate of analysis had a maximum value of 9.5%. The inverse transformation showed an error rate of under 6.5%.

6.1. Methodology

Data-driven models are currently being found for use in manufacturing processes because they can predict the process being analysed quickly and accurately. These models automatically learn the I/O dependencies of the process, based on previous data.

Supervised ML handles labels for input and output data and can be extended to these processes, as the manufacturing process is linked to ultimate goals and certification methods. Predictive modelling helps develop mechanisms that optimize the material joining process when there is no clear relationship between influencing parameters and weld properties such as build geometry, residual stress, and strain. This task uses linear regression, ridge, Random Forest, and Lasso modelling, respectively, to predict build parameters and tensile properties (Yield Strength, % Elongation, and UTS) and identify a hierarchy of process parameters that affect the tensile strength parameters generated during the manufacturing / joining process. The existing literature contains a brief report on the application of ML and its correlation with the manufacturing process. Therefore, this paper focuses on the application of basic linear regression, Lasso and Random Forest models, respectively, and enables better design of welding or FSW processes using ML technology, thus supporting manufacturing, or joining processes. A financially interdisciplinary approach is/can be added. Data analysis using NumPy (Numerical Python) and Pandas, which is an essential package for numerical computation, is espoused. Besides, data visualization using seaborn and Matplotlib is implemented for generating informative results of the data in the form of plots.

6.2. Machine Learning Algorithm Used for this Study

In this case, Supervised Learning algorithm is used to learn the dependent variable for tensile properties prediction, from a given set of predictors. These algorithms help you use Scikit-learn to develop functions that map inputs to outputs based on experimental datasets. The dataset obtained from the experiments is divided into 70% training data and the rest is test data that enables and validates the data predictions obtained using the regression models.

Linear Regression algorithm is used to estimate the actual values of the output parameters driven by the input parameters based on the regression line. Based on the available data, the key value for each input parameter is estimated and the optimal line between the independent and dependent variables is built.

The Ordinary Least Squares method is a technique for training a multi-input regression model in which the coefficients of the line are derived based on the mean squared error (MSE) and the R2 score is determined from the data. This is the amount that needs to be minimized when implementing the algorithm.

The Steepest Descent method is another technique for training a Linear Regression model that uses an iterative process to minimize errors in finding optimal coefficients. Some random values of the coefficients are selected and the sum of the squared errors of each input and output pair error is calculated. The coefficients are then updated to get the minimum sum of squares error. The regularization technique used in the algorithm reduces the complexity of the regression model by focusing on the absolute magnitude of the coefficients or weights assigned to the predictors.

The Feature Importance technique is used to assign weights to the influencing parameters to show

the relative importance of each feature in predicting the dependent variable. These techniques also help improve model performance by allowing dimensionality reduction. Due to the probabilistic nature of the scoring method, it is necessary to perform multiple acquisitions of coefficient values, the average of which is used as the feature importance score. On similar lines, Ridge, Lasso and Random Forest algorithms are sequentially applied and observations are made on its predictions.

6.3. Data Set for Training ML Algorithm

The material joining process has the constraint of restricted experimental data because of various practical limitations in conducting experiments, thus limiting the performance that may be obtained from any of the regression models. Data sets obtained for three different welding processes are used in this work. Data sets contain trials where the variation of input parameters is either continuous or discrete. In some cases, the step variations of the input parameters are non-uniform, and thus restricts the type of study which can be performed using the regression models and prescribes for an alternate and appropriate algorithm.

The execution of these algorithms involves the calling of seaborn, scikit and panda libraries respectively and implementing them in the Jupiter notebook. This entire process involves data collection, preliminary data analysis using the functions called from the panda and seaborn libraries, respectively. This is followed by data pre-processing which involves cleaning of data and removing non-orderly datawhich does not fit into a model. These pre-processed data are then implemented by calling the Scikit libraries that have the algorithm implementation procedures. Subsequently the models are trained with 80% of the available data and the rest of 20% of that is used for testing and validating purposes. Overfitting issues are sorted out using the regularization process and eventually R^2 is calculated which reveals the accuracy and, indirectly, the error percentage.

7. RESULTS AND DISCUSSIONS

The tensile properties as recorded from experiments are shown in table 2. It consists of UTS, YS and % elongation details for varying combinations of process parameters, which entail transverse speed and tool rotation speed.

Table 2. Tensile properties of samples with varying transverse speed and tool rotational speed

No.	Transverse speed [mm/min]	Tool rotational speed [rpm]	UTS [MPa]	YS [MPa]	[%]
1	45	540	104.4	85.5	20.04
2	54	540	88.2	80.1	13.80
3	63	540	93.6	81.9	11.38
4	45	720	85.5	73.8	17.61
5	54	720	95.4	82.8	15.53
6	63	720	100.8	84.6	14.61
7	45	900	76.5	73.8	19.11
8	54	900	87.3	86.4	16.42
9	63	900	111.5	85.5	14.90
10	54	900	124.2	116.1	6.21
11	54	900	127.8	117	6.57
12	81	720	121.5	112.5	3.78
13	81	1080	133.2	122.4	5.22
14	27	1080	178.2	160.2	6.48
15	81	720	123.3	117	4.5
16	54	900	137.7	130.7	5.58
17	54	900	143.1	131.4	6.12
18	54	720	123.3	117	4.5
19	27	720	117	109.8	4.32
20	81	900	130.5	109.8	4.05
21	81	1080	171.9	138.6	5.67
22	72	1080	171	143.1	5.85
23	27	1080	173.7	135.9	5.76
24	27	720	142.2	127.8	4.05
25	54	900	146.7	135.9	4.95
26	27	900	158.4	146.7	5.76
27	54	900	171	143.1	5.85
28	54	900	130.5	121.5	6.57
29	54	900	126	117	6.3

Similarly, tables 4, 5 and 6 shows the segregated values of tensile properties that individually account for UTS, YS and % elongation and are applied specifically to all the algorithms to understand the prediction capability of the algorithms.

