Int J Adv Manuf Technol (2012) 63:1095–1107
DOI 10.1007/s00170-012-3972-z

ORIGINAL ARTICLE

Predicting the grain size and hardness of AZ91/SiC nanocomposite by artificial neural networks

P. Asadi · M. K. Besharati Givi · A. Rastgoo · M. Akbari · V. Zakeri · S. Rasouli

Received: 31 May 2011 / Accepted: 30 January 2012 / Published online: 17 February 2012

Abstract In the present study, SiC nanoparticles were added to as-cast AZ91 magnesium alloy through friction stir processing (FSP) and an AZ91/SiC surface nanocomposite layer was produced. A relation between the FSP parameters and grain size and hardness of nanocomposite using artificial neural network (ANN) was established. Experimental results showed that distribution of nanoparticles in the stirred zone (SZ) was not uniform and SZ was divided into two regions. In the ANN modeling, the inputs included traverse speed, rotational speed, and region types. Outputs were hardness and grain size. The model can be used to predict hardness and grain size as functions of rotational and traverse speeds and region types. To check the adequacy of the ANN model, the linear regression analyses were carried out to compute the correlation coefficients. The calculated results were in good agreement with experimental data. Additionally, a sensitivity analysis was conducted to determine the parametric impact on the model outputs.

Keywords FSP · ANN · Grain size · Hardness · Regression · Sensitivity analysis

P. Asadi
Department of Mechanical Engineering, Tehran Central Branch,
Islamic Azad University,
Tehran, Iran

M. K. B. Givi · A. Rastgoo · M. Akbari
School of Mechanical Engineering, College of Engineering,
University of Tehran,
Tehran, Iran

V. Zakeri · S. Rasouli
Department of Mechanical Engineering, Tabriz Branch,
Islamic Azad University,
Tabriz, Iran

P. Asadi (✉)
College of Engineering, Tehran Central Branch,
Islamic Azad University,
Ashrafi Esfahani Ave.,
Tehran, Iran
e-mail: parvizasadi@ut.ac.ir

P. Asadi
e-mail: p.asadi85@gmail.com

1 Introduction

Increasing demands for improved fuel-efficient vehicles to reduce energy consumption poses a challenge for the automotive industry to produce lighter automotive vehicles; the application of light-weight metals such as magnesium alloys has attracted great attention to replace the conventional metals [1–5]. However, poor formability and ductility and low hardness, strength, and wear resistance have limited use of magnesium alloys [4, 6].

In addition to the grain refinement as an effective method to increase the strength of metals [6], metal matrix composites have gained attractions to modify the structural and functional characteristics of the existing commercial materials through increasing hardness, strength, and wear resistance [7]. Recently, friction stir processing (FSP), invented by Mishra et al. [8], is used for microstructural modification of metallic materials. In this process, based on the principles of FSW, a rotating tool is inserted in a monolithic workpiece for localized microstructural modification to attain specific properties. Severe plastic deformation by the tool pin results in a recrystallized microstructure with fine, equiaxed grains [3, 8]. Although FSP has been basically considered as an advanced process for grain refinement, it is also very attractive to surface composites manufacturers [4, 9–11].

Artificial neural networks (ANNs) are inspired by natural neural networks. The concept of ANN was developed before the advent of computer. ANNs consist of a number of neurons which are the computational units of the ANN. Learning, generalizing, and parallel processing are among the advantages of the ANN. Learning means that the ANN is capable of adjusting the network parameters when new conditions rise. Generalizing means that the ANN is able to adhere to a general rule using limited instances. ANNs are used for modeling and predicting different processes. The field of neural networks is immense and interdisciplinary. Speed of calculation, capability of learning from examples and simplicity are the advantages of ANNs. These features enable ANNs to be used in materials science. Reducing the time and effort for the numerical studies using the approach for unknown mechanical properties within studied integer, desired hardness, and grain size values can be achieved. Okuyucu et al. [12] modeled mechanical properties of butt welding by ANN. They drew the correlation between the FSW parameters of the Al plates and mechanical properties. Acherjee et al. [13] developed an ANN-based model to predict the laser transmission weld quality in terms of lap-shear strength and weld-seam width. Dai et al. [14] established a knowledge base through numerous designs of experiments on beam elements, based on a validated finite element model of a reference vehicle body-in-white, and then applied ANN to extract the correlation between the beam element features and crash dynamic characteristics. They reported that beam element features can be predicted according to expected crash dynamic characteristics. Zhang et al. [15] used back-propagation neural network (BPNN) to predict the circumferential and longitudinal residual stress profiles in components shaped by hard turning processes. Sigh et al. [16] used BPNN predict the flank wear of high-speed steel drill bits for drilling holes on copper workpiece. Park and Kang [17] developed BPNN to predict the fatigue life of spot welds subjected to various geometric factors and loading conditions such as combined tension and shear loading and tensile-shear loading. Óscar Martín et al. [18] estimated the quality of a resistance spot welding joint of 304 austenitic stainless steel from its tensile shear load-bearing capacity where the quality levels were set by ultrasonic nondestructive testing. Pohlak et al. [19] used the response surface method (RSM) for optimization of the structure and manufacturing processes of large composite plastic parts. In RSM, the design surface is fitted to the response values using regression analysis.

In the present study, AZ91/SiC nanocomposite layer was produced. The aim was to adapt the ANN with the FSP parameters to derive the correlations between the FSP parameters and grain size and hardness of the nanocomposite layer. Sensitivity analysis was performed to examine the contribution of input variables to the outputs. Linear regression analyses were carried out to compute the correlation coefficient for the ANN model.

2 Experimental study

2.1 Experimental procedures

The material used was an AZ91 as-cast magnesium alloy with an average grain size of 150 μm and a composition of (in weight percent): Al, 9.1; Zn, 0.68; Mn, 0.21; Si, 0.085; Cu, 0.0097; Ni, 0.001; Fe, 0.0029; and Mg, bal. Commercially available SiC particles with average diameter of 30 nm and 99.98% purity were used as reinforcements. The reinforcing particles were filled into a groove of 0.8× 1.2 mm machined on the AZ91 as-cast plate with the thickness of 5 mm. Two hot working steel tools were used. The pinless tool was employed to cover the top of the grooves after filling with SiC particles to prevent scattering during FSP. The second tool had a square pin with dimensions of 3.54×3.54 mm and a length of 2.5 mm. Both tools' shoulders were 15 mm in diameter. The tool rotational speed was varied from 710 to 1,400 rpm and the traverse speed was varied from 12.5 to 63 mm/min. The FSP tool was rotated in clockwise directions with the tilt angle of 3°.

