

PARAMETRIC STUDY AND OPTIMIZATION FOR WELDING PROCESSES USING MACHINE LEARNING

M. Chaturvedi*, S. Arungalai Vendan

Dayananda Sagar University, Devarakaggalahalli, Harohalli, Kanakapura Road, 562112 Bengaluru, India

*Corresponding author's e-mail address: muktichaturvedi.res-soe-ece@dsu.edu.in

ABSTRACT

Optimization facilitates in attainment of maximum strength, efficiency, reliability, productivity and longevity. In this work, data from three material joining processes - Ultrasonic welding of polymers, arc welding as Metal Inert Gas and Tungsten Inert Gas are analysed for establishing quantitative relationship between the process parameters and for prediction of weld features using Multivariate Linear Regression algorithm. The various dependency coefficients and characteristics generated with the ML algorithms are in agreement with the inherent dependency as obtained from experimental data and simulation results. This investigation is a preliminary attempt with a limited set of data to manifest the suitability of machine learning techniques; nevertheless, the results are far from conclusive owing to small data set and hence may be extended to precisely model joining processes with higher number of process parameters, degree of freedom and responses.

KEYWORDS: ultrasonics, gas metal arc welding, machine learning, regression analysis, parametric analysis.

1. INTRODUCTION

The implementation of artificial intelligence for modelling the industrial processes helps in prediction and establishing relationships between the process parameters and the output features. This section briefly describes the welding processes namely Ultrasonic welding, Friction Stir and Gas Metal Arc Welding considered in this work and the existing literature on the Artificial Intelligence (AI) /Machine Learning (ML) approaches taken up by various researchers for modelling the welding process.

Ultrasonic welding uses mechanical vibrations at a frequency of 20kHz or 40kHz to soften the injection molded components at the intended joint line. The parts which are to be fused are secured together under pressure and are subjected to vibrations created by a welding sonotrode or horn. For this process, the parameters of importance are vibration amplitude, welding mode, down speed, trigger pressure, weld time and hold time. The deposition efficiency in this process is generally in the range of 93-97%. Previous research work on parametric optimization of this process have utilized various branches of AI and ML and have given comparable results with respect to experimental values and FEM based predictions. Li Y. et al. [12] have used the supervised learning mechanism of the Artificial Neural network to predict the joint strength and also to determine the quantitative relationship between the input features and the target parameters. Elangovan S.

et al. [4], utilized the Response Surface Methodology (RSM) with Genetic Algorithm (GA) to determine optimum welding conditions resulting in maximum joint strength for spot and seam welding of Al components and found this as an efficient modelling tool. Anand K. Elangovan et al. [1] compared the performance of ANN with Multiple Regression Analysis (MRA) models for the weld strength prediction in Ultrasonic welded Cu joints in terms of mean prediction error and found ANN models to be more accurate. Mongan P.G. et al. [14] used the ANN with GA and further used Lavenberg-Marquardt algorithm to train the model for a data set of 37 trials and demonstrated high accuracy with mean relative error of 6.79%.

Arc Welding namely Gas Tungsten Arc Welding (GTAW), is a fusion joining process utilizing the heat energy of the arc formed between an electrode and a workpiece to bring the weld materials to melting point and thus fuse together to form the metallurgical bond. The electrode either is used for carrying current or also as a filler material.

Gas Metal Arc Welding (GMAW) is a process in which the electrode tip gets melted with the arc energy and the molten material is transferred through the arc into the weld pool. For the welding of non-ferrous metals, inert shielding gases are used to protect the arc and the weld pool from atmospheric contamination. For GMAW welding, the significant process parameters are weld rate, supply voltage, wire feed

rate, gas flow rate, Nozzle to Surface Distance (NTSD) and torch angle.

Ganjigatti J. P., et al. [5] investigated the impacts of individual and combination of welding parameters on the build geometry using statistical approach of multiple least square regression analysis and ANOVA to compare the correlations results of non-linear and linear regression analysis for MIG welding. Their results indicated higher performance of linear analysis and emphasized that the performance of any regression technique is dependent on the data. Correia D.S. et al. [3] used Genetic Algorithms for predicting optimum settings for a MIG welding process. Weld voltage, wire feed speed and weld speed are considered as the input parameters for prediction of penetration depth, deposition efficiency, bead width and reinforcement. The output parameters showed varying mean errors in the range of 3%-46% which could be due to the limited data population. Sumesh A. et al. [17] used classification algorithms J48 and Random Forest to classify the defects in arc welding into lack of fusion and burn-through based on arc sound using the input parameters of current, voltage and travel speed. Artificial Neural Networks (ANN) and Adaptive Neuro Fuzzy Inference Systems (ANFIS) have been in use for modelling of engineering systems to predict the residual stresses in the form of distribution plots considering the uncertainty bounds [6]. The results from these studies have helped in the characterization of the weld to acceptable confidence levels as compared to the experimental data.

