26th International Congress of Mechanical Engineering
COBEM 2021
ABCM Associação Brasileira de Engenharia e Ciências Mecânicas
November 22-26, 2021, Virtual Congress, Brazil

COB-2021-1917
ROLLING LOAD PREDICTION FOR THICK PLATES VIA ARTIFICIAL NEURAL NETWORKS

Perseu Silva Soares
Yukio Shigaki
Centro Federal de Educação Tecnológica de Minas Gerais, Belo Horizonte, MG, Brasil
perseusoares@hotmail.com, yukio.shigaki@cefetmg.br

Abstract. *Rolling is the most versatile and commonly used mechanical forming process in manufacturing. In reason of this and the great world market competition, it is imperative to produce better quality products and improve constantly the process control. This can be performed if a good rolling load prediction is carried out, that is a fundamental parameter to properly setup the rolling mill. Due to this fact, the conventional mathematical methods are not sufficient to maintain a good prediction ability for the rolling force, for this parameter prediction involves several nonlinear phenomena. A few conventional methods limitations can be avoided if an artificial neural network (ANN) model is applied to rolling load prediction. Therefore, from a comprehensive study of the fundamentals of hot rolling process, the network input variables were selected, namely: rolling speed, temperature, final thickness, width, reduction, equivalent carbon and work roll diameter. And, from the heavy plate rolling process industrial data it was possible to train a multilayer perceptron neural network with variable learning rate, developed in Python programming language. The neural network performance was compared to the Schultz model performance with the same data base and showed better capability for all three evaluation parameters (coefficient of determination, maximum absolute percentage error and the number of predictions with absolute error greater than 10%). However, Schultz model can be more easily applied and adjusted.*

Keywords: *thick plates, hot rolling, rolling load, artificial neural networks.*

1. INTRODUCTION

Rolling is the most widely used metalworking process due to high production and close control of final product. This is a process in which the thickness of workpiece is reduced by compressive stresses exerted by two rolls rotating in opposite direction. Also, the workpiece is submitted to shear stresses due to the frictional force that draw it into the work rolls (Dieter and Bacon, 1988).

Since the world market demands for high-quality products, it is necessary to develop new methods to improve them and avoid waste. This can be performed if a precise rolling load prediction is carried out, for this parameter is highly related to strip thickness precision (Ginzburg, 1989), and it is important to properly setup the rolling mill with specific material properties (Gorni and Silva, 2012).

There are many mathematical formulations to predict hot rolling load. Some of them are semi-analytical, such as Ekelund's and Sims' models. Others are only analytical such as Orowan's model, others are empirical e.g., Schultz model. However, due to the complex nature of the process and difficulty to properly establish the close relationship between the thermal, mechanical and material phenomena, the hot rolling process turns difficult to formulate and solve. The pure analytical models involve several assumptions that are not satisfied in most of the real cases in industry without some degree of adaptation (Bagheripoor and Bisadi, 2012).

As an alternative to those models, the finite element method (FEM) can be used with less simplifications, being a powerful engineering tool. However, it cannot be used in online application, as the simulation of non-linear problems takes a lot of time and computational effort, not being straightforward for setting the hot rolling mill.

Another option is to use the Artificial Neural Network technology (ANN), that has been used to solve many complex nonlinear problems. The ANN is an algorithm of machine learning inspired by human nervous system, able to store experimental knowledge and make it available for use in the future (Haykin, 2009). Once trained, the ANN is able to predict hot rolling load for different input conditions with good accuracy and precision, inside the training domain.

In this paper we develop an ANN model to predict the rolling load of a thick plate hot rolling mill facility.

The application of this algorithm using supervised learning had the support of the Gerdau's thick plate plant located at Ouro Branco city, Minas Gerais state, Brazil. They provided us a set of production data in order to train the ANN and verify its accuracy.

Thus, this paper proposes a low-cost application of artificial neural networks, since the entire model was developed in Python language using open-source libraries. The ANN model uses a relatively new activation function, that showed better results than the conventional one. And finally, in order to compare the precision of the predictions, an analytical model based on Schultz method was programmed as well.

2. THICK PLATES ROLLING PROCESS

The reheating furnace is the first stage of the thick plates rolling, in which the slab is heated up to 1200-1250 °C to remove the dendrite structures that arise in the casting process, and dissolve most of the alloying elements. The temperature is an important parameter here, because if it is higher than necessary more elements will enter into solid solution, but the plate cost will increase and the thickness of the primary scale will become thicker. On the other hand, if the temperature is too low, some chemical elements will not enter into solid solution and this will affect the metallurgical properties of the product (Lenard, 2014).