The transverse sections of the specimens were prepared by standard metallographic techniques and etched with a solution of 5 mL acetic acid, 6 g picric acid, 10 mL water, 100 mL ethanol, 5 mL HCl, and 7 mL nitric acid for 1–2 s. Microstructural observations of the SZ were carried out by optical microscopy. Microhardness of the specimens was measured at 1 mm distance from the top surface in steps of 0.2 mm using a Vickers Microhardness Testing machine by applying a load of 200 g for 15 s.

2.2 Macrostructural and microstructural results

Figure 1a shows the macrograph of SZ for the specimen produced by one-pass FSP. As seen in the figure, distribution of SiC nanoparticles in the AZ91 matrix is not uniform and the SZ is divided into two regions. The bright region is rich in SiC particles, while the dark region is poor in SiC particles. Such non-uniform distribution of the SiC particles makes non-uniform SZ microstructure. Figure 1b and c show the microstructure of bright and dark regions. According to the Zener–Holoman parameter, increasing the volume fraction of SiC particles decreases the grain size [9]. Dense distribution of nanoparticles in the bright region restricts the grain growth due to the pinning effect, while the grains grow more freely in the dark region. Grain size in the bright region is about 2 μm, while it is about 8 μm in the dark region.

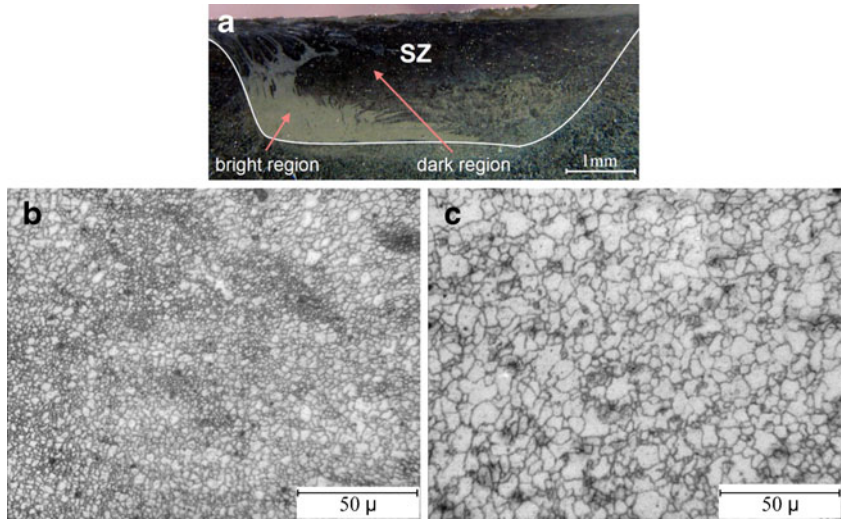

Fig. 1 **a** Macro-image of the SZ for the specimen produced by one-pass FSP; **b** and **c** microstructures of the bright and dark regions, respectively. The rotational and traverse speeds were 710 rpm and 40 mm/min, respectively

2.3 Microhardness

Microhardness profile of the specimen produced by FSP at the rotational and traverse speeds of 900 rpm and 63 mm/min is shown in Fig. 2. The average microhardness value of the as-cast AZ91 was 63 HV. Due to the high hardness of SiC particles and very fine grains, the hardness of SZ increases. As it can be seen in Fig. 2, the hardness profile is divided into two parts due to the non-uniform distribution of the SiC particles discussed earlier (bright and dark regions). Dense distribution of the SiC particles accompanied by very fine grains (~1.8 μm) severely increased the hardness of the bright region up to 115 HV, while the hardness of the dark region, with less distribution of the particles and the grain size of ~7 μm, reached 90 HV.

The experimental results are summarized in Table 1. As shown in the table, the process variations are the rotational speed, traverse speed, and region types, and the outputs are grain size and hardness.

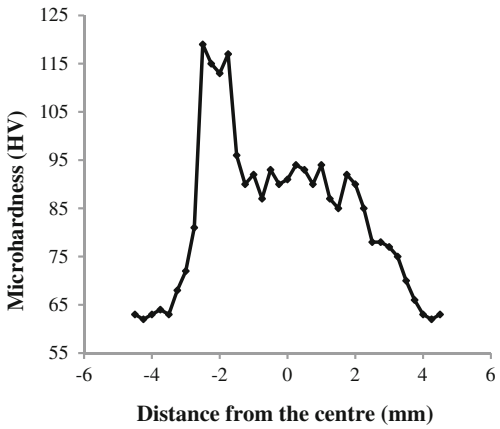

Fig. 2 Microhardness profile of the specimen produced by one FSP pass. The rotational and traverse speeds were 900 rpm and 63 mm/min, respectively

3 Artificial neural network

Artificial neural network is an approach inspired by brain structure and tries to simulate the brain-processing capabilities. Neurons are the constituent parts of neural network. Neural networks consist of different parts. These parts are the inputs, outputs, weight vector, add function, and transfer function. ANN comprise of three layers which are input, hidden, and output layers. The input layer consists of all input factors. Information from the input layer is processed in the hidden layer. Outputs of the hidden layer(s) are also computed in the output layer and result in output vector. Detailed description of the functions of hidden and output layers is presented in [20]. The transfer function of the hidden and output layers was “logsig”.

To model this process, feed-forward neural network with back-propagation algorithm was used. Feed forward is a neural network in which the output of each neuron is only connected to the neurons of the next layer. Feed forward neural networks usually have one or more hidden layers and have an output layer. Inputs and outputs have been normalized in the range of 0–1. Hidden layer with a nonlinear transfer function authorizes the network to learn linear or nonlinear relations between outputs and inputs.

