Assessing the Sensitivity of the Artificial Neural Network to Experimental Noise: A Case Study

Miloš J. Madić
PhD Student
University of Niš
Faculty of Mechanical Engineering

Velibor J. Marinković
Full Professor
University of Niš
Faculty of Mechanical Engineering

This case study deals with modeling of plasma cutting process using artificial neural network (ANN), with the aim of simulation the impact of noise on its performance. The input parameters used in this study were reduced to three cutting parameters which consisted of strength of current (I), cutting speed (V), and material thickness (s). The ten-point height of irregularities (R_z), which is one of the basic characteristics of the surface quality, was adopted as the output parameter (response). The data for this research were gathered from literature. A feed-forward three-layer ANN was created, with backpropagation and algorithm for supervised learning. For the hidden layer neurons sigmoidal type of non-linearity was selected, while a linear activation function was selected for the output layer. ANN training was carried out using Levenberg-Marquardt algorithm with Bayesian regularization. The trained and tested ANN on the original data set showed a satisfactory level of prediction accuracy. In order to simulate an experiment with noise, measured values of the surface roughness were corrected. The correction was performed by adding randomly selected numbers to each measured value, within the range – 0.1 to + 0.1 μm. With the previously selected architecture and the same other parameters, re-training of ANN was carried out. The analysis showed that the ANN model trained on the data with noise has similar performance as the ANN model trained on the original data, which indicates the robustness of this type of ANN.

Keywords: *sensitivity of ANN, experimenatal noise, plasma cutting.*

1. INTRODUCTION

As far as complex (non-linear, "diffusion") processes and systems are concerned, whose structures and laws of action are not well known or known enough, the principle of the "black box" is applied. Basic input (controllable) values are those that can be expressed numerically and selected and varied at will. Other input (uncontrollable, noise) values are those whose action is unknown or can be neglected. Output values (responses, target functions, state characteristics) are those values that can be measured and which are the result of the action of input values.

The theory of experimental design is based on the aforementioned principle, known also as the design of experiment [1-5]. The design of experiment (DoE) has a very broad application across all engineering, natural and social sciences. All this also relates to artificial neural networks [6-10].

In DoE, it is important to choose an adequate mathematical model and experiment design. The choice of appropriate mathematical model is often not straightforward. Also, well chosen experimental designs maximize the amount of information that can be obtained for a given amount of experimental effort.

The realizations of the experiment, regression and dispersion analysis are strictly determined by the chosen mathematical model and design. The application of complex mathematical models and designs implies a more profound knowledge of the theory of experimental design.

In applying the ANN method, neither the mathematical model nor the experiment plan are being set according to the previously determined rules. The results obtained directly from the existing production process can be used in that sense, which significantly reduces research costs.

Another important advantage of ANNs in relation to the classical DoE is reflected in the fact that the input data can be varied on different number of levels. This is not the case with DoE, where certain levels (centre point, "star" points) cannot sometimes be realized experimentally, neither under laboratory nor even less under production conditions.

The basic deficiency of ANNs lies in the need to process a relatively large number of data (samples), in order to secure the sufficient prediction accuracy and the desired level of generalization. However, there are examples where that number of samples is not much larger than the number needed for the application of DoE [11-16].

From the regression analysis, it is known that the accuracy of the selected mathematical model is highly influenced by experimental errors. The aim of this study is to examine to which extent the experimental noise affects the ANN performance, i.e. its prediction accuracy.

Received: September 2010, Accepted: November 2010
Correspondence to: Dr Velibor J. Marinković
Faculty of Mechanical Engineering,
Aleksandra Medvedeva 14, 18000 Niš, Serbia
E-mail: velmar@masfak.ni.ac.rs

2. ARTIFICIAL NEURAL NETWORK

2.1 Overview of ANN

The appearance of ANNs is related to the attempt to form an artificial system based on mathematical models which will be, in its structure, function and information (signal) processing, similar to biological nervous systems, and thus be able to process "intelligently" information (signals), that is, simulate biological intelligence.

The ANNs have found application in many different domains, such as dynamic system identification, complex system modelling, optimization and control, design and classification, speech interpretation, pattern recognition, metal-cutting and metal-forming simulation, robotics and communication and the like.