After the slab slides from the reheating furnace, the next stop is the descaler, where the scale is removed by high-pressure water spray and/or scale brakers. Then, the roughing process begins in the rolling mill where the largest reductions are performed. In this stand, the rolling gap is usually controlled by a screw system, since this step does not require much precision.

When the rough passes have already been performed, the slab goes to the finishing mill. At this stage, several measurements are taken to feed the process automation system that is responsible to control the roll gap. To provide an accurate positioning during the pass, two hydraulic cylinders are installed and their position is controlled by servo valves and corresponding position controllers. The rolling load can be measured by load cells or calculated from the pressure of the hydraulic cylinders (Kucsera and Béres, 2015).

After finishing passes, the thick plate is cooled under controlled conditions on a run-out table. Cooling water is sprayed on the top and the bottom of the steel surface. In addition to reduce the temperature for transportation, precise cooling process is important to control the microstructure of the product and consequently, the mechanical properties (Lenard, 2014).

3. HOT ROLLING LOAD CALCULATION METHODS

3.1 Von Kárman Equation

Von Kárman proposed Eq. (1) to calculate rolling load based in equilibrium of infinitesimal elements in the roll bite as shown in Figure 1. This formulation involves the following assumptions (Helman and Cetlin, 1983):

- Plane strain deformation
- Homogeneous deformation in each plane
- Constant coefficient of friction (Coulomb friction)
- Circular contact arc (deformed radius R')
- Neutral point inside the arc of contact
- Strip elastic deformation negligible

$$hS\frac{d}{d\emptyset}\left(1-\frac{P}{S}\right)+\left(1-\frac{P}{S}\right)\frac{d(hS)}{d\emptyset}=-2R'P(\sin\emptyset\pm\mu\cos\emptyset) \quad (1)$$

Where:
h: sheet thickness along the arc of contact
S: material flow stress in plane strain state
P: rolling load
R': deformed work roll radius
μ: coefficient of friction
$\emptyset$: angular coordinate

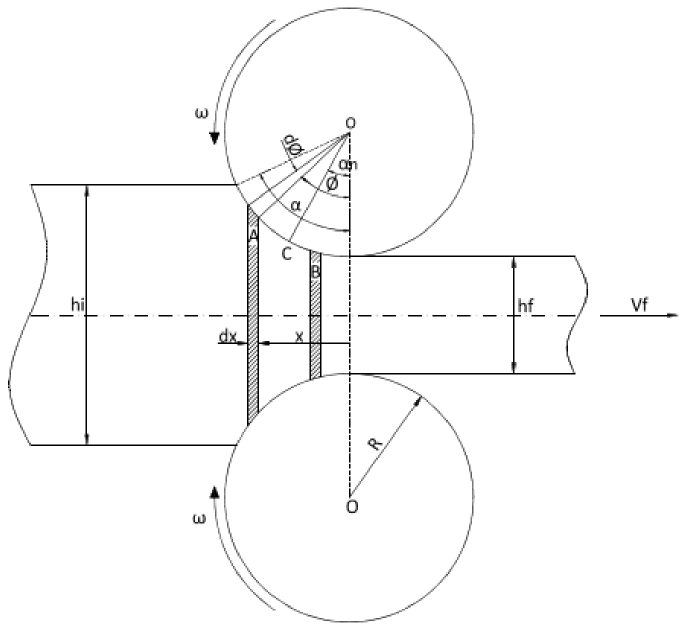

Figure 1. Schematic diagram of rolling process with infinitesimal elements

The high forces involved in the rolling process are transmitted to the workpiece through the rolls, that under these loading conditions suffer two major types of elastic deformations. First, the work rolls tend to bend along their length because of the separating forces generated by workpiece while their displacement are restricted by the bearings on the necks. This implies that the work rolls need to be backed-up by others two more rigid rolls. Second, the rolls flatten in the region where they contact the plate, thus, the rolls curvature radius increase from R to R'. The most famous equation to deformed rolls radius was proposed by Hitchcok in 1950. The Hitchcok's equation, shown below, consider an elliptical distribution of pressure along the arc of contact and that this one remains circular (Dieter and Bacon, 1988).

$$R' = R\left[1 + \left(\frac{c}{h_i - h_f}\right)\left(\frac{P}{w}\right)\right] \tag{2}$$

where $c = 16(1 - v^2)/\pi E$ is a parameter of roll material, v, E, h_i, h_f, w and R are Poisson's ratio of the roll material, Young's modulus of the roll material, initial and final thickness of the plate, plate width and work roll radius, respectively.

Several authors proposed their simplifications and assumptions to solve Eq. (1), as it is only solvable using a numerical procedure.

3.2 Orowan-Pascoe Model

According to Dias *et al.* (2019), Orowan was the first to develop a comprehensive model based in extension of slab method, able to calculate the rolling load, introducing elastic deformation of the work rolls and plastic deformation of the plate. Later Orowan and Pascoe proposed a model to calculate the rolling load, in which consider sticking friction condition.