In this study, the back-propagation is used with a network having an input layer with three neurons for each input factor (rotational speed, traverse speed, and region types) and an output layer with two neurons (grain size and hardness). The performance of neural network depends on the number of hidden layer and the number of neurons in the hidden layer(s), and also activation function in output and hidden layers. Thus, several combinations are tried out to

Table 1 Experimental results with input and output parameters

No.	Inputs			Outputs	
	Region type	Rotational speed (rpm)	Traverse speed (mm/min)	Grain size (μm)	Hardness (HV)
1	Bright region	710	12.5	2.2	111.9
2	Bright region	900	12.5	2.4	111
3	Bright region	1,120	12.5	2.4	110.8
4	Bright region	1,400	12.5	2.7	110.1
5	Bright region	710	25	2	113.6
6	Bright region	900	25	2.1	112.7
7	Bright region	1,120	25	2.2	112
8	Bright region	1,400	25	2.4	111.5
9	Bright region	710	40	1.9	115
10	Bright region	900	40	2	113.9
11	Bright region	1,120	40	2	114
12	Bright region	1,400	40	2.1	112.7
13	Bright region	710	63	1.7	115.9
14	Bright region	900	63	1.8	115.2
15	Bright region	1,120	63	1.8	115
16	Bright region	1,400	63	2	114.1
17	Dark region	710	12.5	10.1	86.1
18	Dark region	900	12.5	11.4	84.3
19	Dark region	1,120	12.5	12.5	82.4
20	Dark region	1,400	12.5	13.9	80.7
21	Dark region	710	25	8.2	88.6
22	Dark region	900	25	9	87
23	Dark region	1,120	25	10.2	84.9
24	Dark region	1,400	25	11.3	83.3
25	Dark region	710	40	7	90.2
26	Dark region	900	40	8	88.5
27	Dark region	1,120	40	9.3	86.6
28	Dark region	1,400	40	10.5	85.2
29	Dark region	710	63	5.7	92.5
30	Dark region	900	63	6.9	90.5
31	Dark region	1,120	63	8.1	88.8
32	Dark region	1,400	63	9	86.9

choose an optimal combination. Several combinations of networks were tested and mean prediction error (%) was calculated for each combination as follows:

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

To choose transfer function for hidden and output layers, and the number of neurons in the hidden layer several combinations, as shown briefly in Table 2, were examined.

As shown in Table 3, the network has the lowest MRE when the transfer function of the hidden and output layers was “logsig” and number of neurons in the hidden layer was six (3-6-2). Schematic of this network is shown in Fig. 3.

The learning-algorithm used was back propagation, one of the most popular learning algorithms. The learning process comprises two passes through different layers of the network. The first one is the forward pass. In the forward pass, the input vector is applied to the network and its effect propagates through the network layer by layer. In this pass, the synaptic weights are assumed to be constant. The second one is the backward pass. In this pass, unlike the forward pass, parameters of the network are not constant and are modified. First the error, the difference between the output of the network and the desired output, is propagated in the backward pass to update the synaptic weights. The weights are continuously updated

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

No.	Activation function		Neurons in hidden layer	Hardness training MRE (%)	Grain size training MRE (%)
	Hidden layer	Output layer			
1	Logsig	Purlin	5	2.11	0.73
2	Logsig	Purlin	6	3.23	0.98
3	Logsig	Logsig	5	1.6024	0.51
4	Logsig	Logsig	6	0.64	0.18
5	Logsig	Tansig	5	1.98	0.64
6	Logsig	Tansig	6	0.82	0.36
7	Tansig	Purlin	5	2.34	1.43
8	Tansig	Purlin	6	0.91	0.59
9	Tansig	Logsig	5	4.67	1.34
10	Tansig	Logsig	6	1.12	0.82
11	Tansig	Tansig	5	1.98	0.42
12	Tansig	Tansig	6	1.23	0.34

every time. The input vector is presented to the network and the process continues till the output of the network comes closer to the desired output [21].

In the present work, 32 patterns were obtained from the experiments by changing the process parameters. Inputs for the ANN (process parameters) were the rotational speed, traverse speed, and region types and the outputs were hardness and grain size.

MATLAB 7.0 was applied in all the stages of developed model including training and testing of the network. The

Table 3 Comparison of actual and predicted grain size for AZ91/SiC nanocomposite produced by FSP at training stage

No.	Region type	Rotational speed (rpm)	Traverse speed (mm/min)	Grain size (μm)		Error	Error%
				Actual	Predicted		
1	Bright region	900	12.5	2.4	2.383127	0.016873	0.703045
2	Bright region	1,120	12.5	2.4	2.386662	0.013338	0.555751
3	Bright region	1,400	12.5	2.7	2.706454	−0.00645	0.239035
4	Bright region	710	25	2	1.994711	0.005289	0.264436
5	Bright region	1,120	25	2.2	2.240124	−0.04012	1.823797
6	Bright region	1,400	25	2.4	2.381186	0.018814	0.783919
7	Bright region	900	40	2	1.986905	0.013095	0.654751
8	Bright region	1,120	40	2	2.014674	−0.01467	0.733705
9	Bright region	1,400	40	2.1	2.107285	−0.00729	0.346919
10	Bright region	710	63	1.7	1.716411	−0.01641	0.965336
11	Bright region	900	63	1.8	1.77156	0.02844	1.580003
12	Bright region	1,400	63	2	2.011424	−0.01142	0.57118
13	Dark region	710	12.5	10.1	10.10113	−0.00113	0.011221
14	Dark region	1,120	12.5	12.5	12.50162	−0.00162	0.012953
15	Dark region	1,400	12.5	13.9	13.89678	0.00322	0.023164
16	Dark region	710	25	8.2	8.130183	0.069817	0.851425
17	Dark region	900	25	9	9.086958	−0.08696	0.966197
18	Dark region	1,400	25	11.3	11.30056	−0.00056	0.004958
19	Dark region	710	40	7	6.966374	0.033626	0.480367
20	Dark region	900	40	8	8.136007	−0.13601	1.70009
21	Dark region	1,120	40	9.3	9.186939	0.113061	1.21571
22	Dark region	1,400	40	10.5	10.49726	0.002744	0.026129
23	Dark region	710	63	5.7	5.718718	−0.01872	0.328382
24	Dark region	1,120	63	8.1	8.033218	0.066782	0.824473
25	Dark region	1,400	63	9	9.043531	−0.04353	0.483682

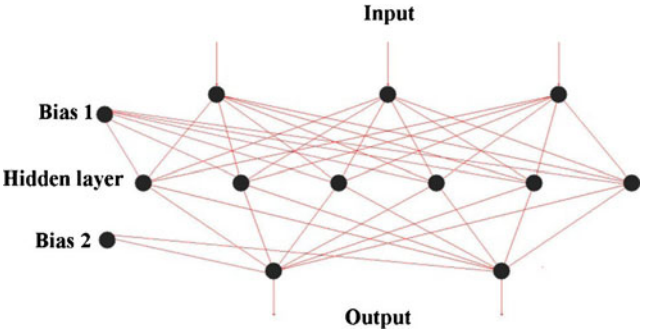

Fig. 3 Schimatic of three layer neural network in the present work

model was trained with "Levenberg–Marquardt optimization" learning algorithm. The Levenberg–Marquardt algorithm is based on approaching second-order training speeds without having the computation of Hessian matrix [22].