Modern ANNs have a parallel-distributed architecture. They consist of a larger number of neurons (as basic processing units) distributed in several layers. Each ANN must have at least three layers: input layer, hidden layer, and output layer. The greatest number of ANNs has one hidden layer though there may be more of them [17-19].

Neurons of one layer are connected by specific synapses to the neurons of their neighbouring layers. Apart from specific types of ANNs, there are no interconnections between the neurons belonging to the same layer. Viewed schematically, the ANN represents an oriented graph in which the nodes are processor units, while the arrows on the lines point to the direction/sense of the signal (information) flow [7] (Fig. 1).

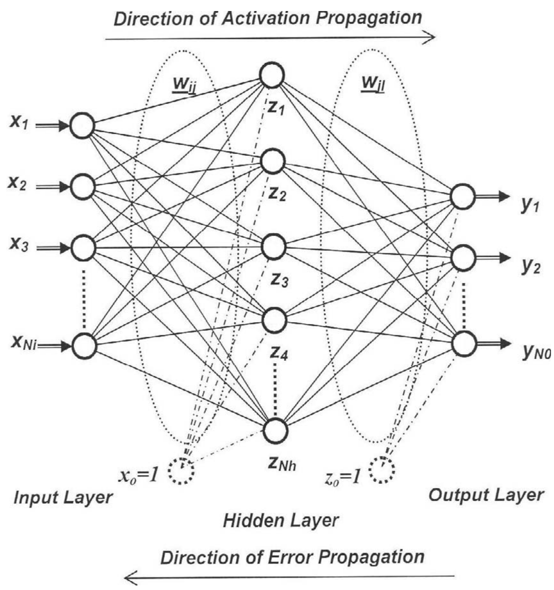

Figure 1. Topology of a typical feed-forward three-layer BP ANN

The interconnections between particular neurons by the layers are characterized by weights, which change during the ANN training.

The number of hidden layers and the number of neurons in each of them are not defined in advance; instead, these numbers can change during the ANN training until the optimum architecture is defined, namely, the one that produces the best performances of ANN. The researchers usually rely on the derived examples and their own experience ("trial and error approach"). This is one of the major obstacles in using ANNs.

The ANN modelling commonly follows these steps: definition of the input and output parameters; collection and analysis of the data base; random dividing of the data set on training and testing data subsets; normalization of the input/output data (optional); designing of an ANN; training of the ANN (choice of architecture, training algorithm, transfer functions, performance criterion, and other ANN parameters); testing the trained ANN.

After adopting the ANN with the best performance, the created ANN may be used for simulation and prediction.

Every ANN is designed for modelling a concrete problem. For an ANN to react appropriately when unknown input data are added to it, it must be trained. ANN training is done on the examples that make up a set of input-output data obtained by the experiment or some other way.

The ANN training represents a process of adjusting weights of interconnections and bias adjoined to every synapsis between neurons on the basis of comparing the output values with the desired (target) ones for the same input ones. Training is a continuous process, which is repeated until the ANN is stabilized or overall error is reduced below a previously defined threshold. Training belongs to the most important parts of ANN designing and procedure.

The trained ANN should be tested in order to assess its ability to predict and make generalization on the basis of the acquired "knowledge" on the selected set of input-output data. Testing of the ANN is carried out by applying a data subset, which is not included in the training data subset.

The trained and tested ANN can be used for modelling and prediction, when it is presented with new (original) input data.

Of all the available types of ANN, multi-layer perceptron (MLP) with back propagation (BP) training procedure, is the most commonly used. The BP ANN is designed to operate as a multilayer fully-connected feed-forward network, with a particular (BP) training algorithm for supervised learning.

For feed-forward BP ANN there are many different standard training algorithms (Batch Gradient Descent, One-step-secant, Resilent Backpropagation, Conjugate Gradient, Levenberg-Marquardt (L-M), etc.). The L-M algorithm belongs to the algorithms that converge very fast (especially for smaller and medium large ANNs), with less danger from entrapment in local minimum, before reaching global minimum at error surface, while at the same time it can provide for high accuracy of prediction. A more detailed approach to ANNs can be found in the referential literature [6-10].