In GTAW, the electrode is used only for carrying current. The heat developed due to an arc established between a tungsten electrode and the base metal causes the base metal and the filler metal to melt and thus fuse together to form the weld on cooling down. An inert gas is used to eliminate the oxidation of weld metal and the filler metal. In this arc welding process, the key parameters considered in some of the literatures [7], [8] are arc efficiency, weld speed and heat input and their influence on the build geometry, mechanical properties and residual stresses generated as part of welding is investigated. For investigation of residual stresses and parametric dependency analysis in welding, support vector regression and neuro evolutionary computing were used [7], [8] They investigated the hybrid models- NN-Genetic Algorithm (GA), NN- Particle Swarm Optimization (PSO) and NN-GA-PSO for the prediction of residual stress. They also developed the hybrid model of Fuzzy support vector regression with Genetic algorithm and yielded results faster and more accurate than the standalone models.

Kesse M.A. et al. [10] proposed an AI algorithm applying fuzzy logic and deep neural network for predicting the weld bead width with current arc length and weld speed as control parameters. The algorithm gave a predictive accuracy of 92.59% with 27 trials of experimental data sets. Priya B.G. [15] used linear regression techniques to predict the tensile strength, hardness and grain size of pulsed current TIG weld

products with peak current, base current, pulse frequency and pulse on time as control parameters. Korat P. and Sama M. [11] utilized ANN with feed forward back propagation algorithm to predict the tensile properties and optimize the weld process parameters for GTAW welding of Ni base superalloys.

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 underfitting of models, techniques of regularization and dropout are to be 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. Virkkunen et al. [20] advocated the use of data augmentation using virtual flaw technology for training of ML analysis of ultrasonic process with limited data, also suggesting proper validation of augmented and training data. Zhu et al. [21] proposed a weld defect recognition based on CNN and random forest classification algorithm. They testified this deep learning method to give a high accuracy to meet the requirement of weld defect classification. Mahadevan et al. [13] experimented with suitable ML techniques for six welding processes. A neural network model and ANFIS was experimented to forecast the weld bead geometry in TIG and Friction stir welding. They worked with CNN and PCA for process monitoring of LASER welding. For resistance welding, they identified electrode degradation as the problem and proposed the use of neuro fuzzy models for controlling the current. For GMAW, they used neural networks and Response surface Methodology (RSM) to determine the feature importance and forecast the weld properties while fuzzy neural networks and sensor systems were found efficient to optimise key hole geometry and for parametric optimization in plasma welding. Bacioiu D. et al. [2] designed a system using image classification, to emulate operators view and established a correlation between the weld pool aspect and the weld quality.

2. METHODOLOGY

Data driven models are now finding application in manufacturing processes as they enable quick and

accurate prediction of the process under analysis. These models automatically learn the process input output dependencies based on previous data. Supervised ML operates with labels of input and output data and can be extended to the manufacturing processes as there are definitive targets and qualification methods involved in these processes. Predictive modelling helps in evolving mechanisms to optimize the material joining process when there is no clear relation between the influencing parameters and the weld features like build geometry, residual stress or distortion. Linear Regression modelling is employed in this work for prediction of build parameters and residual stress and for identification of hierarchy of process parameters influencing the residual stress generated in the manufacturing/ joining process.

The existing literature has trivial reporting on the application of ML and its feature of correlation to the manufacturing processes. This work, thus focuses on applying the basic linear regression model to exhibit the use of ML techniques in enabling better design of the welding processes and in this way promoting the multi-disciplinary approach to the manufacturing or joining processes.

2.1. Machine Learning Algorithm used for the Study

Supervised learning algorithm is used in this case to predict the dependent variable i.e., residual stress from a given set of predictors [9]. These algorithms help develop a function mapping the inputs to the output based on prior experimental or simulation data set using scikit-learn. The data set obtained from either experiments or from FEA is split up into 70 percent training data and the remaining as testing data for enabling and validating the data predictions obtained using any regression model.

Linear Regression Model [18], [19] - This is a supervised ML model and the prediction made by this model are observed to have a constant slope indicating the linear relation of the output with the control variables and are continuous. This algorithm is used to estimate real values of output parameters as controlled by the input parameters based on a regression line. It can be applied to single or multiple variables of both the control and the controlled parameters. The prediction function used in the linear regression model outputs the target variable predictions based upon the independent process parameters. Based on the available data, values of importance of each input parameter are estimated and the best fit line is established between the independent and dependent variables. Ordinary least squares [19] is the technique of training a regression model with multiple inputs in which the coefficients of the line are derived based on Mean Square of Errors (MSE) and the R2 score determined from the data. This is the quantity which is intended to be kept minimum in the implementation of the algorithm.