In this model, the rolling load can be calculated from Eq. (3).

$$P = \bar{S}\, w\, L\, Q_P \tag{3}$$

Where:
$\bar{S}$: material average flow stress in plane strain condition
L: contact arc length
Q_P: geometric factor

$$Q_P = \frac{1}{4}\left[\pi + \sqrt{\frac{R}{h_f}\left(\frac{h_i - h_f}{h_f}\right)}\right] \tag{4}$$

3.3 Ekelund's Model

Ekelund formulated a model to calculate the rolling load, considering that the strip velocity in the roll inlet is smaller than roll tangential speed, gradually increasing, until it exceeds the rolls tangential speed passing after the neutral point (Dias *et al.*, 2019).

The model proposed by Ekelund is defined by the following equations (Helman and Cetlin, 1983):

$$P = \bar{S}\, w\, L\, Q_e \tag{5}$$

Where:

$$Q_e = 1 + \frac{1.6\mu\sqrt{R(h_i - h_f)} - 1.2(h_i - h_f)}{h_i + h_f} \tag{6}$$

And the coefficient of friction can be calculated from Eq. (7).

$$\mu = 0.8(1.05 - 0.0005T) \tag{7}$$

T is the rolling temperature in Celsius degree.

3.4 Sims' Model

Based on Orowan's and Von Kárman's contributions, Sims proposed a model in which considerer that $\sin\emptyset \approx \tan\emptyset \approx 1$ and $1 - \cos\emptyset \approx \emptyset^2/2$. He also assumed that the product of interfacial shear stress and the angular variable is negligible when compared to other terms. In addition to these simplifications, it was considered the hypothesis of sticking friction along the entire contact arc and that the rolled material is characterized as rigid-ideally plastic, which means that the material does not suffer elastic deformation. In other words, its Young's modulus tends to infinity. With these assumptions, it was possible to carry out the integration of equilibrium equation and then obtain the following equations (Lenard, 2014):

$$P = \bar{S}\, w\, L\, Q_s \tag{8}$$

where $L = \sqrt{R'\Delta h}$ is the contact arc and Q_s is the geometric factor, dependent on the radius of the deformed cylinder, R', the thickness reduction, r, the output thickness, h_f, and the thickness of the plate at the neutral point, h_n.

$$Q_s = \frac{\pi}{2}\sqrt{\frac{1-r}{r}}\tan^{-1}\sqrt{\frac{r}{1-r}} - \frac{\pi}{4} - \sqrt{\left(\frac{1-r}{r}\right)\left(\frac{R'}{h_f}\right)}\ln\left(\frac{h_n}{h_f}\right) + \frac{1}{2}\sqrt{\left(\frac{1-r}{r}\right)\left(\frac{R'}{h_f}\right)}\ln\left(\frac{1}{1-r}\right) \tag{9}$$

The location of the neutral point, $\emptyset_n$, is obtained from the Eq. (10).

$$\frac{\pi}{4}\ln(1-r) = 2\sqrt{\frac{R'}{h_f}}\tan^{-1}\sqrt{\frac{R'}{h_f}}\emptyset_n - \sqrt{\frac{R'}{h_f}}\tan^{-1}\sqrt{\frac{r}{1-r}} \tag{10}$$

According to Fonseca *et al.* (2012), the Sims model has been the most used due to good compromise between the precision and simplicity of calculations. However, the original Sims model was developed considering homogeneous deformation along of the workpiece thickness, which satisfies only the rolling of relatively thin materials, with $\frac{L}{h_m} \geq 1$, where h_m is the average thickness of the plate.

In the rolling of thick plates, with small reductions, in which $\frac{L}{h_m} > 1$, the peening effect occurs. In this case, the deformation occurs preferentially on the surface of the material, which characterizes a heterogenous strain model and this effect is responsible to predictions below the measured values (Fonseca *et al*., 2012; Moon and Lee, 2008).

According to Fonseca *et al.* (2012), to consider the peening effect in hot rolling load calculation, it was suggested the Eq. (11), as an adaptation to geometric factor of Sims model. This equation was obtained by regression of process data and suggested by Santos and Giacomin (2010). The last term of the equation represents the peening effect.

$$Q_s = 0.8 + \left(0.45\frac{\Delta h}{h_i} + 0.04\right)\left(\sqrt{\frac{R'}{h_i}} - 0.5\right) + 0.25\frac{h_m}{L} \tag{11}$$

3.5 Schultz's Model

Schultz and Smith Jr. proposed in 1965 a fully empirical model to hot rolling load calculation for online process control. This model stands out due to its simplicity. The constants of the Eq. (12) can be determined by multiple linear regression of the industrial data and therefore, bringing implicit information, such as friction effects, rolls elastic deformation and the resistance deformation of the plate material (Fonseca *et al.*, 2012).

