IOP Conf. Series: Materials Science and Engineering **269** (2017) 012041 doi:10.1088/1757-899X/269/1/012041

Effect of Activation Function and Post Synaptic Potential on Response of Artificial Neural Network to Predict Frictional Resistance of Aluminium Alloy Sheets

T Trzepiecinski[1,] and H G Lemu[2]*

[1]Rzeszow University of Technology, Department of Materials Forming and Processing, al. Powst. Warszawy 12, 35-959 Rzeszów, Poland
[2]University of Stavanger, Department of Mechanical and Structural Engineering, N-4036 Stavanger, Norway

Abstract. Many technological factors affect the friction phenomenon in sheet metal forming process. As a result, the determination of the analytical model describing the frictional resistance is very difficult. In this paper, a friction model was built based on the experimental results of strip drawing tests. Friction tests were carried out in order to determine the effect of surface and tool roughness parameters, the pressure force and mechanical parameters of the sheets on the value of coefficient of friction. The strip drawing friction tests were conducted on aluminium alloy sheets: AA5251-H14, AA5754-H14, AA5754-H18, AA5754-H24. The surface topography of the sheets was measured using Taylor Hobson Surtronic 3+ instrument. In order to describe complex relations between friction and factors influencing tribological conditions of sheet metal forming, the multilayer artificial network was built in Statistica Neural Network program. The effect of activation function and post synaptic potential function on the sensitivity of multilayer neural network to predict the friction coefficient value is presented. It has been found that the difference in the prediction of error of neural network for different approaches can reach 400%. So, the proper selection of activation and post synaptic potential functions is crucial in neural network modelling.

1. Introduction

Friction in metal forming process is an important parameter affecting the possibility to plastically deform different materials. In sheet metal forming, in particular, frictional resistance is a complicated function of mechanical properties of the sheet metal, temperature, lubrication conditions, and topography of both the tools and the workpiece [1-4]. Friction occurred at high pressures may essentially differ from friction in kinematic pairs due to a great influence of plastic deformations on dynamic character of contact [5]. Friction occurs due to some mechanisms which are simultaneously taking place with frictional contact, including the ploughing effect between asperities, the adhesion between contacting asperities, the flattening due to stretching, flattening due to sliding, and the appearance of hydrodynamic friction stresses [6, 7]. Complex interactions between the parameters influencing the frictional resistance and coefficient of friction value make it difficult to analyse and understand this relation. In the last few decades, many investigations were conducted to find real friction models. Challen and Oxley [8], for instance, developed a friction model taking the combined effect of ploughing and adhesion on the coefficient of friction into account. Wilson and Sheu [9] developed an analytical upper bound model to describe the flattening behaviour using wedge-shaped asperities with a constant angle. Trzepiecinski and Lemu [10] noticed that the proper selection of variables that influence the coefficient of friction value can improve the sensitivity of neural network friction model.

Advanced systems of data processing allow for automatic analysis of complex data set and generation of response on interrogated questions. Due to a large number of parameters and phenomena affecting the tribological phenomena, developing analytical relations between input variables and coefficient of friction value is practically impossible. This task is successfully realised by artificial neural networks (ANN) that belong to the artificial intelligence technology. In this paper, the ANN model of friction is built in Statistica Neural Networks program based on the results of strip drawing friction test of aluminium alloy sheets. In the sheet metal forming process, the strip-drawing tests simulate the friction phenomenon that exists between the punch and the wall of the drawpiece. The effect of learning algorithm, network architecture and the method of data selection are topics of recent papers of the authors [10, 11]. In addition, the authors aim to report, in this paper, about the effect of activation function (AF) and post synaptic potential function (PSPF) on the sensitivity of neural network to predict the friction coefficient value.

2. Material and method

2.1. Experiments

To determine the value of the coefficient of friction, the strip drawing tests were performed using a friction simulator (figure 1) mounted in standard tensile machine. During the tests, a strip of the aluminium alloy sheets (AA5251-H14, AA5754-H14 AA5754-H18, and AA5754-H24) was clamped with specified forces (400, 800, 1200, 1600, and 2000 N) between two cylindrical rolls of equal radii r = 10 mm. Next, when the displacement of moving crosshead of the test machine was engaged, the pulling force F_p and clamping force F_c were recorded continuously using a load cell and a computer programme. The value of coefficient of friction has been determined as a ratio of clamping force and pressure (normal) force (figure 1).