After training, the network was tasted with seven patterns. Errors occurring at the learning and testing stages are the root-mean squared (RMS), absolute fraction of variance (R^2), and mean prediction error (Eq. 1).

$$\mathrm{RMS} = \left((1/p)\sum \left|t_j - o_j\right|^2\right)^{1/2} \tag{2}$$

$$R^2 = 1 - \left(\frac{\sum_j \left(t_j - o_j\right)^2}{\sum_j \left(o_j\right)^2}\right) \tag{3}$$

4 Evaluation of results and discussion

Twenty-five patterns of the experimental results were used for training the AAN model and seven patterns were not applied to the model and were used for testing. Tables 3 and 4 show the comparison of actual and predicted values for grain size and hardness. It is clear that the neural network

Table 4 Comparison of actual and predicted hardness for AZ91/SiC nanocomposite produced by FSP at training stage

No.	Region type	Rotational speed (rpm)	Traverse speed (mm/min)	Hardness (HV)		Error	Error%
				Actual	Predicted		
1	Bright region	900	12.5	111	110.662	0.337987	0.304492
2	Bright region	1,120	12.5	110.8	110.6557	0.144336	0.130267
3	Bright region	1,400	12.5	110.1	110.6562	−0.55622	0.505192
4	Bright region	710	25	113.6	113.7531	−0.15312	0.134785
5	Bright region	1,120	25	112	111.8999	0.100141	0.089412
6	Bright region	1,400	25	111.5	111.2893	0.21067	0.188942
7	Bright region	900	40	113.9	113.9467	−0.04672	0.041021
8	Bright region	1,120	40	114	113.7145	0.285473	0.250415
9	Bright region	1,400	40	112.7	113.057	−0.35705	0.316813
10	Bright region	710	63	115.9	115.5514	0.348571	0.300752
11	Bright region	900	63	115.2	115.3825	−0.1825	0.158418
12	Bright region	1,400	63	114.1	114.1437	−0.04371	0.03831
13	Dark region	710	12.5	86.1	86.20448	−0.10448	0.121348
14	Dark region	1,120	12.5	82.4	82.59061	−0.19061	0.231321
15	Dark region	1,400	12.5	80.7	80.54938	0.15062	0.186642
16	Dark region	710	25	88.6	88.67993	−0.07993	0.090213
17	Dark region	900	25	87	86.68799	0.31201	0.358632
18	Dark region	1,400	25	83.3	83.47525	−0.17525	0.210386
19	Dark region	710	40	90.2	90.2957	−0.0957	0.106103
20	Dark region	900	40	88.5	88.48411	0.015895	0.01796
21	Dark region	1,120	40	86.6	86.73067	−0.13067	0.150892
22	Dark region	1,400	40	85.2	84.9473	0.252704	0.296601
23	Dark region	710	63	92.5	92.40487	0.09513	0.102844
24	Dark region	1,120	63	88.8	88.72771	0.072294	0.081412
25	Dark region	1,400	63	86.9	87.01735	−0.11735	0.135042

prediction of FSP parameters follow the experimental results very closely.

Figure 4a and b show the grain size and hardness for the experimental and predicted data at the training stage. As it is clear from the figures, the predicted data by the ANN are in good agreement with the experimental results. Figure 5a and b show the error histograms of the experimental and predicted data for the grain size and hardness. As can be seen, the maximum errors for training of the grain size and hardness are ~1.8 and ~0.5%, respectively.

Tables 5 and 6 show the comparison of actual and predicted values for grain size and hardness at the training stage. The developed ANN can accurately predict the grain size and hardness. Figure 6a and b show the grain size and hardness for the experimental and predicted data at the testing stage. It is clear that the predicted data by the ANN are in good agreement with the experimental results. Therefore, it is possible to predict the grain size and hardness of AZ91/SiC nanocomposite layer produced by FSP without performing the experiment.

4.1 Regression

Linear regression analyses were carried out to compute the R^2 between the experimental and predicted values. Correlation coefficients of 0.9999 and 0.9997 were obtained for the grain size and hardness at the training stage (Fig. 7) and 0.9994 and 0.9987 at the testing stage (Fig. 8). It is clear that the correlation coefficients for both grain size and hardness are close to unity indicating the good accuracy of the developed model. Thus, this ANN model can be used to predict the grain size and hardness of AZ91/SiC nanocomposite layer with adequate accuracy.

4.2 Interpreting the model

To identify the critical parameters and their order of importance on the model outputs, the analysis based on the magnitude of weights and sensitivity analysis were conducted. These analyses show the network output change according to the inputs and provide information about the parameter which should be measured more accurately.

Since the end of the 1980s, different methods have been proposed for interpreting what has been learned by a feed forward neural network composed of input neurons, hidden neurons, and output neurons. These interpretative methods can be divided into two methods: analysis based on the weights magnitude and sensitivity analysis [23].

Analysis based on the weights magnitude are cooperated with those procedures that are solely based on the values stored in the static matrix of weights to extract the relative impact of each input variable on the outputs. Based on the weights magnitude, different equations have been proposed that are characterized by the calculation of the results of 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 hidden neuron, acquiring the sum of the calculated results.

The basis of the sensitivity analysis methods is measurement of the effect occurred in the output y_k due to a change made in the input x_i. Achieving the Jacobian matrix through the calculation of the partial derivatives of the output y_k as a function of the input x_i, $\left(\frac{\partial y_k}{\partial x_i}\right)$, can be considered the analytical translation of sensitivity analysis. It can be so useful to determine the importance of the input, as the calculation of the partial derivatives depends the instant slope of the underlying function between each pair of input x_i *and output* y_k *[23].*

4.2.1 Formulation

Hidden and output layers with "log-sigmoid" transfer function were used to predict hardness and grain size. The log-sigmoid transfer function was:

$$F(x) = \frac{1}{1 + e^{-x}} \tag{4}$$

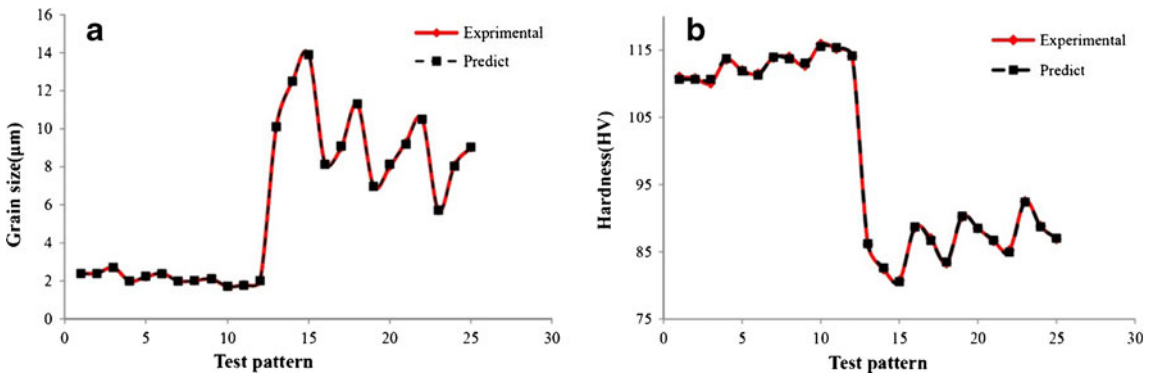

Fig. 4 Experimental and predicted **a** grain size and **b** hardness at the training stage of ANN model