Initially, ML algorithms are used for the analysis of only the UTS, as shown in table 4. The algorithms are executed using Jupiter note. Seaborn library is used for the visualization of the data in the form of plots. These libraries will provide the corresponding function for the data analysis. Figure 5 shows the parametric interdependencies and the heat map. Figure 6 shows the influence percentage of transverse speed and tool rotation speed. Tool rotation speed is the higher governing factor when compared to transverse speed. This is followed by data analysis and implementation of different algorithms. Figures 7 top 10 clearly illustrate the results together with predictions for different algorithms, namely Random Forest, Linear Regression, Lasso and Ridge. Lasso and Ridge work based on Linear Regression. However, it is only used for regularization.

Table 3 presents some observations with respect to the data frame/table such as:

- Count: Count number of non-NA/null observations.
- Meaning: Meaning of the values.
- Std: Standard deviation of the observations.
- Min: Minimum of the values in the object.
- Max: Maximum of the values in the object.
- 25% - The 25% percentile*.
- 50% - The 50% percentile*.
- 75% - The 75% percentile*.

*Percentile meaning: how many of the values are less than the given percentile.

Table 3. Statistical values for data understanding obtained with panda library function

	Transverse speed [mm/min]	**Total rotation speed [rpm]**	**UTS/n [MPa]**
Transverse speed [mm/min]	1.000000	0.055651	-0.091137
Total rotation speed [rpm]	0.055651	1.000000	0,678406
%	-0.091137	0,678406	1.000000

Table 4 represents the pairwise correlation of all the columns in the data which gives information on parametric interdependencies.

The pairs plot (Fig. 5) is built on two basic figures, the histogram and the scatter plot. The histogram on the diagonal allows us to see the distribution of a single variable, while the scatter plots on the upper and lower triangles show the relationship (or lack thereof) between two variables. For example, the left-most plot in the second row shows the scatter plot of tool rotational speed versus transverse speed. These scatter plots clearly demonstrate that there is no multi-collinearity between the input features. If there were any evidence for multi-collinearity between these two input functions, then one of these must be discarded during the analysis. Figure 5 also represents the heat map of the correlation matrix between various columns in the data frame/table. The colour grading represents the correlation between various features. The correlation value increases if the colour tends towards green and the value decreases if the colour tends towards red.

The graph from figure 6 represents the feature importance, which signifies the columns/features that show more influence on ML Model getting developed for prediction.

Table 4. Pairwise correlation of all the columns in the data frame

	Transverse speed [mm/min]	**Total rotation speed [rpm]**	**UTS\n [MPa]**
count	29.000000	29.000000	29.000000
Mean	54.620690	844.137931	127.396552
Std	16.996667	160.301440	29.153001
Min	27.000000	540.000000	76.500000
25%	45.000000	720.000000	104.400000
50%	54.000000	900.000000	126.000000
75%	63.000000	900.000000	143.100000
Max	81.000000	1080.000000	178.200000

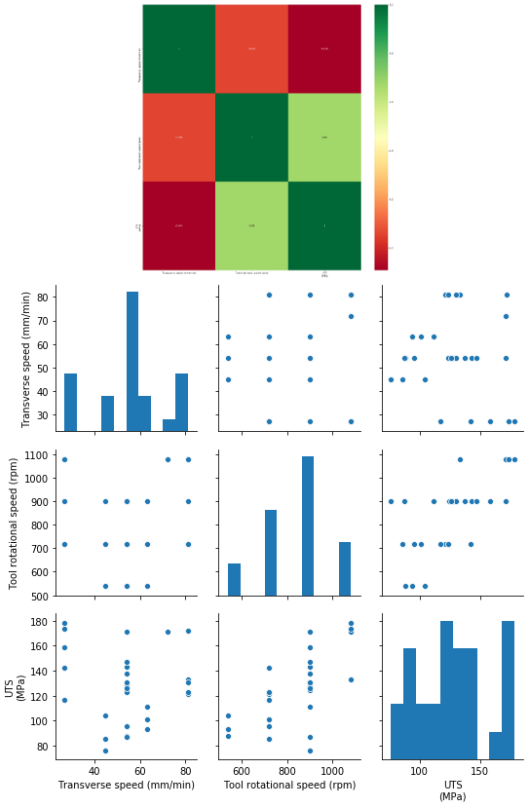

Fig. 5. Parametric interdependency plots and heat map

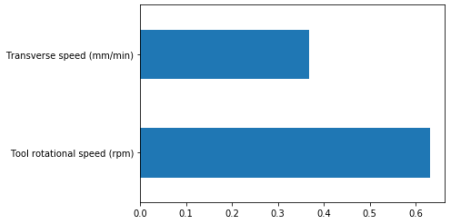

Fig. 6. Influence of rotational speed and traverse speed

7.1. Random Forest Algorithm based UTS Prediction

Random Forest Algorithm uses the ensemble technique to select the best hyper parameters for training the model. The graph shown below represents the distribution plot which usually represents a univariate analysis if you are using one parameter for the analysis. The distribution clearly depicts a non-uniform and inadequate fit. The distribution plot (Fig. 7) represents the prediction in the form of bar graphs and continuous lines, where the continuous line represents the KDE (Kernel Density estimation). The Bar Graph Represents the PDF (Probability Density Function)

Figure 7 also represents a scatter plot for the predicted value of UTS when Random Forest Algorithm is used. The above parameters represent the loss functions, the values of which have to be further minimized to get better prediction and accuracy. The accuracy achieved by the model is not appreciable and hence this algorithm is not appropriate. Therefore, the study attempts to examine the accuracy using Linear Regression ML Algorithm.