2.2 Implementation of ANN

The objective of the plasma cutting process is to concentrate a large amount of energy on a small surface of a workpiece which leads to intensive heating of the material surface. The source of energy is high temperature and high speed ionised gas. The gas is ionised using a direct current passing between the cathode (inside the nozzle) and anode (workpiece). The plasma jet cuts the material by releasing the energy spent for the plasma gas ionisation upon hitting the workpiece surface. The removal of the melted material from the cutting zone is done by the action of plasma jet kinetic energy. The characteristics of plasma jet can be significantly altered by changing the type of gas, gas flow, cutting current, and nozzle size, etc.

Even though it is the case of a complex process, which is characterised by a large number of influential factors, the previous analysis has shown that this number can be reduced to three main influential factors: cutting current (*I*), cutting speed (*V*), and material thickness (*s*). Influential factors were varied on a great number of levels. The ten-point height of irregularities (R_z), which is one of the basic characteristics of the surface quality, was selected as the target function (output value). The experimental data are presented in Table 1 [20].

Table 1. Data for training and testing of the ANN

No.	Cutting current, *I* [A]	Material thickness, *s* [mm]	Cutting speed, *V* [m/min]	Surface roughness, R_z [μm]
1.	80	4	1300	2.13
2.	80	4	1400	2.15
3.	**80**	**4**	**1000**	**2.25**
4.	**80**	**4**	**900**	**2.3**
5.	80	4	1200	2.4
6.	80	4	1700	2.42
7.	80	4	2100	3.2
8.	80	4	2200	3.15
9.	80	4	2300	3.4
10.	**80**	**4**	**2400**	**3.5**
11.	80	4	2500	3.55
12.	80	4	2600	3.58
13.	80	4	2800	3.7
14.	**45**	**4**	**1050**	**3.2**
15.	45	4	1100	3.4
16.	45	4	1150	3.6
17.	45	4	1200	3.67
18.	**45**	**4**	**1250**	**4.1**
19.	45	4	950	3.4
20.	**45**	**4**	**900**	**3.5**
21.	45	4	850	3.3
22.	**45**	**4**	**800**	**3.1**
23.	45	4	1100	3.5
24.	45	4	1300	3.82
25.	45	4	1400	3.8
26.	**45**	**4**	**1500**	**4**
27.	80	6	1225	2.15
28.	**80**	**6**	**1275**	**2.21**
29.	80	6	1300	2.25
30.	80	6	1375	2.25
31.	80	6	1425	2.28
32.	**80**	**6**	**1475**	**2.3**
33.	80	6	1175	2.22
34.	80	6	1125	2.35
35.	**80**	**6**	**1075**	**2.35**
36.	80	6	1025	2.38
37.	80	6	900	2.45
38.	80	6	1700	2.5
39.	**80**	**6**	**1900**	**2.6**
40.	80	6	2100	2.65
41.	80	6	2300	2.8
42.	**45**	**6**	**850**	**2.55**
43.	45	6	900	2.48
44.	45	6	1000	3.1
45.	45	6	1100	3.15
46.	45	6	800	3.1
47.	**45**	**6**	**750**	**3.05**
48.	45	6	700	2.9
49.	45	6	650	2.6
50.	**45**	**6**	**600**	**2.52**
51.	45	6	1300	3.1
52.	80	8	900	3.29
53.	**80**	**8**	**950**	**3.42**
54.	80	8	1000	3.3
55.	80	8	1050	3.25
56.	80	8	1100	3.2
57.	80	8	1150	3.2
58.	**80**	**8**	**1200**	**3.3**
59.	80	8	1250	3.42
60.	80	8	1300	3.6
61.	80	8	1350	4.05
62.	**80**	**8**	**1400**	**4.22**
63.	80	8	1500	4.32
64.	**80**	**8**	**1700**	**4.3**
65.	80	8	2000	4.5
66.	130	12	820	1.79
67.	**130**	**12**	**870**	**1.85**
68.	130	12	920	1.86
69.	130	12	970	1.9
70.	**130**	**12**	**1020**	**2.06**
71.	130	12	1070	2.22
72.	130	12	770	2.1
73.	130	12	720	2.15
74.	**130**	**12**	**670**	**2.12**
75.	130	12	620	2.2
76.	130	12	570	2.25
77.	130	12	1200	2.1
78.	**130**	**12**	**1400**	**2.18**
79.	130	12	1600	2.2
80.	130	12	1800	2.27
81.	130	15	580	2.16
82.	130	15	630	2.2
83.	130	15	680	2.22
84.	**130**	**15**	**730**	**2.3**
85.	130	15	780	2.42
86.	130	15	830	3.05
87.	130	15	530	2.2
88.	**130**	**15**	**480**	**2.42**
89.	130	15	430	2.62
90.	130	15	380	3.23
91.	130	15	330	3.78
92.	130	15	900	3.15
93.	**130**	**15**	**1100**	**3.1**
94.	130	15	1300	3.05
95.	**130**	**15**	**1600**	**2.5**
96.	130	15	1700	2.25