$$\begin{aligned} ln\,F = b_0 + b_1\,ln\left(\frac{R}{h_i}\right) + b_2\,ln\left(\frac{\Delta h}{h_i}\right) + b_3\,ln\left(\frac{R}{h_i}\right)ln\left(\frac{\Delta h}{h_i}\right) + b_4\left(\frac{T}{1000}\right) + b_5\left(\frac{T}{1000}\right)ln\left(\frac{\Delta h}{h_i}\right) \\ + b_6\,ln\left(\frac{\Delta h}{h_i}\right)ln\left(\frac{R}{h_i}\right)^2 + b_7\left(\frac{T}{1000}\right)^2 + ln\,w + \frac{1}{2}[ln\,R + ln\,\Delta h] \end{aligned} \tag{12}$$

Where:
R: average work roll radius (m)
h_i: initial plate thickness (m)
w: plate width (m)
Δh: thickness reduction (m)
T: deformation temperature (°C)
b_0 to b_7: constants

4. ARTIFICIAL NEURAL NETWORK MODELLING

4.1 Selection of input variables

The correct definition of the input variable is one of the most important steps for the correct modeling of the ANN model, and it must have a good correlation between the input variables and the target, otherwise the model will have difficult to converge and find the correct pattern. Known that, a comprehensive study of the hot rolling process was carried out and it was concluded that the rolling force depends upon the flow stress of the material, plate thickness, thickness reduction, plate width, roll diameter and coefficient of friction. The flow stress, on the other hand, mainly depends on the chemical composition, temperature, strain and strain rate. The strain rate, in turn, depends on the rolling speed. The coefficient of friction at roll bite is dependent on the rolling speed and temperature. There are other minor influent parameters, but for the ANN model it is necessary to select only measurable variables.

Therefore, this study considered seven main parameters, namely: plate final thickness, thickness reduction, plate width, rolling average speed, rolling average temperature, work roll diameter and equivalent carbon number.

4.2 ANN model formulation

Due to the nature of the problem, a Multilayer Perceptron (MLP) ANN model was selected. This model structure consists of four layers, which are: one input layer, two hidden layers and one output layer. The input layer consists of 7 nodes, representing the 7 input variables, the hidden layers are of the same dimension, and have m nodes, the output layer has only one node, that is the target. The schematic illustration of the artificial neural network structure is shown in Figure 2. This model was developed in Python programing language, using the open-source web application Jupyter Notebook.

The output values of the nodes for the first hidden layer are given by the following equations:

$$y_k = v_k * \left(\frac{1}{1 + e^{-v_k}}\right) \tag{13}$$

$$v_k = \sum_{j=1}^{n} w_{kj}x_j + b_k \tag{14}$$

where n is the number of inputs of the neuron k, y_k is the output of neuron k, v_k is the induced local field of the neuron k, w_{kj} is the weight referring to input j connected to neuron k, x_j is the input j and b_k is the bias associated to neuron k.

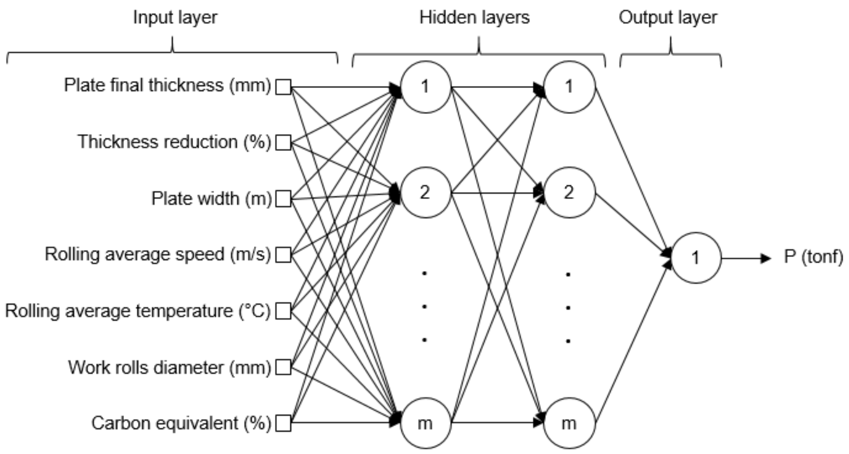

Figure 2. Schematic illustration of the artificial neural network structure

The ANN model was trained by the backpropagation algorithm with gradient descent method optimized by *AdaGrad*. This algorithm minimizes the mean squared error, that is the loss function. The artificial neuron is activated by the *Swish* function, which was proposed by Google researchers in 2017, and got better results than *tanh* function in preliminary tests with the present neural network model and dataset.