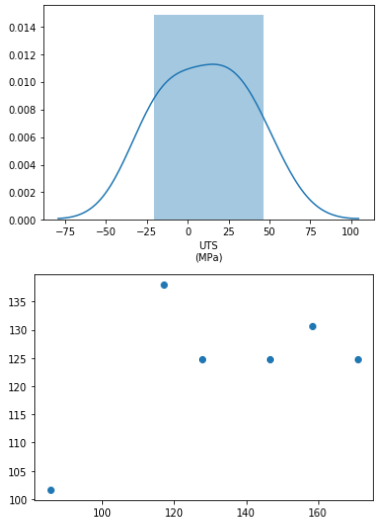

Fig. 7. Random Forest-based predictions

7.2. Linear Regression Algorithm Based UTS Prediction

Linear Regression typically creates a high bias and low variance mode. This bias is a kind of error that is experienced during the training. On adding any number of data into the available dataset, the line will have nil or negligible adjustment for which there is a low variance system. This model, applying Linear Regression Algorithm process, the parameters that includes intercept and coefficients. Regression analysis is applied, assuming that a relation has to be identified between y (output) and x (input). The mean of the residual is assumed to be zero. Furthermore, the error terms are not expected to be correlated. Besides, x and error are not to have any correlation and the error term must showcase a constant variance while also showing a normal distribution. The summation of the residual value yields a minimum value or moves to zero, a fact which is illustrated sequentially in equation 1-3.

$$\sum_{i=1}^{m} r^2 = \sum_{i=1}^{m} \left(y - (mx + c)\right)^2 \quad (1)$$

$$\sum_{i=1}^{m} r^2 = \sum_{i=1}^{m} (y^2 + (mx + c)^2 + 2y(mx + c)) \quad (2)$$

$$\sum_{i=1}^{m} r^2 = \sum_{i=1}^{m} (y^2 + m^2x^2 + c^2 + 2mx(-2ymx - -2c)) \quad (3)$$

The residuals represent the errors. The objective is to find the derivative of R, with respect to m and c, of the straight line, where m is the gradient and c is the intercept coefficient. The equation representing the objective of this regression analysis is presented in equation 4.

$$\frac{dR}{\underset{\longrightarrow 0}{dm}} = \frac{dR}{\underset{\longrightarrow 0}{dc}} \quad (4)$$

where m $_{new}$ and c $_{new}$ are determined in the equations 5 – 6, respectively

$$m_{new} = m_{old} - \eta \frac{1}{m} \left(\frac{\sum_{i=1}^{m} (y - \hat{y})}{r} \right) \quad (5)$$

$$c_{new} = c_{old} - \eta \frac{1}{m} \left(\frac{\sum_{i=1}^{m} (y - \hat{y})}{r} \right) \quad (6)$$

These equations may be further deduced to equations 7-8 in terms of learning rate.

$$m_{new} = m_{old} - \eta \nabla E_m \quad (7)$$

$$c_{new} = c_{old} - \eta \nabla E_c \quad (8)$$

Equations 7 and 8 provide the best values of m and c, where residuals are going to be either minimum or zero, and wherein η is the learning rate (hyper parameters). It tries to control the changes. According to figure 8 (univariate analysis with Linear Regression Algorithm) it may be observed that PDF and KDE are improved in compared with that of Random Forest Algorithm prediction (Fig 7). The scatter plot (Fig. 8) represents the output predicted results using Linear Regression Algorithm. The distribution plots show higher points within the curve, indicating a better prediction. R^2 is calculated to determine the accuracy matrix and can be considered one of the assumptions for feature selection. R^2 is given by the equation (9).

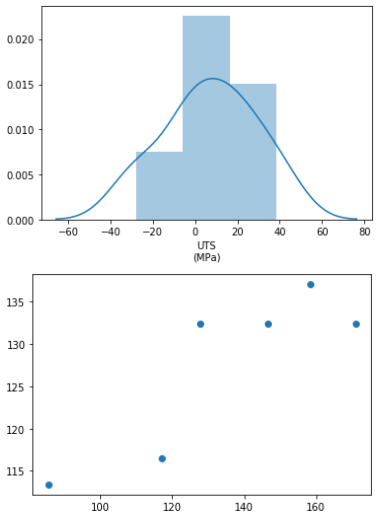

Fig. 8. Linear Regression-based algorithmic predictions

$$R^2 = \left(1 - \frac{Rss}{Tss}\right) \quad (9)$$

where RSS is the residual summation of square;
TSS – the total summation of squares.

Initially, using the scikit learn library, a statistical model is built, which follows a statistical approach. Secondly, OLS (Ordinary Least Square) model is called. Although the R^2 value increases, it may not be reliable, as it does not necessarily mean that this model is accurate, as it can also indicate overfitting. Hence, an adjusted R^2 (equation 10) that accounts for number of samples and number of features is calculated. It is expected to yield much more reliable results.

$$Adjusted\ R^2 = 1 - \left(\frac{(1-R^2)(N-1)}{N-p-1}\right) \quad (10)$$

The R^2 predicted through this result is 81.6% accuracy. Adjusted R^2 yields 87.6% accuracy. Though there is considerable improvement of the results with respect to all the features compared results attained from the Random Forest, it may further be improved by regularization of the Linear Regression Algorithm by employing Lasso and Ridge Algorithms which also fundamentally work on the Linear Regression concept. The loss functions parameters are higher than Random Forest Algorithm based predictions and require more minimization. Controlling the error term will facilitate generalization of the model in the best possible way and is called the regularization. Two of the prominent techniques used to regularize are Lasso (Least Absolute Shrinkage and Selection Operator) also known as L1.

7.3. Lasso Algorithm based UTS Prediction

This algorithm builds a model that explains the variance into the data set. The regularization is given by the equations 11-12.

$$L_1 = Rss + \lambda \sum_{j=1}^{N} |\beta_j| \quad (11)$$

where, $Rss = (Y - \hat{y})^2$ (12)

λ is the shrinkage factor;
β is m.