Gradient Descent is another technique [19] of training the linear regression model which uses an iterative process for the minimizing of the error for finding the optimum values of the coefficients. Some random values of coefficients are selected and for error in each input-output pair, squared error sum is calculated. Coefficients are then updated to have minimum sum squared error. Regularization methods used in the algorithm focus on the absolute size of the coefficients or weights assigned to the predictors, thus reducing the complexity of regression models.

Feature importance techniques [20] are used to assign weights to influencing parameters to indicate the relative importance of each feature in making a prediction of the dependent variable. These techniques also help to improve the performance of a model by enabling dimensionality reduction. Due to the stochastic nature of the evaluation procedure, the retrieval of coefficient value is to be performed for a few times and the average is then considered to be the score of feature importance.

2.2. Data Set for Training ML Algorithm

Availability of experimental welding data in large scale is difficult owing to the several constraints in doing the experimentation work like lack of resources or financial constraints, thus limiting the performance that may be obtained from any of the Machine learning models.

Previous experimental work by other researchers provided the data sets for three different welding processes and the same is used in this work using ML models for estimating the feature importance and predicting the weld characteristics. Selection of data sets was made using Design of Experiments for control of weld characteristics by the process parameters.

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.

2.2.1. Process & Data Set - Butt Joining

Datasets from the previous research work [7], [8] on evaluation of residual stresses obtained from Finite Element Analysis (FEA) based on the process parameters for the butt joining of ASTM A36 mild steel plates was used for training the ML regression model. Parameters affecting the residual stresses i.e., arc efficiency, weld voltage, current and welding speed are given as input parameters and residual stress as an effect of these parameters is considered to be the output for the Linear Regression model. 43 sets of data obtained from the FEA of butt-joining process are used for machine learning. 70 % of the data is used for training the model and 30% is used for testing the predictions.

2.2.2. Process & Data Set - GMAW

Ganjigatti et al. [5] used the regression analysis -both linear and non-linear for establishing the parameter interdependencies for the bead-on-plate type MIG welding process. The experiment was conducted on a semi mechanized welding station for MIG welding with structural mild steel as the base material and electrode wire steel ER70S-6 as the welding consumable.

The input parameters chosen were weld speed, input voltage, wire feed rate, gas flow rate, Nozzle to top surface distance (NTSD) and the torch angle. For the 6 input parameters, 2 levels of values were selected, thus giving 64 combinations of input process parameters. The output variables considered are Build Height (BH), Build Width (BW) and Build Penetration (BP) (Fig. 1). They formed pareto graphs, figure 2 showing the relative impact of main and interaction factors on the target parameters.

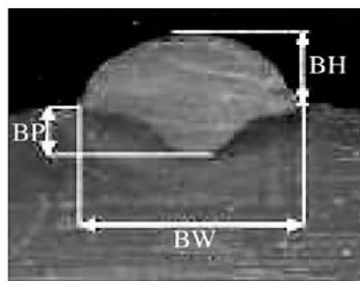


Fig. 1. Weld Bead after macro etching. Geometry parameters: BH-Build Height, BW- Build Width, BP- Build Penetration [5]

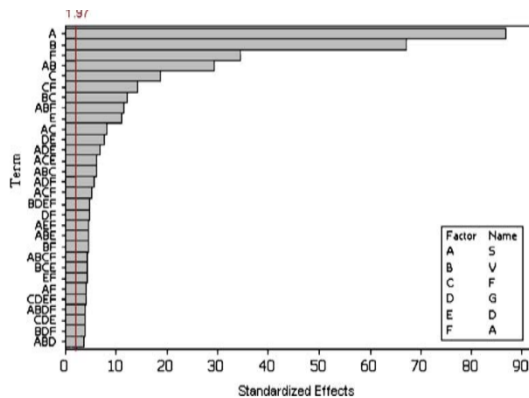


Fig. 2. Pareto charts of the effects on response: S - Speed, V - Voltage, F - Wire Feed Rate, G - Gas Flow rate, D - Nozzle to Substrate distance, A - torch angle [6]

2.2.3. Process & Data Set - Ultrasonic Welding

Satpathy M. P. et al. [16] analysed and compared the Fuzzy Logic and Genetic Algorithm approach to find the optimal combinations of process parameters for ultrasonic welding of Aluminum and brass for its use in the electrical industry. Vibration amplitude, weld

pressure and weld time are considered as control parameters and their impact on tensile shear stress, peel stress and weld area has been investigated using conventional optimization techniques. The influence of all input parameters on the weld samples was investigated with the experiments and this data is used as testing and training data for the ML algorithm.

3. RESULTS AND DISCUSSIONS

3.1. Butt Joining Parameters

The variation of Arc Efficiency (AE), Weld Speed (WS), Arc Voltage (AV), Arc Current (AC), Heat Input (HI) and Heat Flux (HF) are considered in this analysis for prediction of output parameter values, generation of correlation matrix and heat map, feature importance plot, regression plots and regression equation describing the impact of each parameter in the generation of Residual Stresses (RS).