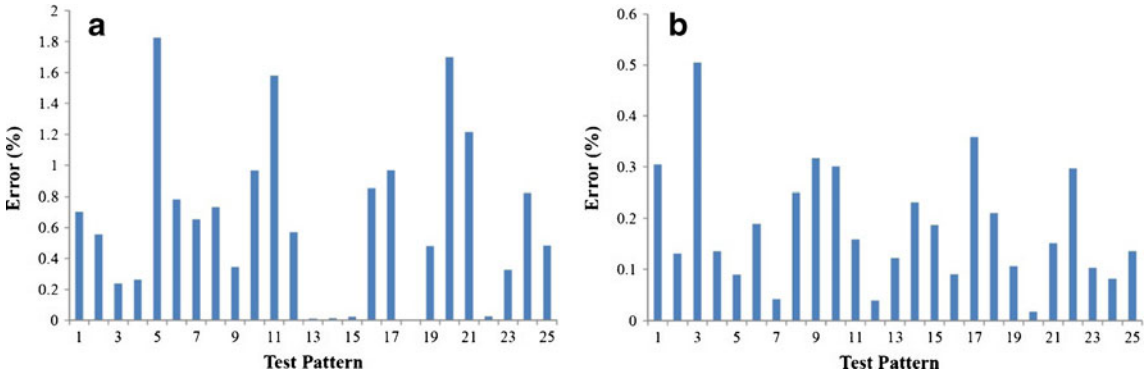

Fig. 5 Error histogram of the experimental and predicted data for the **a** grain size and **b** hardness; at the training stage of the ANN model

where, x is the weighted sum of the input. To determine the grain size and hardness of the nanocomposite layer produced by FSP, Eqs. 6 and 7 were derived from ANN.

$$\text{Grain size} = \frac{1}{1+e^{-(-495.4445f_1+495.7971f_2+5.1233f_3+0.2828f_4+495.4909f_5-8.0142f_6+0.86883)}} \tag{5}$$

$$\text{Hardness} = \frac{1}{1+e^{-(1026.8846f_1-1027.7274f_2-0.35732f_3+0.0019929f_4-1027.6086f_5+2.1075f_6+1.565)}} \tag{6}$$

where, f_1 to f_6 are the weights, calculated as follows:

$$f_1 = \frac{1}{1+e^{-(-85.8039T+3.7552R+53.6996A+48.1241)}} \tag{7}$$

$$f_2 = \frac{1}{1+e^{-(-32.2399T+3.7486R-1.4443A+17.5183)}} \tag{8}$$

$$f_3 = \frac{1}{1+e^{(-(-96.7354T+2.9658R-76.6959A+90.1754))}} \tag{9}$$

$$f_4 = \frac{1}{1+e^{-(-12.0599T+41.1793R+4.3362A-36.1781)}} \tag{10}$$

$$f_5 = \frac{1}{1+e^{-(32.3637T-3.7488R+36.0293A-52.174)}} \tag{11}$$

$$f_6 = \frac{1}{1+e^{-(0.6663T-2.5005R-10.155A+9.1859)}} \tag{12}$$

where, T, R, and A are the traverse speed, rotational speed, and region type, respectively. It should be noticed that in

Table 5 Comparison of actual and predicted grain size for AZ91/SiC nanocomposite produced by FSP at testing stage

No.	Region type	Rotational speed (rpm)	Traverse speed (mm/min)	Grain size (μm)		Error	Error%
				Actual	Predicted		
1	Bright region	710	12.5	2.2	2.380853	−0.18085	8.220577
2	Bright region	900	25	2.1	2.135706	−0.03571	1.700273
3	Bright region	710	40	1.9	2.016804	−0.1168	6.147604
4	Bright region	1,120	63	1.8	1.849962	−0.04996	2.775668
5	Dark region	900	12.5	11.4	11.32775	0.072247	0.63375
6	Dark region	1,120	25	10.2	9.886165	0.313835	3.076812
7	Dark region	900	63	6.9	6.920494	−0.02049	0.297013

Table 6 Comparison of actual and predicted hardness for AZ91/SiC nanocomposite produced by FSP at testing stage

NO.	Region type	Rotational speed (rpm)	Traverse speed (mm/min)	Hardness (HV)		Error	Error%
				Actual	Predicted		
1	Bright region	710	12.5	111.9	110.6684	1.231555	1.100585
2	Bright region	900	25	112.7	112.7161	−0.01609	0.014275
3	Bright region	710	40	115	113.7893	1.210681	1.052766
4	Bright region	1,120	63	115	115.0441	−0.04413	0.038377
5	Dark region	900	12.5	84.3	84.41421	−0.11421	0.135483
6	Dark region	1,120	25	84.9	84.95368	−0.05368	0.063225
7	Dark region	900	63	90.5	90.51442	−0.01442	0.015935

Eqs. 7–12, traverse speed should be divided into 70 and traverse speed should be divided into 1,500 and region type (A) is equal to 1 in dark region and 0 in bright region. Grain size should be multiplied to 15 and hardness should be multiplied at 120.

The coefficients from the neurons of hidden layer were multiplied by the output of the input layers. The last constant values in Eqs. 7–12 indicate the bias term associated with hidden layer. The weights of the neural network were extracted and made into mathematical formulation. Six neurons were used in the hidden layer, providing six equations for hidden layer. The weights of hidden layer were multiplied by the input layer and added to the bias and then were passed to the "log-sigmoid" transfer function (f_1–f_6). The output of the logsig transfer function is multiplied by the weight of the output layer to predict grain size and hardness (Eqs. 5 and 6).

4.2.2 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 [24], is representative of this type of analysis [23]:

$$Q_{ik} = \frac{\sum_{j=1}^{L}\left(\frac{w_{ij}}{\sum_{r=1}^{N} w_{rj}} v_{jk}\right)}{\sum_{i=1}^{N}\left(\sum_{j=1}^{L}\left(\frac{w_{ij}}{\sum_{r=1}^{N} w_{rj}} v_{jk}\right)\right)} \tag{13}$$

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, in such a way that the sum of this index must give a value of 100% for all of the input variables [23].