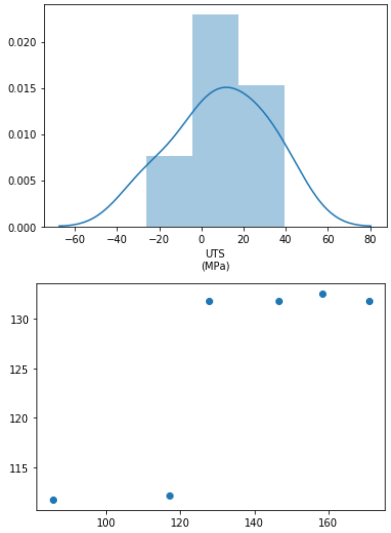

Fig. 9. Lasso Based Algorithmic Prediction

From figure 9 (univariate analysis using lasso algorithm) it may be observed that PDF and KDE have improved considerably as compared to that of Linear Regression Algorithm and Random Forest Algorithm-based predictions, as shown in figures 8 and 7, respectively.

The scatter plot (Fig. 9) represents the output predicted results using Lasso Algorithm. Though there is an improvement in the results with respect to all the features compared to the results obtained from the Random Forest and Linear Regression, it is negligible and may further be improved by using other regularization ML Algorithms. The loss functions parameters are lesser than Linear Regression Algorithm and Random Forest Algorithm based predictions and require more minimization.

7.4 Yield Strength Prediction based on ML Algorithms

Secondly, ML algorithms are employed for the analysis of only the YS and uses Table 5 for this analysis. Figure 11 shows the parametric interdependencies and the heat map. Figure 12 shows/depicts the influence percentage of transverse speed and tool rotation speed. Figure 13 to 16 clearly illustrates the results along with predictions for different algorithms, namely Random Forest, Linear Regression, Lasso and Ridge. Ridge Mean Line covers a major part of the plot, indicating its higher efficacy and being followed by Lasso.

Table 5. Yield Strength Analysis

No.	Transverse speed [mm/min]	Total rotational speed [rpm]	YS [MPa]
1	45	540	85.5
2	54	540	80.1
3	63	540	81.9
4	45	720	73.8
5	54	720	82.8
6	63	720	84.6
7	45	900	73.8
8	54	900	86.4
9	63	900	85.5
10	54	900	116.1
11	54	900	117
12	81	720	112.5
13	81	1080	122.4
14	27	1080	160.2
15	81	720	117
16	54	900	130.7
17	54	900	131.4
18	54	720	117
19	27	720	109.8
20	81	900	109.8
21	81	1080	138.6
22	72	1080	143.1
23	27	1080	135.9
24	27	720	127.8
25	54	900	135.9
26	27	900	146.7
27	54	900	143.1
28	54	900	121.5
29	54	900	117

Table 6 represents some observations with respect to the data frame/table such as:

- Count: Count number of non-NA/null observations.
- Mean: Mean of the values.
- Std: Standard deviation of the observations.
- Min: Minimum of the values in the object.
- Max: Maximum of the values in the object.
- 25% - The 25% percentile*.
- 50% - The 50% percentile*.
- 75% - The 75% percentile*.

*Percentile meaning: how many of the values are less than the given percentile.

Table 6. Statistical values for data understanding obtained with panda library function

	Transverse speed [mm/min]	Total rotation speed [rpm]	UTS\n [MPa]
count	29.000000	29.000000	29.000000
Mean	54.620690	844.137931	113.368966
Std	16.996667	160.301440	24.680922
Min	27.000000	540.000000	73.800000
25%	45.000000	720.000000	85.500000
50%	54.000000	900.000000	117.000000
75%	63.000000	900.000000	131.400000
Max	81.000000	1080.000000	160.200000

Table 7 represents the pairwise correlation of all the columns in the data frame/table.

Table 7. Pairwise correlation of all the columns in the data frame

	Transverse speed [mm/min]	Total rotation speed [rpm]	UTS/n [MPa]
Transverse speed [mm/min]	1.000000	0.055651	0.124082
Total rotation speed [rpm]	0.055651	1.000000	0.650302
%	-0.124082	0.650302	1.000000

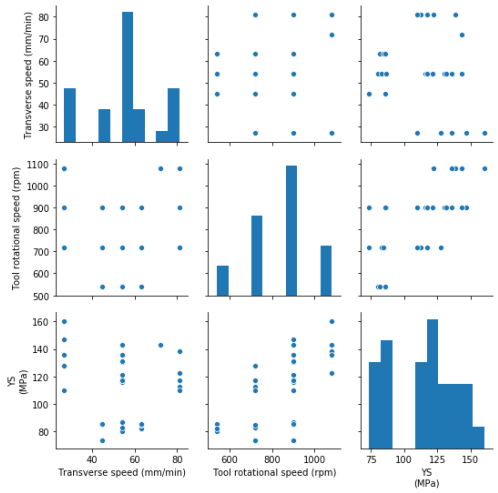

Fig. 10. Parametric interdependency and heat map

Figure 10 also represents the heat map of the correlation matrix between various columns in the data frame/table. The colour grading represents the correlation between various features. The correlation value increases if the colour tends towards green, and the value decreases if the colour tends towards red.

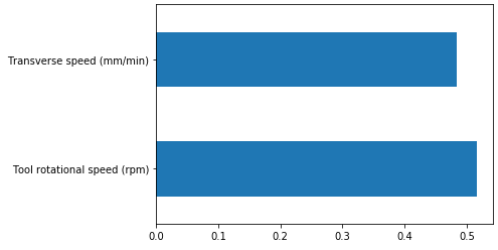

Fig. 11. Influence of traverse speed and tool rotational speed

The pairs plot (Fig. 11) is built on two basic figures, the histogram and the scatter plot. The histogram on the diagonal allows us to see the distribution of a single variable, while the scatter plots on the upper and lower triangles show the relationship (or lack thereof) between two variables. For example, the left-most plot in the second row shows the scatter plot of the tool rotational speed versus transverse speed.