To perform the ANN training, it is necessary to initialize the weights randomly, but according to Mishking and Matas (2015) it is known that arbitrary initializations can slow down or even completely paralyze the convergence process. This happens because arbitrary initializations can result in the deeper layers receiving inputs with small variances, which in turn slows down backpropagation, and retards the overall convergence process. In order to initialize deep networks and maintain the activation variance and backpropagated gradients variance, as one moves up or down the network, Glorot and Bengio (2010) proposed the following initializing procedure.

$$W \sim U\left[-\frac{\sqrt{6}}{\sqrt{n_j + n_{j+1}}}, \frac{\sqrt{6}}{\sqrt{n_j + n_{j+1}}}\right] \tag{15}$$

where W is the weight matrix, $U[-a, a]$ is the uniform distribution in the interval $(-a, a)$, n_j is the size of the previous layer and n_{j+1} is the size of the actual layer. The biases were initialized to be 0.

According to Poliak *et al.* (1998), in hot rolling load calculation models it is acceptable that the average load varies up to $\pm$ 10%. One of the reasons for this, is the temperature variation along the plate. At the edges of the plate, the temperature is lower than at its center and, thus, the rolling load at the center of the plate will be relatively lower. Therefore, the performance of the model was evaluated from the scatter plots of the calculated load versus measured load, coefficient of determination (R^2), maximum absolute percentage error (MaxAPE) and the number of predictions with absolute error greater than 10%.

5. PRODUCTION DATA

The data used in this study was collected at Gerdau Ouro Branco’s thick plates plant. This rolling line produces plates for manufacturing products in many sectors, such as: civil construction, wind energy industry, oil and gas industry, naval plants, machine and equipment, etc.

Gerdau’s plant has one 4HI rolling mill, that is responsible to roughing and finishing operations. The Figure 3 shown the picture of the rolling mill.

Figure 3. Thick plates rolling mill (Fraga et al., 2016)

From the above-mentioned equipment and the automation system involved, data were obtained from thirty sketches of ASTM A36 steel, total of 382 samples, with chemical composition varying according to Figure 4. The data collected from the process were: final thickness for each pass, plate width, thickness reduction, work roll diameter, average rolling speed, average rolling temperature, chemical composition and the rolling load which is measured by load cells.

Element	C	Mn	Si	P	S	Al	Nb	Ti	Cu	V	Ni	Cr	Mo
Minimum value	0,14	0,8	0,17	0,011	0,005	0,029	0	0,001	0	0,002	0,01	0,01	0
Maximum value	0,17	0,88	0,21	0,023	0,012	0,043	0,001	0,002	0,01	0,003	0,02	0,02	0,01

Figure 4. Chemical composition of ASTM A36 steel

After data collection, it was executed the data cleansing and then the normalization (or scaling) of these. The normalization process is essential to avoid larger number from overriding the smaller ones and the premature saturation of the hidden nodes, which spoil the learning process (Basheer and Hajmeer, 2000). To perform the scaling was used the Eq. (16).

$$\hat{x}_{\mathrm{i}} = 0.1 + 0.8\left(\frac{x_{\mathrm{i}} - x_i^{\min}}{x_i^{\max} - x_i^{\min}}\right) \tag{16}$$

where $\hat{x}_i$ is the normalized value of x_i, $x_i^{\max}$ and $x_i^{\min}$ are the maximum and minimal value of x_i in dataset, respectively.

In order to train the artificial neural network and later test it with fresh data, the dataset was divided into two subsets, which are: training and test sets. In this case, 75% of the data was used for training and, consequently, 25% of these for testing. The effective distribution was performed using the train_test_split function from the scikit-learn library.

There are n possibilities of datasets divided in the proportion defined above. However, for a particular dataset chosen for training, the network will be able to learn better and that means higher performance during the test. Thus, preliminary tests were carried out with 43 different data sets in the proportion 75-25, and the best distribution was selected to this problem by setting the seed of the random state.

6. RESULTS AND DISCUSSION

With the purpose of the network training, it was necessary to adjust the initial learning rate value and determine the number of epochs. The learning rate value was manually adjusted in simulations of one thousand iterations. The number of epochs was defined from simulations, in which this number varied from one thousand to fourteen thousand, using the optimal learning rate value previously found. In these simulations, the influence of this hyperparameter on the three evaluation parameters was analyzed, and it was found that, as the number of epochs increases, the evaluation metrics tend to experience significant improvements – although the MSE doesn't experience significant change -, but stabilize at approximately ten thousand epochs. From this point, it is quite probable that the random variables involved in the process will be more decisive than the number of epochs.