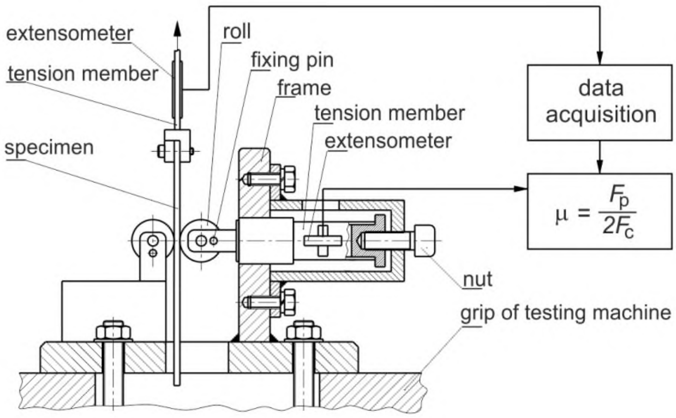

Figure 1. The schematic of friction simulator.

The specimens for the friction tests were prepared as the strips with 20 mm width and about 200 mm length, which were cut along the rolling direction of the sheet. Four sets of rolls (counter samples) with different surface roughness values (0.32, 0.63, 1.25 and 2.5 μm) were used. The tests were conducted using machine oil LAN-46 as lubricant. Surface roughness parameter measurements were carried out using Taylor Hobson Surtronic 3+ instrument to determine arithmetic average height along the rolling (Ra^0) and transverse directions (Ra^{90}) of the sheet metal. The value of yield the stress σ_y of the tested sheets was determined in the uniaxial tensile test.

2.2. ANN modelling

The friction model, based on the multilayer perceptron (MLP) network, was built in Statistica Neural Networks program. The methodology of creating computations in this program was described in the recent works of authors [10, 11]. Based on the training data, the network architecture MLP 5:5-11-1:1 was found as the optimal structure which ensures the minimal prediction of error of the ANN [11]. The following parameters were selected as input set of variables: yield stress of the sheets materials σ_y, arithmetical mean deviation of the assessed profile of sheet surface Ra^0 and Ra^{90} measured along and transverse to sheet rolling direction respectively, clamping force F_c and average surface roughness of rolls Ra^r.

The significance of the influence of input parameters on the output variable (coefficient of friction) is confirmed by genetic algorithm [10]. The total number of experiments was 80. The training set (64 training pairs) consists of values of input variables and corresponds to value of the friction coefficient. For the purpose of keeping an independent check on the progress of the learning algorithm from all observations, a validation set containing 8 training pairs (i.e. 10% of the all training samples) was randomly separated. Furthermore, the independent set of 10% training samples was selected and assigned to the test set. This set was used for assessment of the quality of ANN prediction.

To study the effect of activation function and post synaptic potential functions on response of MLP 5:5-11-1:1 network, seven activation functions (table 1) and two post synaptic potential functions (linear and radial) were considered. In the case of MLP network, the highest performance is received for the network trained using back propagation (BP) algorithm. This BP algorithm was used with the following settings: learning rate value was 0.1 and momentum value was set to 0.3 [11].

Comparison of the ANNs with different activation functions and post synaptic potentials was carried out in reference to MLP in which the logistic activation function and linear post synaptic potential were used. As a stop criterion of the training process, the moment at which the value of Root-Mean-Square (RMS) error [12] of the validation set tends to decrease was assumed. Based on this criterion, learning process of the MLP with AF and PSPF was stopped after about 1200 epochs. So, the analysis of effect of the type of AF and PSPF on the quality of learning process was performed after 1200 epochs.

Table 1. The analysed activation functions of the neurons.

Function	Logistic	Hiperbolic	Softmax	Unitsum	Sine	Ramp	Step
Formula	$x = \frac{1}{1-e^{-1}}$	$\frac{e^x - e^{-x}}{e^x + e^{-x}}$	$\frac{e^x}{\sum_i e^{x_i}}$	$\frac{x}{\sum_i x_i}$	sin(x)	-1 $x \leq 1$ x -1< x <+1 +1 $x \geq 0$	0 $x < 0$ +1 $x \geq 0$
Range	(0, +1)	(-1, +1)	(0, +1)	(0, +1)	[0, +1]	[-1, 1]	[0, +1]

3. Results and discussion

Application of *step* AF ensures the highest rate of convergence of training process of the MLP in conjugation with *linear* PSPF. However, it was observed that use of the *unitsum* function considerably increases the time of learning process and the network error decreased just after 450 epochs. The network error for the validation set, for all the used activation functions, except for the *unitsum* AF, is stabilized after about 300 epochs. The highest network training error after 1200 epochs is observed for *unitsum* AF, and the lowest for *step* AF function.