3.1.1. Prediction of Residual Stress Values

Residual stresses field is induced in the welding process due to high thermal gradient between the weld interface and the base metal. Prior knowledge of dependency of residual stress on the process parameters enables optimum parameter selection for an efficient weld. Actual and Predicted values of residual stress from the ML model are tabulated in table 1. The model gave an accuracy of 85.54% with 70:30 ratio data segregation for training and testing.

Table 1. Predictions for Residual Stress

Actual	Predicted
236.3	234.20
74.5	157.86
60.7	42.38
58.9	59.87
64.2	93.05
178.1	191.44
60.36	62.44
275	250
300	383.85
322.3	337.45
289	256
60.36	60.54
58.93	65.78

3.1.2. Correlation Matrix

Correlation matrix helps estimate the interdependencies of target and process parameters. The diagonal elements in the heat map represent the correlation of the element with itself and thus has value 1. The scale on the right of the heat map is the legend for identification of the colour codes represented in the

map. The matrix from table 2 gives the correlation coefficients from which weld speed is observed to have strong positive correlation and heat input negative correlation to the residual stress. The process parameters are also seen to have correlation between them indicating interdependence and thus should be further analysed for dimensionality reduction. Correlation of arc voltage, arc current and weld speed gives scope for further analysis and removal of redundant factors. All the parameters monitored in the weld experiment, irrespective of feature importance are considered here to get a broad idea of the ML analysis as applied to the manufacturing/ joining processes. This explains the non-optimum correlation as the feature selection is not considered.

3.1.3. Correlation Heat Map

The correlation Heat Map (Fig. 3) suggests strong

correlation of weld speed, heat input and thus the heat flux on the residual stress. Welding voltage and welding current show negative correlation amongst the two. Interdependencies amongst the process parameters give an idea of the complexities involved in designing an optimum weld parameter setup and should be analysed to find the maximum correlation. One parameter from the pair of parameters having the maximum correlation can be thus ignored in the design sequence.

3.1.4. 3D-Plot of Influential Parameters

The 3D plot shows the isolated impact of heat input and welding speed on the residual stress, instead of including all other parameters of lesser/indirect impact on the residual stress. The darker points are nearer than the lighter ones and this helps to interpret the interactivity of the two process parameters with the output parameters (Fig. 4).

Table 2. Correlation matrix

	AE	WS	AV	AC	HI	HF	RS
AE	1	0.46069	0.2298	-0.1766	0.8207	0.8207	0.1281
WS	0.041	1	-0.008	-0.02079	0.046	0.046	0.877
AV	0.2296	-0.0084	1	-0.888	0.3904	0.3904	0.0925
AC	-0.1766	0.02079	0.046	1	-0.0817	-0.0817	0.0407
HI	0.8207	0.04606	0.3904	-0.0817	1	1	0.1647
HF	0.8207	0.0460	0.3904	-0.0817	1	1	0.1647
RS	-0.1281	0.8779	-0.0925	0.0407	-0.1647	-0.1647	1

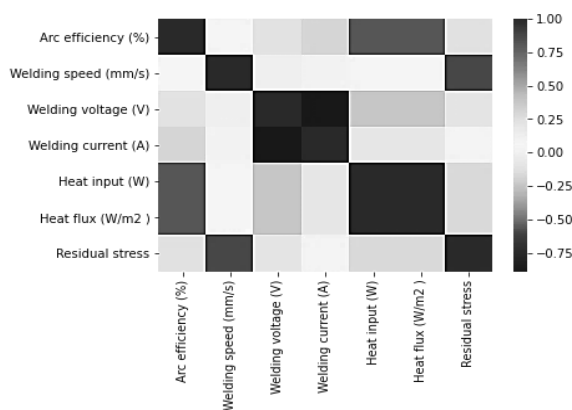


Fig. 3. Correlation heat map

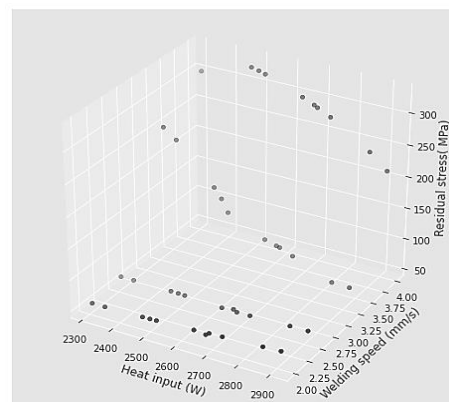


Fig. 4. 3D dependency plot

3.1.5 Hierarchy of influential parameters:

Feature importance is the score assigned to the input features based on their effect in predicting the target variable [9]. This score is of importance in predictive modelling as it provides insight into the data and the process model.

The feature importance chart figure 5 identifies the significant features based on the coefficients (weights) assigned to parameters in the regression model trained with the given data set [7], [8].