The graph from figure 11 represents the feature importance, which signifies the columns/features which show more influence on the ML Model, getting developed for prediction. Tool rotation speed is a higher governing factor when/if compared to transverse speed.

7.5. Random Forest Algorithm based YS Prediction

Random Forest Algorithm uses ensemble technique to select the best hyper parameters for training the model. The graph shown below represents the dist. plot which usually represents a univariate analysis if you are using one parameter for the analysis.

The distribution plot from figure 13 represents the prediction in the form of bar graphs and continuous lines, where the continuous line represents the KDE (Kernel Density Estimation). The Bar Graph Represents the PDF (Probability Density Function).

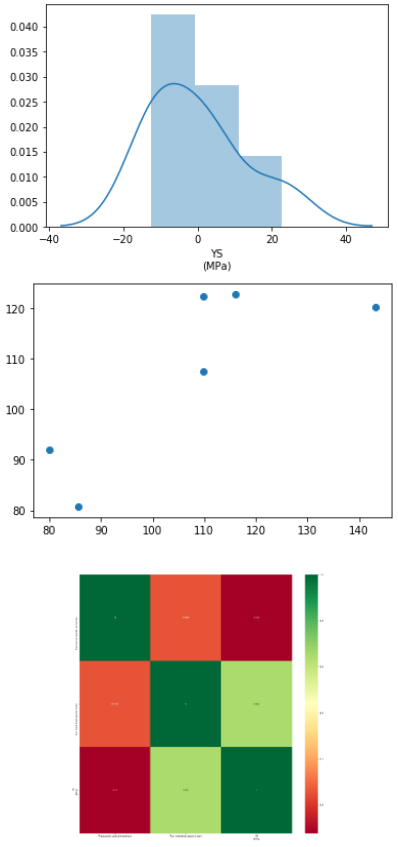

Fig. 12. Random forest (Decision Making Tree) algorithm-based prediction

Figure 12 also represents a scatter plot for predicted the value of YS when Random Forest Algorithm is used. The above parameters represent the loss functions, the values of which have to be further to get better a prediction and accuracy. The points within the curve are considerably fewer and hence the prediction is less accurate.

7.6 Linear Regression Algorithm based YS Prediction

From figure 13 (univariate analysis with Linear Regression Algorithm) it may be observed that PDF and KDE are improved, when compared to that of Random Forest Algorithm prediction (Fig. 13). The scatter plot (Fig. 14) represents the output predicted results using Linear Regression algorithm. Although there is considerable improvement of the results with respect to all the features that are compared with the results obtained from the Random Forest, it may be further improved by other lasso and ridge Algorithms through regularization. The loss functions parameters are higher than Random Forest Algorithm based predictions and require more minimization. The bigger number of points under the curve indicate the better prediction of the model compared to that of the Random Forest.

The R^2 predicted through this result is of 83.4% accuracy, and Adjusted R^2 yields 84.7% accuracy. The accuracy of R^2 is 83.4 and adjusted R^2 is 84.7%, which can be further improved with more training of the data.

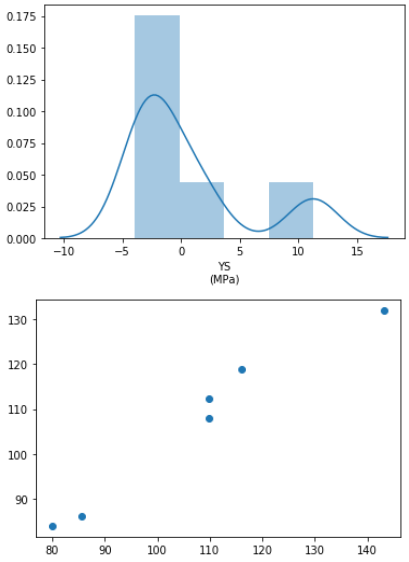

Fig. 13. Linear Regression Algorithm-based prediction

7.7. Lasso Algorithm based YS Prediction

From figure 14 (univariate analysis using Lasso algorithm) it may be observed that PDF and KDE have improved considerably when compared to that of Linear Regression Algorithm and Random Forest

algorithm-based predictions, as shown in figures 13 and 12, respectively.

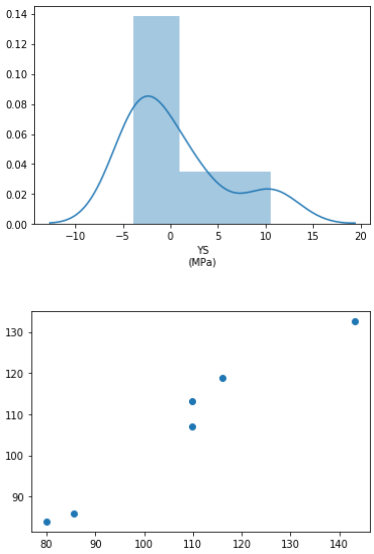

Fig. 14. Lasso Regression Algorithm Based Prediction

The scatter plot (Fig. 14) represents the output predicted results using Lasso algorithm. Although there is considerable improvement in the results with respect to all the features compared to the results attained from the Random Forest and Linear Regression, it may further be improved by using other ML Algorithms. The bigger number of points under the curve indicate the better prediction of the model compared to that of Random Forest and Linear Regression. The loss functions parameters are lesser than Linear Regression Algorithm and Random Forest Algorithm based predictions and require more minimization.

7.8. % Elongation Prediction based on ML Algorithms

Finally, ML algorithms are employed for the analysis of only the % elongation and uses Table 8 for this analysis. Figure 17 shows the parametric interdependencies and the heat map. Figure 18 depicts the influence percentage of the transverse speed and the tool rotation speed. Figures 19 to 22 clearly illustrate the results together with predictions for different algorithms, namely Random Forest, Linear Regression, Lasso and Ridge, respectively. Ridge mean line covers a major part of the plot, indicating its higher efficacy and being followed by Lasso.