The series of 29 input-out data for ANN testing is marked with bold numbers.

For the needs of training and testing the created ANN the whole experimental data set ($N_{tot} = 96$) is randomly divided into a data subset for training ($N_1 = 67$) and a data subset for testing the ANN ($N_2 = 29$). Approximately, two-thirds of the whole data set have been employed for training and one-third of the whole data set has been used for testing the trained ANN.

Bayesian regularization with L-M algorithm ("trainbr")[1] was chosen for ANN training in order to obtain an ANN which generalizes well. The non-linear hyperbolic tangent transfer function ("tansig") for the hidden layer and the linear transfer function ("purelin") for the output layer are applied. Numerous cases were proven that such transfer functions make ANNs capable to perform successfully the approximation of different non-linear functions. The training data subset is presented to the ANN by the batch method.

Mean absolute percentage error (*MAPE*), achieved by the training and testing of ANNs, represents the criterion for optimization of the network interconnecting weights.

According to the available sample size, the created architecture of ANN (3 $[11]_1$1) has turned out to be the optimal solution (after trade-off).

Figure 2 shows topology of this ANN, according to the notation system of the software package MATLAB.

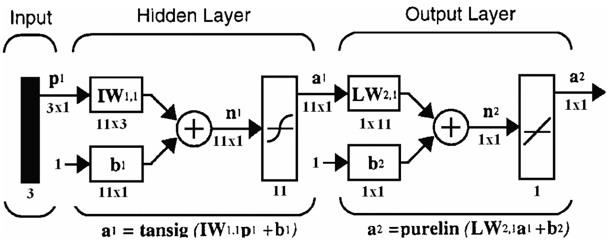

Figure 2. The created feed-forward BP ANN 3 $[11]_1$1

The responses of the trained ANN are shown in Figure 3.

After that, the retraining of the selected ANN with modified measuring data for surface quality (R_z') was performed. This procedure was conducted in the following manner. To each value of parameter R_z a randomly chosen value $\Delta R_z \in [-0.1, +0.1]$, which follows the law of uniform distribution, was added. By checking, it was determined that the maximal percentage errors, entered into the original data, were: $\delta R_z \approx \pm 4.5$ %. These errors can be treated as regular errors in the engineering measurements.

The results of the prediction of such ANN in relation to the experimental data with noise are given in Figure 4.

The most important aspect of this study was to determine the prediction capability of a retrained ANN in the view of agreement of the desired (target) values of parameter R_z', with the original values of this parameter without noise R_z.

The comparative results of this analysis are given in Figure 5.

