Research Article

Sensitivity Analysis of the Artificial Neural Network Outputs in Friction Stir Lap Joining of Aluminum to Brass

Mohammad Hasan Shojaeefard,[1] Mostafa Akbari,[1] Mojtaba Tahani,[1,2] and Foad Farhani[3]

[1] *Department of Automotive Engineering, Iran University of Science and Technology, Tehran 16846-13114, Iran*
[2] *Faculty of New Sciences and Technologies (FNST), University of Tehran, Tehran, Iran*
[3] *Department of Mechanical Engineering, Iranian Research Organization for Science and Technology (IROST), Tehran, Iran*

Correspondence should be addressed to Mostafa Akbari; mr.mostafaakbari@yahoo.com

Al-Mg and CuZn34 alloys were lap joined using friction stir welding while the aluminum alloy sheet was placed on the CuZn34. In addition, the mechanical properties of each sample were characterized using shear tests. Scanning electron microscopy (SEM) and X-ray diffraction analysis were used to probe chemical compositions. An artificial neural network model was developed to simulate the correlation between the Friction Stir Lap Welding (FSLW) parameters and mechanical properties. Subsequently, a sensitivity analysis was performed to investigate the effect of each input parameter on the output in terms of magnitude and direction. Four methods, namely, the "PaD" method, the "Weights" method, the "Profile" method, and the "backward stepwise" method, which can give the relative contribution and/or the contribution profile of the input factors, were compared. The PaD method, giving the most complete results, was found to be the most useful, followed by the Profile method that gave the contribution profile of the input variables.

1. Introduction

Friction Stir Welding (FSW), a solid state welding invented by Welding Institute in 1991, has been employed to join wrought aluminum alloys, steel, Mg, Ti, and Cu, some of which have been classified as practically unweldable by traditional welding methods. In general, joining dissimilar materials by the conventional fusion welding is difficult, because normally poor weldability arises due to the different chemical, mechanical, and thermal properties of the welded materials as well as the formation of hard and brittle intermetallic compounds (IMCs) at the weld interface. This technique results in low distortion and high joint strength in comparison to other techniques and is capable of joining all aluminum alloys including dissimilar ones. It uses a nonconsumable rotating tool to generate frictional heat and deformation at the welding zone leading to the formation of a joint while the materials are still in the solid state. At the joint, the material is frictionally heated to temperatures in which it is easily plasticized. The current literature on FSW of aluminum and brass alloys mainly concerns research into friction stir butt joints. Recently, Esmaeili et al. [1] studied the feasibility of FSW of dissimilar butt joints of an aluminum plate to a brass one. They concluded that structure of the sound joint at the nugget zone of aluminum is made of a composite structure, consisting of intermetallic and brass particles, especially at the upper region of the weld cross-section.

ANNs are inspired by natural neural networks. The concept of ANN was developed before the emergence of computers. ANNs consist of a number of neurons which are the computational units of the ANN. Learning, generalization, and parallel processing are some advantages of the ANN. Learning means that the ANN is capable of adjusting the network parameters when new conditions are experienced. Generalization means that the ANN is able to attain a general rule using limited instances. ANNs are used for modeling and predicting different processes. Okuyucu et al. [2] modeled mechanical properties of butt welding by ANN. They derived the correlation between the FSW parameters of the Al plates and mechanical properties. Shojaeefard et al. [3] modeled

Table 1: Chemical composition and thickness of welded materials.

Materials	Chemical composition (in wt.%)										
	Al	Mg	Zn	Ti	Fe	Si	Mn	Cu	Ni	Pb	Sn
AA5083	Balance	4.27	0.02	0.026	0.31	0.1	0.61	0.04	—	—	—
CuZn34	0.02	0.05	Balance	—	0.05	0.25	0.05	64.2	0.03	0.1	0.1

and optimized mechanical properties of friction stir welding of AA7075/AA5083 butt joints using neural network and particle swarm algorithm. Asadi et al. [4] established a relation between the FSP parameters and grain size and hardness of nanocomposite using artificial neural network. They used back-propagation feed forward neural network to predict hardness and grain size.