Table 8. Percentage Elongation Analysis

No.	Transverse Speed [mm/min]	Tool rotational speed [rpm]	Elongation [%]
1	45	540	20.0403
2	54	540	13.8006
3	63	540	11.3805
4	45	720	17.6121
5	54	720	15.5322
6	63	720	14.6115
7	45	900	19.1088
8	54	900	16.4205
9	63	900	14.9022
10	54	900	6.21
11	54	900	6.57
12	81	720	3.78
13	81	1080	5.22
14	27	1080	6.48
15	81	720	4.5
16	54	900	5.58
17	54	900	6.12
18	54	720	4.5
19	27	720	4.32
20	81	900	4.05
21	81	1080	5.67
22	72	1080	5.85
23	27	1080	5.76
24	27	720	4.05
25	54	900	4.95
26	27	900	5.76
27	54	900	5.85
28	54	900	6.57
29	54	900	6.3

Table 9 represents some observation with respect to the data frame/table such as:

- Count: Count number of non-NA/null observations.
- Mean: Mean of the values.
- Std: Standard deviation of the observations.
- Min: Minimum of the values in the object.
- Max: Maximum of the values in the object.
- 25% - The 25% percentile*.
- 50% - The 50% percentile*.
- 75% - The 75% percentile*.

*Percentile meaning: how many of the values are less than the given percentile.

Table 9. Observation with respect to the data frame.

	Transverse speed [mm/min]	Total rotation speed [rpm]	UTS\n [MPa]
count	29.000000	29.000000	29.000000
Mean	54.620690	844.137931	8.672369
Std	16.996667	160.301440	5.218111
Min	27.000000	540.000000	3.780000
25%	45.000000	720.000000	5.220000
50%	54.000000	900.000000	6.120000
75%	63.000000	900.000000	13.800600
Max	81.000000	1080.000000	20.040300

Table 10 represents the pairwise correlation of all the columns in the data frame/table which provides an idea on the interdependency of parameters.

Table 10. Pairwise correlation of all the columns in the data frame

	Transverse speed [mm/min]	Total rotation speed [rpm]	%
Transverse speed [mm/min]	1.000000	0.055651	-0.112210
Total rotation speed [rpm]	0.055651	1.000000	-0.402077
%	-0.112210	-0.402077	1.000000

The pairs plot from figure 15 is built on two basic figures, the histogram and the scatter plot. The histogram on the diagonal allows us to see the distribution of a single variable, while the scatter plots on the upper and lower triangles show the relationship (or lack thereof) between two variables. For example, the left-most plot in the second row shows the scatter plot of tool rotational speed versus transverse speed.

Figure 15 also represents the heat map of the correlation matrix between various columns in the data frame/table. The colour grading represents the correlation between various features. The correlation value increases if the colour tends towards green and the value decreases if the colour tends towards red.

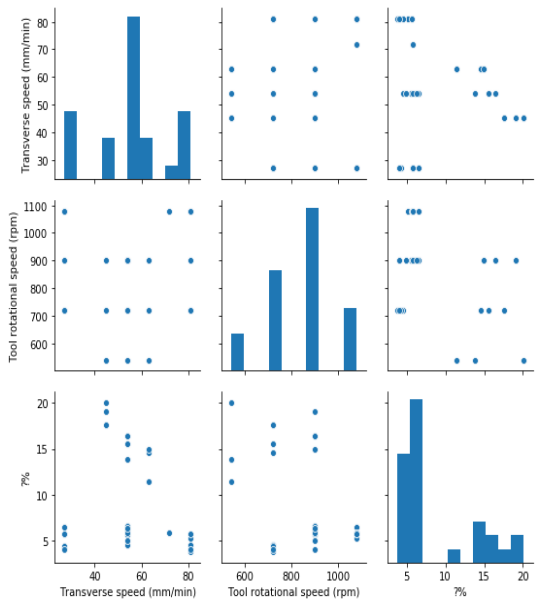

Fig. 15. Parametric Interdependencies and Heat Map

Graph from figure 16 represents the feature importance, which signifies the columns/features which shows more influence on the ML Model getting developed for prediction. Transverse speed governs the % elongation more than the tool rotational speed.

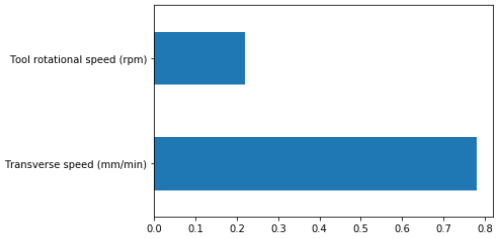

Fig. 16. Influence of traverse speed and tool rotation speed

7.9. Random Forest Algorithm based Elongation prediction

Random Forest Algorithm uses ensemble technique to select the best hyper parameters for training the model. The graph shown below represents the dist. plot which usually represents a univariate analysis if you are using one parameter for the analysis. The distribution plot from figure 16 represents the prediction in the form of bar graphs and continuous lines, where the continuous line represents the KDE (Kernel Density Estimation). The Bar Graph Represents the PDF (Probability Density Function). Figure 16 also represents a scatter plot for the predicted value of % Elongation when Random Forest Algorithm is used. The above parameters represent the loss functions, the values of which have to be minimized further to get a better prediction and accuracy. Random Forest Line partially covers the plot and the prediction is not satisfactory.

