Material Forming - ESAFORM 2023 Materials Research Forum LLC
Materials Research Proceedings 28 (2023) 1067-1074 https://doi.org/10.21741/9781644902479-117

An ANN based approach for the friction stir welding process intrinsic uncertainty

QUARTO Mariangela[1,a *], BOCCHI Sara[1,b], GIARDINI Claudio[1,c] and D'URSO Gianluca[1,d]

[1]Department of Management, Information and Production Engineering, University of Bergamo, Via Pasubio 7/b, 24044 Dalmine (BG), Italy

[a]mariangela.quarto@unibg.it, [b]sara.bocchi@unibg.it, [c]claudio.giardini@unibg.it, [d]gianluca.d-urso@unibg.it

Keywords: Artificial Neural Network, Process Variability, Friction Stir Welding

Abstract. Friction Stir Welding is a solid-state bonding process that during last years has caught the researcher's attention for the mechanical characteristics of the welded joints that are quite similar to the properties of the base material. The Friction Stir Welding is affected by several process parameters leading to intrinsic variability in the process. The present paper would introduce a new approach for predicting the surface hardness in different areas of the welded parts. Specifically, this method is based on the hypothesis that multiple Artificial Neural Networks, characterized by the same architecture but different weights, can be used for forecasting both the punctual value of the local hardness and its confidence interval, resulting in taking into account the intrinsic variability of the process.

Introduction

Friction Stir Welding (FSW) is one of the most significant technological developments in metal joining of the last thirty years and it is also considered a sustainable technology thanks to its energy efficiency. Indeed, the energy involved is on average 30% - 40% lower than conventional welds. In addition, the emission of greenhouse gases in FSW is about 30% lower than in traditional welds.

Friction Stir Welding is a welding technology patented by The Welding Institute (TWI) in 1991 [1]. This solid state mechanical joining technique is used to make joints between similar or dissimilar materials that are difficult to be welded with conventional fusion techniques, such as sintered materials, magnesium, copper, Inconel, titanium, metal matrix composites and thermoplastics. In particular, it is used for high-strength aluminum alloys, which are difficult to join using traditional techniques due to their typical microstructural alteration during aging hardening.

From a structural point of view, the FSW introduces welded joints without overlapping edges and allows the elimination of connecting parts, such as rivets, with a consequent reduction in the weight and costs of the structures. This process, still relatively young, has produced great interest in the industrial world so as to achieve enormous development, especially in naval applications, but also in the aerospace, railway and automotive industries. The difficulty of the traditional welding processes is linked to the problems developed due to the high temperatures during processing. Conversely, in Friction Stir Welding, which is a friction process by plastic deformation, fusion temperature is never reached during processing, therefore, the formation of the problems that occur in conventional welds is avoided.

In general, welds made using FSW have a different microstructure with respect to the joints welded with fusion processes because the maximum temperature reached in the heat-affected zone is significantly lower than the melting temperature [2]. By basing the microstructural characterization of the joints on the size of the grains and precipitates, it is possible to divide the structure of the FSW joint into four characteristic zones: the nugget, the thermo-mechanically affected zone, the heat affected zone and the base material. The nugget is the area of completely recrystallized material close to the welding line; near to the nugget, the thermo-mechanically affected zone (TMAZ) is detectable. In the TMAZ the material undergoes both plastic deformation and heating, whilst only alterations of the microstructure and mechanical properties due to the welding thermal cycle are present in the heat affected zone (HAZ) [3].

For these FSWed joints characteristics, it is evident that the process parameters greatly influence the flow of the plastically deformed material [4]. In particular, there is a strong relationship between the microstructure and the quality of the joints is mainly due to the closed connection between the mechanical properties of the FSWed joints and the microstructural variations of the above-mentioned different zones of the welds. Therefore, particular attention must be paid in the choice of parameters to ensure adequate processing conditions, in order to avoid the formation of potential defects or weakness points in the identified welding areas [2].