ANNs are powerful methods in tasks involving pattern classification and forecasting. However, the major shortcoming of ANNs is the difficulty of interpreting the knowledge gained by model. In short, an ANN model functions like a "black-box" package, giving no clue on (1) how the answers or model outputs are obtained and (2) how the input parameters affect the output [5]. Since the credibility of an artificial intelligence program frequently depends on its ability to explain its conclusions, therefore for verification of such models, as well as accuracy measuring of ANN-based models with available data, a methodology should be adopted to extract the meaningful rule from the trained network, which is comparable with trends inferred from experiments. Several methods, commonly called sensitivity analysis, have been proposed to overcome this disadvantage. Sensitivity analysis is used to determine how much "sensitive" a model is to the changes in the value of the parameters of the model and to the changes in the structure of the model. The sensitivity coefficients describe the change in the system's outputs due to variations in the parameters that affect the system. A large sensitivity to a parameter suggests that the system's performance can drastically change with small variation in the parameter. Vice versa, a small sensitivity suggests little change in the performance. There are few methods to investigate the sensitivity of the ANN model.

In this study, four different methods that allow contribution analysis were used: (1) the "PaD" (for Partial Derivatives) method consists in calculating the partial derivatives of the output according to the input variables [6]; (2) the "Weights" method is a computation using the connection weights [7]; (3) the "Profile" method is a successive variation of one input variable while the others are kept constant at a fixed value [8]; (4) the "backward stepwise" method is an observation of the change in the error value when an elimination (backward) step of the input variables is operated [9].

2. Experimental Procedure

The materials selected in the current study were 5083 aluminum alloy and brass sheets with thicknesses of 2.5 mm. Their chemical compositions are given in Table 1.

A set of tensile shear tests is carried out in order to characterize the mechanical properties of the joint under various welding conditions. The width of each specimen was
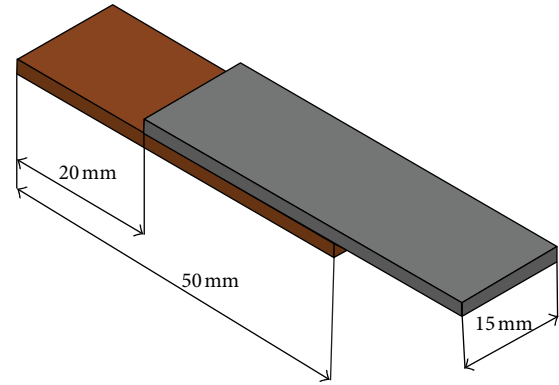

Figure 1: Dimensions of the prepared specimens for tensile shear test.

20 mm, and the shape of the test specimen was rectangular. The dimensions of the prepared specimens employed in the tensile shear tests are given in Figure 1. Scanning electron microscopy (SEM) and X-ray diffraction analysis were utilized to investigate the phases and microstructures.

3. Experimental Results and Discussion

Lap joints may be loaded primarily either by peel or in shear. In this study, the strength of the lap joints, which are nominally loaded in overlap shear, was examined. The failure loads of the joints in the tensile shear tests are given in Figure 2. The failure loads of all the joints are found to be lower than those of the base materials. They ranged from 2.5 ± 0.3 kN to 5.4 ± 0.4 kN. As illustrated in Figure 2, increasing the rotational speed of the tool at constant traverse speeds of 6.5, 12, 25, 40, and 60 mm/min resulted in an increase of the failure load to maximum, and then a decrease in failure load. The maximum load for the traverse and rotational speeds, which are the two important parameters affecting the FSW, was investigated at 6.5 mm/min and 1120 rpm, respectively. From the tensile shear tests results, fracture in the entire weld specimens occurred in the HAZ region except for the specimens welded with tool rotational speed of 1400 rpm.