The effect of number of learning epochs on the training and validation sets of an MLP network error for the cases of linear and radial PSPF assumptions are plotted in figure 2 and 3 respectively. As depicted in the figures, the MLPs errors for validation set are characterized by high noise (figures 2b, 3b), while normal effect is observed when the number of training pairs in the validation set is significantly smaller than the number of training pairs in the training set. The *radial* PSPF (figure 3a) eliminates the initial lack of convergence in *linear* PSPF (figure 2a). In this case, the *step* AF also ensures the lowest training error after 1200 epochs.

IOP Conf. Series: Materials Science and Engineering **269** (2017) 012041 doi:10.1088/1757-899X/269/1/012041

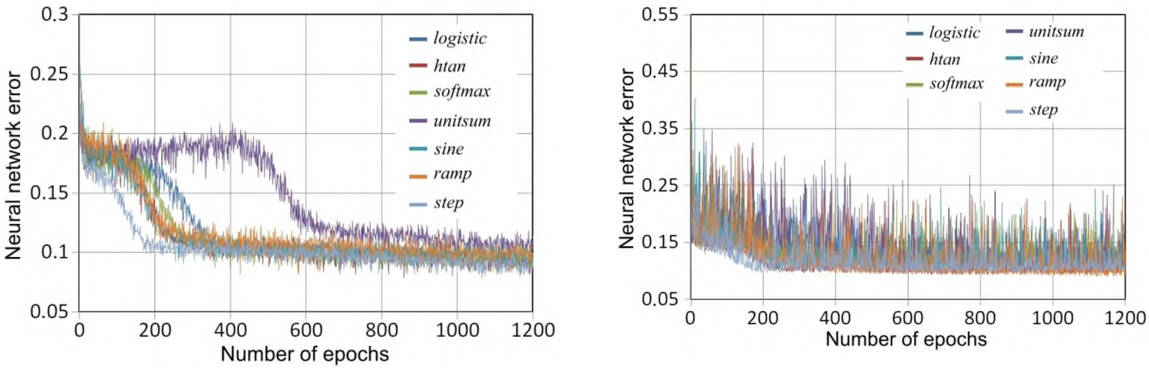

Figure 2. The effect of number of learning epochs on MLP's error of
(a) training set and (b) validation set, in the case of assumption *linear* PSPF.

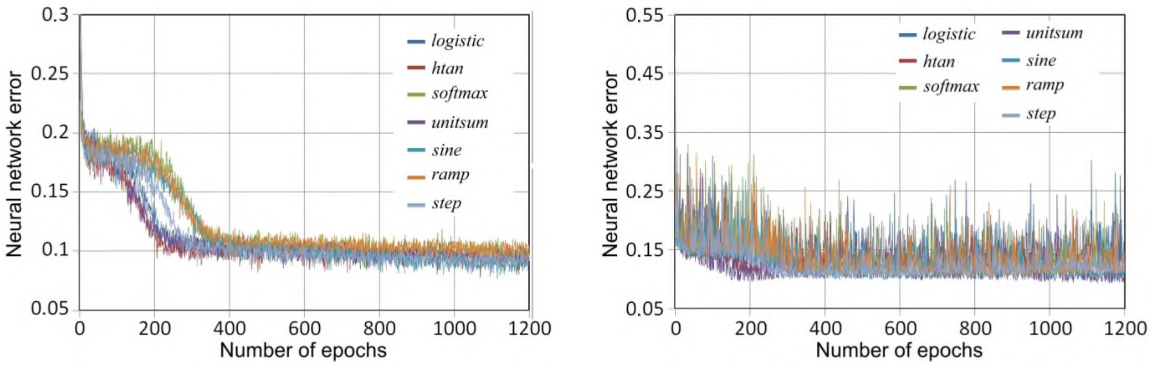

Figure 3. The effect of number of learning epochs on MLP's error of
(a) training set and (b) validation set (b), in the case of assumption *radial* PSPF.

The activation function is used for transformation of the activation level of the neuron in output signal. Typically, activation functions include a "flattening" effect. Together with the PSPF, it defines a neuron type. Similar to the logistic function, the hyperbolic tangent function (*tanh*) is an S-shaped curve, with the difference that the output values are in the range (-1, +1). Because of the symmetry of the *tanh* function, it often works better than the *logistic* function. The *tanh* function is preferred for activation of neurons in multilayer neural networks. The trigonometric AF activation functions (i.e., *sine*) is useful for recognizing radially distributed data. *Ramp* AF is a piecewise linear version of a sigmoid function which ensures relatively low efficiency of learning process, but enabling a very fast operation of the neural network.