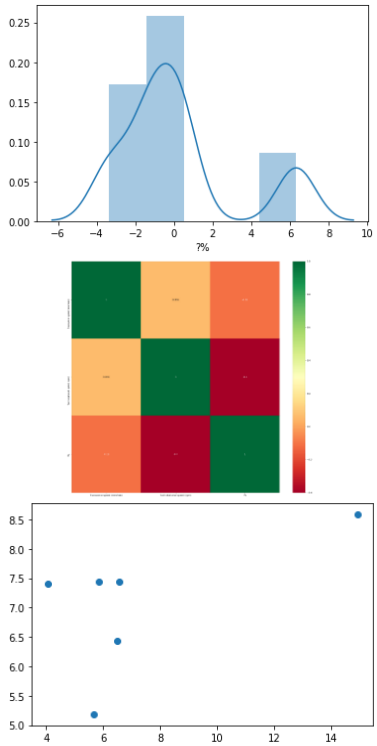

Fig. 16. Random forest (Decision tree) algorithm-based prediction

7.10. Linear Regression Algorithm based Elongation Prediction

From figure 17 (univariate analysis with linear regression algorithm) it may be observed that PDF and KDE are improved when compared to that of the Random Forest Algorithm prediction (Fig. 18). The scatter plot (Fig. 17) represents the output predicted results using Linear Regression Algorithm. Although there is considerable improvement in the results with respect to all the features that are compared to the results obtained from the Random Forest, it may further be improved by other ML Algorithms. The loss functions parameters are higher than Random Forest Algorithm based predictions and requires more minimization.

The R^2 predicted through this result is of 89.1% accuracy. Adjusted R^2 yields 89.3% accuracy. The accuracy of R^2 is 89.1% and of the adjusted R^2 is 89.3 which can be further improved with more training of the data. However, in this case, there is no significant improvement in adjusted R^2. Linear regression line partially covers the plot, and the prediction is not satisfactory, but better than that of the Random Forest.

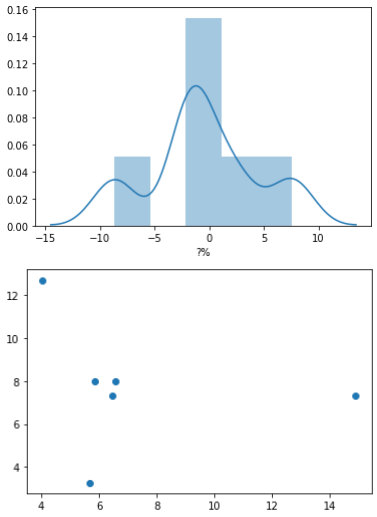

Fig 17. Linear Regression Algorithm Based Prediction

7.11. Lasso Algorithm Based YS prediction

From figure 18 (univariate analysis using lasso algorithm) it may be observed that PDF and KDE have improved considerably when compared to those of the Linear Regression Algorithm and Random Forest Algorithm-based predictions as shown in fig's 16 and 17, respectively.

The scatter plot (Fig. 18) represents the output predicted results using Lasso Algorithm. Although there is considerable improvement in the results with respect to all the features that are compared to the results obtained from the Random Forest and Linear Regression, it may further be improved by using other Algorithms that use regularizations. The loss functions parameters are lesser than Linear Regression Algorithm and Random Forest Algorithm based predictions and require more minimization. Lasso line covers a major portion of the plot and the prediction is more satisfactory and better than that of the Random Forest and Linear Regression.

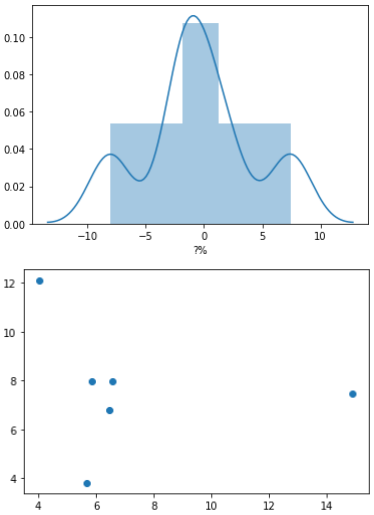

Fig. 18. Lasso Regression Algorithm based prediction

8. CONCLUSIONS

This study investigates the feasibility of the FSW process for the joining of 6061AL mmc's reinforced with SiC/B_4C/Mg. The following conclusions are drawn from this study:

1. Tensile strength is directly proportional to the wt percentage of SiC and inversely proportional to the wt percentage B_4C. Higher wt percentages of SiC resulted in more Al-SiC particle precipitates during the solidification of the homogeneous solution, resulting in the maximum composite strength.
2. % Elongation is directly proportional to the wt percentage of SiC and inversely proportional to the wt percentage B_4C
3. The FSW process leads to homogenous dispersion of these particles at the joint, due to friction and grinding action.
4. SEM micrographs with EDAX analysis show Al and reinforcing particle peaks. Small peaks of Fe were seen, indicating that the reinforcing particles abraded the surface. The particles are also evenly dispersed, and there is no wear on the worn surface at greater weights.
5. As observed from the radiography test, most of the welded specimens are free from flaws and discontinuities.
6. Si and Al phase grains of varying sizes are dispersed in the small gaps between B_4C particles. The Al phase controls the solid solubility of various reinforcing particles as observed from the micrographs.

7. Data visualization and analysis using panda and seaborn libraries, while implementing in Jupiter notebook, reveals that the tool rotation speed has higher influence on UTS and YS, and % Elongation is highly influenced by the transverse speed. Linear Regression Algorithm-based model yields improved results. R^2 value gets considerably improved when adjusted R^2 is calculated for the case of UTS and YS. However, for % Elongation, the improvement is negligible. Lasso facilitates regularization and, also improves the predictions when compared to Linear Regression. However, for all the three parameters Lasso Algorithm-based model provides the best fit.

ACKNOWLEDGEMENTS

The authors thank Matrix NDT Hosur and SITRA Coimbatore for improving the testing and characterization facilities.

REFERENCES

[1] **Threadgill P. L.**, **Leonard A. J.**, **Shercliff H.R.**, **Withers P.J.**, *Friction Stir Welding of aluminium alloys*, International Materials Reviews, 2009, vol. 54, pp. 49–93.

[2] **Rosso M.**, *Ceramic and metal matrix composites: routes and properties*, Journal of Materials Processing Technology, 2006, vol. 175, pp. 364-375.