[1] MATLAB command for the corresponding function

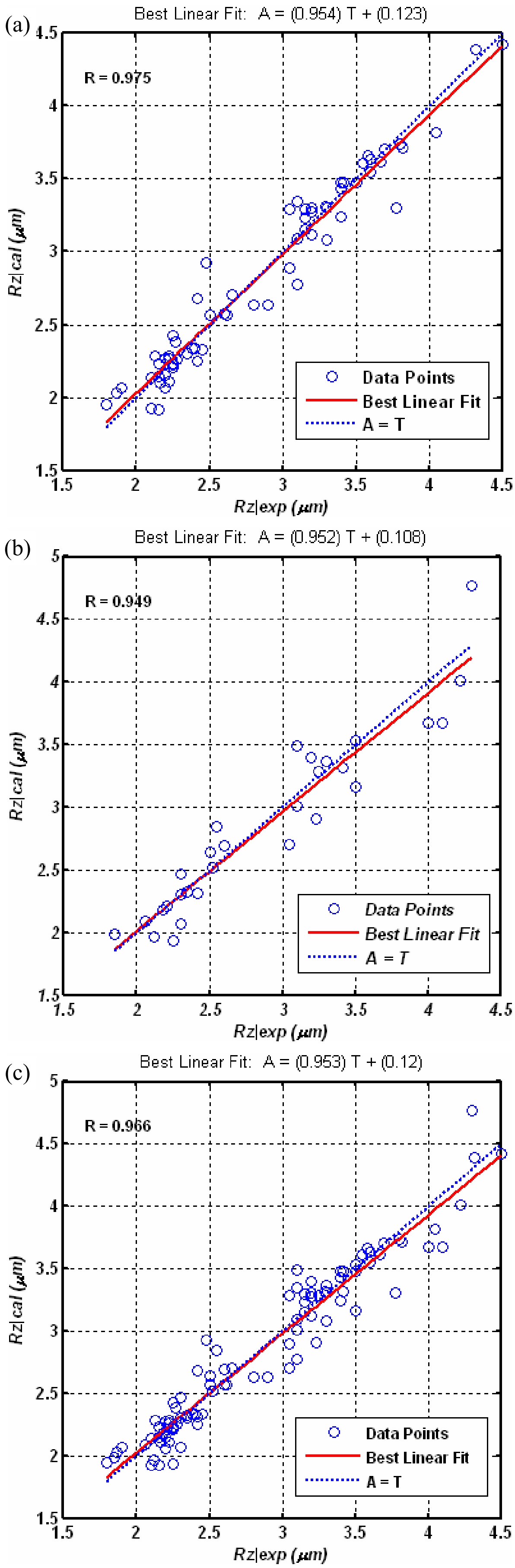

Figure 3. The performance of ANN model for surface roughness in plasma cutting process: (a) for training data set, (b) for testing data set and (c) for whole data set; The ANN was trained without noise ($R_{z|cal}$). The experimental values ($R_{z|exp}$) are values without noise

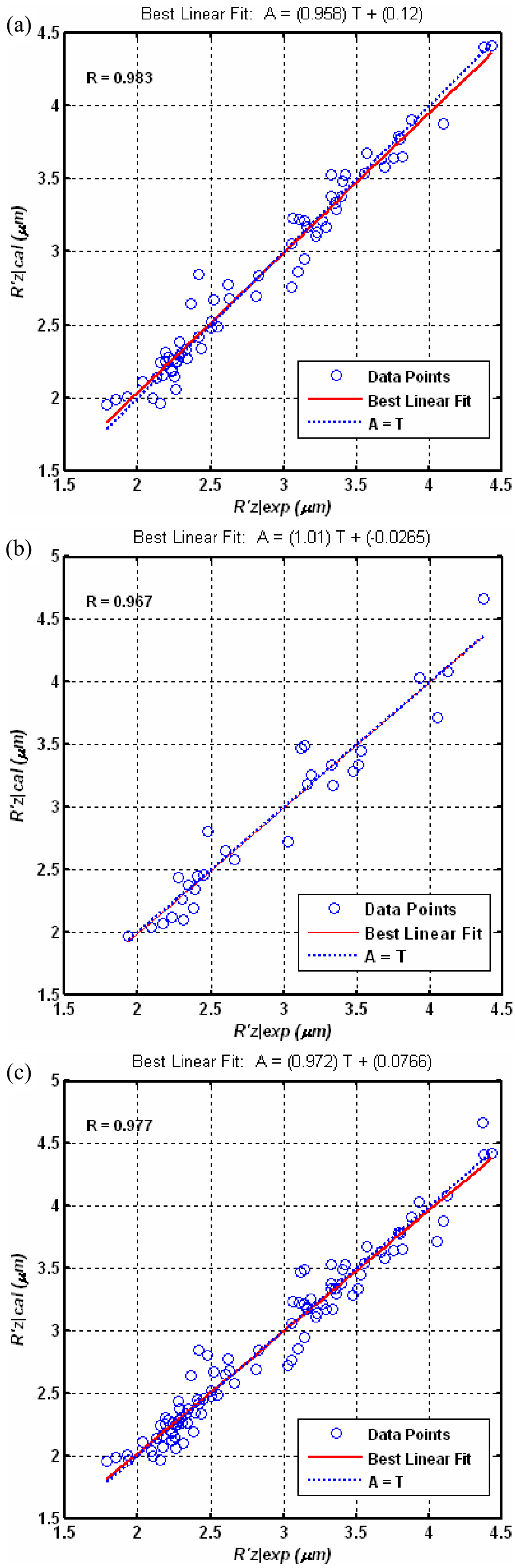

Figure 4. The performance of ANN model for surface roughness in plasma cutting process: (a) for training data set, (b) for testing data set and (c) for whole data set; The ANN was trained with added noise ($R'_{z|cal}$). The experimental values ($R'_{z|exp}$) are values including noise