The shear load of the joint was probably affected by two factors: (1) the amount of the brittle and hard intermetallic compounds and (2) the cold weld condition, which was performed at a low rotational and a medium or high translational speed. By increasing the rotational speed (or decreasing the welding speed) of the tool, the frictional heat generation increases, leading to more intense stirring and mixing of the materials (see Figure 3) and, consequently, increasing the size of the fine-grained zone (nugget). Therefore, the
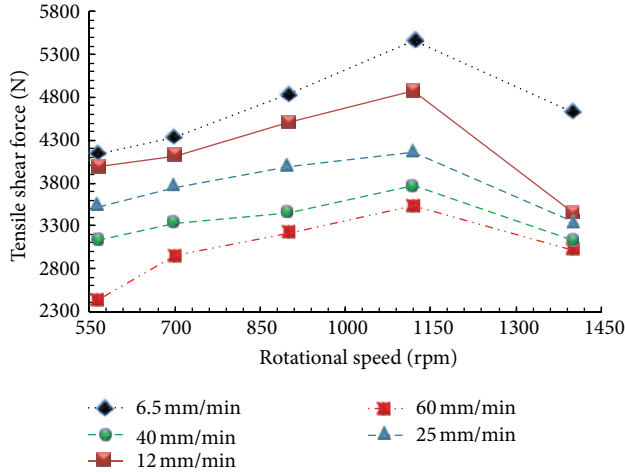
Figure 2: Effect of rotational and traverse speeds on tensile shear strength.

Table 2: Summary of different networks evaluated to yield the criteria of network performance.

No.	Activation function		Neurons in hidden layer	Training MRE	Testing MRE
	Hidden layer	Output layer			
1	logsig	logsig	4	1.6	2.8
2	logsig	logsig	5	1.2	3.7
3	**logsig**	**logsig**	**6**	**0.45**	**2.1**
4	logsig	purline	5	2	4.5
5	logsig	purline	6	1.63	3.52
6	logsig	purline	7	1.27	3.81
7	logsig	tansig	6	1.54	2.87
8	logsig	tansig	7	2.67	2.34
9	purlin	logsig	6	0.89	2.9
10	purlin	logsig	7	0.75	2.3
11	tansig	purline	6	1.35	2.74
12	tansig	purline	7	1.72	3.37

mechanical strength of the joint is improved with increasing the rotational speed or by reducing the traverse speed. A subsequent increase of rotational speed resulted in a large amount of intermetallic compounds (larger intermetallic compound region) at the interface between aluminum and copper (see Figure 3); so, the shear load was decreased. The reason for the increase in the amount of intermetallic compounds is that higher rotational speed gives rise to a higher temperature at the interface because the formation process of intermetallic compounds is thermally activated [10].

The X-ray diffraction (XRD) analysis results shown in Figure 4 indicate that the intermetallic compound region should mainly contain the intermetallic compounds. The XRD results show that the intermetallic compound region contains aluminum, brass, Al_2Cu, Al_4Cu_9, and CuZn phases. The presence of Al_2Cu and Al_4Cu_9 phases in the weld indicates that a certain amount of stir occurs at the interface and entraps some Cu into the stir zone.

4. Artificial Neural Network

In the current work, feed forward neural network with back-propagation algorithm was employed. In feed forward neural network, the output of each neuron is only connected to the neurons of the next layer. Inputs and outputs have been normalized in the range of 0-1.

In this study, the back-propagation is used with a network having an input layer with two neurons for each input factor and an output layer with one neuron. One of the most important tasks in ANN studies is to choose the optimal network architecture which is related to the activation function and the number of neurons in hidden layer. Generally, the trial-and-error approach is used. In this study, the optimal architecture of the network was obtained by trying different activation function and number of neurons shown briefly in

Table 2. The performance of each network was checked by MRE shown in the following equation:

$$
\mathrm{MRE} = \frac{1}{P}\sum_{1}^{P}\left(\frac{\left|\text{experimental value} - \text{predicted value}\right|}{\text{experimental value}} \times 100\right). \tag{1}
$$

The goal is to minimize MRE to obtain a network with the best generalization. MRE values were calculated for many different network models. Based on this analysis, the optimal architecture of the ANN was constructed as 2–6–1 NN and activation function in hidden layer and output layer both were "logsig."