After the simulations, the learning rate and the number of epochs was defined to 0.0305 and 10000, respectively. The different possibilities of network architecture considered were trained and tested and the results achieved are shown in Figure 5. The network number 7, containing 8 nodes, had the best performance in all three evaluation parameters. The learning graph is shown in Figure 6.

		Training					Test		
ANN Number	m	R^2	MaxAPE	$f_{>10\%}$	MSE	Time (s)	R^2	MaxAPE	$f_{>10\%}$
1	2	0,960	27,01	21	5,40E-04	78,49	0,969	17,13	8
2	3	0,965	26,38	17	4,90E-04	75,05	0,974	12,69	2
3	4	0,970	23,66	13	4,10E-04	80,86	0,977	10,32	3
4	5	0,967	20,68	14	4,60E-04	77,20	0,975	11,23	1
5	6	0,976	17,49	6	3,36E-04	85,33	0,981	11,24	1
6	7	0,971	19,04	9	3,95E-04	84,60	0,980	10,32	1
7	**8**	**0,976**	**16,30**	**3**	**3,34E-04**	**86,82**	**0,982**	**9,05**	**0**

Figure 5. Tried ANN models and their validation metrics for defining the optimal neural network architecture

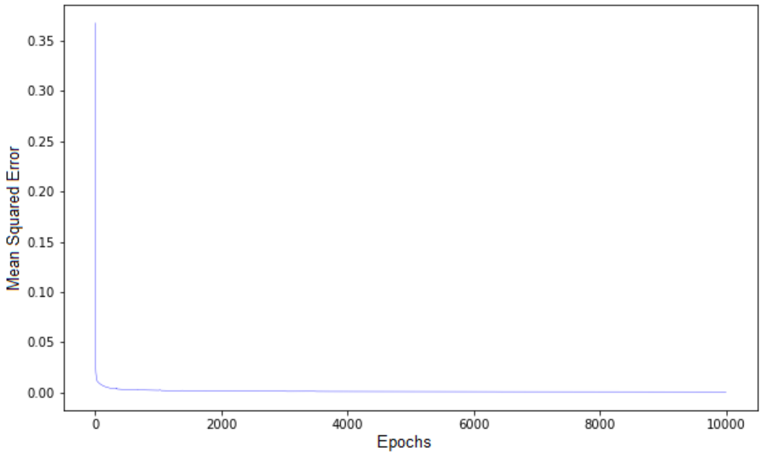

Figure 6. Plot of MSE curve on training set

In order to obtain the rolling load by Schultz model and compare the result with the ANN model, the same training dataset was used to obtain the constants b_0 to b_7 and same test dataset was used to validate the model. The predictions were carried out and the evaluation parameters are shown in Figure 7.

Regression				Test		
R^2	MaxAPE	$f_{>10\%}$	Time (s)	R^2	MaxAPE	$f_{>10\%}$
0,949	22,41	24	0,24	0,960	20,80	8

Figure 7. Results of the validation parameters for the Schultz model

The best network architecture reached a coefficient of determination equal to 0.976, maximum absolute percentage error of 16.30 and a number of 3 predictions that exceeded an absolute error of 10% for the training dataset. The scatter plot is shown in Figure 8(a). For the test dataset, the results were even better, reaching a coefficient of determination of 0.982, maximum absolute percentage error of 9.05, indicating that there were no predictions with absolute error greater than 10%, as shown in Figure 9(a).

The Schultz model reached a coefficient of determination equal to 0.949, maximum absolute percentage error of 22.41 and a number of 24 predictions that exceeded an absolute error of 10% for the training dataset, as shown in Figure 8(b). For the test dataset, the model reached a coefficient of determination of 0.960, maximum absolute percentage error of 20.80 and 8 predictions with absolute error greater than 10%. The scatter plot is shown in Figure 9(b).

This way, analyzing the results and the scatter plots, it is evident that the ANN predictions proved to be more accurate and precise than the Schultz model with both training and test dataset. The ANN training processing time was less than two minutes, but this value is about of 360 times greater than the time required by the computer to perform the multiple linear regression and obtain the constants of the Schultz model.