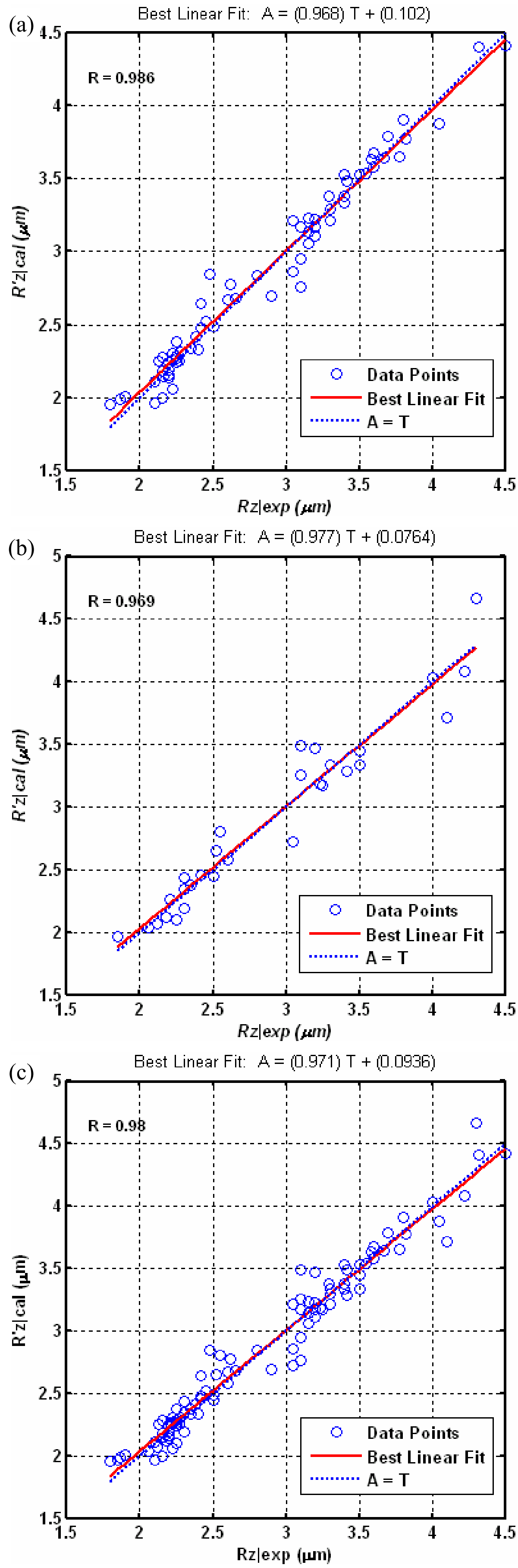

Figure 5. The performance of ANN model for surface roughness in plasma cutting process: (a) for training data set, (b) for testing data set and (c) for whole data set; The ANN was trained with added noise ($R'_{z|cal}$). The experimental values ($R_{z|exp}$) are values without noise

3. RESULTS AND DISCUSSION

The prediction accuracy of an ANN can be displayed in various ways. The correlation coefficient is a statistical measure of the strength of correlation between actual versus predicted values. For example, the value of + 1 indicates perfect correlation. In that case, all points should lie on the line passing through the origin and inclined at 45°. In all cases investigated in this study most of the points are close to this line and correlation coefficients are very high (Table 2).

Table 2. Correlation coefficients for all ANN models

Model	*R*		
	Training	Testing	Training + Testing
The ANN (original data)	0.975	0.949	0.966
The retrained ANN (data with noise)	0.983	0.967	0.977
The retrained ANN (original data)	0.986	0.969	0.980

Also, the histogram is the appropriate statistical tool for the analysis of prediction accuracy. The distribution of the percentage error in the form of histogram in eleven equal ranges, between minimum and maximum values, is presented in Figure 6. The histogram refers onto re-trained ANN and whole data set, without the noise (see Fig. 5c).