To come up with precise weights, two ANN models were generated with both having single output. The relative importance of each input parameter on the outputs is shown in Tables 7 and 8. To follow the procedure to calculate the relative importance, see Appendix. As indicated in the tables, the relative importance of the input parameters on the grain size is different from that

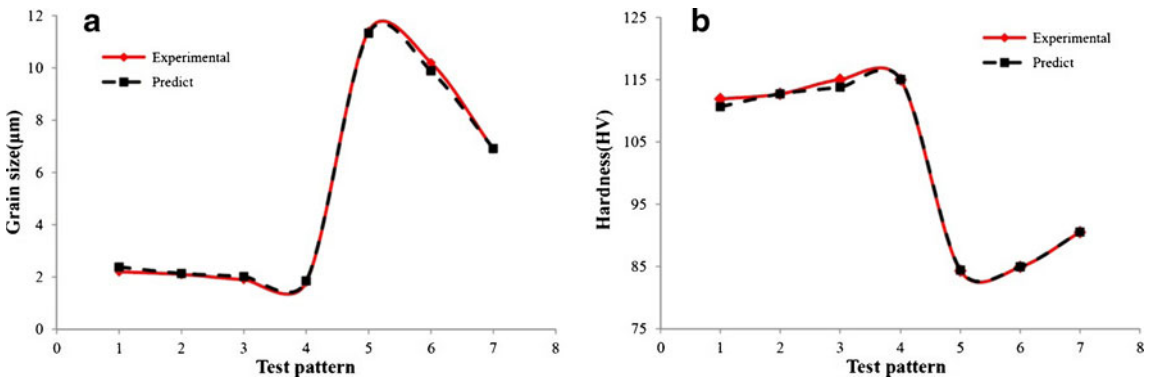

Fig. 6 Experimental and predicted **a** grain size and **b** hardness at the testing stage of ANN model

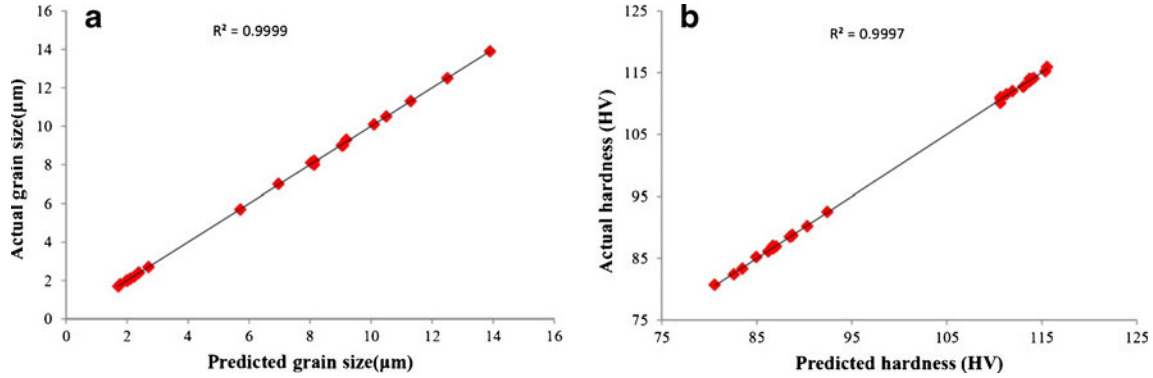

Fig. 7 **a** Predicted grain size by ANN versus the experimental grain size (*grain size regression*) at the training stage and **b** predicted hardness by ANN versus the experimental hardness (*hardness regression*) at the training stage

on the hardness. However, the impact of rotational speed on the grain size and hardness is higher than that of the other parameters.

4.2.3 Pad method

In some works, Pad method is employed to apply sensitivity analysis to ANN [25, 26]. The Pad method calculates the Pad of a specific ANN output with respect to each input at each vector of the inputs. In the present method, the sensitivity of input can be defined by the following equation [25, 27, 28].

$$S_i = \frac{1}{N}\sum_p \frac{\partial o_k^p}{\partial x_i^p} \tag{14}$$

where, N is the total number of data variables, p the pattern number, o_k^p, the output value of the ANN for the pattern p, and x_i^p the input value from the pattern p. In this case,

$$o_k = f_2\left(\sum_j w_{jk} o_j\right) \tag{15}$$

and

$$o_j = f_1\left(\sum_i w_{ij} x_i\right) \tag{16}$$

where, w_{jk} 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, o_j is the output of the hidden neuron j, and f_1 and f_2 are the activation functions.

Applying the chain rule to Eq. 14, and combining Eqs. 15 and 16 results in:

$$S_I = \frac{1}{N}\sum_p \frac{\partial o_k}{\partial o_j}\frac{\partial o_j}{\partial x_i} = \frac{1}{N}\sum_p f_2'\left(\sum_j w_{jk} o_j\right)\sum_j w_{jk} f_1'\left(\sum_i w_{ij} x_i\right) w_{ij} \tag{17}$$

Because f_1 and f_2 are the sigmoid functions, then $f'=f(1-f)$. Thus, Eq. 18 will become:

$$S_i = \frac{1}{N}\sum_p o_k(1-o_k)\sum_j w_{jk} o_j(1-o_j) w_{ij} \tag{18}$$

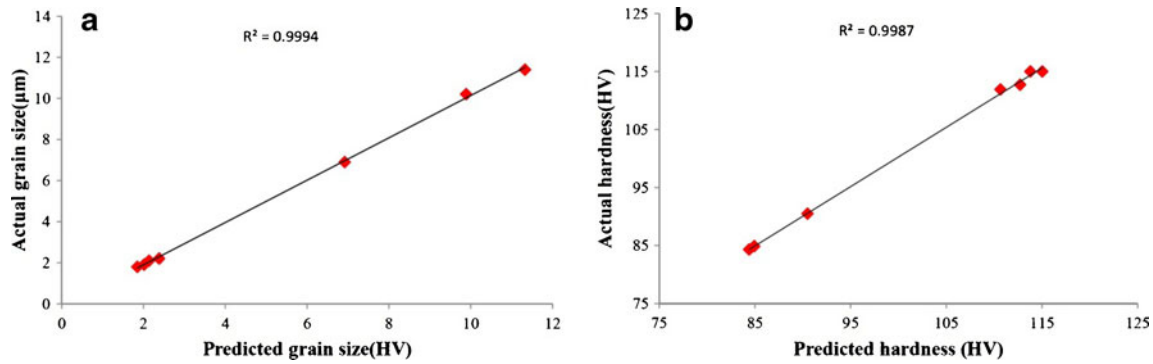

Fig. 8 **a** Predicted grain size by ANN versus the experimental grain size (*grain size regression*) at the testing stage and **b** predicted hardness by ANN versus the experimental hardness (*hardness regression*) at the testing stage