When using *linear* PSPF, the highest correlation for training set is observed for *tanh* AF. When using *radial* PSPF, the highest correlation coefficient assures *sine* activation function (figure 4a). The highest difference in the correlation coefficient, when using both considered PSPFs, is noticed in the case of *tanh* and *ramp* AF. The quality of the network prediction is described by the correlation coefficient for the test set, which is the highest for conjugation of *unitsum* AF and *radial* PSPF, and in conjugation of *logistic* AF and *linear* PSPF (figure 4b). It is visible that selection of the PSPF function is a crucial element of ANN modelling. PSPF is a function which evaluates the value of neuron activation based on the input values, weights and threshold value. In the linear PSPF activation, the activation is expressed as a difference between a sum of weighted inputs minus and threshold value. In radial PSPF, the activation is a scaled square of an input vector and vector of weights.

The importance of proper selection of neuron activation function and post synaptic potential function is clearly visible in figure 5. In the case of *softmax* AF the difference in mean error of prediction between *linear* and *radial* PSPFs differs in about 400%.

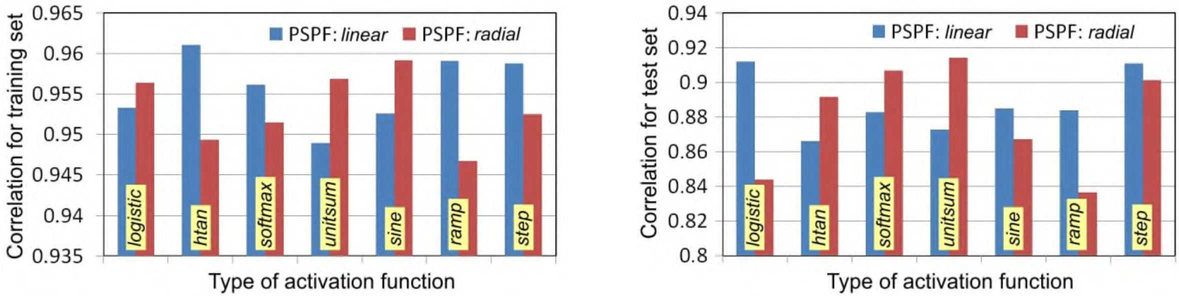

Figure 4. The effect of activation function on correlation coefficient of (a) training set and (b) test set.

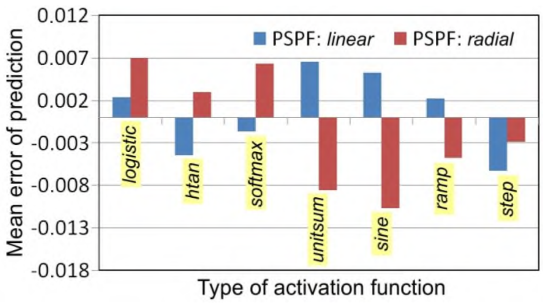

Figure 5. The effect of activation function on value of mean error of MLP prediction.

4. Conclusions

The results of modelling the frictional phenomena in sheet metal forming using artificial neural networks showed that the appropriate choice of the parameters of neuron activations is an important parameter, and this determines the rate of learning process and ability of prediction of ANN. Although, in our research, the conjugation of *softmax* AF and *linear* PSPF assured the minimal error of prediction, both types of the activation and post synaptic potential functions should be selected separately for individual set of training data. The analyses of MLP 5:5-11-1:1 network simulated with seven types of neuron activation functions showed that the difference in prediction error between all approaches may reach about 400%.

5. References

[1] Figueiredo L, Ramalho A L, Oliveira M C and Menezes L F 2011 *Wear* **271** 1651.
[2] Gil I, Galdos L, Mugarra E, Mendiguren J and de Argandona E S 2016 *IOP Conf. Series: Mater. Sci. Eng.* **159** 012019.
[3] Sigvant M, Pilthammar J, Hol J, Wiebenga J H, Chezan T, Carleer B and van den Boogard A H 2016 *IOP Conf. Series: Mater. Sci. Eng.* **159** 012021.
[4] Wang C, Ma R, Zhao J and Zhao J 2017 *J. Mater. Proc.* **27** 126.
[5] Trzepieciński T, Bazan A and Lemu H G 2015 *Int. J. Automot. Technol.* **16** 849.
[6] Wilson W and Chang D 1996 *J. Tribol.* **118** 83.
[7] Hol, J, Cid Alfaro M V, de Rooij M B and Meinders T 2012 *Wear* **286-287** 66.
[8] Challen J and Oxley P 1979 *Wear* **53** 229.
[9] Wilson W and Sheu S 1988 *Int. J. Mech. Sci.* **30** 475.
[10] Trzepieciński T and Lemu H G 2012 *J. Mech. Eng. Autom.* **2** 69.
[11] Lemu H G and Trzepieciński T 2014 *Adv. Mater. Res.* **1051** 204.
[12] STATISTICA Neural Networks. Addendum for version 4. Tulsa: StatSoft, 1999.