Vignesh et al. 2018. *Int. J. Vehicle Structures & Systems*, 10(2), 98-102
ISSN: 0975-3060 (Print), 0975-3540 (Online)
doi: 10.4273/ijvss.10.2.05

International Journal of Vehicle Structures & Systems
Available online at www.maftree.org/eja

Comparison of ANN Training Algorithms for Predicting the Tensile Strength of Friction Stir Welded Aluminium Alloy AA1100

R.V. Vignesh[a] and R. Padmanaban[b]

Dept. of Mech. Engg., Amrita School of Engg., Amrita Vishwa Vidyapeetham, Amrita University, Coimbatore, India
[a]Email: r.vairavignesh@gmail.com
[b]Corresponding Author, Email: dr_padmanaban@cb.amrita.edu

ABSTRACT:

Aluminium alloy AA1100 finds application in light weight structures due to its high strength to weight ratio. Friction stir welding is a solid state welding process, in which the materials are joined in the plasticized state. The quality of the friction stir welded joints depends on the process parameters used and tool parameters. In this study, four process parameters were varied at five levels and experimental trials were performed as per face centered central composite design. Artificial neural network model was developed with cascade forward propagation network architecture and trained with LM algorithm and BFGS QN algorithm. The models were used to predict the tensile strength of the joints and the error in prediction was used to judge the accuracy of the developed models. It is observed that BFGS QN algorithm trains the ANN efficiently and results in accurate predictions.

KEYWORDS:

Aluminium alloy; Friction stir welding; Friction stir welding; Artificial neural network; Tensile strength

CITATION:

R.V. Vignesh and R. Padmanaban. 2018. Comparison of ANN Training Algorithms for Predicting the Tensile Strength of Friction Stir Welded Aluminium Alloy AA1100, *Int. J. Vehicle Structures & Systems*, 10(2), 98-102. doi:10.4273/ijvss.10.2.05.

1. Introduction

High strength to weight ratio is a unique property of aluminium alloys, which enabled their use predominantly in aerospace, automotive and other lightweight applications. The formation of alumina over the surface of the alloy at high temperatures reduced the conventional weld ability of the aluminium alloys. But the aluminium alloys could be joined at the plasticizing temperature to prevent the formation of alumina. Friction stir welding (FSW) is a solid state welding process in which the metal/alloy to be joined do not melt. In this process, a rotating tool is plunged into the abutting face of the joint under the action of load and traversed along the joint line [1-2]. The frictional heat generation and extreme plastic deformation induces dynamic re-crystallization and allows flow of material in the solid state [3]. A sound weld is formed at optimum level of heat input to the material during FSW process.

The quality of the welds is governed by the process parameters [4-5] such as tool rotation speed (TRS), welding speed (WS), shoulder diameter (SD) [6], pin diameter (PD), pin profile [7-8], backing plate temperature etc. Sato et al [9] improved the TS of friction stir welded (FSW) AA1100 producing welds with fine grained structure. Khorrami et al [10] found that decrease in TRS decreased the tensile strength (TS) of FSW joints in AA1100. Hussein et al [11] found that the FSW process parameters significantly influence the mechanical properties of the dissimilar weld between aluminium alloy and steel. Kartsonakis et al [12] reinforced nano-phase cerium molybdate in the friction stir welds of dissimilar aluminium alloys and enhanced its corrosion resistance. Sajed et al [13] found that the dwell time had minimal effect in improving the strength of the FSW AA1100. Artificial neural network (ANN) is a strategy for predicting the responses of an event by training the network with the experimental data [14-15]. Ghetiya et al [16] predicted the TS of FSW aluminium alloy 8014 using ANN model.

Narayanan et al [17] compared the response surface methodology technique with ANN model for predicting the TS of FSW aluminium alloy AA7039. Okuyucu et al [18] developed ANN models to predict the TS, yield strength and hardness of FSW aluminium alloy. In this study, aluminium alloy AA1100 was FSW by varying the FSW process parameters TRS, WS, SD and PD as per face centered central composite design. The TS of the joints was measured. Artificial neural network model was developed to predict the TS of the joints based on the FSW process parameters. The developed artificial neural network was trained with Levenberg Marquardt algorithm as well as BFGS - Quasi Newton algorithm. The training efficiency of each algorithm was compared. It is inferred from the correlation coefficient and mean squared error values that BFGS - Quasi Newton algorithm efficiently trains the network.

2. Materials and methods

2.1. Materials

Aluminium alloy AA1100 plate with composition given in the Table 1 was used in the study. The plate of

thickness 5mm was machined into work pieces of length 150mm and width 75mm. The edges of the work pieces were paralleled using vertical milling center to ensure proper clamping in the fixture during FSW process.

Table 1: Composition of aluminium alloy AA1100

Element	Mn	Zn	Cu	Fe	Si	Al
Composition (weight %)	0.05	0.10	0.15	0.95	0.95	Rest

2.2. Response surface methodology

One of the efficient statistical methods to interpret the dependency of a response on process parameter is response surface methodology. The efficiency and quality of friction stir welds are dependent on the FSW process parameters such as TRS, WS, SD, PD, temperature of the backing plate, cooling rate of the welds post FSW process etc. In this study, four FSW process parameters were considered with five levels of variation in each process parameter. The levels of process parameters chosen are given in Table 2.

2.3. Friction stir welding

FSW trials were performed in a numerically controlled vertical milling center. The work pieces were clamped in the fixture, which was designed exclusively for performing FSW trials. The experimental trials were performed as per face centered central composite design and the experimental layout is given in Table 3. A dwell time of 60s was adopted for each FSW trial.

Table 2: Process parameters levels and their corresponding implicit values

Implicit value	Explicit value			
	TRS (rpm)	WS (mm/min)	SD (mm)	PD (mm)
-2	750	30	12	4
-1	900	45	15	5
0	1050	60	18	6
+1	1200	75	21	7
+2	1350	90	24	8

2.4. Tensile strength measurement

The beginning and ending regions of the welds were removed off and not used for metallurgical or mechanical characterization studies. Two tensile test specimens were prepared from the transverse section of each of the welded joints and prepared as outlined by the standard ASTM E-8M-08. The tensile tests were performed in a universal testing machine (Make: Tinius Olsen). The average of the tensile test results is given in Table 3.

Table 3: Experimental and predicted TS of FSW AA1100

Implicit value				Experimental TS (MPa)	Predicted TS (MPa) using ANN model		% Error in prediction	
TRS	WS	SD	PD		LM algorithm	BFGS QN algorithm	LM algorithm	BFGS QN algorithm
-1	-1	-1	-1	45.1	34.07	41.67	24.45	7.60
+1	-1	-1	-1	41.5	36.34	44.01	12.44	-6.04
-1	+1	-1	-1	48.2	47.92	47.09	0.57	2.30
+1	+1	-1	-1	42.2	42.03	42.46	0.40	-0.62
-1	-1	+1	-1	40.3	33.86	41.77	15.97	-3.65
+1	-1	+1	-1	36.2	36.74	36.11	-1.48	0.25
-1	+1	+1	-1	42.2	42.23	39.27	-0.08	6.95
+1	+1	+1	-1	36.9	36.70	37.76	0.55	-2.33
-1	-1	-1	+1	59.3	65.03	61.79	-9.67	-4.20
+1	-1	-1	+1	54.3	53.51	48.98	1.46	9.79
-1	+1	-1	+1	60.3	59.99	67.77	0.51	-12.38
+1	+1	-1	+1	54.3	54.19	55.08	0.19	-1.43
-1	-1	+1	+1	55.3	54.97	67.43	0.59	-21.94
+1	-1	+1	+1	50.1	70.69	50.69	-41.10	-1.17
-1	+1	+1	+1	59.4	65.20	58.87	-9.77	0.88
+1	+1	+1	+1	54.3	69.35	53.94	-27.71	0.66
-2	0	0	0	60.3	59.33	60.26	1.61	0.07
+2	0	0	0	48.2	47.75	46.97	0.93	2.55
0	-2	0	0	59.3	53.36	53.83	10.01	9.22
0	+2	0	0	60.3	60.12	62.82	0.30	-4.18
0	0	-2	0	42.2	41.89	40.73	0.74	3.47
0	0	+2	0	37.9	40.60	44.79	-7.13	-18.17
0	0	0	-2	29.0	30.17	37.33	-4.04	-28.73
0	0	0	+2	54.3	54.63	56.48	-0.60	-4.01
0	0	0	0	71.1	70.20	67.40	1.27	5.21
0	0	0	0	69.3	70.20	67.40	-1.30	2.75
0	0	0	0	69.9	70.20	67.40	-0.43	3.58
0	0	0	0	72.4	70.20	67.40	3.04	6.91
0	0	0	0	70.5	70.20	67.40	0.43	4.40
0	0	0	0	72.4	70.20	67.40	3.04	6.91
0	0	0	0	70.9	70.20	67.40	0.99	4.94

2.5. Artificial neural network

Artificial neural network (ANN) model is constructed by interconnecting the artificial neurons. The three layers that constitute ANN are input layer, hidden layer and output layer. The number of hidden layers is increased with increase in complexity or non-linearity of the input variables with the output. Each layer of ANN has interconnected artificial neurons, which frames the network architecture. Cascade feed forward network architecture of ANN developed in the study is shown in Fig. 1. The neurons in the previous layer are not only connected to the next adjacent layer, but also to the proceeding layers in ANN. This improves the efficiency of the developed model. Cascade forward propagation architecture with p neurons has increasing number of neurons in successive layers. The neuron of the first layer is connected to two input nodes from $x_1, x_2, \dots, x_m$. The neuron in the successive layer is linked to the first input and the output of previous neuron. Thus r^{th} neuron is linked to two input nodes and also to the output of previous neurons. The output of the r^{th} neuron with $p = r + 1$ is given by,

$$z_r = f(u, w) = \frac{1}{1+e^{-\left(\sum_i^p u_i w_i\right)}} \quad (1)$$

Where z is the output, r is the number of layer, f is the activation function, u is the input vector and w is the weight vector.

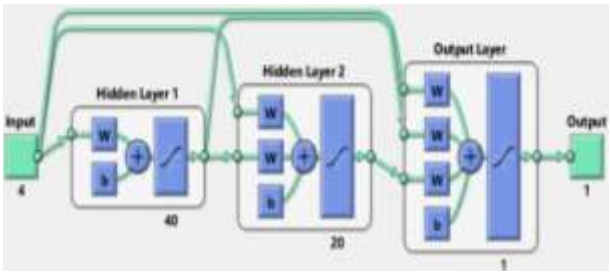

Fig. 1: Architecture of the ANN model

In this study, cascade forward propagation network architecture was used. The architecture has four neurons in the input layer (corresponding to four inputs: TRS, WS, SD and PD); two hidden layers with forty neurons in the first hidden layer and twenty neurons in the second hidden layer; one neuron in the output layer (corresponding to one output: tensile strength). The prediction efficiency of the ANN model could be improved by proper fitting of the experimental data with the ANN. During the process of fitting, ANN reads the input and output data, modifies the initially assumed weights. This process of fitting the ANN is known as training. ANN loses the ability to generalize the output, if the network is over trained. The data that arrive in the artificial neurons are weighted and summed to form activation function. The activation function is transformed by transformation function and fed as data for the proceeding neuron. One of the widely used transformation function, namely hyperbolic tangent sigmoid function, is given by,

$$f(x) = \frac{2}{1+e^{-2x}} - 1 \quad (2)$$

2.5.1. Levenberg Marquardt algorithm

Levenberg Marquardt (LM) algorithm is used to solve non-linear least squares problem. In L.M. algorithm, the learning rate, ε is set to unity and an additional term e^{λ} is introduced in the second derivative error. This minimizes the difference between the output and the predicted value. Let us consider a second order function $F(w)$, $\vec{g}$ be its gradient vector and $[H]$ be its Hessian matrix. In L.M. algorithm, an optimum adjustment in the weight vector $(\vec{w})$ is given by the Eqn. (3). The feed forward network with single output neuron is trained by minimizing the cost function which is given by Eqn. (4).

$$\Delta w = [H + \lambda I]^{-1} g \quad (3)$$

$$\xi_{av} = \frac{1}{2N}\sum_{i=1}^{N}[d(i) - F(x(i); w)]^2 \quad (4)$$

Where Δw optimum adjustment is applied for $\vec{w}$, I is identity matrix with dimensions of $[H]$, λ is regularizing parameter that forces $[H + \lambda I]$ to be positive definite; $\begin{Bmatrix} x(i), \\ d(i) \end{Bmatrix}_{i=1}^{N}$ is the training data set, $F(x(i); w)$ is the approximating function. The gradient and approximated Hessian of ξ_{av} are given by the Eqn. (5) and Eqn. (6) respectively.

$$g(w) = \frac{\partial \xi_{av}(w)}{\partial w} = \frac{-1}{N}\sum_{i=1}^{N}[d(i) - F(x(i); w)]^2 \frac{\partial F(x(i);w)}{\partial w} \quad (5)$$

$$H(w) \approx \frac{1}{N}\sum_{i=1}^{N}\left[\frac{\partial F(x(i);w)}{\partial w}\right]\left[\frac{\partial F(x(i);w)}{\partial w}\right]^T \quad (6)$$

In L.M. method, λ is chosen automatically (starting from a value), until a downhill step is produced for each epoch. The training of network terminates prior to the number of epochs specified, if the following conditions are reached,

$$\lambda > 10 * \Delta\lambda + Max\,[H] \quad (7)$$

$$\frac{MSE_{w_m} - MSE_{w_{m+1}}}{MSE_{w_m}} \leq MSE_{min} \quad (8)$$

Where MSE is mean squared error. In this study, the training parameters for L.M. algorithm were chosen as $Min.gradient = 10^{-7}$ and $No.ofepochs = 10$.

2.5.2. BFGS - Quasi Newton algorithm

Broyden - Fletcher - Goldfarb - Shanno Quasi Newton (BFGS QN) algorithm is used to solve unconstrained non-linear optimization problems in numerical optimization. In this method, the Hessian matrix is approximated using the updates specified by approximate gradient evaluation. The relation between the B_k and $\nabla f(x_k)$ is given by,

$$B_k \boldsymbol{p}_k = -\nabla f(x_k) \quad (9)$$

Where B_k is an approximation to Hessian matrix which is updated iteratively, $\nabla f(x_k)$ is the gradient function evaluated at x_k. A line search (one of the iterative approach in finding the local minimum of the objective function) is then used to find x_{k+1} in the direction of $\boldsymbol{p}_k$. The quasi – Newton update imposed on B_k is given by the Eqn. (10). Let $y_k = \nabla f(x_{k+1}) - = \nabla f(x_k)$ and $s_k = x_{k+1} - x_k$. Then B_{k+1} satisfies the Eqn. (3) which is a secant equation. Instead of requiring the full Hessian matrix at the point x_{k+1} to be computed as B_{k+1}, two

matrices are added to get the approximate Hessian at stage k as given by the Eqn. (12).

$$B_{k+1}(x_{k+1} - x_k) = \nabla f(x_{k+1}) - \nabla f(x_k) \quad (10)$$

$$B_{k+1} \times s_k = y_k \quad (11)$$

$$B_{k+1} = B_k + U_k + V_k \quad (12)$$

Where U_k and V_k are symmetric rank-one matrices, but their sum is a rank-two update matrix. The symmetry and positive definiteness of B_{k+1} is maintained by forming the update form as given by the Eqn. (13). The values of $\alpha$and β are obtained by imposing the secant condition in Eqn. (13) and substituting $u = y_k$ and $v = B_k s_k$. The values of $\alpha$and β are given by Eqns. (14) and (15) respectively.

$$B_{k+1} = B_k + \alpha u u^T + \beta v v^T \quad (13)$$

$$\alpha = \frac{1}{{y_k}^T s_k} \quad (14)$$

$$\beta = \frac{1}{{s_k}^T B_k s_k} \quad (15)$$

The final Eqn. for updating B_{k+1} is given by Eqn. (16), which is obtained by substituting the $\alpha$and β values in Eqn. (13)

$$B_{k+1} = B_k + \frac{y_k {y_k}^T}{{y_k}^T s_k} - \frac{B_k s_k {s_k}^T B_k}{{s_k}^T B_k s_k} \quad (16)$$

In this study, the training parameters for BFGS QN algorithm were chosen as follows: $Min.gradient = 10^{-7}$ and $No.ofepochs = 10$, $\alpha = 0.001$, $\beta = 0.1$.

3. Results and discussions

The experimentally determined TS of the FSW AA1100 joints are given in the Table 3. From the Table 3, it is observed that the specimen joined at TRS of 1050rpm, WS of 60mm.min-1, SD of 18mm and PD of 6mm resulted in high TS of 72.40MPa. The specimen joined at TRS of 1050rpm, WS of 60mm.min-1, SD of 18mm and PD of 4mm resulted in low TS of 29.0MPa. The effect of each process parameter on the TS of the FSW joints are discussed elsewhere [19].

3.1. Error histogram

The error histogram for the ANN model trained with L.M. algorithm is shown in Fig. 2 (a). The error value varied between -19.540 and +9.975. Though many instances of error were closer to zero, the model is not reliable. The error histogram for the ANN model trained with L.M. algorithm is shown in Fig. 2 (b). The error value varied between -11.250 and +4.588. The maximum error value for the predictions made with ANN model trained with BFGS QN algorithm was found to be less than the maximum error value for the predictions made with the ANN model trained with L.M. algorithm.

3.2. Percentage error in prediction

The comparison of the percentage error in prediction for the training data, testing data and the validation data for the ANN models trained with L.M. algorithm and BFGS Q.N. algorithm is shown in the Fig. 3. It is observed from the graph that the percentage error in prediction is lower for the ANN trained with BFGS QN algorithm than the ANN trained with L.M. algorithm. From the Fig. 2 and Fig. 3 it is concluded that the ANN trained with BFGS QN algorithm has higher efficiency in predicting the TS the ANN trained with L.M. algorithm.

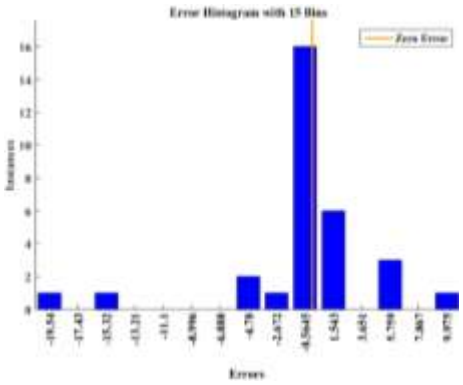

Fig. 2(a): Error histogram of predicted TS using ANN model trained with L.M. algorithm

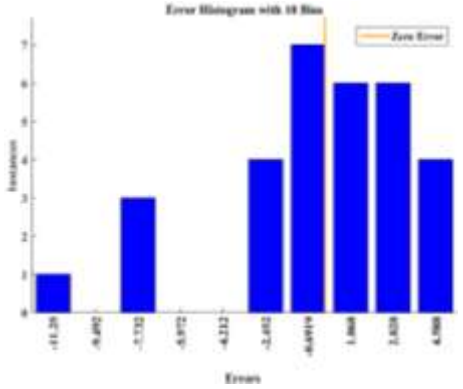

Fig. 2(b): Error histogram of predicted TS using ANN model trained with BFGS Q.N. algorithm

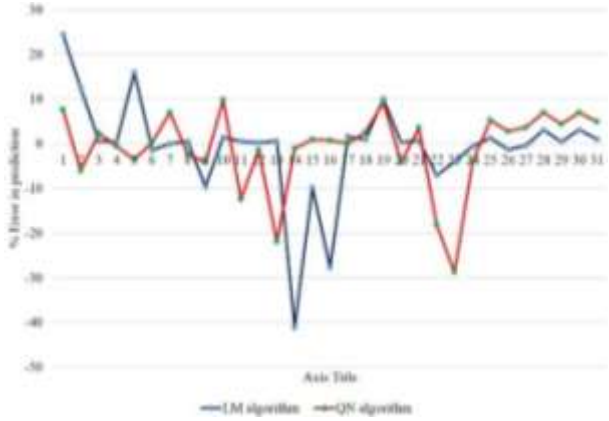

Fig. 3: Comparison of ANN models trained with L.M. algorithm and BFGS Q.N. algorithm based on % error in prediction

4. Conclusions

Aluminium alloy AA1100 was FSW by varying the FSW process parameters TRS, WS, SD and PD. Artificial neural network was framed with cascade forward propagation architecture. The network was trained with L.M. algorithm and BFGS Q.N. algorithm using 80% of the experimental data. The remaining data were used for testing and validation. The results demonstrated the following. The maximum TS of the AA1100 joint was found to be 72.40 MPa for the specimen welded with TRS of 1050rpm, WS of 60mm/min, SD of 18mm and PD of 6mm. The percentage error in prediction was found to be lower for

the ANN model trained with BFGS QN algorithm than the ANN model trained with L.M. algorithm. The maximum error in the predictions made by ANN model trained with BFGS QN algorithm was found to be -28% and ANN model trained with L.M. algorithm was found to be -41%.

REFERENCES:

[1] K.J. Colligan. 2010. The friction stir welding process: an overview, *Friction Stir Welding*, Woodhead Publishing, 15-41.

[2] R.S. Mishra and H. Sidhar. 2017. Friction stir welding, *Friction Stir Welding of 2XXX Aluminum Alloys Including Al-Li Alloys*, Butterworth-Heinemann, 1-13.

[3] L.E. Murr, G. Liu and J.C. McClure. 1997. Dynamic recrystallization in friction-stir welding of aluminium alloy 1100, *J. Materials Sci. Letters*, 16, 1801-1803. https://doi.org/10.1023/A:1018556332357.

[4] R. Padmanaban, V. Balusamy and V.R. Kishore. 2012. Effect of axial pressure and tool rotation speed on temperature distribution during dissimilar friction stir welding, *Advanced Materials Research*, 418-420, 1934-1938. https://doi.org/10.4028/www.scientific.net/AMR.418-420.1934.

[5] R. Padmanaban, V. Balusamy and K.N. Nouranga. 2015. Effect of process parameters on the tensile strength of friction stir welded dissimilar aluminum joints, *J. Engg. Sci. and Tech.*, 10, 790-801.

[6] R.V. Vignesh, R. Padmanaban, M. Arivarasu, K.P. Karthick, A.A. Sundar and J. Gokulachandran. 2016. Analysing the strength of friction stir spot welded joints of aluminium alloy by fuzzy logic, *IOP Conf. Series: Materials Sci. and Engg.*, 149.

[7] R. Zettler. 2010. Material deformation and joint formation in friction stir welding, *Friction Stir Welding*, Woodhead Publishing, 42-72.

[8] R. Zettler, T. Vugrin and M. Schmücker. 2010. Effects and defects of friction stir welds, *Friction Stir Welding*, Woodhead Publishing, 245-276.

[9] Y.S. Sato, Y. Kurihara, S.H.C. Park, H. Kokawa and N. Tsuji. 2004. Friction stir welding of ultrafine grained Al alloy 1100 produced by accumulative roll-bonding, Scripta Materialia, 50, 57-60. https://doi.org/10.1016/j.scriptamat.2003.09.037.

[10] M.S. Khorrami, M. Kazeminezhad and A.H. Kokabi. 2012. Mechanical properties of severely plastic deformed aluminum sheets joined by friction stir welding, *Materials Sci. and Engg.: A*, 543, 243-248. https://doi.org/10.1016/j.msea.2012.02.082.

[11] S.A. Hussein, A.S.M. Tahir and A.B. Hadzley. 2015. Characteristics of aluminum-to-steel joint made by friction stir welding: A review, *Materials Today Communications*, 5, 32-49. https://doi.org/10.1016/j.mtcomm.2015.09.004.

[12] I.A. Kartsonakis, D.A. Dragatogiannis, E.P. Koumoulos, A. Karantonis and C.A. Charitidis. 2016. Corrosion behaviour of dissimilar friction stir welded aluminium alloys reinforced with nanoadditives, *Materials & Design*, 102, 56-67. https://doi.org/10.1016/j.matdes.2016.04.027.

[13] M. Sajed. 2016. Parametric study of two-stage refilled friction stir spot welding: Part 1, *J. Mfg. Processes*, 24, 307-317. https://doi.org/10.1016/j.jmapro.2016.09.011.

[14] J.M. Zurada. 1992. *Introduction to Artificial Neural Systems*, 8, West St. Paul.

[15] B. Kröse, B. Krose, P. Van der Smagt and P. Smagt. 1993. *An Introduction to Neural Networks*.

[16] N.D. Ghetiya and K.M. Patel. 2014. Prediction of tensile strength in friction stir welded aluminium alloy using artificial neural network, *Proc. Tech.*, 14, 274-281. https://doi.org/10.1016/j.protcy.2014.08.036.

[17] A.K. Lakshminarayanan and V. Balasubramanian. 2009. Comparison of RSM with ANN in predicting tensile strength of friction stir welded AA7039 aluminium alloy joints, *Trans. of Nonferrous Metals Society of China*, 19, 9-18. https://doi.org/10.1016/S1003-6326(08)60221-6.

[18] H. Okuyucu, A. Kurt and E. Arcaklioglu. 2007. Artificial neural network application to the friction stir welding of aluminum plates, *Materials & Design*, 28, 78-84. https://doi.org/10.1016/j.matdes.2005.06.003.

[19] R.V. Vignesh and R. Padmanaban. 2017. Modelling tensile strength of friction stir welded aluminium alloy 1100 using fuzzy logic, *Proc. 11th Int. Conf. Intelligent Systems and Control*, 449-456.