The assigned coefficient values and the

intercept value thus help generate the regression equation for the used process model. Regression equation for residual stress (RS) as generated from the Linear regression model is:

$$RS = 1224 * AE + 109 * S + 37.08 * AV + 6.03 * AI - 1.176 * 10^{-8} * HI - 9.36 * 10^{-5} * HF - 1648.71 \quad (1)$$

where, AE- Arc Efficiency, S-weld Speed, AV-weld voltage, AI-weld current, HI-Heat Input, HF-Heat Flux.

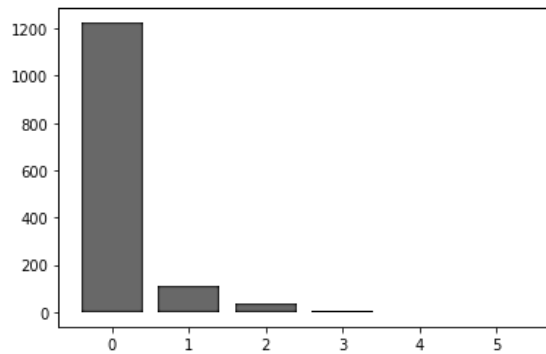


Fig. 5. Feature importance chart: 0 - Arc efficiency, 1- welding speed, 2- welding voltage, 3- welding current, 4-heat input, 5- heat flux

3.2 GMAW PARAMETERS

3.2.1. Multi Output Regression

The dataset used for the second analysis described the variation in the build geometry- i.e., Build Height (BH), Build Width (BW) and Build Penetration (BP) due to the variation in weld speed, voltage, wire feed rate, gas flow rate, nozzle to surface distance and the torch angle [9]. This dataset did not include the continuous variation of input parameters and their impact on the target, instead considers the variation of the build geometry for two levels of the process parameters. The lesser range of data variation caused restricted results from the ML analysis.

3.2.2. Prediction of Build Geometry Values

The data set of 64 trials was used to train the linear regression model to predict the build height, width and penetration. The result of the prediction of the three build geometry parameters (Table 3) is attained with an accuracy of 82.8%.

3.2.3. Correlation Matrix

Correlation matrix, Table 4 shows no correlation between the input parameters indicating no redundancy of factors considered for study and clearly quantifies the dependence of the input and target variables and also amongst the three target variables.

3.2.4. Correlation Heat Map

Visual representation of the correlation matrix in the form of heat map figure 6 helps broad identification of the dependencies, here indicating strong negative correlation of weld speed on the build height, width and penetration. Weld voltage is observed to have negative correlation with build height and medium correlation to build width and build penetration. Torch angle has strong negative correlation with build penetration and medium on build height. Correlation exists between the target variables as was observed in the correlation matrix. The correlation heat map can be compared with the pareto curves generated in [9] to correlate the results shown with the two analyses with the heat map having the advantage of indicating positive or negative correlation as well.

Table 3. Predicted and actual values of build geometry

Data Order No.	Actual BH	Predicted BH	Actual BW	Predicted BW	Actual BP	Predicted BP
24	3.165	2.761	8.002	9.019	1.991	2.06
39	3.934	3.697	8.904	10.345	1.865	2.082
52	3.259	3.614	12.457	11.93	3.335	2.717
27	3.367	3.330	8.298	7.87	1.935	1.949
44	2.447	2.476	10.083	10.03	1.914	2.061
2	3.134	3.419	11.447	11.31	1.797	2.451
21	3.786	3.611	7.821	9.59	1.679	1.726
62	3.328	3.283	12.216	12.15	2.202	2.411
41	3.806	3.652	8.258	9.55	1.626	1.774
50	3.159	2.999	7.435	8.094	1.448	1.643
38	4.348	4.083	8.885	10.06	2.48	2.344
54	3.267	3.242	11.822	12.20	2.06	2.363
35	3.257	3.514	11.565	11.21	2.109	2.454
26	3.204	3.317	8.948	7.89	1.976	2.042
57	2.729	2.902	9.34	9.704	2.596	2.371
49	3.047	2.986	8.238	8.112	1.455	1.736
3	4.434	3.987	8.935	10.173	2.504	2.341
40	4.109	3.711	8.726	10.327	1.702	1.989
36	2.969	3.129	12.726	11.48	2.182	2.192
10	3.219	3.2308	8.029	7.146	1.86	1.686
56	4.21	3.752	8.885	10.282	1.551	2.037

Data Order No.	Actual BH	Predicted BH	Actual BW	Predicted BW	Actual BP	Predicted BP
19	4.01	3.983	8.556	9.334	2.175	2.081
34	3.229	3.187	12.582	12.26	2.413	2.408
60	2.346	2.472	10.21	9.19	1.842	1.801
17	4.425	4.042	9.498	10.11	2.508	2.296