Due to the large number of variables that must be considered simultaneously, it is not easy to optimize and predict with a good level of accuracy the process performance and the weld quality as a function of both the process conditions and the weld zones considered. For this reason, optimizing process performance by considering each process parameter individually is not an applicable approach. To overcome this problem, various artificial intelligence techniques have recently been used, among which the Artificial Neural Network stands out [5,6].

An Artificial Neural Network (ANN) is a mathematical model capable of reproducing the reasoning mode of the human brain, simulating its processing capabilities through the connection of neurons, which are the elementary computational elements of any neural network. The ANN, once it has been adequately trained, is able to solve even non-linear problems easily and quickly. To do this, the ANN is composed by an oriented network formed by different layers: the input layer, which collects the external data, the hidden layers, which process the information through the activation functions and the weights of each connection, and the outputs layer.

The training of the neural network is based on the use of a dataset thanks to which the network is able to determine the weights of the interconnections and of the activation function. These weights define the reliability of the prediction and can vary as a function of trainings; in fact, repeating the training defines different weights configurations which generate different results in the prediction. Consequently, it is evident that if the same weights are always considered, the ANN produces a constant predictive response with respect to a specific set of inputs. However, this aspect is a distorted version of reality. Indeed, it is known that every process is characterized by an intrinsic variability caused by external conditions that can influence systems in different and, often unpredictable, ways. The possibility of considering this uncertainty in the analysis and prediction of complex systems is a topic of considerable importance [7]. For all these reasons, the main objective of this work is to develop a model capable of predicting not only the punctual mechanical characteristics reached in the different areas of the weld starting from the welding parameters, but also of being able to evaluate the intrinsic variability of the process.

The proposed approach aims at forecasting the welding properties considering the process' natural variability by means of a group of Artificial Neural Network (ANN) having the same architecture, but different weights. The hypothesis of this work is that predicting the output using this approach allows to define an upper and a lower limit in which the proper results are placed. This idea was applied to the FSW process for the prediction of the hardness value reached by an aluminum alloy (AA 2024-T3) as a function of the main welding parameters (considered as input nodes of the ANN).

Materials Research Proceedings 28 (2023) 1067-1074 https://doi.org/10.21741/9781644902479-117

Methodology

Experimental tests.

AA2024-T3 sheets were friction stir welded perpendicularly to the rolling direction. The FSW process was performed using a CNC machine tool with a tool characterized by a smooth plane shoulder (16-mm diameter) and a frustum of cone pin shape (maximum and minimum diameters equal to 6 and 4 mm, height equal to 3.9 mm). The tool was realized from a drawn bar of carbon steel C40 without any heat treatment. Table 1 reports the welding parameters considered for the experimental campaign. The difference in the feed rate combined to 1200 rpm and 1500 rpm is linked to the presence of macro and micro defects in the joints welded with 1500 rpm-40 mm/min and 1500 rpm-100 mm/min. These defects led the authors to eliminate these combinations of parameters from further evaluation, considering them not suitable for the material considered.

Table 1. Parameters set-up for experimental tests.

Rotational speed (rpm)	Feed rate (mm/min)
1200	40
1200	70
1200	100
1500	70

All the other process parameters, such as the clamping system for the sheets, the tool inclination (fixed at 3°), and the tool penetration depth into the sheets (fixed at 3.99 mm), were kept constant.

Rockwell B hardness tests (HRB) were performed for all the welding conditions, following a 5mm spaced grid in the central zone of the top of the specimens, according to ISO 6508. For each sample, three profiles of indentations were carried out moving from the joint axis to the base material.

Developed approach.

One of the main issues in the FSWed joints is represented by the desire to obtain welds having mechanical properties as close as possible to those of the base material. In this context, one of the most important characteristics is the surface hardness of the joints.

