Journal of Dynamics, Monitoring and Diagnostics, 2022, 1, 176-187
https://doi.org/10.37965/jdmd.2022.118

ISTP
RESEARCH ARTICLE

Industry 4.0 Application in Manufacturing for Real-Time Monitoring and Control

Debasish Mishra,[1] Ashok Priyadarshi,[2] Sarthak M Das,[3] Sristi Shree,[4] Abhinav Gupta,[5] Surjya K Pal,[6] and Debashish Chakravarty[7]

[1]Advanced Technology Development Centre, Indian Institute of Technology Kharagpur, Kharagpur, West Bengal 721302, India

[2]Mathematics and Computing, Birla Institute of Technology Mesra, Ranchi, Jharkhand 835215, India

[3]Information Technology, Birla Institute of Technology Mesra, Ranchi, Jharkhand 835215, India

[4]Computer Science and Engineering, Birla Institute of Technology Mesra, Ranchi, Jharkhand 835215, India

[5]Electronics and Communication Engineering, Birla Institute of Technology Mesra, Ranchi, Jharkhand 835215, India

[6]Department of Mechanical Engineering, Indian Institute of Technology Kharagpur, Kharagpur, West Bengal 721302, India

[7]Department of Mining Engineering, Indian Institute of Technology Kharagpur, Kharagpur, West Bengal 721302, India

(Received 13 June 2022; Revised 18 July 2022; Accepted 07 September 2022; Published online 07 September 2022)

Abstract: Modern manufacturing aims to reduce downtime and track process anomalies to make profitable business decisions. This ideology is strengthened by Industry 4.0, which aims to continuously monitor high-value manufacturing assets. This article builds upon the Industry 4.0 concept to improve the efficiency of manufacturing systems. The major contribution is a framework for continuous monitoring and feedback-based control in the friction stir welding (FSW) process. It consists of a CNC manufacturing machine, sensors, edge, cloud systems, and deep neural networks, all working cohesively in real time. The edge device, located near the FSW machine, consists of a neural network that receives sensory information and predicts weld quality in real time. It addresses time-critical manufacturing decisions. Cloud receives the sensory data if weld quality is poor, and a second neural network predicts the new set of welding parameters that are sent as feedback to the welding machine. Several experiments are conducted for training the neural networks. The framework successfully tracks process quality and improves the welding by controlling it in real time. The system enables faster monitoring and control achieved in less than 1 s. The framework is validated through several experiments.

Key words: cloud; edge; deep neural networks; friction stir welding; Industry 4.0; internet of things; machine learning; manufacturing; process control; process monitoring; signal processing

I. INTRODUCTION

Modern manufacturing aims to reduce downtime and track process anomalies to make profitable business decisions. This ideology is strengthened by the introduction of Industry 4.0, the fourth industrial revolution. It aims at continuous monitoring of high-value manufacturing assets [1,2]. This article reports an Industry 4.0 application developed and implemented on a welding process. Welding leads to the joining of discrete materials into one component. This attribute makes quality assurance in welding essential. Conventionally, weld quality is determined by post-weld tests. These tests may be destructive or nondestructive; however, they are time- and capital-intensive and are performed offline [3]. Instead, a network of sensors can be engaged for indirect monitoring, thereby creating a closed-loop feedback system from the sensory data [4]. Industry 4.0 has brought network connectivity into manufacturing [5]. It aims at telemonitoring, i.e., to transmit data to remote/cloud servers enabling the Internet of Things (IoT) [2]. This article presents an Industry 4.0-enabled framework for continuous monitoring and feedback-based real-time control of an advanced welding technique named friction stir welding (FSW). In the following, the theory of FSW process is discussed first, and then state of the prior research is elaborated. The major contributions of this work are presented next.

FSW joins materials by frictional heating, mechanical deformation, and stirring [6]. It involves a nonconsumable tool mounted onto the spindle of the welding machine. The rotating tool plunges into base materials, generating frictional heat and deforming the base materials plastically. Later, it travels over the weld joint line stirring the plasticized material. This fuses the materials, and a weld joint is formed. Because FSW does not melt the base materials, it is recommended for soft alloys such as aluminum, magnesium, and the like materials in both similar and dissimilar configurations [7–10]. This welding process has multiple applications in ship building, railways, aerospace, and automobile sectors. Owing to the vast range of applications, real-time monitoring of the process is essential to ensure quality of the welds.

Corresponding author: Surjya K Pal (e-mail: skpal@mech.iitkgp.ac.in).

The prior research in this regard reports strategies developed for monitoring the welding process. Methods were devised to classify defective and defect-free welds using information derived from acoustic emission (AE), and longitudinal and transverse force signals [11,12]. A few works presented methods to localize welding defects using AE, axial force, power, and spindle torque signals [12–18]. Information derived from AE signal was also utilized in detecting voids in the weld joint line [19,20]. A significant change in the energy of the signal was observed for welds with and without voids. Prior research also communicates about offline prediction of weld quality using features computed from force, spindle, and feed motor current signals [21]. The outputs include prediction of ultimate tensile strength (UTS) and yield strength of welded samples. In addition to the data-driven methods proposed for weld quality prediction, prior research also presents use of data obtained from the physics-driven model to predict the weld quality. Grain size of the welded samples was predicted using strain, strain rate, and temperature modeled in a neural network [22]. Grain size prediction was also achieved using a thermomechanical model with the same set of inputs, i.e., strain, strain rate, and temperature [23]. While these studies present methods to judge the weld quality after the weld is fabricated, they do not provide a requisite measure that can be taken to avoid rejection of the weld. In other words, the prior research cites methods to monitor the welding process; however, they overlook the control action required to reduce the rejection or improve a weld in real time.

To address this challenge, a few authors proposed controlling the plunge depth *(p)* of the tool considering axial force as a feedback signal [24,25]. The plunge depth marks the penetration of the tool's shoulder into the base metals to ensure a proper contact between them. A proportional integral derivative (PID) controller was employed for this task. A PI controller was implemented where rotational speed maintained the temperature [26]. However, PID controllers do not perform well in uncertainties and often have a narrow range of stability. Nonlinearity exists in FSW, because the process involves transient events. The heat generated during welding process is influenced by speed at which the tool is rotating and also the linear speed of the machine bed. Increasing rotation of the tool or lesser linear speeds lead to excessive deformation of base metals. Thus, the reaction of the base metals to these movements decreases. Contrary, the reaction force increases if the tool rotation is insufficient or welding occurs at higher linear speeds. Therefore, these events are not steady and might change with fluctuations in the welding parameters. Hence, an adaptive PID controller was utilized to control welding speed by taking force as a feedback signal [27]. The control systems tracking plunge depth of the tool may be unstable, leading to sudden tool plunging or withdrawal, resulting in welding defects [6]. FSW is governed by welding speed and tool rotational speed which decide the amount of frictional heat and plastic deformation of the base materials. Hence, controlling both is crucial for the welding process. Besides, these studies rely on a single sensor, which may form a threat to the IoT systems. In this regard, data from multiple sensors were collected for real-time control of FSW [28]. Discrete wavelet transform was applied on the signals to locate the welding defects and predict the weld quality using wavelet features. The developed framework was executed on a cloud server, which had a processing time of 8.53 s per batch of data for weld monitoring and control. This article presents an enhanced and accurate system that monitors and controls the welding process. Specifically, the analyses and results demonstrate meaningful information derived from the as-received signals and welding parameters in contrast to the wavelet coefficients. It also demonstrates the application of an edge device in addition to a commercial cloud server that reduces the computational time. Following are the major contributions of this article:

- Application of as-received raw signals in FSW process for accurate prediction of UTS using a deep neural network (DNN). A high coefficient of determination, *R-squared = 0.9999*, is achieved by the deep learning model with a mean absolute error *(MAE)* of *0.10* only.
- Application of a cost-effective machine learning model like random forest in predicting UTS is also shown. This model is also developed using the as-received raw signals. Its performance is compared with that of the neural network. The random forest model achieved a *MAE* of *0.1557* only.
- To lower the variance in the as-received signals, several filtering methods are applied on them. Random forest models are built with each set of filtered data. These models are compared based on their performance with the processed data. The performance of random forest model developed with the unfiltered data is found to be equivalent with other models.
- An edge-cloud coupled monitoring and control system is developed and successfully applied to the welding process. Process monitoring task is accomplished at the edge device located closer to the welding machine. Process control happens via the cloud server. With this improved framework, the feedback to the welding machine is achieved in less than 1 s. Thus, reliability and security are higher.

Base materials in FSW are placed side by side on the machine bed without any gap using suitable fixtures. The materials' surface may have undulations at the mating edge or inhomogeneity. Applications of FSW involve joining large structures. However, it is impossible to check their integrity. Besides, problems may occur in the welding machine. These may deviate the welding speed or tool rotational speed. Therefore, it is essential to monitor and control the process continuously.

The remainder of the paper is organized as follows. Section II discusses the experimental details and the signals acquired during welding process. It also presents the monitoring and control system developed in this work. The results obtained from the analyses and justifications are presented in Section III. Results of the test experiments that validated the developed system are also presented in this section. Finally, the conclusions are presented in Section IV.

II. MATERIALS AND METHODS

This section lists the details of the welding experiments conducted for understanding the domain knowledge of FSW process. It also discusses sensor data collection and analytics. Following is the Industry 4.0 framework developed in this work.

A. TEST MATERIAL AND POST-WELD STUDY

Figure 1 depicts a schematic diagram of the welding process with workpieces fixed on the machine bed in a butt-joint

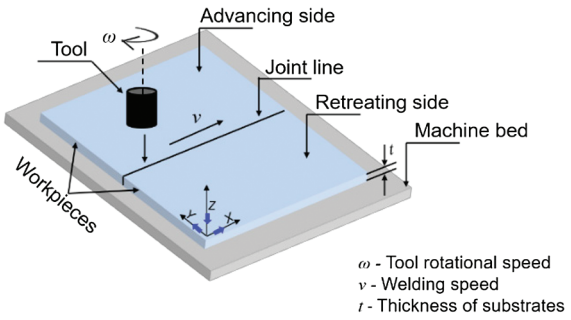


Fig. 1. Schematic picture of FSW.

configuration. The rotating tool (shown in black color) plunges into the workpieces (shown in blue color) resulting in frictional heating and plastic deformation of the workpieces. It is followed by the translation of the tool leading to the transport of plasticized material from one side to the other. For the present work, weld joints were fabricated on a numerically controlled FSW machine (WS004, ETA Bangalore). This machine operates in a displacement-controlled mode, wherein the plunging position is fed to the machine as an input, along with welding speed *(v)* and tool rotational speed *(ω)*. The test material is an aluminum alloy, AA6061, measuring 100 × 80 × 3 mm. H13 steel-alloy tool was used to weld test plates in butt joint. The tool had a flat shoulder of 18 mm diameter, a conical-shaped pin of 6 and 4 mm top and bottom diameters, with a pin length of 2.7 mm.

The experimental combinations (ω and v) are mentioned in Table I. A full factorial approach was carried out, where each value of ω was combined with each value of v resulting in 42 experiments. During the experiments, the tool plunging depth *(p)* was kept constant for all the experiments. The value was 0.2 mm, and it ensured a proper contact between the tool and workpieces. The post-weld study determined the UTS of 42 welds.

Tensile test was performed as per ASTM-E8 standard, where a subsize specimen was selected. The test was performed in a universal tensile testing machine (Instron, 8862) at 1 mm/min strain rate. The UTS values obtained for the 42 welded samples are shown as a scatter plot in Fig. 2.

Table I. Range of ω and v

ω values (in rpm)	600, 1000, 1400, 1800, 2200, and 2600
v values (in mm/min)	40, 50, 60, 80, 100, 150, and 200

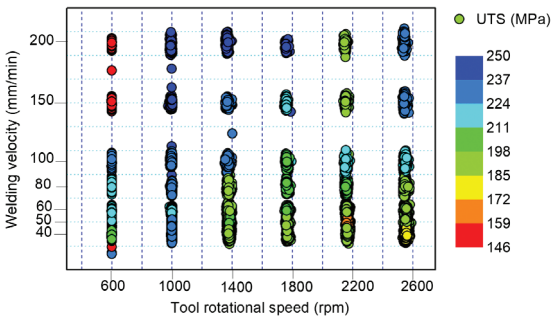


Fig. 2. Variation of UTS values along with parametric combinations.

All the experimental observations are also listed in Table II along with the selected welding parameters. The UTS of the base material was 270 MPa. In the 42 welded samples, the highest UTS value was obtained as 249.3 MPa with the parametric combination of 1000 rpm (ω) and 200 mm/min (v). The lowest UTS value was 146.2 MPa with the parametric combination of 600 rpm (ω) and 200 m/min (v). It was observed that there is an increase in

Table II. Experimental observations

Expt#	Welding parameters: ω (rpm)–v (mm/min)	UTS (MPa)	Joint strength efficiency (%)
1.	600–40	210.09	77.81
2.	600–50	214.61	79.48
3.	600–60	215.47	79.80
4.	600–80	221.08	81.88
5.	600–100	228.15	84.50
6.	600–150	150.71	55.82
7.	600–200	146.20	54.15
8.	1000–40	225.60	83.55
9.	1000–50	224.58	83.18
10.	1000–60	221.50	82.04
11.	1000–80	229.44	84.98
12.	1000–100	235.41	87.19
13.	1000–150	244.82	90.67
14.	1000–200	249.31	92.34
15.	1400–40	191.06	70.76
16.	1400–50	199.81	74
17.	1400–60	192.90	71.44
18.	1400–80	188.44	69.79
19.	1400–100	233.26	86.39
20.	1400–150	231.33	85.68
21.	1400–200	241.68	89.51
22.	1800–40	190.33	70.49
23.	1800–50	195.14	72.27
24.	1800–60	197.97	73.32
25.	1800–80	205.06	75.95
26.	1800–100	208.25	77.13
27.	1800–150	218.36	80.88
28.	1800–200	241.68	89.51
29.	2200–40	185.76	68.80
30.	2200–50	171.48	63.51
31.	2200–60	194.19	71.92
32.	2200–80	207.40	76.81
33.	2200–100	211.94	78.50
34.	2200–150	190	70.37
35.	2200–200	190.02	70.38
36.	2600–40	178.40	66.08
37.	2600–50	186.09	68.92
38.	2600–60	188.53	69.82
39.	2600–80	193.89	71.81
40.	2600–100	210.84	78.09
41.	2600–150	226.22	83.79
42.	2600–200	225.66	83.58

the UTS value when ω takes a lower value of 600 rpm, while v keeps increasing, starting from 40 mm/min to 100 mm/min. However, as the v value increases beyond 150 mm/min, the UTS values decrease remarkably. This is due to the poor material flow, which signifies that the available heat was insufficient for deforming the material plastically. The welded sample ($\omega = 600$, $v = 150$, and $\omega = 600$, $v = 200$) had several surface defects, and thus, lower UTS.

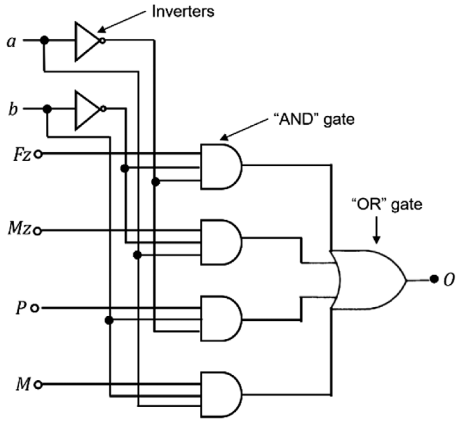
Further, a few experiments were also conducted at 500 rpm; however, the fabricated welds had several surface defects. Thus, the range of parameter selection was restricted to 600 rpm. For the other parametric combinations, for instance, at a higher value of ω, i.e., 2600 rpm, the UTS of the welds which are fabricated with different values of v is said to increase. In comparison to other values of ω, 2600 rpm is the highest, which indicates that frictional heat availability at this value would have been the highest as compared to other values. A few experiments were repeated for ensuring the results.

B. SENSING AND DATA COLLECTION

The welding machine is integrated with multiple sensors such as speed sensors for capturing the variation in v and ω, a load cell for capturing axial force (F_z) and spindle torque (M_z), and a power sensor for recording the total power (P) consumed during welding. The speed sensors installed on the machine are encoders that record rotational speed of the spindle and linear speed of the machine bed. The load cell is of strain gauge type, and a Hall effect-type power sensor was integrated externally. As communicated, the weld joints are created because of the pressure exerted by the tool resulting in heat and deformation. Therefore, Fz forms a crucial variable. In the event of any disturbance in the process, Fz is expected to vary. With variation in Fz, the torque experienced by the spindle will also vary. Therefore, these data were captured during welding experiments. Likewise, the power consumed by the welding machine would differ for a faulty setting. For FSW, welding defect is a fault which might occur because of improper settings. The details about the working principle of strain gauge-type load cell and Hall effect-type power sensor can be read from the cited work [29]. For data collection, a virtual instrument (VI) consisting of a time-controlled 4:1 multiplexer was built. The VI refers to a logic circuit designed to select one of the several input lines to a common output line. Figure 3(a) shows the circuit of the multiplexer utilized in this study. The truth table and Boolean logic expression are shown in Fig. 3(b).

The data from the force, torque, and power sensors together with a marker value are inputs to the multiplexer. The marker signifies the completion of one complete set of data. The output from the multiplexer is linked with the input of the transmission control protocol (TCP) write block. From here, the data is sent to the next block for further processing. The machine input parameters, i.e., ω and v, are returned to the machine via the same protocol. LABVIEW VI obtains ω and v along with some markers to differentiate them, i.e., firstly, a 0 is sent as a marker followed by ω, and similarly, 1 is sent as a marker followed by v. The actual data acquisition rate is 10 Hz. However, there is a delay of ~31 milliseconds in the multiplexer, which leads to an effective sampling rate of 8 Hz.

For 42 experiments, these sensors' data was collected. Figure 4 presents the variation in the acquired raw signals.



(a)

Select lines		Inputs				Output
a	*b*	*Fz*	*Mz*	*P*	*M*	*O*
0	0	1	0	0	0	*F*
0	1	0	0	1	0	*P*
1	0	0	1	0	0	*Mz*
1	1	0	0	0	1	*M*

Fz : force, *Mz* : torque, *P* : power, *M* : marker

$$\text{Boolean expression: } O = \bar{a}\,\bar{b}\,F_z + a\,\bar{b}\,M_z + \bar{a}\,b\,P + a\,b\,M$$

(b)

Fig. 3. Multiplex build in the VI: (a) circuit diagram and (b) Boolean expression and truth table.

To represent ω and v together, feed per revolution (FPR) is considered, which is the ratio of v and ω. Figure 4(a) shows the correlation among UTS and rotational speed and FPR. Lower FPR produces welds with high UTS (> 210 MPa) and moderate UTS (180–201 MPa), whereas higher FPR gives rise to welds with lower UTS (< 150 MPa). FPR is low when v is less, and ω is high, the combination results in desirable frictional heating and plastic deformation of the base materials. However, the opposite condition results in low heat and deformation, resulting in defective welds. For these situations, the sensors' data also vary. F_z lies in 6000–8400 N (Fig. 4(b)) and M_z in 9–16 NM (Fig. 4(c)) for higher UTS. The variation is also expressed with FPR and UTS values, which shows sufficient segregation of welds. These plots indicate that the raw signals can be utilized to model UTS due to their close relationship.

C. MONITORING AND CONTROL FRAMEWORK

Figure 5 depicts the schematic representation of the framework consisting of welding machine and coupled edge-cloud system. The major highlights of the framework are as follows: (a) use of raw signals, (b) use of an edge device and

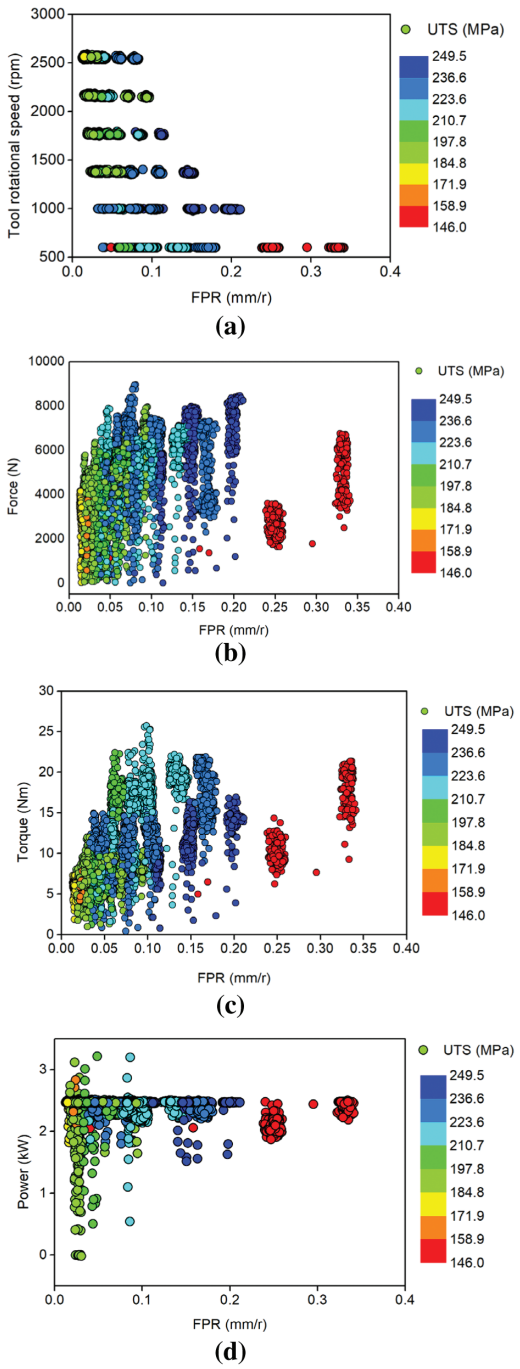


Fig. 4. Scatter plot of UTS with: (a) tool rotational speed and FPR, (b) force and FPR, (c) torque and FPR, and (d) power and FPR.

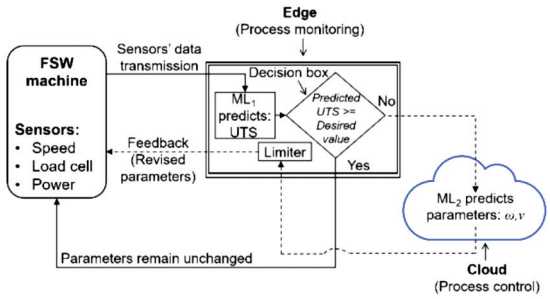


Fig. 5. Schematic representation of the framework for real-time monitoring, prediction, and feedback-based control of weld quality.

cloud server, and (c) use of a limiter. The real-time data acquired in the FSW machine is continuously fed to an edge device. The edge device is connected to the FSW machine via *WebSocket* using the TCP/internet protocol (IP). As the effective sampling rate is 8 Hz, data to the edge device is received in each millisecond interval. Each interval's data is fed to the DNN model (ML_1) residing in the edge device that predicts the UTS of the weld in real time.

As data is sampled in millisecond intervals, UTS prediction also occurs in millisecond intervals. The decision box compares the predicted UTS value with a desired UTS value. If the predicted value is less than the desired UTS, the data is sent to the cloud. For making it more relevant to the industrial application, an instance of *Google cloud server* was utilized. The cloud server consists of another DNN model (ML_2), which predicts the modified parameters to achieve the desired UTS. The modified parameters form the feedback to the FSW machine. Instead of directly feeding the parameters, they are routed via the edge device. The edge device consists of a limiter that checks and determines whether the values are within the safe limits of the machine. In the next section, the results achieved are discussed.

III. RESULT AND DISCUSSION

This section discusses the results obtained by applying the proposed methodology.

A. DATA ANALYSIS – CORRELATION

The correlation matrix was determined to explore the signals' association with UTS. Correlation measures the extent to which two variables tend to change together. Two methods were considered for determining the correlation among the parameters, namely (a) *Pearson* and (b) *Spearman*.

While the *Pearson correlation* evaluates the linear relationship between two variables, the *Spearman correlation* identifies the monotonic relationship. It is helpful because the variables need not change together at a constant rate. As the aim here was to determine the association of the variables/parameters with UTS, the *Spearman correlation* will be more helpful. This method is based on the ranked values for each variable rather than the raw data.

Figure 6 depicts the correlation maps obtained by employing the two methods. From the figure, ω, v, F_z, and M_z can be found to have a moderate relationship with UTS, and P has a weak relationship. The highest correlation was found with FPR. Despite the weak relationship exhibited by power, it can act as an alternative solution for the control scheme if the load cell fails or starts behaving erroneously.

B. IMPORTANCE OF DATA – NOISE AND VARIANCE IN SIGNALS

It is imperative to study the presence of noise in the signals. Raw signals might have random errors that imply unpredictable variations in the measured signals. This error is also referred to as noise, whose sources could be vibrations of the machine, fluctuations in incoming electrical power, electromagnetic radiation from nearby electrical equipment, etc. Therefore, it was imperative to study the presence of noise, i.e., variance and bias in the signals. Various filters were applied to the raw sensors' data: *average filter, median*

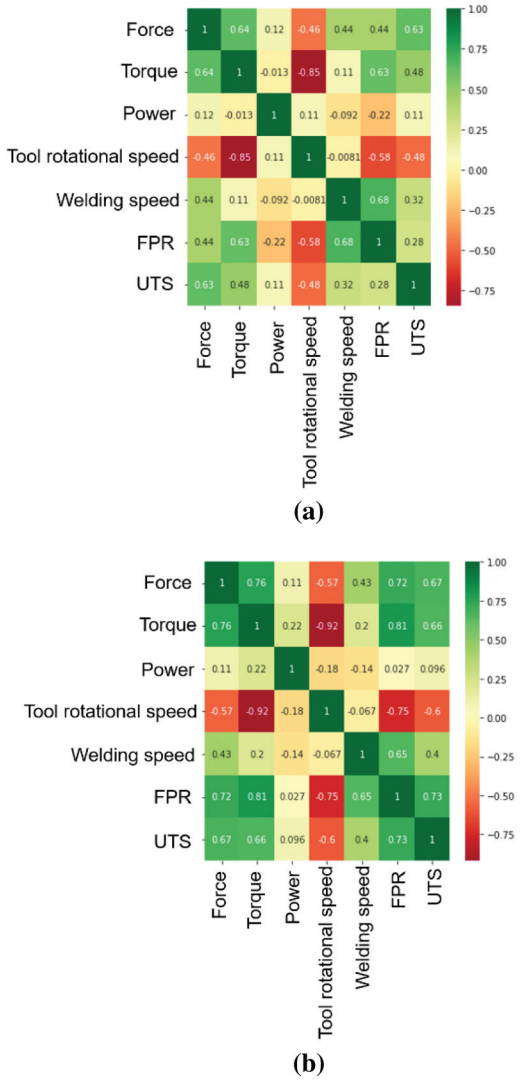


Fig. 6. Correlation in the dataset: (a) *Pearson* and (b) *Spearman*.

filter, Gaussian filter, Wiener filter, Fourier denoising, wavelet denoising, and Whittaker smoothing. It was assumed that the noise was additive. The *average filter* is a linear filter used to reduce additive shot noise. It takes a window sliding along the signal replacing the data value with the mean value of all the neighboring data values. The size of the sliding window *(N)* is an odd number. The *median filter* is a nonlinear filtering technique utilized for removing the additive noise of impulsive nature. It is similar to the *average filter*, excluding the fact that the median of the neighboring values alters the data value in this case. *A Gaussian filter* is a linear low-pass filter in the frequency domain, which convolves a *Gaussian* function over the signal. It is effective in removing *Gaussian* noise. The degree of smoothing is controlled by the standard deviation *(σ)*.

Wiener filter is based on the statistical approach, and it works well for both additive and multiplicative noises. It assumes that the spectral properties (cross-spectral density between observed and original signal and power spectral density of the original signal) are known. The noise is considered to be uncorrelated with the signal. This filter tries to reduce the overall mean square error of the denoised data by comparing it with the original data. It is effective in removing *Poisson noise*, *Gaussian noise,* and *Speckle noise* from the signal. *Fourier denoising* is another filtering technique in the frequency domain where *Fast Fourier Transform* is applied to the data, and thresholding is done based on *Donoho's* method. *Wavelet denoising* refers to the time–frequency domain transform using wavelets to localize the signal features at different scales. Lastly, *Whittaker smoothing* was utilized, which smoothens the time series based on optimization of penalized least squares. It attempts to fit a curve that characterizes original data. The parameter λ is the smoothing factor. A cost-effective machine learning model, random forest, is developed with each set of filtered data to predict the UTS. The performance of these models developed with the filtered data is compared with the model built using the unfiltered data. Random forest regression model is chosen as it is popular for nonlinear modeling and takes less computational time with no local minima/maxima problems. The performance is compared in terms of *MAE* as presented in Table III.

Except for *Fourier* denoising, there does not exist a significant difference in *MAE*. Least *MAE* was obtained with *Whittaker smoothing*, which is not a substantial improvement from the *MAE* obtained using raw data. If this filter were utilized in real time, the cost and time invested in filtering signal would weigh more than the improvement in accuracy it provides. *Fourier denoising* provided a substantial improvement in *MAE*. But it is a frequency domain filter, so utilizing it in the model was not possible for denoising in real time.

It is important to identify the reason why the raw signal gives a better performance in the case of FSW. The probable reasons for the same could be (a) negligible vibration during

Table III. MAE obtained in the ML model by using the filtered and raw datasets

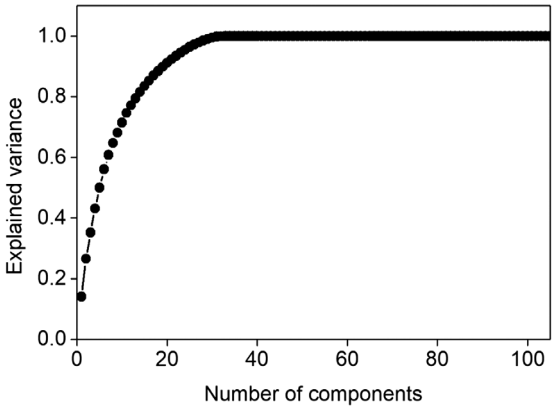
Raw signal	Whittaker smoothing		Wavelet denoising	Fourier denoising	Weiner filter		Gaussian filter		Median filter		Average filter	
MAE	MAE	λ	MAE	MAE	MAE	*N*	MAE	σ	MAE	N	MAE	*N*
0.1557	0.145	0.01	0.1349	0.068	0.144	3	0.121	2	0.135	3	0.134	3
	0.134	0.1			0.135	5	0.113	4	0.134	5	0.137	5
	0.133	1			0.133	7	0.127	6	0.132	7	0.151	7
	0.131	10			0.129	9	0.139	8	0.122	9	0.148	9
	0.129	100			0.132	11	0.118	10	0.117	11	0.137	11
	0.124	1000			0.132	13	0.112	12	0.113	13	0.140	13
	0.116	10,000			0.133	15	0.100	14	0.113	15	0.131	15
	0.107	100,000			0.137	17	0.098	16	0.113	17	0.126	17

the welding and (b) negligible effect of other electronic components, etc.

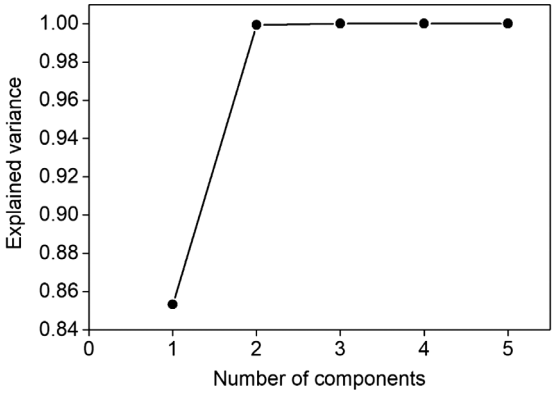
There might be other reasons for this as well, but they have not been the focus of this article. Further to explore the efficacy of the raw data over the processed data, principal component analysis was also carried out. As mentioned in the introduction, earlier research has shown the utility of the *wavelet decomposition* technique for localizing the welding defects in real time, where the detail coefficients were extracted from the signal [28]. It comprised of a set of 35 detail coefficients, each of force, torque, and power, a total being 105. The raw data utilized in this study consists of five components, i.e., force, torque, power, ω, and v.

Figure 7(a) shows the plot of variance in the case of the *wavelet* coefficients, and Fig. 7(b) shows the same with respect to the raw data. It can be observed that close to 90% of the variance can be explained by only 2 components of raw data, while the same requires more than 20 components in the case of the processed data. Thus, the weld quality prediction can be better performed by using the raw data instead of performing further signal processing. This will ensure the timeliness of the monitoring and control system

Thus, the database consists of the historical data of various welding jobs (42 numbers of experiments) performed at different process parameters, i.e., ω and v, corresponding sensors' data (i.e., force, torque, and power), and UTS values of each weld. FPR was not taken into consideration as the input parameters for welding consisted



(a)



(b)

Fig. 7. Plot of explained variance in case of: (a) processed data and (b) raw data.

of ω and v only, and this was beneficial for real-time application.

C. MACHINE LEARNING

The application of a cost-effective ML model like random forest for predicting the weld quality is established in Section III.B. This model is developed using the unfiltered and unprocessed data, and its performance is determined to be equivalent to the models developed with the filtered data. This section discusses the development of DNN model for predicting the UTS using the unfiltered and unprocessed data. The performance of this model is compared with that of the random forest model.

Two DNN models are built: one (ML_1) for predicting the UTS and the second (ML_2) for predicting the modified parameters. These models are trained and tested offline and then are deployed in the scheme for real-time decision-making. ML_1 is deployed in the edge device and ML_2 in the cloud server. The dataset for training ML_1 consisted of five features, i.e., F_z, M_z, P, ω, and v, and UTS was the output. From the correlation map, it is observed that all the sensors' data are having a moderate degree of correlation with UTS. Therefore, all the contributing factors were considered in modeling. The dataset for ML_2 consisted of four features, i.e., F_z, M_z, P, and desired UTS, and ω and v were the outputs. A neural network tries to generalize a problem by learning a mathematical transformation between given inputs and outputs. The learning occurs in deep networks of the model having several layers and neurons. Because of their capability in generalizing a given problem, they are considered to be more robust than the shallow ML models. Therefore, the performance of random forest model and DNN model is compared. Also, an earlier research work developed neural network model for UTS prediction using filtered and processed data [28]. Therefore, to compare the usefulness of employing unfiltered and unprocessed data, the architecture of the models proposed in [28] is considered. This includes hidden layers and hidden neurons in each layer. The modification was in the input layer only. The UTS prediction model consisted of 8 hidden layers and the other model for predicting the modified parameters consisted of 10 hidden layers, as presented in Fig. 8(a) and (b), respectively, with multiple neurons in each layer. The optimized architectures of ML_1 and ML_2 were obtained by applying genetic algorithm.

These models were trained with the unfiltered data recorded in this study by utilizing the *backpropagation* algorithm that employs the *gradient descent* method. The learning rate for ML_1 was varied from 0.001 to 0.00001, and for ML_2, it was varied from 0.01 to 0.00001. The models were tuned using the *Adam* optimizer [30]. The momentum coefficient was fixed as 0.9. The activation function is a scaled exponential linear unit that helps preserve the mean and the activation function. It preserves the mean and variance information from the previous layer making the neural net self-normalizing. It prevents the problem of vanishing gradients.

For training and testing of the models, a random split of 70% and 30% of the dataset was performed. The validation set was a part of the training data. In the model development process, a part of the data was always held back and the *MAE* was computed to determine the models' performance. The performance of the models was ensured to be adequate by testing on several portions of the training data. Finally,

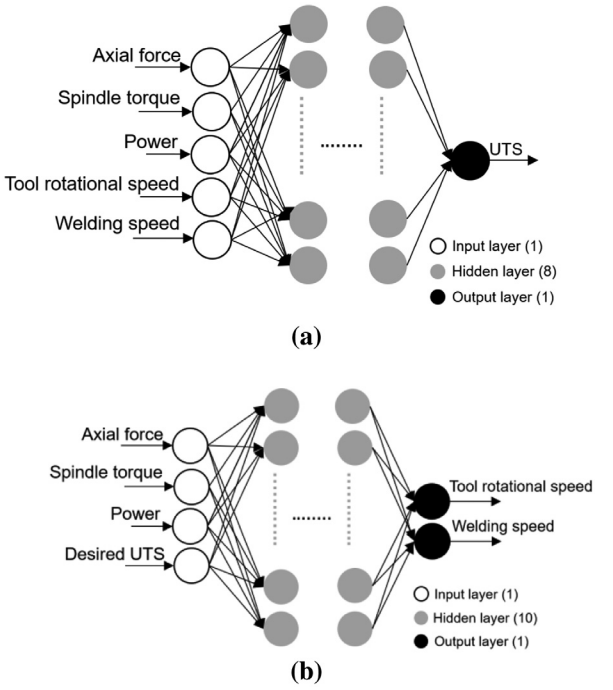


Fig. 8. Architecture of DNN models: (a) UTS prediction model and (b) modified parameters prediction model.

the performance was ensured by computing on the unseen test data. The test data was completely new to the model and, therefore, a separate data for validation was not considered. The training and testing loss for the models predicting UTS and modified parameters are presented in Fig. 9(a) and (b).

These plots show the minimization of the loss with the increasing number of epochs during the training. The testing results for the model predicting UTS is depicted in Fig. 10, where the predicted UTS values are shown to correspond to the actual values. The metrics considered for evaluating the performance of the model are listed in Table IV.

A satisfactory *MAE* with a high R^2 value is obtained for the model. However, since multiple features are input to the DNN, the R^2 value can give a false impression. It is validated by evaluating the *Adjusted* R^2, which in this case is 0.9994. It suggests the ability of the model to make accurate predictions. In comparison to prior research, the model developed in [28] has a mean absolute percentage prediction error of ~4%, which is less than 1% with the DNN model developed in this work.

1) EDGE DEVICE – PROCESS MONITORING. An edge device is located at the edge of communication between a source of data generation and a server. In this work, it is between the FSW machine and the *Google cloud* server. This device usually has a low hardware configuration and is attached to the machine via some communication protocol. The purpose is to filter, analyze, or process the data locally, close to the point of generation. Manufacturing machines will generate a vast amount of data, and the data has to be analyzed to make intelligent decisions. The analyses can be entirely performed in a cloud server; however, the data must be continuously sent and received, requiring an uninterrupted high bandwidth. Therefore, an edge device proves beneficial to having a reliable system and preventing loss of data. It accelerates data streaming and helps in eliminating

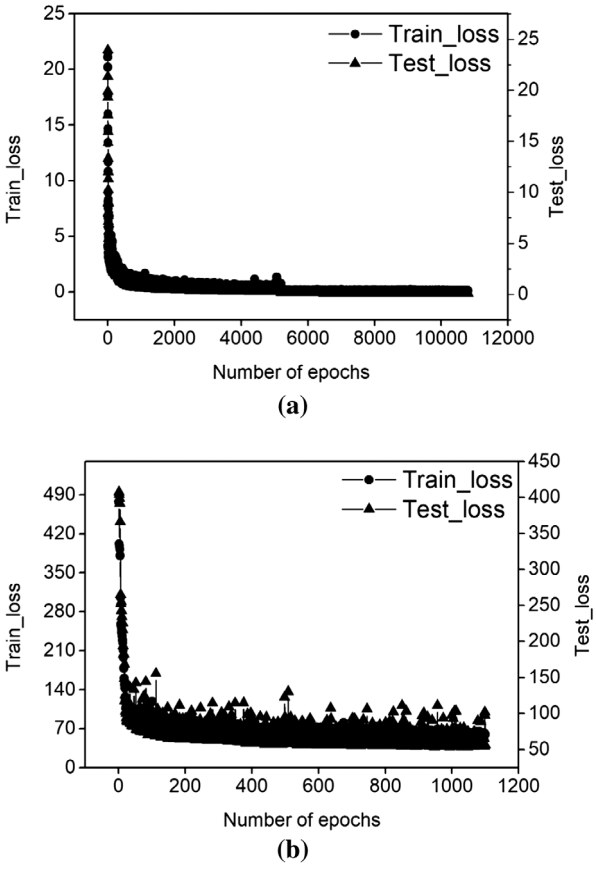


Fig. 9. Plots depicting the training and testing loss: (a) UTS prediction model and (b) modified parameters prediction model.

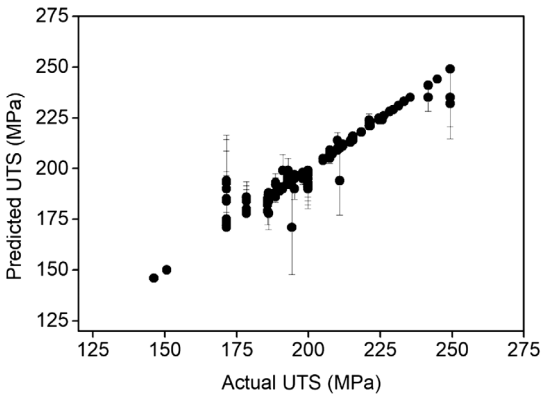


Fig. 10. Plot depicting actual and predicted UTS values.

Table IV. Model evaluation

MAE	R-squared	Adjusted R-squared
0.10	0.9999	0.9994

time lag. The edge device used in this work has a random-access memory (RAM) of 8 GB and 500 GB disk space and a Windows operating system. This computer accessed the FSW machine via the TCP/IP protocol, and they exchange data through *WebSocket*. As the welding starts, this device receives the sensors' data (F_z, M_z, and P) and the input parameters (ω and v) and feeds them to ML_1 for predicting the UTS.

2) CLOUD SERVER – PROCESS CONTROL. The cloud server is an instance of *Google Cloud* with a RAM of 16 GB and 500 GB of disk space. It was connected with the edge device via TCP/IP but on a different port number. The cloud server consists of the ML_2 to predict the modified parameters. The sensory data is received here when the predicted UTS is lower than the desirable UTS value. While controlling the FSW machine for achieving a better weld is taken in the edge, the action is performed through this cloud. The model exploits the correlation to find the values of ω and v needed to achieve the desired value of the UTS. The predicted ω and v are sent back to the edge using the socket connection between the cloud and the edge based on the TCP/IP protocol. The edge acts as a relay between the cloud and the FSW machine, i.e., to transfer the feedback received from the cloud to the FSW machine. Edge performs a safety check using the predefined constraints on the parameters received as feedback from the cloud and ensures that the predicted values do not compromise the machine's health.

The employed FSW machine can accommodate v of 1000 mm/min and ω of 3000 rpm. The edge bounds ω in 600–2600 rpm and v in 35–200 mm/min. These were identified from the experiments. A set of experiments in this study was conducted at 2600 rpm, which produced satisfactory welds. Thus, 2600 rpm has been set as the upper limit. $\omega < 600$ rpm is insufficient to plasticize the workpiece and, therefore, has been ignored. At the same time, no welding was performed at $v > 200$ mm/min. The lower limit (35 mm/min) avoids making the process slow, affecting efficiency. If ω is predicted to < 600 or > 2600 rpm, then 600 and 2600 will be sent to the machine, respectively. Similarly, if v is predicted < 35 mm/min or > 200 mm/min, then 35 and 200 will be sent, respectively. These constraints ensure that the motors of the machine are rotating within the safe values.

D. VALIDATION OF THE DEVELOPED SCHEME

Several experiments were performed to test the proposed framework. The results are presented in the following subsections. Tensile specimens were cut from the welded sample, and their UTS was determined to validate the scheme's results. However, one challenge here was that as the model detected the predicted UTS to be lesser than the desired value, it sent the data to the cloud to control the weld quality, which occurred in less than a second. But, the tensile specimen (as per ASTM E8) needs to have a width of 10 mm (subsize sample). Thus, it was decided to deal with the data in batches for validating the scheme to address this challenge. However, a constant batch size may not be capable of deriving acceptable results each time. Considering the continuous and constant length of the samples to be welded, it may work fine when v is chosen in small values. Still, when v is considerably high, the welding will finish instantly without even letting the scheme control the process during any anomaly. Therefore, an equation has been written for the dynamic batch selection, as shown in Eq. 1.

$$Batch\ size = \frac{60 \times 10 \times 8}{Initial\ velocity\ (v)} \tag{1}$$

where 60 is used for conversion of time to seconds, 10 corresponds to the width of the tensile sample for which the dynamic batch is being selected, 8 refers to the sampling rate of 8 Hz, and the initial velocity is the v with which the welding process begins, and the batch size is always an integer value.

1) REAL-TIME CONTROL OF WELD QUALITY – STARTING WITH OPTIMUM CONDITIONS. A sample was welded by selecting the following input parameters: 1400 rpm (ω) and 50 mm/min (v). These parameters were chosen randomly from the database.

As per the database, this parametric combination produced a weld with a UTS of 199 MPa, which corresponds to ~70% joint strength efficiency (shown as "Initial selection" in Fig. 11). However, to validate the developed scheme, 80% joint strength efficiency was fed to the system, which formed the desirable value. As the welding begins, the data is received in the edge device, which continuously predicts the UTS of the weld being fabricated. The picture of the welded sample is depicted in Fig. 11. As per Eq. 1, the batch size for this weld was 96.

Therefore, after collecting 96 data points, the scheme was fed with the next set of data points of F_z, M_z, P, ω, and v for predicting the UTS, which was less than the desired value (80% joint strength efficiency). Therefore, that corresponding batch of sensors' data was sent to the cloud server, which predicted the modified parameters as shown in Fig. 11: 1006 rpm (ω) and 49 mm/min (v) (shown as "Achieved result"). The predicted UTS was 220 MPa which corresponds to 81.5% joint strength efficiency. This result is validated from the experimental knowledge base developed in this study. The closest parametric combination is 1000 rpm (ω) and 50 mm/min (v), which produced a weld sample with joint strength efficiency of 83%.

Figure 12 shows the corresponding sensors' data during the entire process. The maximum amplitude of F_z and

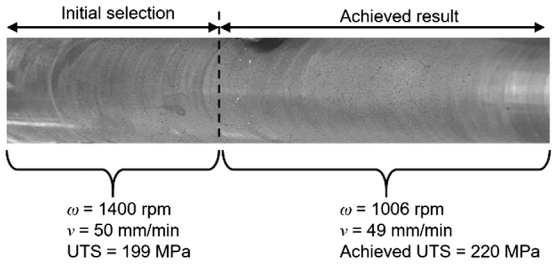


Fig. 11. Picture of the welded sample with achieved joint strength efficiency of 80% from the initial selection of 70% using the developed monitoring and control scheme.

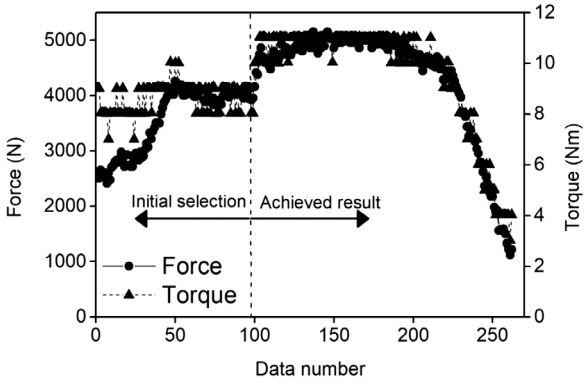


Fig. 12. Picture depicting variation in sensors' data (force and torque) corresponding to Fig. 11.

M_z during the stage of "Initial selection" is ~4000 N and ~10 NM, respectively, which was a result of the selected process parameters, 1400 rpm (ω) and 50 mm/min (v). Upon detection, this resulted in modified parameters in real time, i.e., 1006 rpm (ω) and 49 mm/min (v), while the magnitude of force was ~5000 N and torque of ~12 NM.

In this manner, the framework is able to control the weld quality. Variation in power is not shown, since there was no significant change in the power signal. The reason for this could be its lower correlation with UTS, ω and v.

2) REAL-TIME CONTROL OF WELD QUALITY – STARTING WITH DEFECTIVE CONDITIONS. Figure 13 depicts the picture of another welded sample, the fabrication of which was initiated at 600 rpm (ω) and 250 mm/min (v).

This set was chosen to produce a defective weld, which was known from the created knowledge base. The defects can be seen in the figure (shown as "Initial selection"). For this experiment, the scheme was fed with the desirable UTS to be 220 MPa, i.e., 80% joint strength efficiency. In this case, the batch size was 19 (Eq. 1). As the welding began and the data was received, the model fed the machine with modified parameters, 1008 rpm (ω) and 60 mm/min (v), which has a UTS of 220 MPa (shown as "Achieved result"). This predicted parametric combination avoided the welding defects in contrast to the initial set conditions. Figure 14 shows the corresponding variation in the sensors' data (F_z and M_z) during the entire process. Huge fluctuations in the data can be observed for the selected initial combination. As highlighted, the combination of ω and v decides the frictional heating and plastic deformation. When these quantities are improper, defects will be produced. The sensors capture the effect of the welding parameters; thus, huge fluctuations are observed. However, post the control over the process in real time, there is negligible variation in the sensors' data, and a sound weld is produced.

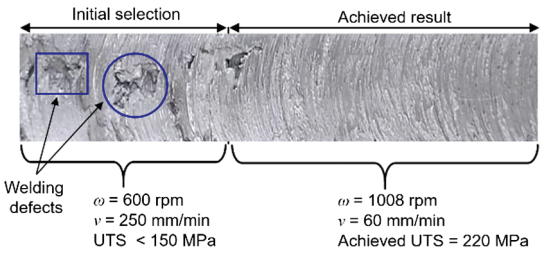


Fig. 13. Picture of the welded sample with achieved joint strength efficiency of 80% from the initial selection of 50% using the developed monitoring and control scheme.

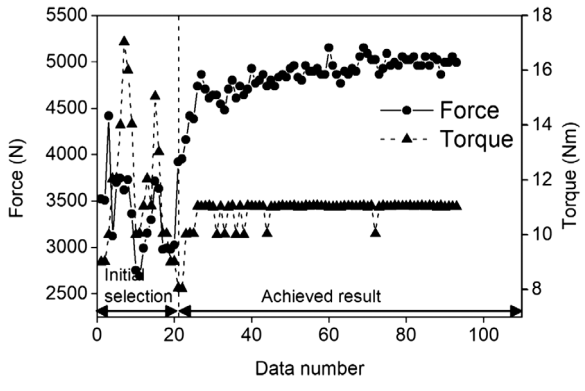


Fig. 14. Picture depicting variation in sensors' data (force and torque) corresponding to Fig. 13.

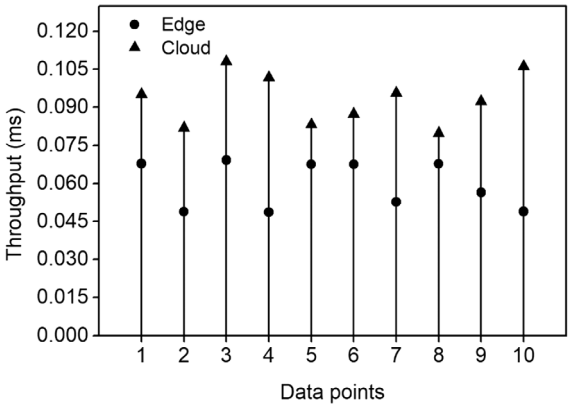


Fig. 15. Data transmission between edge and cloud.

As highlighted, the developed scheme always attempts to maintain the desired UTS. Industries may not have a fixed UTS value. In that case, the system can be fed with an upper and lower band of UTS values catering to industry needs. If the predicted UTS is out of that band, the batch of sensors' data and the mean of the lower and upper bounds can be sent to the cloud server for predicting v/ω. In this way, the system can maintain the UTS of the sample being welded in certain bounds. As the framework works on each data point, the control is achieved in less than a second. However, to test and validate the scheme model, a batch size has been selected here. For industrial applications, this batch size can be removed.

E. BENEFIT OF USING EDGE ALONG WITH THE CLOUD SERVER

The edge-cloud architecture provides a trade-off between the qualities of edge and cloud computations and attempts to derive the best from both. It utilizes the high computational power of the cloud system, which may simultaneously control many machines in a manufacturing industry. At the same time, it also uses the edge system attached to the individual machine for real-time data streaming and processing, which eliminates time lag while making the process real time.

In this manner, the presented architecture derives optimal results by harnessing the qualities of the blocks mentioned above. Figure 15 shows the throughput of the real-time data transmission between edge and cloud. It can be observed that the time taken for transmitting data to the edge device is always lesser than that of the cloud. Hence, it benefits by reducing time lag. The other advantage could be from the security perspective. Industries are concerned about sharing their data on a public cloud server. As the manufacturing data is limited within the edge device for monitoring, it ensures the security of the data. The data will only be sent over a public network in an anomaly in the process.

IV. CONCLUSION

An intelligent Industry 4.0 framework for continuous monitoring and feedback-based real-time control of defects in

the FSW process is presented in this article. The following are the major conclusions:

- Raw signals carry adequate information that can be leveraged for online monitoring of the weld quality. In this study, the variance and bias in the data is assured by modeling the UTS using filtered and unfiltered data.
- A cost-effective ML model like random forest predicts the UTS values using unfiltered or unprocessed data with a *MAE* of *0.1557* only. The neural network model trained on the unfiltered data is superior to random forest model and predicts UTS values with a *MAE* of *0.10* only.
- Axial force and spindle torque are strong indicators of weld quality. Severe fluctuations are observed in these signals while welding with improper welding parameters resulting in defective welds.
- An on-off type controller, as proposed in this study, can avoid defective welding conditions. It operates on a desired UTS value which is compared with the predicted UTS value.
- Edge devices are essential for addressing time-critical manufacturing decisions. The edge-cloud system developed in this study reduces the throughput time compared to the use of cloud infrastructure only.

Acknowledgments

The authors are grateful for getting the opportunity to submit their research to this Special Issue. They also announce the start of an international collaborative project on "Smart Manufacturing and Industry 4.0" to be supported by *The Efficiency and Performance Engineering Network International Collaboration Fund Award 2021"* (ID: TEPEN-ICF2021-04).

CONFLICT OF INTEREST STATEMENT

The authors declare no conflicts of interest.

References

[1] R. Ortt, C. Stolwijk, and M. Punter, "Implementing Industry 4.0: assessing the current state," *J. Manuf. Technol. Manage.*, vol. 31, pp. 825–836, 2020. https://doi.org/10.1108/JMTM-07-2020-0284.

[2] S. K. Pal, D. Mishra, A. Pal, S. Dutta, D. Chakravarty, and S. Pal, "Evolution of manufacturing and its journey towards digital twin," pp. 1–20, 2022. https://doi.org/10.1007/978-3-030-81815-9_1.

[3] D. Mishra, S. K. Pal, and D. Chakravarty, "Industry 4.0 in welding," pp. 253–298, 2021. https://doi.org/10.1007/978-3-030-63986-0_8.

[4] S. K. Pal, D. Mishra, A. Pal, S. Dutta, D. Chakravarty, and S. Pal, "Signal processing for digital twin," pp. 117–187, 2022. https://doi.org/10.1007/978-3-030-81815-9_3.

[5] J. W. Veile, D. Kiel, J. M. Müller, and K. I. Voigt, "Lessons learned from Industry 4.0 implementation in the German manufacturing industry," *J. Manuf. Technol. Manage.*, vol. 31, pp. 977–997, 2020. https://doi.org/10.1108/JMTM-08-2018-0270.

[6] D. Mishra, R. B. Roy, S. Dutta, S. K. Pal, and D. Chakravarty, "A review on sensor based monitoring and control of friction stir welding process and a roadmap to Industry 4.0," *J. Manuf. Process.*, vol. 36, pp. 373–397, 2018. https://doi.org/10.1016/j.jmapro.2018.10.016.

[7] O. Mypati et al., "Enhancement of joint strength in friction stir lap welding between AA6061 and AISI 304 by adding diffusive coating agents," *Proc. Inst. Mech. Eng. Part B J. Eng. Manuf.*, pp. 1–14, 2019. https://doi.org/10.1177/0954405419838379.

[8] O. Mypati, D. Mishra, S. Sahu, S. K. Pal, and P. Srirangam, "A study on electrical and electrochemical characteristics of friction stir welded lithium-ion battery tabs for electric vehicles," *J. Electron. Mater.*, vol. 49, pp. 72–87, 2020. https://doi.org/10.1007/s11664-019-07711-8.

[9] S. K. Sahu et al., "Friction stir welding of polypropylene sheet," *Eng. Sci. Technol. An. Int. J.*, vol. 21, pp. 245–254, 2018. https://doi.org/10.1016/j.jestch.2018.03.002.

[10] S. K. Sahu, K. Pal, R. P. Mahto, and P. Dash, "Monitoring of friction stir welding for dissimilar Al 6063 alloy to polypropylene using sensor signals," *Int. J. Adv. Manuf. Technol.*, vol. 104, pp. 159–177, 2019. https://doi.org/10.1007/s00170-019-03855-3.

[11] V. Soundararajan, H. Atharifar, and R. Kovacevic, "Monitoring and processing the acoustic emission signals from the friction-stir-welding process," *Proc. Inst. Mech. Eng. Part B J. Eng. Manuf.*, vol. 220, pp. 1673–1685, 2006. https://doi.org/10.1243/09544054JEM586.

[12] E. Boldsaikhan, E. M. Corwin, A. M. Logar, and W. J. Arbegast, "The use of neural network and discrete Fourier transform for real-time evaluation of friction stir welding," *Appl. Softw. Comput. J.*, vol. 11, pp. 4839–4846, 2011. https://doi.org/10.1016/j.asoc.2011.06.017.

[13] C. Chen, R. Kovacevic, and D. Jandgric, Wavelet transform analysis of acoustic emission in monitoring friction stir welding of 6061 aluminum. Int J Mach Tools Manuf 2003; vol. 43, pp. 1383–1390. https://doi.org/10.1016/S0890-6955(03)00130-5.

[14] M. S. Orozco, E. J. Macias, A. S. Roca, H. C. Fals and J. B. Fernandez, "Optimisation of friction-stir welding process using vibro-acoustic signal analysis," *Sci. Technol. Weld Join*, vol. 18, 532–540, 2013. https://doi.org/10.1179/1362171813Y.0000000134.

[15] S. Kumari et al., "Defect identification in friction stir welding using continuous wavelet transform," *J. Intell. Manuf.*, pp. 1–12, 2016. https://doi.org/10.1007/s10845-016-1259-1.

[16] U. Kumar, I. Yadav, S. Kumari, K. Kumari, and N. Ranjan, "Defect identification in friction stir welding using discrete wavelet analysis," *Adv. Eng. Softw.*, vol. 85, pp. 43–50, 2015. https://doi.org/10.1016/j.advengsoft.2015.02.001.

[17] R. B. Roy et al., "Weld defect identification in friction stir welding through optimized wavelet transformation of signals and validation through X-ray micro-CT scan," *Int. J. Adv. Manuf. Technol.*, vol. 99, pp. 623–633, 2018. https://doi.org/10.1007/s00170-018-2519-3.

[18] D. Mishra et al., "Weld defect localization in friction stir welding process," *Weld World*, vol. 65, pp. 451–461, 2021. https://doi.org/10.1007/s40194-020-01028-5.

[19] P. Fleming et al., "In-process gap detection in friction stir welding," vol. 1, pp. 62–67, 2008. https://doi.org/10.1108/02602280810850044.

[20] Y. Yang, P. Kalya, R. G. Landers, and K. Krishnamurthy, "Automatic gap detection in friction stir butt welding operations," *Int J Mach Tools Manuf.*, vol. 48, pp. 1161–1169, 2008. https://doi.org/10.1016/j.ijmachtools.2008.01.007.

[21] B. Das, S. Pal, and S. Bag, "Weld quality prediction in friction stir welding using wavelet analysis," *Int. J. Adv.*

Manuf. Technol., vol. 89, pp. 711–725, 2017. https://doi.org/10.1007/s00170-016-9140-0.

[22] L. Fratini, G. Buffa, and D. Palmeri, "Using a neural network for predicting the average grain size in friction stir welding processes," *Comput. Struct.*, vol. 87, pp. 1166–1174, 2009. https://doi.org/10.1016/j.compstruc.2009.04.008.

[23] M. P. Iqbal, A. Tripathi, R. Jain, R. P. Mahto, S. K. Pal, and P. Mandal, "Numerical modelling of microstructure in friction stir welding of aluminium alloys," *Int J Mech Sci.*, p. 185, 2020. https://doi.org/10.1016/j.ijmecsci.2020.105882.

[24] G. E. Cook, R. Crawford, D. E. Clark, and A. M. Strauss, "Robotic friction stir welding," *Ind. Robot. An Int. J.*, vol. 31, pp. 55–63, 2004. https://doi.org/10.1108/01439910410512000.

[25] X. Z. X. Zhao, P. Kalya, R. G. Landers, and K. Krishnamurthy, "Design and Implementation of a Nonlinear Axial Force Controller for Friction Stir Welding processes," in *2007* Am Control Conf., vol. 130, pp. 5553–5558, 2007. https://doi.org/10.1109/ACC.2007.4282731.

[26] A. Bachmann, J. Gamper, M. Krutzlinger, A. Zens, and M. F. Zaeh, "Adaptive model-based temperature control in friction stir welding," *Int. J. Adv. Manuf. Technol.*, vol. 93, pp. 1157–1171, 2017. https://doi.org/10.1007/s00170-017-0594-5.

[27] T. A. Davis, Y. C. Shin, and B. Yao, "Observer-based adaptive robust control of friction stir welding axial force," in IEEE/ASME Int. Conf. Adv. Intell. Mechatron., AIM, pp. 1162–1167, 2010. https://doi.org/10.1109/AIM.2010.5695824.

[28] D. Mishra et al., "Real time monitoring and control of friction stir welding process using multiple sensors," *CIRP J. Manuf. Sci. Technol.*, 2020. https://doi.org/10.1016/j.cirpj.2020.03.004.

[29] S. K. Pal, D. Mishra, A. Pal, S. Dutta, D. Chakravarty, and S. Pal, "Sensor electronics for digital twin," 2022, pp. 21–115. https://doi.org/10.1007/978-3-030-81815-9_2.

[30] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," in *3rd Int. Conf. Learn Represent*, 2015, p. arXiv. https://doi.org/10.48550/arXiv.1412.6980.