Table 4. Correlation Matrix

	WS	V	WFR	GFR	NTSD	TA	BH	BW	BP
WS	1	0	0	0	0	0	-0.69	-0.559	-0.37
V	0	1	0	0	0	0	0.5341	0.65	0.457
WFR	0	0	1	0	0	0	0.15	0.1690	0.323
GFR	0	0	0	1	0	0	-0.012	0.0172	0.077
NTSD	0	0	0	0	1	0	0.088	-0.08	-0.032
TA	0	0	0	0	0	1	-0.2752	0.0271	-0.4767
BH	-0.69	-0.53	0.15	-0.012	0.06	-0.27	1	-0.0763	0.1577
BW	-0.5597	0.6560	0.1690	0.0172	-0.0851	0.0271	0.0763	1	0.6416
BP	-0.3763	0.457	0.3239	0.077	-0.032	-0.4767	0.1577	0.6416	1

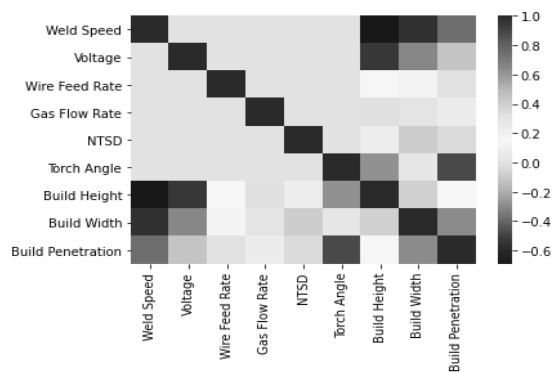


Fig 6. Correlation heat map

3.2.5. Hierarchy of Influential Parameters

Hierarchical importance of features on all the 3 build geometry parameters can be identified with the feature importance graph. The graph from the figure 7 shows the impact of input parameters on the build geometry in the form of strong negative coefficients (weights) of weld speed and voltage and positive coefficient of wire feed rate, with impact of other parameters also shown in their reduced strength.

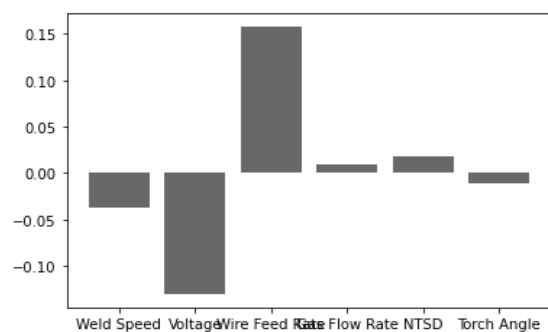


Fig. 7. Feature importance chart

3.3. UW PARAMETERS

Multi output regression is used for finding the dependencies of Peel Stress (PS), Weld Area (WA) and Tensile Shear stress (TS). The available data set was obtained from previous experimental works using DOE considering three levels for weld pressure and time respectively and five levels of variation in vibration amplitude to estimate the effects of the factors over the complete range of experimental conditions. The data set is split up to train the linear regression model with 75 % data used as training data and 25% as test data resulting in the following performance parameters:

- Root Mean Square Error (RMSE) value = 1.59
- R2 score = 0.082
- Model Accuracy obtained = 90.44%

Based on training the regression model, the regression equations generated for the three output parameters are:

$$TS = 0.06 * A + 2.93 * P + 0.92 * t - 3.05 \quad (2)$$

$$PS = 0.01 * A + 0.46 * P + 0.038 * t - 0.25 \quad (3)$$

$$WA = 0.79 * A + 17.58 * P + 29.71 * t - 18.33 \quad (4)$$

where: A-vibration amplitude, P-pressure, t- vibration effect time duration

3.3.1. Prediction of Output Parameters

With the mentioned performance metrics, the comparison of the output parameter values from the data set and those predicted by the regression model is given in Table 5.

3.3.2. Correlation Matrix

The dependency coefficients between the input parameters of amplitude, pressure and time influencing

the output parameters of tensile shear stress, tensile peel stress and the weld area are listed as correlation matrix in Table 6.

Static pressure exerted on the weld stack and the vibration amplitude are the factors that control the weld characteristics and is accordingly quantized in the correlation matrix.

3.3.3. Correlation Heat Map

Figure 8 generated from the coefficients of the correlation matrix helps in visualizing the effects of vibration amplitude, pressure and time on the output parameters. The weld area and time and also peel stress and shear stress have correlation between 0.5 to 1 while other parameters show lower correlation.