Due to the importance of the prediction of these kind of properties without the need to execute experimental tests, the prediction models have become very useful in the scientific research. Nevertheless, these forecasting models are characterized by a limitation in the prevision since they supply a deterministic constant result for the same inputs, not considering the intrinsic variability of the process. This is a huge limitation typical also of the trained artificial neural networks (ANN). Indeed, the hypothesis behind the development of this models is that, even if the same architecture is considered, the ANN can be characterized by different connection weights and activation functions, obtainable through different training iterations. This means that same architecture, but different weights, generate different, even if still reliable, forecasts. This means that it might be possible to develop a model able to predict not only the required punctual value but also its confidence interval. This kind of model has been developed following the flow-chart reported in Fig. 1.

Materials Research Proceedings 28 (2023) 1067-1074 https://doi.org/10.21741/9781644902479-117

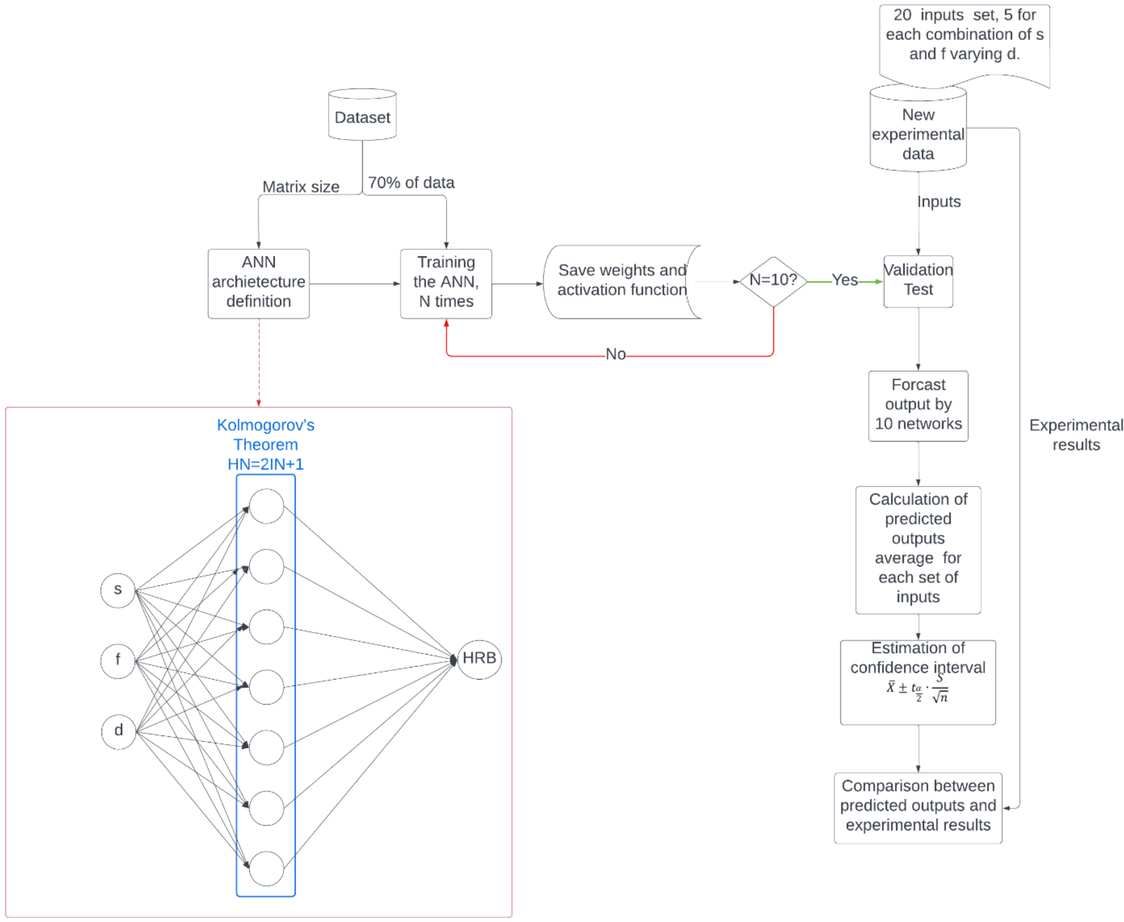

Fig. 1. Algorithm flow-chart.

The following stages are considered:

1. <u>ANN architecture definition.</u> The inputs and the outputs layers were defined according to the welding parameters and the final characteristic to be evaluated. Specifically, the input nodes (IN) were three and represented respectively the rotational speed of the tool (s), the feed rate of the tool (f) and the distance from the welding line (d) where the hardness, which represented the output node (ON), has to be known. A single hidden layer was introduced since the literature supports the idea of introducing few hidden layers to avoid unnecessary complexity of the network. The number of hidden nodes (HN) was selected based on one of the most applied heuristic techniques found in the literature: the Kolmogorov's Theorem, which demonstrates that a network with a single hidden layer can compute any arbitrary function of its input [8]. This theorem allows to estimate the size of the hidden layer as reported in Eq.1.

$$HN = 2 \cdot IN + 1 \quad (1)$$

2. <u>ANN training.</u> The dataset used for the ANN training contained 470 inputs combinations and the relative outputs, considered as target values. The dataset dimension fit with the requirements defined by Alwosheel A. et al. [9]. The dataset was based on the data of experiments performed to evaluate the surface hardness of the welded joints. Through a Matlab code, the ANN was trained to identify the weights, defining N=10. N indicated the number of ANN, with fixed architecture, considered for the multiple ANN approach trained.

This allows to obtain 10 ANNs with the same size, but with the layers differentiated from the weights of the connections and the activation functions. Levenberg–Marquardt algorithm was chosen as the training function for all the ANNs.

3. Forecasting tests. A new group of experimental data were considered for the verification of the applicability of the algorithm. Specifically, 20 combinations of input parameters were selected, and the relative experimental results were collected, ensuring to consider inputs always included in the training interval of the network. These sets of inputs were introduced in the group of the trained ANNs obtaining 10 predictions for each combination, for which the average value was calculated. At this point, the idea of confidence interval was introduced to verify the reliability of the previsions and to forecast the intrinsic variability of the process.
4. Confidence interval definition. Since the standard deviation of the data distribution was unknown, it was necessary to estimate the interval of the population average (μ) using, in addition to the sample mean ($\bar{X}$), the sample standard deviation (S), allowing to estimate the population standard deviation (σ). This can be done assuming that the random variable was normally distributed and, as a consequence, that the confidence interval can be defined by introducing the t-distribution. The t-distribution is similar to the normal distribution (symmetrical, bell-shaped, with mean and median equal to 0), but with a higher probability to fall into its queues than in the central part of the bell. Moreover, increasing the number of degrees of freedom (n), the random variable t tends to the normal distribution. Thus, considering α as the confidence level and n as the sample dimension (represented by the number of repetitions of inputs combination introduced into the dataset), the confidence interval is estimated as reported in Eq. 2. Considering the degrees of freedom equal to 3 and an α equal to 5%, from the statistical table, $t_{\frac{\alpha}{2}} = 4.3027$. The confidence interval was estimated based on the training dataset.

$$\bar{X} - t_{\frac{\alpha}{2}} \cdot \frac{S}{\sqrt{n}} \leq \mu \leq \bar{X} + t_{\frac{\alpha}{2}} \cdot \frac{S}{\sqrt{n}} \tag{2}$$

In this way, the estimated upper and lower bounds allowed to predict and define the interval where the results, in this case the surface hardness, can be placed if the experiments will be repeat more than once. Specifically, from the practical point of view, it was possible to predict HRB with a certain level of variability reproducing the intrinsic variability of the process, linked to external and uncontrollable factors.

Results and Discussion

The validation was performed by comparing the experimental and forecasted data and verifying their placement considering the predicted confidence interval. 20 combinations of inputs, indicating different welding conditions and hardness locations, were selected from the experimental tests (not used for ANN training): five hardness measurements for each combination of welding parameters by randomly selecting the distance from the welding line (d). Table 2 reports the combination used for the validation and the related results obtained from the experiments (performed as reported in previous section) and the average of previsions. It is important to underline that the data used for the validation are not related to a unique welding joint, thus resulting in a nonlinear distribution of the obtained hardness values in the typical FSW W-shape. This choice was made because the main aim of this step was to demonstrate the validity of the developed method considering the totality of the dataset and not limiting it to a determined couple of inputs and their related output. Each ANN responds with a slightly different result from the others and this ANN behavior allows to assimilate these simulation repetitions to the intrinsic variability of the process, such as possible to observe by the standard deviation and the dispersion of the data is coherent with a stable and feasible process. Furthermore, the error in prevision is very low showing a maximum error equal to 6.25%. As it is possible to observe more clearly in Figure 2, all the average values of the predictions are contained in the estimated confidence interval, the same for the ten predicted values. The validation process allows to demonstrate the reliability of the developed approach not only in the prediction performed by a single ANN, but also in the identification of a confidence interval in which the process results can vary due to the intrinsic variability.

Table 2. Inputs combinations and ANN previsions.

INPUTS			TARGET	PREDICTED OUTPUT		
S	**f**	**d**	**HRB exp**	**MEAN**	**Std. Deviation**	**%Error**
1200	40	75	79	79.5	1.4	0.58%
1200	40	60	79	78.9	0.5	-0.12%
1200	40	10	75	70.3	1.2	-6.25%
1200	40	-2.5	70	67.9	2.3	-2.93%
1200	40	-35	77	77.4	1.0	0.56%
1200	70	50	80	78.4	0.7	-2.05%
1200	70	20	72	74.9	0.8	4.05%
1200	70	2.5	75	71.6	1.2	-4.50%
1200	70	-17.5	74	75.1	0.8	1.45%
1200	70	-45	80	78.4	1.0	-2.03%
1200	100	20	73	75.5	0.6	3.46%
1200	100	2.5	76	73.0	0.6	-3.98%
1200	100	0	76	73.0	0.6	-3.93%
1200	100	-5	71	73.4	0.7	3.31%
1200	100	70	79	78.8	0.5	-0.29%
1500	70	25	79	77.0	1.0	-2.53%
1500	70	10	78	74.5	0.9	-4.46%
1500	70	0	76	73.4	1.3	-3.46%
1500	70	-10	77	74.4	0.8	-3.36%
1500	70	-12.5	79	77.4	1.2	-2.06%

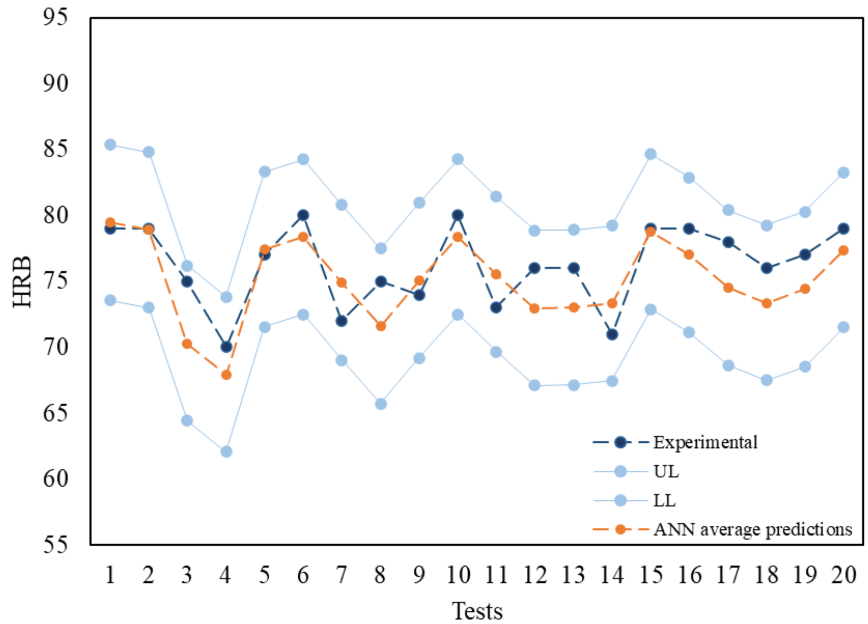

Fig. 2. Validation results.

Summary

A new approach for the prevision of Rockwell B hardness values as a function of the distance from the center of the welding line of joints obtained by the FSW process was developed. The novelty of the model regarded the possibility to predict the results considering the intrinsic variability of the process linked to uncontrollable factors of the process.

This means that, through the proposed approach, it was possible to correctly predict the surface hardness obtainable after a welding process, performed with the welding parameters used for the training. Furthermore, the approach was also able to show the confidence interval of the previsions.

The present approach identifies a new method for the use of ANN in manufacturing processes. In this way, with the same input, it is possible to obtain different outputs, placed in a confidence interval acceptable from the statistical point of view thanks to the different distribution of the weights of the connections, reflecting more a real process that is always characterized by a slight variability in the results.

References

[1] W.M. Thomas, E.D. Nicholas, J.C. Needham, M.G. Murch, P. Templesmith, C.J. Dawes, GB Patent application no. 9125978.8, 1991. Avalible: https://patentimages.storage.googleapis.com/66/60/ad/5f784d0b8653b7/US5460317.pdf (accessed 28 February 2019).

[2] P. Kah, R. Rajan, J. Martikainen, R. Suoranta, Investigation of weld defects in friction-stir welding and fusion welding of aluminium alloys, Int. J. Mech. Mater. Eng. 10 (2015) 1-10. https://doi.org/10.1186/s40712-015-0053-8

[3] M. Jariyaboon, A.J.J. Davenport, R. Ambat, B.J.J. Connolly, S.W.W. Williams, D.A.A. Price, R. Ambat, A.J.J. Davenport, The effect of welding parameters on the corrosion behaviour of friction stir welded AA2024–T351, Corros. Sci. 49 (2007) 877-909. https://doi.org/10.1016/j.corsci.2006.05.038

[4] M. Aissani, S. Gachi, F. Boubenider, Y. Benkedda, Design and optimization of friction stir welding tool, Mater. Manuf. Process. 25 (2010) 1199-1205. https://doi.org/10.1080/10426910903536733

[5] K. Deepandurai, R. Parameshwaran, Multiresponse Optimization of FSW Parameters for Cast AA7075/SiCp Composite, Mater. Manuf. Process. 31 (2016) 1333-1341. https://doi.org/10.1080/10426914.2015.1117628
[6] K. Mallieswaran, R. Padmanabhan, V. Balasubramanian, Friction stir welding parameters optimization for tailored welded blank sheets of AA1100 with AA6061 dissimilar alloy using response surface methodology, Adv. Mater. Process. Technol. 4 (2018) 142-157. https://doi.org/10.1080/2374068X.2017.1410690
[7] W.L. Oberkampf, J.C. Helton, C.A. Joslyn, S.F. Wojtkiewicz, S. Ferson, Challenge problems: Uncertainty in system response given uncertain parameters, in: Reliab. Eng. Syst. Saf., Elsevier, 2004: pp. 11–19. https://doi.org/10.1016/j.ress.2004.03.002
[8] R. Hecht-Nielsen, Kolmogorov's Mapping Neural Network Existence Theorem, in: SOS Printing, 1987.
[9] A. Alwosheel, S. van Cranenburgh, C.G. Chorus, Is your dataset big enough? Sample size requirements when using artificial neural networks for discrete choice analysis, J. Choice Model. 28 (2018) 167-182. https://doi.org/10.1016/J.JOCM.2018.07.002