4.1. ANN Results and Discussion. Twenty patterns of the experimental results were employed to train the ANN model, and 5 patterns were used for testing. Linear regression analysis was done to compute the correlation coefficient (R^2) between the experimental and predicted values. Figures 5 and 6 illustrate the comparison between the experimental and predicted data. According to these figures, at the training and testing stage, correlation coefficient of 0.9972 and 0.9715 was obtained, respectively. It is clear that the neural network prediction of FSLW parameters follows the experimental results very closely, and the developed ANN can accurately predict the tensile shear force.

5. Sensitivity Analysis

To identify the critical parameters and their degree of importance on the model outputs, a sensitivity analysis was done. The results show that the network output changes according to the inputs, providing information about the more sensitive
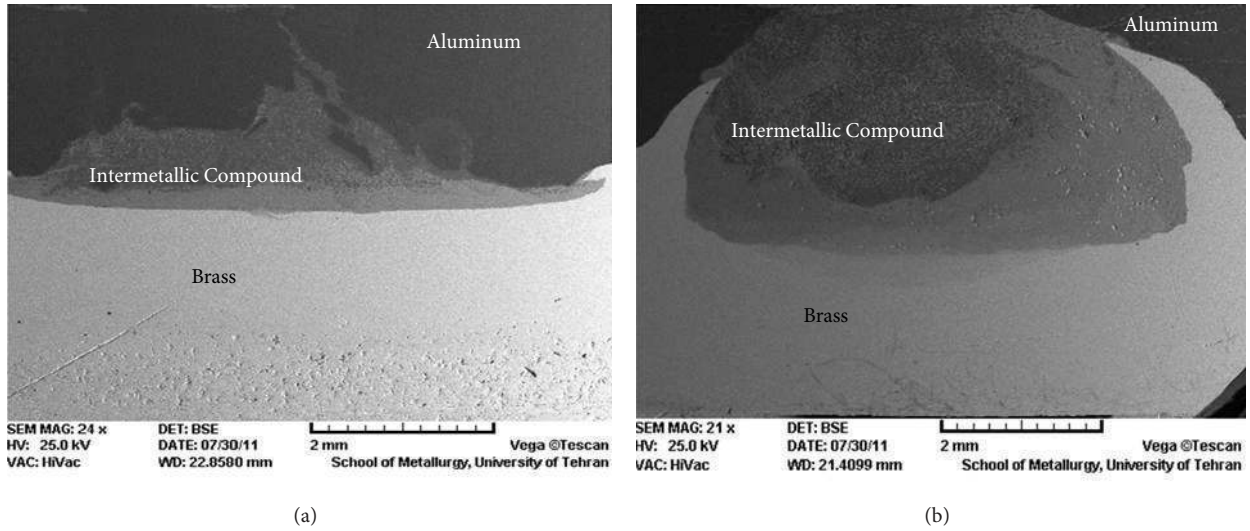
(a)

(b)

Figure 3: SEM images of the weld zone under various conditions: (a) 900 rpm, 6.5 mm/min; (b) 1400 rpm, 6.5 mm/min.
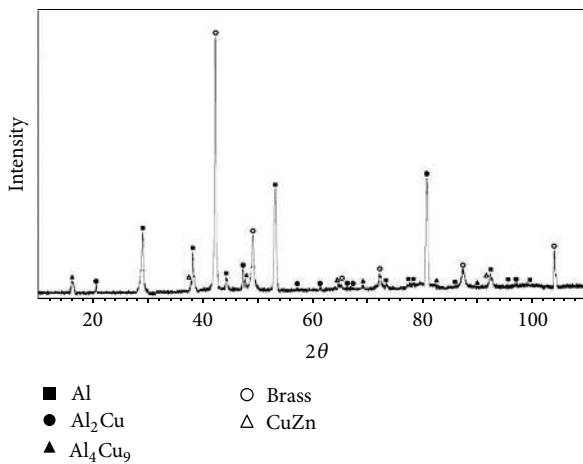