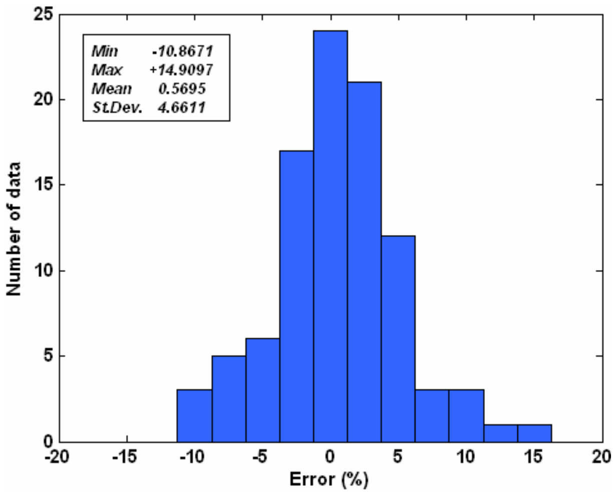

Figure 6. Distribution of the percentage error (δ) for modelling the surface roughness (R_z) in plasma cutting process

As it is known, for the normal distribution, two standard deviations from the mean account 95 % of the whole set of data. In this case, it is: $\mu \pm 2\sigma < 10$ %. In other words, the number of errors greater than 10 % is practically negligible.

Furthermore, in all ANN models the mean absolute percentage errors are very small (Table 3). Similar results were obtained for other cases examined in this study.

Table 3. Mean absolute percentage errors for all ANN models

Model	*MAPE* [%]		
	Training	Testing	Training + Testing
The ANN (original data)	4.061	5.835	4.597
The retrained ANN (data with noise)	3.356	4.640	3.744
The retrained ANN (original data)	3.127	4.480	3.515

4. CONCLUSION

In this study, the effects of experimental noise on prediction ability of the ANN by imposing random noise, similar to those found experimentally, are examined.

For that purpose, a BP ANN model for the analysis and prediction of the relationship between process parameters and characteristic of surface quality in plasma cutting was developed.

A very good performance of an ANN, trained on noisy experiment, in terms of agreement with original experimental data was achieved. That indicates the robustness of this type of ANN.

ACKNOWLEDGEMENT

The paper is a part of the research done within the project "Modelling of the correlations between the parameters of the plasma arc cutting and heat balance analysis using methods of artificial intelligence". The authors would like to thank the Ministry of Science and Technological Development of the Republic of Serbia.

REFERENCES

[1] Montgomery, D.C.: *Design and Analysis of Experiments*, John Wiley & Sons, Inc., New York, 2001.

[2] Novik, F.S. and Arsov, J.B.: *Optimization of Technological Processes Using Methods of Experimental Design*, Mashinostoenie, Moscow, 1980, (in Russian).

[3] Stanić, J.: *Method of Engineering Measurements*, Faculty of Mechanical Engineering, University of Belgrade, Belgrade, 1990, (in Serbian).

[4] Marinković, V.: Application of some non-linear mathematical models to the theory of experimental design, Facta Universitatis – Series: Mechanical Engineering, Vol. 1, No. 1, pp. 103-117, 1994.

[5] Marinković, V.: Rationalization of experimental investigations using theory of experimental design, IMK-14 - Istraživanje i razvoj, Vol. 13, No. 1-2, pp. 23-36, 2007, (in Serbian).

[6] Haykin, S.: *Neural Networks: A Comprehensive Foundation*, Prentice Hall, New York, 1998.

[7] Bose, N.K. and Liang, P.: *Neural Network Fundamentals with Graphs, Algorithms, and Applications*, McGraw-Hill, New York, 1996.

[8] Dreyfus, G.: *Neural Networks*, Springer Verlag, Berlin, 2005.

[9] Miljković, Z.: *Systems of Artificial Neural Networks in Production Technologies*, Faculty of Mechanical Engineering, University of Belgrade, Belgrade, 2003, (in Serbian).

[10] Miljković, Z. and Aleksendrić, D.: *Artificial Neural Networks – Collection of Resolved Tasks with Excerpts From the Theory*, Faculty of Mechanical Engineering, University of Belgrade, Belgrade, 2009, (in Serbian).