Table 7 Relative importance of input parameters on grain size

Relative importance (%)		
Traverse speed	Rotational speed	Region type
34.01885	41.68061	24.30054

The relative contribution of each input variable on a specific output can be determined by calculating the sum of the squares of partial derivatives (SSD) [25].

$$\mathrm{SSD}_i = \sum_{P=1}^{N} \left(\frac{\partial y_k}{\partial x_i} \right)_P^2 \tag{19}$$

where $\left(\frac{\partial y_k}{\partial x_i}\right)$ is the PAD for Pth observation. The contribution of each input variable is given by:

$$\text{Contribution of } P\text{th variable} = \frac{\mathrm{SSD}_i}{\sum_i \mathrm{SSD}_i} \times 100\% \tag{20}$$

The variable having the highest SSD affects the output most. On this basis, the inputs can be ranked in the order of their contribution on the outputs. Tables 9 and 10 show sensitivity results for hardness and grain size.

Sensitivity analysis shows that the order of importance of input parameters on the grain size and hardness is the same. As cited by Hall–Petch relationship, hardness depends on grain size [9]. Therefore, each parameter has an impact on the grain size effects on the hardness as well.

Comparison of the results of weight method with those of the backward stepwise indicates that results are not compatible. It is clear from Figs. 1, 2 and Table 1 that the impact of region type on the grain size and hardness of the nanocomposite is much higher. 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 [23, 24, 29].

Table 8 Relative importance of input parameters on hardness

Relative importance (%)		
Travers speed	Rotational speed	Region type
23.5608	46.14119	30.29802

Table 9 Sensitivity analysis for grain size

No.	Variables	SSD	Contribution (%)	Rank
1	Region type	17.67933	53.22412	1
2	Rotational speed	5.912697	17.80034	3
3	Traverse speed	9.624735	28.97554	2

5 Conclusions

This study was an attempt to show the possibility of using of neural networks for calculation of the grain size and hardness of AZ91/SiC nanocomposite produced by FSP. Results showed that the networks can be used as an alternative way in these systems. The validation of the model was evaluated quantitatively, using the MRE (%). The best architecture, obtained for the present work, was 3-6-2, using back-propagation rate. The correlation coefficients (R^2) for training and test patterns for hardness and grain size were close to unity, indicating excellent agreement between experimental data and predicted values. Expressions for grain size and hardness were developed from the neural network model. The model can be used to predict hardness and grain size as functions of rotational and traverse speeds and region types. Sensitivity analysis conducted to determine the parametric impact on the model outputs showed that the region types was the most effective parameter on the grain size and hardness of produced nanocomposite layer.

Table 10 Sensitivity analysis for hardness

No.	Variables	SSD	Contribution (%)	Rank
1	Region type	3.851351	65.80826	1
2	Rotational speed	0.602763	10.29944	3
3	Traverse speed	1.398269	23.8923	2

Acknowledgment Special thanks go to the Iranian Nanotechnology Initiative for partial financial support.

Appendix

The computation process is as follows:

1. For each hidden neuron, the absolute value of the output layer weight multiply by the absolute value of the hidden layer weight. This process implies for each input. Then, the following tables are obtained:

Weight of various layer

Hidden layer neuron no.	Traverse speed	Rotational speed	Region type	Grain size
1	−55.3178	−95.5384	3.3025	−2.3188
2	−4.1112	4.3033	3.9569	1.3847
3	1.4662	−9.1465	25.9532	1.6541
4	6.6795	25.6189	1.3834	0.17792
5	−113.018	3.3554	5.0194	23.9326
6	−2.4343	6.6717	4.376	1.9267

Weight of various layer

Hidden layer neuron no.	Traverse speed	Rotational speed	Region type	Hardness
1	39.9372	−61.4851	−3.8629	0.092301
2	7.0352	56.0055	15.0031	−0.17464
3	−3.4478	−3.9941	−3.4619	−0.76524
4	0.19705	19.4762	23.4907	−3.3633
5	−1.8155	23.5069	−14.1	−0.99639
6	−7.0957	−0.6522	−4.542	−0.64619

Grain size calculation.

Hidden layer neuron no.	Traverse speed	Rotational speed	Region type	Sum
1	P_{11}=128.2709	P_{12}=221.5344	P_{13}=7.6578	357.4632
2	P_{21}=5.6927	P_{22}=5.9587	P_{23}=5.4791	17.1306
3	P_{31}=2.4251	P_{31}=15.1292	P_{33}=42.9291	60.4836
4	P_{41}=1.1884	P_{42}=4.5581	P_{43}=0.2461	5.9926
5	P_{51}=2704.819	P_{52}=80.3034	P_{53}=120.1273	2,905.25
6	P_{61}=4.6901	P_{62}=12.8543	P_{63}=8.4312	25.97577

Hardness calculation

Hidden layer neuron no.	Traverse speed	Rotational speed	Region type	Sum
1	P_{11}=3.6862	P_{12}=5.6751	P_{13}=0.3565	9.7179
2	P_{21}=1.2286	P_{22}=9.7808	P_{23}=2.6201	13.6295
3	P_{31}=2.6383	P_{32}=3.0564	P_{33}=2.6491	8.3440
4	P_{41}=0.6627	P_{42}=65.5043	P_{43}=79.0062	145.1733
5	P_{51}=1.8089	P_{52}=23.42204	P_{53}=14.0491	39.28009
6	P_{61}=4.58517	P_{62}=0.4214	P_{63}=2.9349	7.9416

2. To obtain Q_{ij} for each hidden neuron, P_{ij} was divided into the sum; for all the input variables. For example for 1, $Q_{11} = P_{11}/(P_{11} + P_{12} + P_{13})$.

3. For each input neuron, assume S_j to be the sum of Q_{ij}. For example, $S_1 = Q_{11} + Q_{21} + Q_{31} + Q_{41} + Q_{51} + Q_{61}$. Then, the following tables are obtained.

Grain size calculation

Hidden layer neuron no.	Traverse speed	Rotational speed	Region type
1	Q_{11}=0.3588	Q_{12}=0.6197	Q_{13}=0.021423
2	Q_{21}=0.33231	Q_{22}=0.347843	Q_{23}=0.319843
3	Q_{31}=0.040097	Q_{32}=0.250137	Q_{33}=0.709765
4	Q_{41}=0.198312	Q_{42}=0.760616	Q_{43}=0.041073
5	Q_{51}=0.931011	Q_{52}=0.02764	Q_{53}=0.041348
6	Q_{61}=0.180559	Q_{62}=0.49486	Q_{63}=0.324581
Sum	S_1=2.041131	S_2=2.50083	S_3=1.458032

Hardness calculation

Hidden layer neuron no.	Traverse speed	Rotational speed	Region type
1	Q_{11}=0.379324	Q_{12}=0.583986	Q_{13}=0.03669
2	Q_{21}=0.090144	Q_{22}=0.717616	Q_{23}=0.192239
3	Q_{31}=0.316202	Q_{32}=0.366303	Q_{33}=0.317495
4	Q_{41}=0.004565	Q_{42}=0.451214	Q_{43}=0.54422
5	Q_{51}=0.046052	Q_{52}=0.596283	Q_{53}=0.357665
6	Q_{61}=0.57736	Q_{62}=0.053068	Q_{63}=0.369572
Sum	S_1=1.413648	S_2=2.768471	S_3=1.817881

4. By dividing each S_j into the sum, the relative importance of each input parameter can be calculated. For example, for the first input parameter (traverse speed), the relative importance is equal to $(S_1 \times 100)/(S_1 + S_2 + S_3)$.

Relative importance of input parameters on grain size.

Relative importance (%)		
Traverse speed	Rotational speed	Region type
34.01885	41.68061	24.30054

Relative importance of input parameters on hardness.

Relative importance (%)		
Travers speed	Rotational speed	Region type
23.5608	46.14119	30.29802

References

1. Liao J, Yamamoto N, Nakata K (2009) Effect of dispersed intermetallic particles on microstructural evolution in the friction stir weld of a fine-grained magnesium alloy. Metall Mater Trans A 40:2212–2219
2. Hassan SF, Gupta M (2006) Effect of particulate size of Al2O3 reinforcement on microstructure and mechanical behaviour of solidification processed elemental Mg. J Alloy Comp 419:84–90
3. Asadi P, Besharati Givi MK, Faraji G (2010) Producing ultrafine-grained AZ91 from As-Cast AZ91 by FSP. Mater Manuf Process 25:1219–1226
4. Faraji G, Asadi P (2011) Characterization of AZ91/alumina nanocomposite produced by FSP. Mater Sci Eng A 528:2431–2440
5. Huang W, Hou B, Pang Y, Zhou Z (2006) Fretting wear behavior of AZ91D and AM60B magnesium alloys. Wear 260:1173–1178
6. Kang SH, Lee YS, Lee JH (2008) Effect of grain refinement of magnesium alloy AZ31 by severe plastic deformation on material characteristics. J Mater Process Technol 201:436–440
7. Lai MO, Lu L, Laing W (2004) Formation of magnesium nanocomposite via mechanical milling. Compos Struct 66:301–304
8. Mishra RS, Ma ZY (2005) Friction stir welding and processing. Mater Sci Eng R 50:1–78
9. Barmouz M, Asadi P, BesharatiGivi MK, Taherishargh M (2011) Investigation of mechanical properties of Cu/SiC composite fabricated by FSP: effect of SiC particles' size and volume fraction. Mater Sci Eng A 528:1740–1749
10. Woo W, Choo H (2008) Microstructure, texture and residual stress in a friction-stir-processed AZ31B magnesium alloy. Acta Mater 56:1701–1711
11. Barmouz M, Besharati Givi MK, Seyfi J (2010) On the role of processing parameters in producing Cu/SiC metal matrix composites via friction stir processing: investigating microstructure, microhardness, wear and tensile behavior. Mater Charact 62:108–117
12. Okuyucu H, Kurt A, Arcaklioglu E (2007) Artificial neural network application to the friction stir welding of aluminum plates. Mater Des 28:78–84
13. Acherjee B, Mondal S, Tudu B, Misra D (2011) Application of artificial neural network for predicting weld quality in laser transmission welding of thermoplastics. Appl Soft Comput 11:2548–2555
14. Dai Y, Duan C (2009) Beam element modelling of vehicle body-in-white applying artificial neural network. Appl Math Model 33:2808–2817
15. Zhang JY, Liang SY, Zhang G, Yen D (2006) Modeling of residual stress profile in finish hard turning. Mater Manuf Process 21:39–45
16. Singh AK, Panda SS, Pal SK, Chakraborty D (2006) Predicting drill wear using an artificial neural network. Int J Adv Manuf Technol 28:456–462
17. Park JM, Kang HT (2007) Prediction of fatigue life for spot welds using back-propagation neural networks. Mater Des 28:2577–2584
18. Martín Ó, De Tiedra P, López M, San-Juan M, García C, Martín F, Blanco Y (2009) Quality prediction of resistance spot welding joints of 304 austenitic stainless steel. Mater Des 30:68–77
19. Pohlak M, Majak J, Karjust K, Küttner R (2010) Multicriteria optimization of large composite parts. Compos Struct 92:2146–2152
20. Ates H (2007) Prediction of gas metal arc welding parameters based on artificial neural networks. Mater Des 28:2015–2023
21. Muthukrishnan N, Paulo Davim J (2009) Optimization of machining parameters of Al/SiC-MMC with ANOVA and ANN analysis. J Mater Process Technol 209:225–232
22. Demuth H, Beale M (1998) Neural network toolbox for use with MATLAB, Users Guide. Version 3. The MathWorks, Inc., Massachusetts
23. Montano JJ, Palmer A (2003) Numeric sensitivity analysis applied to feed forward neural networks. Neural Comput Appl 12:119–125
24. G. D. Garson (1991) Interpreting neural-network connection weights. AI Expert 47–51
25. Dutta S, Gupta JP (2010) PVT correlations for Indian crude using artificial neural networks. J Petrol Sci Eng 72:93–109
26. Chiang WK, Zhang D, Zhou L (2006) Predicting and explaining patronage behavior toward web and traditional stores using neural networks: a comparative analysis with logistic regression. Decis Support Syst 41:514–531
27. Gevrey M, Dimopoulos I, Lek S (2003) Review and comparison of methods to study the contribution of variables in artificial neural network models. Ecol Model 160:249–264
28. Tchaban T, Taylor MJ, Griffin JP (1998) Establishing impacts of the inputs in a feed forward neural network. Neural Comput Appl 7:309–317
29. Wang W, Jones P, Partridge D (2000) Assessing the impact of input features in a feed forward neural network. Neural Comput Appl 9:101–112