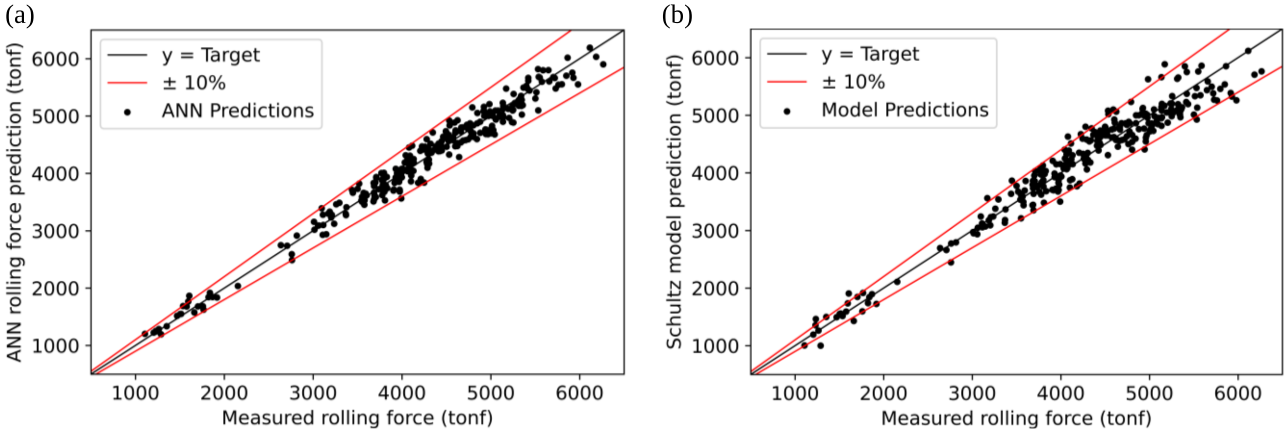

Figure 8. Training performance of the proposed (a) ANN model and (b) Schultz model in prediction of rolling force

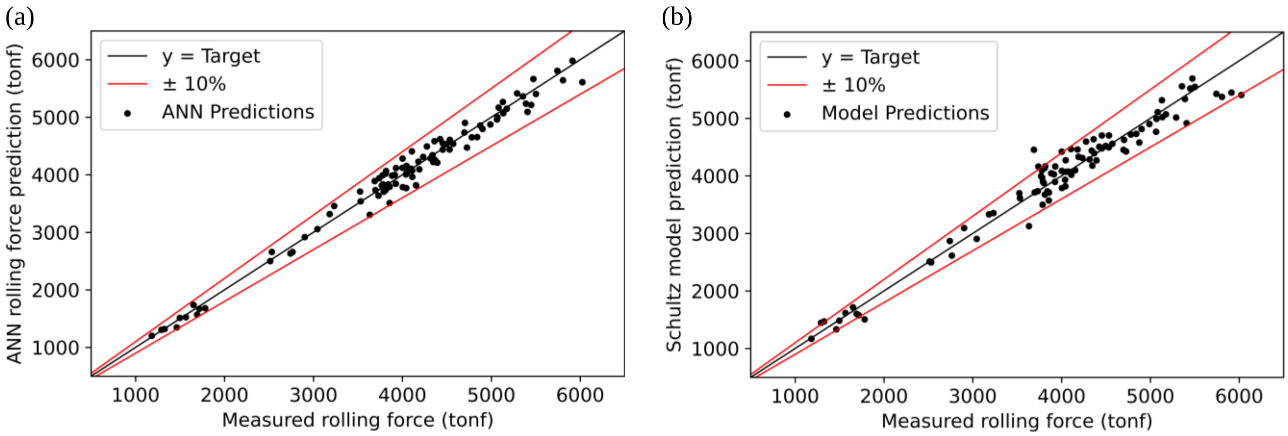

Figure 9. Test performance of the proposed (a) ANN model and (b) Schultz model in prediction of rolling force

Since it is a problem of complex nature, a large amount of data may be necessary to achieve even better results. According to Lek and Guégan (1999), the dataset must be large enough to be representative of the problem. However, estimating the optimal sample size is a very complex task. The references consulted used diffents samples sizes ranging from tens to thousands.

Pereira and Centeno (2017) conducted a study evaluating the size of data samples for supervised learning of artificial neural network for image classification. The authors concluded that as the training samples are increased, there is a tendency to increase the classification accuracy. Although this study was carried out on a classification problem, it is believed that the knowledge extracted can be applied to regression problems as well. Therefore, it is expected that by increasing the training dataset, the ANN model will achieve even more accurate and precise results.

7. CONCLUSIONS

This paper focuses on developing an artificial neural network to predict the hot rolling load, using data from the thick plates rolling process. After training a MLP network with backpropagation algorithm with adaptative learning rate, the evaluation parameters were calculated for different network architectures and selected the optimal one to the problem. The network with eight nodes in hidden layers achieved the best performance for all three evaluations metrics.

The ANN predictions showed a good correlation between the input variables and the rolling force. Furthermore, it can be noted in Figure 8 and Figure 9 that there is a strong match between the predicted and measured roll force, which shows that this model is capable to consider the effect of the seven input variables in the rolling load and, so, can be used as an accurate and precise method to obtain this parameter in online and offline applications.

Compared to the Schultz model, the artificial neural network showed best results for all defined evaluation parameters, reinforcing its precision and accuracy. However, when comparing the ANN training time with the time to calculate the coefficients of the Schultz model, it is observed that the processing time of the ANN algorithm is about 360 times greater, even though this time is small (less than two minutes). But it must be noted that a continuously trained ANN may be implemented, thus being more flexible when facing changes.