[11] Calcaterra, S., Campana, G. and Tomesani, L.: Predicrion of machanical properties in spheroidal cast iron by naural networks, Journal of Materials Processing Technology, Vol. 104, No. 1-2, pp. 74-80, 2000.

[12] Mani, V. and Omkar, S.N.: Understanding weld modelling processes using a combination of trained neural networks, International Journal of Production Research, Vol. 40, No. 3, pp. 547-559, 2002.

[13] Altinkok, N. and Koker, R.: Neural network approach to prediction of bending strength and hardening behaviour of particulate reinforced (Al-Si-Mg)-aluminium matrix composites, Materials & Design, Vol. 25, No. 7, pp. 595-602, 2004.

[14] Marinković, V.: Determination of steel formability for warm forming by applying artificial neural network, in: *Proceedings of the 7th International Conference "Research and Development in Mechanical Industry" – RaDMI 2007*, 16-20.09.2007, Belgrade, Serbia, pp. 217-224.

[15] Tsao, C.C. and Hocheng, H.: Evaluation of thrust force and surface roughness in drilling composite material using Taguchi analysis and neural network, Journal of Materials Processing Technology, Vol. 203, No. 1-3, pp. 342-348, 2008.

[16] Marinković, V.: Application of artificial neural network for modeling the flash land dimensions in the forging dies, Strojniški vestnik – Journal of Mechanical Engineering, Vol. 55, No. 1, pp. 64-75, 2009.

[17] Eyercioglu, O., Kanca, E., Pala, M. and Ozbay, E.: Prediction of martensite and austenite start temperatures of the Fe-based shape memory alloys by artificial neural networks, Journal of Materials Processing Technology, Vol. 200, No. 1-3, pp. 146-152, 2008.

[18] Elangovan, K., Sathiya Narayanan, C. and Narayanasamy, R.: Modelling of forming limit diagram of perforated commercial pure aluminium sheets using artificial neural network, Computational Materials Science, Vol. 47, No. 4, pp. 1072-1078, 2010.

[19] Miljković, Z., Bojović, B. and Babić, B.: Application of artificial neural network and fractals in biomedical materials surface behavior prediction, Tehnika – Novi materijali, Vol. 19, No. 4, pp. 5-14, 2010.

[20] Lazarević, A.: *Modelling of the Correlations between the Parameters of the Plasma Arc Cutting and Heat Balance Analysis Using Method of Artificial Intelligence*, PhD thesis, Faculty of Mechanical Engineering, University of Niš, Niš, 2010, (in Serbian).

ПРОЦЕНА ОСЕТЉИВОСТИ ВЕШТАЧКИХ НЕУРОНСКИХ МРЕЖА НА ШУМ ЕКСПЕРИМЕНТА: СТУДИЈА СЛУЧАЈА

Милош Ј. Мадић, Велибор Ј. Маринковић

Ова студија се бави моделовањем процеса резања плазмом применом вештачке неуронске мреже (ВНМ), са циљем симулације утицаја шума експеримента на њене перформансе. Улазни параметри коришћени у овој студији су били редуковани на три главна параметра резања: јачину струје, брзину резања и дебљину материјала. Као излазни параметар (одзив) била је усвојена средња висина неравнина профила у десет тачака (R_z), једна од базичних карактеристика квалитета обрађене површине. Подаци за ово истраживање су били узети из литературе. Креирана је директна трослојна ВНМ, са пропагацијом грешке уназад и алгоритмом за надгледано учење. За неуроне скривеног слоја изабран је сигмоидални тип нелинеарности, а за излазни слој линеарна активациона функција. ВНМ тренинг је извршен употребом Левенберг-Маркеовог алгоритма са Бајесовом регуларизацијом. Тренирана и тестирана ВНМ на оригиналном сету података показала је задовољавајући ниво тачности предикције. Да би се симулирао експеримент са шумом, мерне вредности површинске храпавости су биле кориговане. Корекција је изведена додавањем случајно изабраних бројева свакој измереној вредности, унутар граница – 0,1 и + 0,1 μm. Анализа је показала да ВНМ тренирана на подацима са шумом има сличне перформансе као и ВНМ тренирана на оригиналним подацима, што указује на робустност овог типа ВНМ.