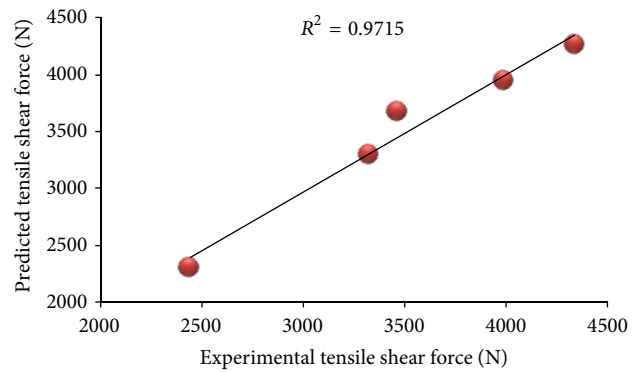Figure 4: X-ray diffraction analysis of the weld zone.

Figure 6: Predicted shear strength by ANN versus the experimental data and the regression line at the testing stage for tensile shear force.
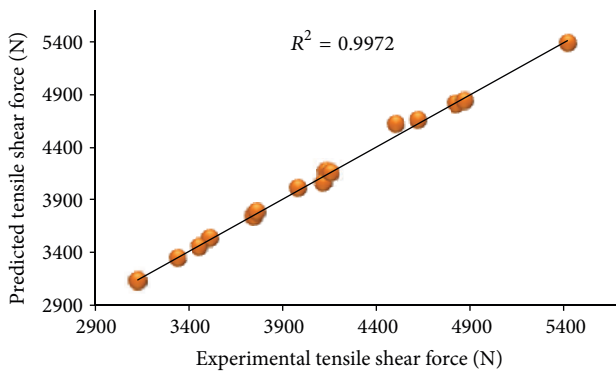
parameters, which should be measured more accurately. The results of such an analysis would also provide useful details about the "robustness" of the model parameters, leading to a better decision making process.

5.1. PaD Method. This method is employed so far to apply sensitivity analysis to ANN in some articles [11, 12]. PaD method is another method employed to calculate the effect of the input value on the output. It defines the sensitivity of the input by the following equation [13, 14]:

$$S_i = \frac{1}{N} \sum_p \frac{\partial o_k^p}{\partial x_i^p}, \tag{2}$$

where N is the total number of data variables, p is the pattern number, o_k^p is the output value of the ANN for the pattern p,

Figure 5: Predicted shear strength by ANN versus the experimental data and the regression line at the training stage for tensile shear force.

TABLE 3: Sensitivity analysis for tensile shear force.

Input parameter	S_i	SSD	Contribution
Rotational speed	0.259482	5.75	0.096308
Traverse speed	-0.98147	53.954	0.903692

and x_i^p is the input value from the pattern p. In a three-layered feed forward network:

$$o_k = f_2\left(\sum_j w_{kj}o_j\right), \tag{3}$$

$$o_j = f_1\left(\sum_i w_{ij}x_i\right), \tag{4}$$

where w_{kj} is the weight between the output neuron k and the hidden neuron j, w_{ij} is the weight between the input neuron i and the hidden neuron j, and o_j is the output of the hidden neuron j, f_1 and f_2 are the activation functions.

Applying the chain rule to (2), and combining (3) and (4) produces

$$
\begin{aligned}
S_i &= \frac{1}{N}\sum_P \frac{\partial o_k}{\partial o_j}\frac{\partial o_j}{\partial x_i} \\
&= \frac{1}{N}\sum_P f_2'\left(\sum_j w_{kj}o_j\right)\sum_j w_{kj}f_1'\left(\sum_i w_{ij}x_i\right)w_{ij}.
\end{aligned} \tag{5}
$$