According to Dopico *et al.* (2009), several authors have implemented parallel computing using a GPU (Graphics Processing Units) to processing ANN algorithms, and achieved significant reductions in computing time when compared to processing the same algorithm in the CPU (Central Processing Unit). Therefore, the ANN performance can be easily improved using a GPU for the training process and a large data sample, and can become more generic if the range of the input variables increases. In this way, the model will be able to meet a wider application range and become more efficient.

8. ACKNOWLEDGEMENTS

The authors are grateful to Gerdau Ouro Branco for providing the industrial data for this paper. They are also grateful to Emanuelle Garcia Reis for her support.

9. REFERENCES

Bagheripoor, M., Bisadi, H., 2013. "Application of artificial neural networks for the prediction of roll force and roll torque in hot strip rolling process". Applied Mathematical Modelling, v. 37, n. 7, p. 4593-4607.

Basheer, I. A.; Hajmeer, M., 2000. "Artificial neural networks: fundamentals, computing, design, and application". Journal of microbiological methods, v. 43, n. 1, p. 3-31.

Dias, E. D. *et al.* 2019. "Análise comparativa de cargas de laminação a quente industriais com as obtidas através de modelamento matemático". Tecnologia em Metalurgia, Materiais e Mineração, v. 16, n. 3, p. 325-333.

Dieter, G. E., Bacon, D. J, 1988. "Mechanical Metallurgy". New York: McGraw-Hill Book Company.

Dopico, J. R. R. *et al*., 2009. "Encyclopedia of artificial intelligence". New York: Information Science Reference.

Fonseca, N. M. *et al.*, 2012. "Modelo matemático para cálculo da carga de laminação de chapas grossas processadas por resfriamento acelerado". In: Seminário de Laminação – Processos e Produtos Laminados e Revestidos, 49, 2012, Rio de Janeiro, RJ. São Paulo: ABM, 2012. p. 640-649.

Fraga, R. A. *et al.*, 2016. "Laminação de chapas grossas da Gerdau". In: Seminário de Laminação – Processos e Produtos Laminados e Revestidos, 53, 2016, Rio de Janeiro, RJ. São Paulo: ABM, 2016. p. 253-262.

Ginzburg, V. G, 1989. "Steel Rolling Technology – Theory and Practice". New York: Marcel Dekker Inc.

Glorot, X., Bengio, Y., 2010. "Understanding the difficulty of training deep feedforward neural networks". In: Proceedings of the thirteenth international conference on artificial intelligence and statistics. JMLR Workshop and Conference Proceedings, 2010. p. 249-256.

Gorni, A. A., Silva, M. R. S. da. 2013. "Comparação entre os modelos para o cálculo de carga na laminação a quente industrial". Tecnologia em Metalurgia, Materiais e Mineração, v. 9, n. 3, p. 197-203.

Haykin, S. S., 2009. "Neural networks and learning machines". Pearson Education. Upper Saddle River, NJ.

Helman, H., Cetlin, P. R., 1983. "Fundamentos da conformação mecânica dos metais". 2 ed. Rio de Janeiro: Guanabara Dois.

Kucsera, P., Béres, Z., 2015. "Hot rolling mill hydraulic gap control (HGC) thickness control improvement". Acta Polytechnica Hungarica, v. 12, n. 6, p. 93-106.

Lek, S., Guégan, J. F., 1999. "Artificial neural networks as a tool in ecological modelling, an introduction". Ecological modelling, v. 120, n. 2-3, p. 65-73.

Lenard, J. G., 2014. "Primer on flat rolling". 2nd edition, Elsevier.

Mishkin, D., Matas, J., 2015. "All you need is a good init". arXiv preprint arXiv:1511.06422.

Moon, C. H., Lee, Y., 2008. "Approximate model for predicting roll force and torque in plate rolling with peening effect considered". ISIJ international, v. 48, n. 10, p. 1409-1418.

Pedregosa, Fabian *et al*., 2011. "Scikit-learn: Machine learning in Python". the Journal of machine Learning research, v. 12, p. 2825-2830.

Pereira, G. H. de A., Centeno, J. A. S., 2017. "Avaliação do tamanho de amostras de treinamento para redes neurais artificiais na classificação supervisionada de imagens utilizando dados espectrais e laser scanner". Boletim de Ciências Geodésicas, v. 23, n. 2, p. 268-283.

Poliak, E. I. *et al.*, 1998. "Application of linear regression analysis in accuracy assessment of rolling force calculations". Metals and Materials, v. 4, n. 5, p. 1047-56.

Santos, A. A., