International Journal of Modern Manufacturing Technologies
ISSN 2067–3604, Vol. III, No. 2 / 2011

SOME POSSIBILITIES OF USING DOE IN SETTING ANN PARAMETERS: AN APPLICATION IN MODELING OF ABRASIVE WATERJET CUTTING

Milos Madic [1] & Miroslav Radovanovic [1]

[1] Faculty of Mechanical Engineering, University of Nis, Serbia

Corresponding author: Miroslav Radovanovic, mirado@masfak.ni.ac.rs

Abstract: Artificial neural networks (ANNs) have been successfully applied for solving a wide variety of problems. However, determining of ANN architectural and training parameter values still remains a difficult task. This paper is concerned with the usage of design of experiment (DOE) method in order to determine parameter settings of multilayer feedforward (MLFF) ANN trained with backpropagation (BP) algorithm with momentum for modeling purposes. In this paper, a case study of abrasive waterjet (AWJ) cutting was used to find ANN parameters settings in order to develop high performance prediction model. For evaluating the predictive performance of ANN models, the combined mean absolute percentage error ($MAPE_{comb}$) is used as performance criterion. The ANN model was used for prediction of traverse rate of the separation cut based on material thickness, water pressure, and abrasive rate as input cutting parameters. The selected 3-8-1 ANN model showed high prediction accuracy with correlation coefficient of 0.999. Finally, the developed ANN model is given in the form of mathematical equation. The results indicated that ANN model is able to learn the relationships between AWJ process parameters. Also, in searching for ANN model of high performance, DOE method can be efficiently used in setting ANN architectural and training parameters.

Key words: Artificial neural networks, design of experiments, abrasive waterjet cutting, modeling.

1. INTRODUCTION

Artificial neural networks (ANNs) are massive parallel systems made up of simple processing units (neurons) that are linked with weighted connections where the knowledge possessed by the networks is held. By tuning a set of weights in the training process, ANN learns the relationship between process parameters. In past 20 years a number of researchers successfuly applied ANN modeling for various processes. The use of ANNs in modeling of machining processes has been extensive and multifaceted. The literature reveals that ANNs find many applications to predict surface finish, tool wear, cutting forces, etc. through different machining processes, but very little effort is reported on the use of ANNs in modeling of AWJ cutting process. The applications of ANNs in modeling abrasive waterjet (AWJ) cutting is reported in literature (Srinivasu & Ramesh Babu, 2008; Lu et al., 2005; Çaydaş & Hasçalik, 2008; Ergür et al., 2006; Parikh & Lam, 2009). ANN architectural and training parameters are summarized in table 1.

The learning ability of nonlinear relationship without going deep into the mathematical complexity of the relationship between process parameters makes ANN an attractive choice for process modeling. However, there is a disadvantage of ANN for process modeling that is the lack of guidance in selecting ANN architectural and training parameters. The most typical method is a repetitive trial-and-error method, where a large number of different models are examined and compared to one another. This method is very time-consuming and is mainly based on the past experience and some literature guidance, and above all there is no prior guarantee that the model will perform well for the problem at hand. Nevertheless, various techniques and methods to help developers select the optimum ANN architecture and

Table 1. ANN architectural and training parameters used for modeling of AWJ

Author, year	ANN architecture	Data preprocessing	Data for modelling	Training algorithm	Learning rate	Momentum	Training epochs	Performance function
Srinivasu & Ramesh Babu, 2008	4-[1÷50]-1	Yes	Experimental 81 data	backpropagation	varied in the range [0,1]	varied in the range [0,1]	Not stated	Not stated
Lu et al., 2005	3-12-1	Yes	Experimental 300 data	Levenberg-Marquardt	0.1	adaptive	26	MSE
Çaydaş & Hasçalik, 2008	13-22-1	Not stated	Experimental 27 data	backpropagation	0.9	0.2	331	MSE
Ergür et al., 2006	2-11-1 2-12-1	Not stated	Simulated 42 data	Levenberg-Marquardt	Not stated	Not stated	Not stated	MAPE
Parikh & Lam, 2009	3-4-4	Not stated	Experimental 78 data	backpropagation	0.1	0.4	500	PMAE

training parameters have been proposed. Application of Taguchi's design of experiment (DOE) method for optimization of ANN parameters were reported in literature (Sukthomya & Tannock, 2005; Lin & Tseng, 2000; Kim & Yum, 2004; Packianather et al., 2000; Yang & Lee, 1999; Khaw et al., 1995). In reference (Laosiritaworn & Chotchaithanakorn, 2009) the 2^4 full factorial experimental design was used to find optimal settings of ANN trained to model ferromagnetic material data. Factors that often found to have the significant effect on ANN model performances are number of neurons in hidden layers, learning rate and momentum.

This paper is concerned with the design of the multilayer feedforward (MLFF) neural network trained by the backpropagation (BP) algorithm with momentum, which is one of the most popular ANNs. In order to develop ANN model of high performance, the parameters related to ANN training (learning rate and momentum) as well as the network architecture (number of hidden neurons) must be considered carefully. In this paper, design of experiments (DOE) was used to find the optimum setting of ANNs' parameters in order to develop accurate ANN prediction model regarding $MAPE_{comb}$ as proposed performance criterion. A case study of AWJ cutting modeling was used to illustrate the proposed method.

2. ANN MODELING OF AWJ CUTTING PROCESS

In this study was cutting data of KMT Waterjet Systems Inc. was used. Material of workpiece is carbon steel A36HR, cutting head is Autoline with orifice size of 0.23mm and focusing tube of 0.76mm, grit garnet is w/80. By changing the levels of three process parameters, namely material thickness **s**, water pressure **p**, and abrasive rate **q**, values for traverse rate of the separation cut (**v**) were obtained. The mathematical equation of this process is not known, and therefore methodology of ANN modeling with the help of DOE was used to model the process. The generated 216 data constituted the input/output data set for training and testing the ANN model (table 2).

Generally, the ANN model development process include the following basic steps:

- Data pre-processing,
- ANN model design,
- ANN model training, and
- ANN model testing.

2.1 Data pre-processing

Data pre-processing usually speeds up the learning process and prevents stacking of training process. Pre-processing can be in the form of data scaling (normalization) and transformation. Prior to the ANN training, the input and output data set was scaled. The dataset was scaled to range between -1 and 1. The scaled value (x_i) for each raw input and output data (d_i) was calculated as:

$$x_{scaled} = 2 \cdot \frac{(x_i - x_{min})}{(x_{max} - x_{min})} - 1 \qquad (1)$$

where x_{max} and x_{min} are the maximum and minimum values of the raw data. The scaled input and output data were then partitioned into two subsets consisting training data set, 75% (162 data), and the test data set, 25% (54 data).

Table 2. Cutting data

s [mm]	p [MPa]	v [mm/min]						s [mm]	p [MPa]	v [mm/min]					
		q [g/min]								q [g/min]					
		250	300	350	400	450	500			250	300	350	400	450	500
3	350	904	972	1033	1089	1140	1189	**5**	350	503	540	574	605	634	661
	360	952	1023	1087	1146	1201	1252		360	529	569	604	637	667	696
	370	1001	1076	1143	1205	1263	1316		370	556	598	635	670	702	731
	380	1051	1130	1201	1266	1326	1382		380	584	628	667	703	737	768
	390	1103	1185	1259	1327	1390	1449		390	613	659	700	738	773	805
	400	1155	1241	1319	1390	1456	1518		400	642	690	733	773	809	844
7	350	341	367	390	411	430	449	**10**	350	226	243	259	273	286	298
	360	359	386	410	433	453	472		360	238	256	272	287	301	313
	370	378	406	432	455	477	497		370	251	269	286	302	316	330
	380	397	426	453	478	500	522		380	263	283	301	317	332	346
	390	416	447	475	501	525	547		390	276	297	315	332	348	363
	400	436	468	498	525	550	573		400	289	311	330	348	365	380
15	350	142	153	162	171	179	187	**20**	350	102	110	117	123	129	134
	360	150	161	171	180	189	197		360	107	115	123	129	136	141
	370	157	169	180	189	198	207		370	113	121	129	136	142	149
	380	165	178	189	199	208	217		380	119	128	135	143	150	156
	390	173	186	198	209	218	228		390	124	134	142	150	157	164
	400	181	195	207	218	229	239		400	130	140	149	157	164	171

2.2 ANN model design

In this study, MLFF neural network models were designed with software package MATLAB. The ANN models consisted of three layers: input layer; one hidden layer; and output layer. There were three input neurons representing material thickness, water pressure and abrasive rate and one neuron in the output layer representing traverse rate of separation cut. Number of hidden neurons is data dependent and is determined with the help of DOE. The upper limit of number of hidden neurons was determined knowing that the number of weights (sum of the product between the numbers of neurons in each layer) doesn't exceed one fifth of the number of training samples. It is easy to calculate that for three inputs and one output, the maximum allowed number of hidden neurons is 8. For all ANN models linear transfer function "purelin"[1] and tangent sigmoid transfer function "tansig"[1] were used in the output and hidden layer, respectively.

2.3 ANN model training

The ANN models were trained with BP algorithm with momentum, "traingdm"[1], which is essentially a gradient steepest descent method. A complete description of the BP algorithm can be found in numerous sources, including (Haykin, 1999; Tarassenko, 1998; Bishop, 1995; Zurada, 1992).

The most important training parameters are: learning rate, momentum, training epochs and initial weights. Learning rate and momentum control the speed and efficiency of the training process. Learning rate is the rate, at which the network adjusts its weights during training, hence primarily affects the training speed. A high learning coefficient provides faster convergence but training process becomes unstable and divergent oscillations may occur. With a small learning rate, training time is increased, but the probability of reaching the global minimum is increased. Momentum is a training parameter used to reduce training time of the BP algorithm and to enhance the stability of the training. A high momentum reduces the risk of the network being stuck in local minimum, but it increases the risk of skipping over the solution. Using a small value for momentum will lead to prolonged training. Reviewing the literature (Haykin, 1999; Tarassenko, 1998; Bishop, 1995; Zurada, 1992) it can be seen that typical ranges of these two training parameters is between 0 and 1.

The training epochs of the training cycle is the number of times the training data has been presented to the network. The BP algorithm guarantees that total error in the training set will continue to decrease as the number of training epochs increases. But, on the other side, excessive training results in phenomenon called over-training or memorization. Error on training set can result in near zero while error on test set dramatically increases. The criterion for the training process terminatation is when maximum number of 10000 epochs was reached since the training process is finished in relative short period. BP network is sensitive to initial values of weights. The initialization of the weights has a great impact on the network training time and generalization performance. Too small initial weights will the training time and difficulties in converging to an optimal solution may occur. If initial weights are too large the network may get unstable weights. The initial connection weights are specified according to Nguyen-Widrow initialization method (Nguyen & Widrow, 1990).

[1] Notation in the software package MATLAB

2.3 ANN model testing

Once the ANN models are developed, the performance of the trained ANN models should be checked. The test data set was used to test the generalization ability of the developed ANN models. In order to assess the prediction capability of the ANN models on both training and test data sets, the combined mean absolute percentage error ($MAPE_{comb}$) measure is proposed:

$$MAPE_{comb} = \frac{N_{tr} \cdot MAPE_{tr} + N_{ts} \cdot MAPE_{ts}}{N_{tr} + N_{ts}} \qquad (2)$$

where N_{tr} and N_{ts} are the number of data points for training and testing respectively, and $MAPE_{tr}$ and $MAPE_{ts}$ are mean absolute percentage error on training and testing data sets respectively.

3. APPLICATION OF DOE IN SETTING ANN PARAMETERS

DOE is a scientific approach of planning and conducting experiments to model, analyze and interpret data so that valid conclusions can be drawn efficiently (Montgomery, 2001). In this paper, the factors that affect the ANN models prediction accuracy, which is measured by $MAPE_{comb}$, were analyzed. Three factors were selected: learning rate, momentum and number of hidden neurons. The levels of the factors are selected so that stable training process is assured. These factors and their levels are given in table 3.

Table 3. Factors and their levels

Factor	Low level	High level
Learning rate	0.01	0.1
Momentum	0.5	0.9
Number of hidden neurons	2	8

The DOE was achieved using the 2^3 full factorial design. The design was twice replicated since some noise is entered in ANN design by setting initial weights. As a result, a total of 16 runs were conducted. The experimental design and results are given in table 4.

Table 4. Coded design matrix and results

Factor	Learning rate	Momentum	Number of hidden neurons	$MAPE_{comb}$	
Trial No.	A	B	C		
1	-1	-1	-1	12.32	11.445
2	+1	-1	-1	11.623	4.917
3	-1	+1	-1	12.608	12.104
4	+1	+1	-1	5.606	11.075
5	-1	-1	+1	22.141	8.314
6	+1	-1	+1	6.729	6.238
7	-1	+1	+1	10.514	7.865
8	+1	+1	+1	6.912	4.72

4. ANALYSIS OF RESULTS

In order to identify the significant main/interaction effects which affects the mean response, normal probability plot (NPP) and Pareto plot were constructed. Inactive main and interaction effects tend to fall roughly along a straight line whereas active effects tend to appear as extreme points falling off each end of the straight line. These active effects are judged to be statistically significant. The Pareto plot allows one to detect the factor and interaction effects which are most important to the process. It displays the absolute values of the effects, and draws a reference line on the chart. Any effect that extends past this reference line is potentially important **[20]**. Figure 1 show NPP and Pareto plot for DOE results.

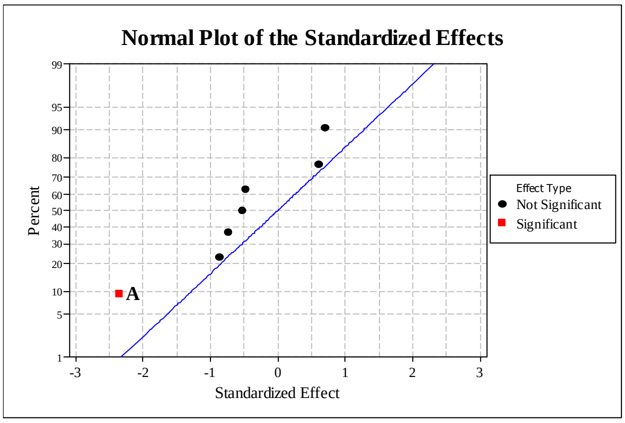

Fig. 1. NPP plot

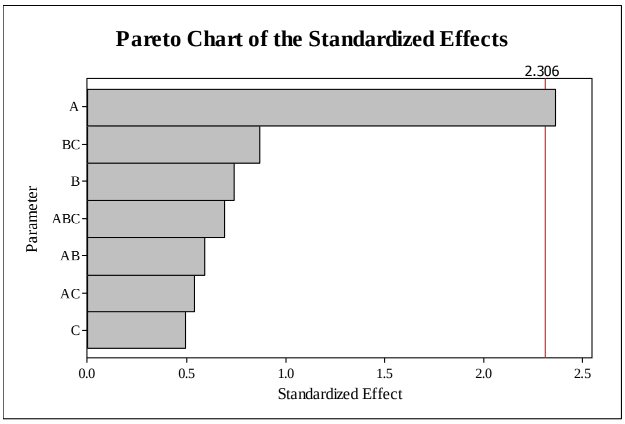

Fig. 2. Pareto plot

The graph illustrates that factor A (learning rate) is statistically significant at 5 percent significance level. In other words, this effect has large impact on the mean $MAPE_{ekv}$, while other two factors and interactions have only little impact. This finding can be further supported by considering the main effects plot (figure 3). From the figure 3 it is obvious that learning rate has huge impact on $MAPE_{ekv}$ whereas momentum and number of hidden neurons have very little impact.

Since the ultimate goal was to minimize $MAPE_{comb}$, the ANN model with settings: A(+1), B(+1), C(+1) was selected as best model trained at 10000 training epochs.

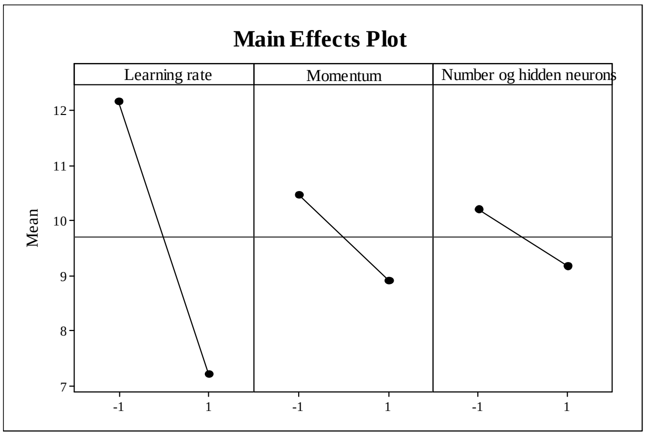

Fig. 3. Main effects plot

In order to analyze the ANN model performances, referring to the training period, it was decided to shorten/prolong the training process. In such a way, ANN model with settings: A(+1), B(+1), C(+1) was trained at 5000 and 15000 training epochs. The results are given in table 5. ANN with settings: A(+1), B(+1), C(+1) could be trained even better by using more training epochs which means prolonging training period. However, good ANN model performances have already been developed.

Table 5. Performance comparison of selected ANN model

ANN model	Number of training epochs		
A(+1), B(+1), C(+1)	5000	10000	15000
$MAPE_{ekv}$	5.7035	4.7195	3.916

5. ANN MODEL FOR THE PREDICTION OF TRAVERSE RATE OF SEPARATION CUT IN AWJ CUTTING PROCESS

Using the ANN modeling with the help of DOE for selecting the ANN training and architectural parameters, the ANN model was proposed. ANN model with parameter settings: A(+1), B(+1), C(+1), trained at 15000 training epochs is used for modeling the AWJ cutting process. The ANN modeling methodology was successfully implemented for the prediction of traverse rate of separation cut, by using material thickness, water pressure and abrasive rate as model inputs. Prediction results on training and testing, in the form of regression analysis, for optimal ANN model are shown in figure 4 and 5 respectively. As can be seen, the prediction accuracy is very high on both training and testing set with the correlation coefficient of 0.999.

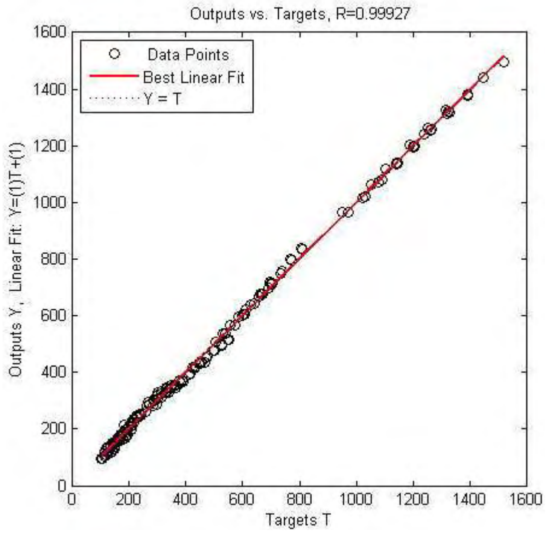

Fig. 4. Prediction accuracy on training data

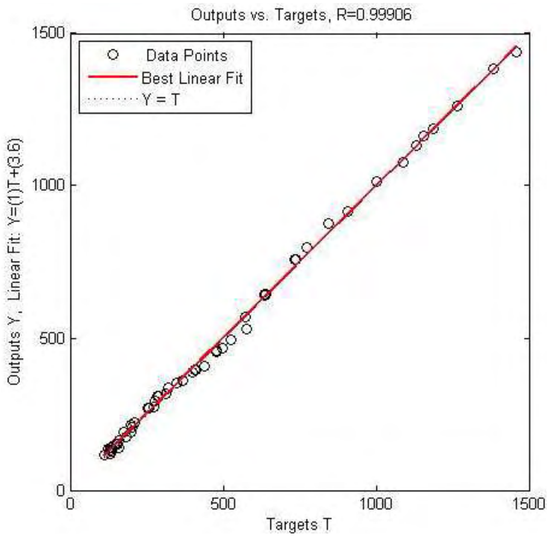

Fig. 5. Prediction accuracy on testing data

The percentage error of the model prediction was also calculated as the percentage difference between the predicted and calculated value relative to the calculated value. The error distribution of the ANN model for the prediction of traverse rate of separation cut using the entire dataset is shown in figure 6.

The error has a uniform distribution pattern about zero with a mean value and standard deviation of 0.3583 and 5.071%, respectively. The result shows that 91% of the entire dataset have the percentage error ranging between ±8.75%.

The ANN model with 3-8-1 architecture proved to be able to capture the underlying knowledge of the simulated process. The representation of knowledge is accomplished by the weights and biases between the layers. The values of these weights are given in table 6 and 7.

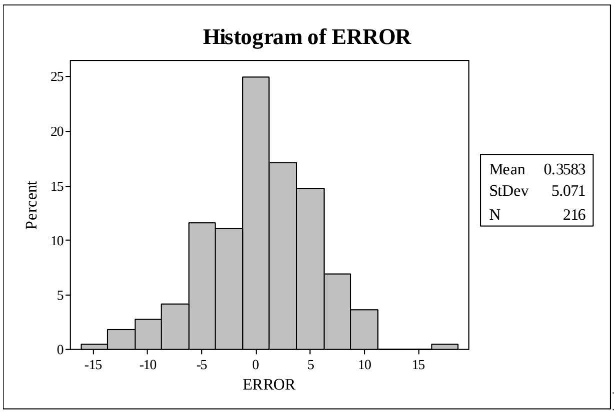

Fig. 6. Error distribution of the ANN model for the entire dataset

Table 7. The weights and biases corresponding to the input and hidden layer

	Weights			Bias
i	w_1	w_2	w_3	b_1
1	-2.449	0.85541	0.28374	2.9199
2	1.4503	-1.6859	-1.7668	-1.8355
3	2.782	-0.31751	0.45904	-0.63957
4	2.3487	-1.0828	-0.6225	-0.36545
5	-1.4232	-1.594	-1.5978	-0.55287
6	0.59178	1.6572	1.8495	-1.5936
7	-0.92183	-1.6197	-1.9895	-2.0504
8	2.6074	-0.12357	-0.12955	2.8229

Table 8. The weights and biases corresponding to the hidden and output layer

Weights w_j								Bias
w_1	w_2	w_3	w_4	w_5	w_6	w_7	w_8	b_2
2.9199	-1.8355	-0.63957	-0.36545	-0.55287	-1.5936	-2.0504	2.8229	0.7074

The traverse rate of separation cut v can be calculated by using the weights and biases of the trained 3-8-1 ANN model and rescaling the ANN outputs to get actual values of traverse rate of separation cut by using following equations:

$$v_{ANN} = \left[\frac{2}{1+e^{-2\left(U\cdot IW_{i1}+b_{i1}\right)}}-1\right]\cdot IW_2 + b_2 \qquad (3)$$

$$v_{actual} = \frac{1}{2}\cdot\left(v_{ANN}+1\right)\cdot\left(x_{max}-x_{min}\right)+x_{min} \qquad (4)$$

where U is the row vector which contain scaled values of material thickness, water pressure and abrasive rate respectively, IW_1 is the transpose matrix of weights corresponding to the input and hidden layer given in table 6, IW_2 is the transpose matrix of weights corresponding to the hidden layer and output layer given in table 7, v_{ANN} is the ANN scaled value for traverse rate of separation cut, v_{actual} is actual traverse rate of separation cut, and x_{max} and x_{min} are the maximum and minimum values of the raw data.

6. CONCLUSIONS

This paper discussed some possibilities of using DOE for setting the ANN architectural and training parameters in order to develop prediction model of AWJ cutting process. DOE analysis of ANN parameters clearly showed that only learning rate was main influencing factor on mean $MAPE_{ekv}$. The variability on ANN model predictions was highly influenced by random weights initialization. By setting the ANN parameters with the help of DOE, and by prolonging the ANN training until 15000 training epoch, the 3-8-1 ANN model of high performance was developed. It was tested on both training and testing set, and showed high prediction accuracy with correlation coefficient of 0.999. Therefore, the values of traverse rate of separation cut are accurately determined by the ANN model, by using 3 input parameters: material thickness, water pressure, and abrasive rate and are defined by proposed equation. In the authors' opinion, the setting of ANNs parameters are largely problem dependent and particular attention should be placed on ANN architectural and training parameters. Further research would include Taguchi's DOE with particular attention to weights initialization, training epochs, and comparing different training algorithms. In this way, development of high performance ANN models that can model more complex relationships in AWJ cutting would be provided.

ACKNOWLEDGEMENT

The authors would like to thank to the Ministry of Science of the Republic of Serbia.