Table 5. Predicted and actual values

Actual Shear Stress [MPa]	Predicted Shear Stress [MPa]	Actual Peel Stress [MPa]	Predicted Peel Stress [MPa]	Actual Weld Area [mm ²]	Predicted weld area [mm ²]
1.57	1.53	0.58	0.51	48.24	51.76
1.27	1.35	0.52	0.5	42.1	45.81
1.15	1.83	0.5	0.65	44.4	45.03
3.23	1.82	0.74	0.61	50.03	46.39
3.39	2.01	0.76	0.62	52.22	52.33
2.8	2.6	0.79	0.75	52.45	54.49
0.78	2.085	0.47	0.58	58.45	62.45
2.05	2.12	0.7	0.69	48.95	46.79
0.85	2.56	0.44	0.68	68.25	68.8
1.7	2.38	0.58	0.67	65.95	62.85
2.1	3.15	0.55	0.77	70.53	72.31
1.4	2.2	0.56	0.66	55.45	56.91

Table 6. Correlation Matrix

	Amp [μm]	Pressure [MPa]	Time [s]	Shear Stress [MPa]	Peel Stress [MPa]	Weld Area [mm ²]
Amp [μm]	1.00E+0	-2.11E-17	2.43E-17	0.239664	0.371456	0.466432
Pressure [MPa]	-2.11E-17	1.00E+0	7.68E-18	0.415227	0.4351	0.161338
Time(s)	2.43E-17	6.68E-18	1.00E+0	0.13948	0.050514	0.806178
Shear Stress [MPa]	2.4E-1	4.15E-01	1.39E-01	1.00E+0	0.915303	0.40341
Peel Stress [MPa]	3.71E-01	4.35E-01	-5.05E-02	0.915303	1.00E+0	0.494841
Weld Area [mm ²]	4.66E-01	1.61E-01	8.06E-01	0.40341	0.294821	1.00E+0

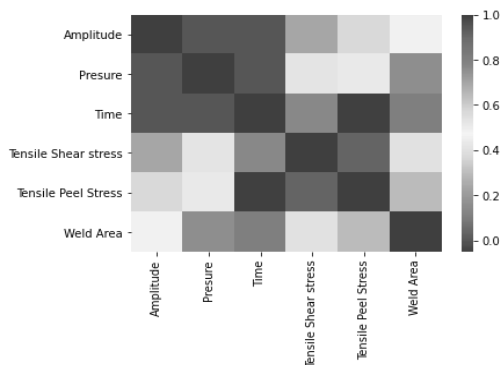


Fig. 8. Correlation heat map

The data set utilized for this analysis does not indicate continuous variation of the input parameters. Different levels of variation are considered for the three input parameters to observe the behavior of the output parameters when one or more of the process parameters is held constant and the others are varied. Although the data pattern is different, the performance metrics

remains unaltered and the accuracy recorded is 90.44%.

4. CONCLUSIONS

The aim of this work was to examine the robustness of ML based regression analysis to the different techniques and data sets of welding process. The different data sets for various processes that were used in this work for training the regression models and the results obtained herewith, reinforce the idea that the performance of ML models adopted for any system depends on the available data. The data analysis and visualization techniques for a process analysis depends on the scale and type of data set. The conclusions drawn from this study on applying linear regression analysis for joining processes are as follows:

a. Butt Joint

Predictions of residual stress obtained with an accuracy of 85%. Strong positive correlation is shown between weld speed and residual stress while negative between

the heat input and stress. Feature importance graph shows arc efficiency followed by weld speed to be significant features impacting the residual stress. The data set that is available may be used to augment the data to enhance a restricted data set or can be used for prediction of residual stress for theoretical analysis purposes.

b. Gas Metal Arc Welding

Predictions of weld bead geometry obtained with an accuracy of 82.8% accuracy in the predictions. The data set is limited to two levels of input parameters. Parameters of build geometry – width, height and penetration are taken as the output variable and 6 input parameters are considered. As per the analytics, the weld speed has negative correlation to the three output parameters, weld voltage has strong negative correlation with the build height and positive with build width and penetration. Wire feed rate has positive correlation to all the build parameters.

c. Ultrasonic Welding

Predictions of Stress values and weld area obtained with an accuracy of 90.44%, but the data set used in this case has different number of levels of input parameters. Regression equation is obtained for the three output parameters of tensile shear stress, peel stress and the weld area. Correlation matrix indicates positive correlation between vibration time and weld area and also between peel stress and shear stress.

The analysis conducted as part of this work were focused on prediction of residual stress, weld geometry, tensile shear stress and on establishing their dependency on the process parameters. The correlation study has given observations of importance of the controlling parameters and their interdependencies. The study is of exploratory nature and thus, firm establishment/ generalization of the regression and correlation as appropriate tool for analyzing varying data types and format may be injudicious. However, the favorable results though for a limited set of data' institute the need for further investigations on the viability of this algorithm with unmethodical variation in collected data, large data sets, and multiple governing parameters. Extension of this work with larger data set - obtained experimentally or through FEA with data analytics using ML techniques may result in better planning and design of the joining process with optimum parameter selection.

ML modeling and analysis also enables dimensionality reduction which helps reduce redundancy and complexity involved in the prediction modelling. Accurate results for feature prediction using the ML models will depend on appropriate algorithm selected for the given set of data.