[3] **Miracle D.,** *Metal matrix composites – from science to technological significance*, Composites Science and Technology, 2005, vol. 65, pp. 2526-2540.

[4] **Kumar A., Kumar S., Mukhopadhyay N. K., Yadav A., Winczek J.,** *Effect of SiC reinforcement and its variation on the mechanical characteristics of AZ91 composites*, Materials, 2020, vol. 13, iss. 21, 4913.

[5] **Kumar A., Kumar S., Mukhopadhyay NK., Yadav A., Kumar V., Winczek J.,** *Effect of variation of SiC reinforcement on wear behaviour of AZ91 alloy composites*, Materials, 2021, vol. 14, 990.

[6] **Kumar S., Yadav A., Patel. V., Nahak B., Kumar A.**, *Mechanical behaviour of SiC particulate reinforced Cu alloy-based metal matrix composite*, Materials Today: Proceedings, 2021, vol. 41, pp. 186–190.

[7] **Morozova I., Obrosov A., Naumov A., Królicka A., Golubev I., Bokov D.O., Doynov N., Weib S., Michailov V.**, *Impact of impulses on microstructural evolution and mechanical performance of Al-Mg-Si alloy joined by impulse friction stir welding*, Materials, 2021, vol. 14, 347.

[8] **Chen X. G., da Silva M., Gougeon P., St-Georges L.,** *Microstructure and mechanical properties of friction stir welded AA6063–B4C metal matrix composites*. Materials Science and Engineering: A, 2009, vol. 518, pp. 174–184.

[9] **Ramnath V. B., Abhishek Subramanian S., Rakesh R., Sharun Krishnan S., Ashwin Ramanathan A. L.,** *A Review on Friction Stir Welding of Aluminium Metal Matrix Composites*, IOP Conference Series: Materials Science and Engineering, vol. 2018, 390: 12103.

[10] **Yadav A., Ghosh A., Gupta P., Kumar A.**, *Mathematical modelling of heat affected zone width in submerged arc welding process*, 2018 International Conference on Computational and Characterization Techniques in Engineering & Sciences (CCTES), IEEE, 2018.

[11] Suban K. A. A., Perumal M., Ayyanar A., Subbiah A.V., *Microstructural analysis of B_4C and SiC reinforced Al alloy metal matrix composite joints*, International Journal of Advanced Manufacturing Technology, 2017, vol. 93, pp. 515–525.

[12] **Palanivel R., Mathews P. K., Murugan N., Dinaharan I.,** *Prediction and Optimization of Wear Resistance of Friction Stir Welded Dissimilar Aluminum Alloy*, Procedia Engineering, 2012, vol. 38, pp. 578–584.

[13] **Storjohann D., Barabash O. M., David S. A., Sklad P.S., Bloom E. E., Babu S. S.,** *Fusion and friction stir welding of aluminum-metal-matrix composites*, Metallurgical and Materials Transactions A, 2005, vol. 36, iss. 11, pp. 3237-3247.

[14] **Vijay S. J., Murugan N.,** *Influence of Tool Pin Profile on The Metallurgical and Mechanical Properties of Friction Stir Welded Al–10wt.% TiB_2 Metal Matrix Composite*, Materials and Design, 2010, vol. 31, pp. 3585–3589.

[15] **Xu W. F., Liu J. H., Chen D. L., Luan G. H., Yao J. S.**, *Improvements of Strength and Ductility in Aluminum Alloy Joints Via Rapid Cooling During Friction Stir Welding*, Materials Science and Engineering: A, 2012, vol. 548, pp. 89–98.

[16] **Praga-Alejo R., Torres-Treviño L., Piña-Monarrez M.,** *Prediction in Welding Process Using Multiple Linear Regression and Neural Network*, International Journal of Industrial Engineering, 2008, pp. 481-488.

[17] **Jamrozik W., Górka J., Kik, T.**, *Temperature-based prediction of joint hardness in TIG welding of inconel 600, 625 and 718 nickel superalloys*, Materials, 2021, vol. 14, iss. 2, pp. 442.

[18] **Sumesh A., Rameshkumar K., Mohandas K., Babu R. S.**, *Use of machine learning algorithms for weld quality monitoring using acoustic signature*, Procedia Computer Science, 2015, vol. 50, pp. 316-322.

[19] **Chen C., Lv N., Chen S.,** *Welding penetration monitoring for pulsed GTAW using visual sensor based on AAM and random forests*, Journal of Manufacturing Processes, 2021, vol. 63, pp. 152-162.

[20] **Mishra A., Sefene E. M., Tsegaw A. A.**, *Process parameter optimization of 6061AA Friction Stir Welded Joints using Supervised Machine Learning Regression-based Algorithms*, 2021, arXiv preprint arXiv:2109.00570.

[21] **Liu Q., Song J., Hao P.,** *Automatic Reading and Writing Model of Welding Parameters Predicted Based on PSO-RFR*, International Conference on Intelligent Computing, Automation and Systems (ICICAS), 2019, pp. 387-391.

[22] **Kim S., Hwang I., Kim D. Y., Kim Y. M., Kang M., Yu J.**, *Weld-Quality Prediction Algorithm Based on Multiple Models Using Process Signals in Resistance Spot Welding*, Metals, 2021, vol. 11, iss. 9, pp. 1459.

[23] **Singh L. M., Saha A.,** *Optimization of welding parameters for maximization of weld bead widths for submerged arc welding of mild steel plates*, International Journal of Engineering Research and Technology, 2012, vol. 1, iss. 4.

[24] **Lalitnarayan,K., Sarcar M. M. M., Rao K. M., Kameswaran K.,** *Prediction of weld bead geometry for CO_2 welding process by multiple regression analysis*, International Journal of Mathematics and Scientific Computing, 2011, vol. 1, iss. 1, pp. 52-57.