If f_1 and f_2 are the sigmoid functions, then $f' = f(1-f)$. Thus, (6) will become

$$S_i = \frac{1}{N}\sum_P o_k(1-o_k)\sum_j w_{kj}o_j(1-o_j)w_{ij}. \tag{6}$$

The downside of this method is that the values of the sensitivities are affected by the output value. The sensitivities of the patterns with small and large output values will be influenced due to the terms o_k and $(1 - o_k)$ in (6) [13]. If the partial derivative is negative, it means that the output will decrease with increase in the input variable and vice-versa. The relative contribution of each input variable on a specific output can be determined by computing the sum of the squares of the partial derivatives

$$\mathrm{SSD}_i = \sum_P \left(\frac{\partial o_k^p}{\partial x_i^p}\right)^2. \tag{7}$$

The contribution of each input variable is given by

$$\text{contribution of } i\text{th variable} = \frac{\mathrm{SSD}_i}{\sum_i \mathrm{SSD}_i}. \tag{8}$$

The variable with the highest SSD has the most effect on the output. Based on this fact, the inputs may be ranked in order of their influence on the output. Table 3 shows sensitivity results for tensile shear force.

As shown in Table 3, the traverse speed is the most significant parameter on tensile shear force. S_i for the traverse speed is negative because by increasing traverse speed, tensile shear force decreases.
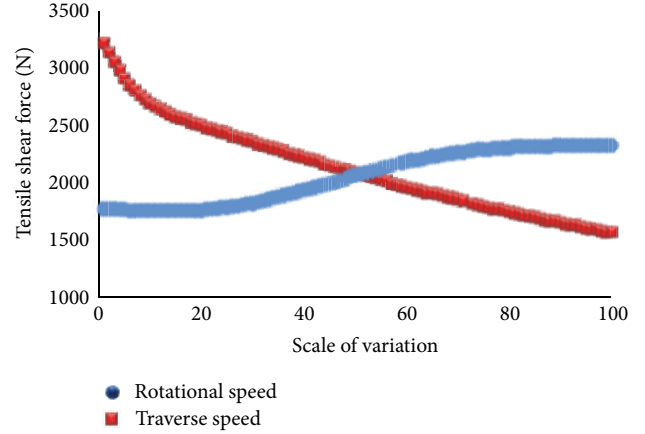

FIGURE 7: Contribution of variables used in the ANN model for tensile shear force, by the Profile algorithm with scale of 100.

5.2. Profile Method. This method was proposed by Lek et al. [8]. The general idea is to study each input variable successively when the others are then blocked at fixed values. The principle of this algorithm is to construct a fictitious matrix pertaining to the range of all input variables. In greater detail, each variable is divided into a certain number of equal intervals between its minimum and maximum values. The chosen number of intervals is called the scale. All variables except one are set initially (as many times as required for each scale) at their minimum values, then successively at their first quartile, median, third quartile, and maximum. The studies have shown that five values for each of the scale's points are obtained. These five values are reduced to the median value. Then, the profile of the output variable can be plotted for the scale's values of the variable considered. The same calculations can then be repeated for each of the other variables.

For each variable, a curve is then obtained. This gives a set of profiles of the variation of the dependent variable according to the increase of the input variables. Figures 7 and 8 represent the Profile method, respectively, for 100 and 20 scale intervals of input variables between their minimum and maximum. Each graph represents a different scale. It is interesting to notice the stability of the method whatever the scale is. In fact, the profiles of the different variables always have the same shape.

In Figures 7 and 8, traverse speed is the variable, which has the greatest effect on the output as it can be seen through the large range, that is, which is the most important. An increase of traverse speed leads to decrease of tensile shear force. As shown in Figures 7 and 8, the increase of rotational speed leads to increase tensile shear force.