REFERENCES

1. Antony, J. (2003). *Design of Experiments for Engineers and Scientists*, Butterworth-Heinemann, Oxford.
2. Bishop, C. (1995). *Neural Networks for Pattern Recognition*, Clarendon Press, ISBN 978-0-19-853864-6, Oxford.
3. Çaydaş, U., Hasçalik, A. (2008). *A study on surface roughness in abrasive waterjet machining process using artificial neural networks and regression analysis method.* Journal of Materials Processing Technology, 202, 1-3, ISSN 0924-0136, pp. 574-582.
4. Ergür, H.S., Gölcü, M., Pancar, Y., Altan Dombayci, Ö. (2006). *Analysis of the cutting power on abrasive waterjet system applications.* International Symposium on Intelligent Manufacturing Systems, 475-483, Sakarya, Turkey.
5. Haykin, S. (1999). *Neural Networks: A Comprehensive Foundation*, Prentice Hall, ISBN-13-978-0132733502, New Jersey.
6. Khaw, J.F.C., Lim, B.S., Lim, L.E.N. (1995). *Optimal design of neural networks using the Taguchi method*. Neurocomputing, 7, 3, ISSN 0925-2312, pp. 225-245.
7. Kim, Y.S., Yum, B.J. (2004). *Robust design of multilayer feedforward neural networks: an experimental approach*. Engineering Applications of Artificial Intelligence, 17, 3, ISSN 0952-1976, pp. 249–263.
8. Laosiritaworn, W., Chotchaithanakorn, N. (2009). *Artificial neural networks parameters optimization with design of experiments: an application in ferromagnetic materials modeling.* Chiang Mai Journal of Science, 36, 1, ISSN 0125-2526, pp. 83-91
9. Lin, T.Y., Tseng, C.H. (2000). *Optimum design for artificial neural networks: an example in a bicycle derailleur system*. Engineering Applications of Artificial Intelligence, 13, 1, ISSN 0952-1976, 3-14.
10. Lu, Y., Li, X., Jiao, B., Liao, Y. (2005). *Application of artificial neural networks in abrasive waterjet cutting process*. Proceedings of Second International Symposium on Neural Networks ISNN 2005, LNCS 3498. Wang, J., Liao, X., Yi, Z., Springer, 877-882, ISSN 0302-9761, Chongqing, China.
11. Montgomery, D.C. (2001). *Design and Analysis of Experiments*, John Wiley & Sons, New York.
12. Nguyen, D., Widrow, B. (1990). *Improving the Learning Speed of 2-layer Neural Networks by Choosing Initial Values of the Adaptive Weights*. Proceedings of the International Joint Conference on Neural Networks (IJCNN), Edward Brothers, volume 3, pp. 21-26, San Diego.
13. Packianather, M.S., Drake, P.R., Rowlands, H. (2000). *Optimizing the parameters of multilayered feedforward neural networks through Taguchi design of experiments.* Quality And Reliability Engineering International, 16, 6, ISSN 0748-8017, pp. 461-473
14. Parikh, P.J., Lam, S.S. (2009). *Parameter estimation for abrasive water jet machining process using neural networks.* International Journal of Advanced Manufacturing Technology, 40, 5-6, ISSN 1433-3015, pp. 475-483.
15. Srinivasu, D.S., Ramesh Babu, N. (2008). *A neuro-genetic approach for selection of process parameters in abrasive waterjet cutting considering variation in diameter of focusing nozzle.* Applied Soft Computing, 8, 1, ISSN 1568-4946, pp. 809–819.
16. Sukthomya, W., Tannock, J. (2005). *The optimisation of neural network parameters using Taguchi's design of experiments approach: an application in manufacturing process modelling*. Neural Computing and Applications, 14, 4, ISSN 0941-0643, pp. 337-344.
17. Tarassenko, L. (1998). *A Guide to Neural Computing Applications*, Arnold Publishers, ISBN 0-340-70589-2, London.
18. Yang, S.M., Lee, G.S. (1999). *Neural network design by using Taguchi method.* Journal of Dynamic Systems, Measurement and Control - Transactions of the ASME, 121, 1, ISSN 0022-0434, pp. 560-563
19. Zurada, J. (1992). *Introduction to Artificial Neural Systems*, PWS Publishing Company, ISBN 0-534-95460-X, Boston.
20. KMT Watejet Systems – Abrasive Waterjet Cutting Speed Calculator.

Received: July 3, 2011 / Accepted: November 30, 2011 / Paper available online: December 10, 2011