This work is an illustration of a multidisciplinary approach to the study of joining processes using ML that partially reveal benefits and detrimental aspects of these techniques to industrial applications. It emphasizes that with the availability of larger data set, the machine

learning can be effectively employed for optimum design of any manufacturing / joining process.

REFERENCES

- [1] **Anand K., Elangovan S., Rathinasuriyan C.,** *Modeling and prediction of weld strength in ultrasonic metal welding process using artificial neural network and multiple regression method.* Materials Science & Engineering International Journal, 2(2), 2018, pp. 40-47.
- [2] **Bacioiu D., Melton G., Papaelias M., Shaw R.,** *Automated defect classification of SS304 TIG welding process using visible spectrum camera and machine learning.* NDT & E International, vol. 107, 2019, p. 102139.
- [3] **Correia D. S., Gonçalves C. V., Junior S. S., Ferraresi V. A.,** *GMAW welding optimization using genetic algorithms,* Journal of the Brazilian Soc. of Mech. Sci. and Eng., 26(1), 2004, pp. 28-32.
- [4] **Elangovan S., Anand K., Prakasan K.,** *Parametric optimization of ultrasonic metal welding using response surface methodology and genetic algorithm.* The International Journal of Advanced Manufacturing Technology, 63(5-8), 2012, pp. 561-572.
- [5] **Ganjigatti J. P., Pratihari D. K., Roy Choudhury A.,** *Modelling of the MIG welding process using statistical approaches.* The International Journal of Advanced Manufacturing Technology, vol. 35(11), 2008, pp. 1166-1190.
- [6] **Mathew J., Moat R. J., Paddea S., Fitzpatrick M. E., Bouchard P. J.,** *Prediction of residual stresses in girth welded pipes using an artificial neural network approach,* Int. J. Press. Vessel. Pip., vol. 150, 2017, pp. 89-95.
- [7] **Dhas J. E. R., Kumanan S.,** *Evolutionary fuzzy SVR modeling of weld residual stress,* Appl. Soft Comput. 42, 2016, pp. 423-430.
- [8] **Dhas J. E. R., Kumanan S.,** *Neuro evolutionary model for weld residual stress prediction,* Appl. Soft Comput. 14, 2014, pp. 461-468
- [9] **Brownlee J.,** *Machine Learning Mastery,* 2020, <https://machinelearningmastery.com/feature-selection-subspace-ensemble-in-python/>
- [10] **Kesse M. A., Buah E., Handroos H., Ayetor G. K.,** *Development of an Artificial Intelligence Powered TIG Welding Algorithm for the Prediction of Bead Geometry for TIG Welding Processes using Hybrid Deep Learning.* Metals, 10(4), 2020, p.451.
- [11] **Korat P., Sama M.,** *Implementation of Artificial Intelligence in TIG.* Available at SSRN 3462445, 2019
- [12] **Li Y., Lee T. H., Wang C., Wang K., Tan C., Banu M., Hu S. J.,** *An artificial neural network model for predicting joint performance in ultrasonic welding of composites,* Procedia CIRP, 76 2018, pp.85-88.
- [13] **Mahadevan R., Jagan A., Pavithran L., Shrivastava A., Selvaraj, S. K.,** *Intelligent welding by using machine learning techniques,* Materials Today: Proceedings, 2021
- [14] **Mongan P. G., Hinchy E. P., O'Dowd N. P., McCarthy C. T.,** *Optimisation of Ultrasonically Welded Joints through Machine Learning,* Procedia CIRP, 93, 2020, pp.527-531.
- [15] **Priya B. G.,** *Automizing Pulsed Current Tig Welding Process Using Regression,* IJSRR, vol 7., iss. 8, 2019
- [16] **Satpathy M. P., Moharana B. R., Dewangan S., Sahoo, S. K.,** *Modeling and optimization of ultrasonic metal welding on dissimilar sheets using fuzzy based genetic algorithm approach,* Engineering Science and Technology, an International Journal, vol. 18(4), 2015, pp. 634-647.
- [17] **Sumesh A., Rameshkumar K., Mohandas K., Babu R. S.,** *Use of machine learning algorithms for weld quality monitoring using acoustic signature,* Proc. Comp. Sci., 50, 2015, pp. 316-322.
- [18] *** <https://www.analyticsvidhya.com/blog/2017/09/common-machine-learning-algorithms/>
- [19] *** <https://machinelearningmastery.com/linear-regression-for-machine-learning/>
- [20] **Virkkunen I., Koskinen T., Jessen-Juhler O., Rinta-Aho J.,** *Augmented ultrasonic data for machine learning,* Journal of Nondestructive Evaluation, vo.1 40(1), 2021, pp. 1-11.
- [21] **Zhu H., Ge W., Liu Z.,** 2019. *Deep Learning-based classification of weld surface defects,* Applied Sciences, vol. 9, iss. 16, 2019, p.3312.