5.3. Stepwise. One of the most important methods in sensitivity analysis is the backward stepwise method. Stepwise method consists of step by step adding or rejecting one input variable and examining the effect on the output results. Based on the changes in MSE, the input variables can be sorted according to their importance in several different ways. For
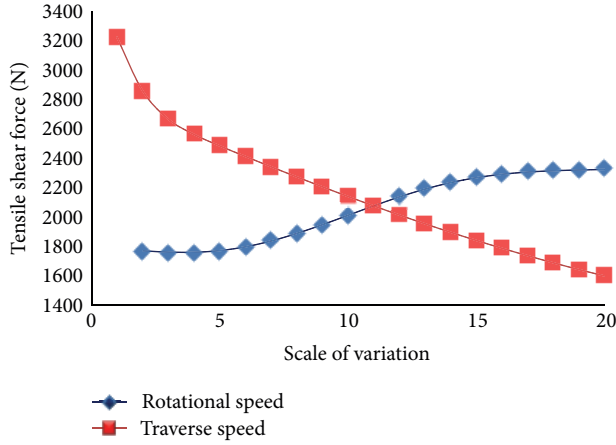
FIGURE 8: Contribution of variables used in the ANN model for tensile shear force, by the Profile algorithm with scale of 20.

TABLE 4: Sensitivity analysis for tensile shear force.

ANN model no.	Rotational speed	Traverse speed	Tensile shear force
1	Yes	—	0.626
2	—	Yes	0.373

instance, the largest value in RMS due to one input omission shows the most important input [13].

In the backward stepwise, used in this study, two models were generated, using only one of the variables as inputs. The omitted variable for which the resulting models gave the largest error is the most important one. The order of omission of the input variables is the order of the importance of their contribution [13].

Table 4 presents the backward stepwise result in which two models were generated using one of the available variables. The RMS of each model is given in the table. As shown, RMS in the model generated by rotational speeds is the largest one. Therefore, traverse speed is the most influential factor on tensile shear force.

5.4. *Weight Method*. Based on the weight magnitude, different equations have been proposed which share common characteristics: calculation of the product of the weights w_{ij} (connection weight between the input neuron i and the hidden neuron j) and v_{jk} (connection weight between the hidden neuron j and the output neuron k) for each of the hidden neurons of the network, which gives the sum of calculated products. The following, proposed by Garson [7], is representative of this type of analysis [15]:

$$Q_{ik} = \frac{\sum_{j=1}^{L} \left(\left(w_{ij} / \sum_{r=1}^{N} w_{rj} \right) v_{jk} \right)}{\sum_{i=1}^{N} \left(\sum_{j=1}^{L} \left(\left(w_{ij} / \sum_{r=1}^{N} w_{rj} \right) v_{jk} \right) \right)}, \qquad (9)$$

where $\sum_{r=1}^{N} w_{rj}$ denotes the sum of the connection weights between the input neurons N and the hidden neuron j. Q_{ik} represents the percentage of impact of the input variable x_i on the output y_k, in relation to the rest of the input variables,

TABLE 5: Relative importance of input parameters on grain size.

Relative importance (%)	
Traverse speed	Rotational speed
54.99	45.001

in such a way that the sum of this index must give a value of 100% for all of the input variables [15].

The relative importance of each input parameter on the outputs is shown in Table 5. As indicated in the tables, the relative importance of rotational speed on the tensile shear force is higher than traverse speed.

Comparison of the results of weight method with those of previous method indicates that results are not compatible. Therefore, it can be concluded that the weight method is not an effective method to calculate the relative importance of input parameters. Different experimental studies have demonstrated that analysis based on the weight method is not effective to determine the relative importance of input variables on the outputs [7, 15, 16].

6. Conclusion

In this research, Al-Mg and CuZn34 alloys were lap joined using friction stir welding. Tensile shear tests were used to investigate weld properties. In the aluminum near the Al/CuZn34 interface, an intermetallic compound area was observed. Al_2Cu, Al_4Cu_9, and CuZn compounds were produced at the intermetallic compound area. The results indicated that using a high rotational speed or a low traverse speed increased the size of the intermetallic compound area. Several sensitivity analysis methods have been used in this study. Comparing the PaD and Profile methods, the first is more coherent from a computation point of view. Comparison of the results of weight method with those of previous method indicates that results are not compatible. The classical stepwise methods gave exactly the same result from PAD and profile method but the contributions were not sufficiently expressed.

References

[1] A. Esmaeili, M. K. B. Givi, and H. R. Z. Rajani, "A metallurgical and mechanical study on dissimilar Friction Stir welding of aluminum 1050 to brass (CuZn30)," *Materials Science and Engineering A*, vol. 528, no. 22-23, pp. 7093–7102, 2011.

[2] H. Okuyucu, A. Kurt, and E. Arcaklioglu, "Artificial neural network application to the friction stir welding of aluminum plates," *Materials & Design*, vol. 28, no. 1, pp. 78–84, 2007.

[3] M. H. Shojaeefard, R. A. Behnagh, M. Akbari, M. K. B. Givi, and F. Farhani, "Modelling and Pareto optimization of mechanical properties of friction stir welded AA7075/AA5083 butt joints using neural network and particle swarm algorithm," *Materials & Design*, vol. 44, pp. 190–198, 2013.

[4] P. Asadi, M. K. B. Givi, A. Rastgoo, M. Akbari, V. Zakeri, and S. Rasouli, "Predicting the grain size and hardness of AZ91/SiC nanocomposite by artificial neural networks," *International Journal of Advanced Manufacturing Technology*, vol. 63, pp. 1095–1107, 2012.

[5] M. Lu, S. M. AbouRizk, and U. H. Hermann, "Sensitivity analysis of neural networks in spool fabrication productivity studies," *Journal of Computing in Civil Engineering*, vol. 15, no. 4, pp. 299–308, 2001.

[6] Y. Dimopoulos, P. Bourret, and S. Lek, "Use of some sensitivity criteria for choosing networks with good generalization ability," *Neural Processing Letters*, vol. 2, no. 6, pp. 1–4, 1995.

[7] G. D. Garson, "Interpreting neural-network connection weights," *AI Expert*, vol. 6, pp. 46–51, 1991.

[8] S. Lek, A. Belaud, P. Baran, I. Dimopoulos, and M. Delacoste, "Role of some environmental variables in trout abundance models using neural networks," *Aquatic Living Resources*, vol. 9, no. 1, pp. 23–29, 1996.

[9] G. R. Balls, D. Palmer-Brown, and G. E. Sanders, "Investigating microclimatic influences on ozone injury in clover (*Trifolium subterraneum*) using artificial neural networks," *New Phytologist*, vol. 132, no. 2, pp. 271–280, 1996.

[10] M. Akbari and R. Abdi Behnagh, "Dissimilar friction-stir lap joining of 5083 aluminum alloy to CuZn34 brass," *Metallurgical and Materials Transactions B*, vol. 43, pp. 1177–1186, 2012.

[11] W. Y. K. Chiang, D. Zhang, and L. Zhou, "Predicting and explaining patronage behavior toward web and traditional stores using neural networks: a comparative analysis with logistic regression," *Decision Support Systems*, vol. 41, no. 2, pp. 514–531, 2006.

[12] S. Dutta and J. P. Gupta, "PVT correlations for Indian crude using artificial neural networks," *Journal of Petroleum Science and Engineering*, vol. 72, no. 1-2, pp. 93–109, 2010.

[13] M. Gevrey, I. Dimopoulos, and S. Lek, "Review and comparison of methods to study the contribution of variables in artificial neural network models," *Ecological Modelling*, vol. 160, no. 3, pp. 249–264, 2003.

[14] T. Tchaban, M. J. Taylor, and J. P. Griffin, "Establishing impacts of the inputs in a feedforward neural network," *Neural Computing and Applications*, vol. 7, no. 4, pp. 309–317, 1998.

[15] J. J. Montaño and A. Palmer, "Numeric sensitivity analysis applied to feedforward neural networks," *Neural Computing and Applications*, vol. 12, pp. 119–125, 2003.

[16] W. Wang, P. Jones, and D. Partridge, "Assessing the impact of input features in a feedforward neural network," *Neural Computing and Applications*, vol. 9, no. 2, pp. 101–112, 2000.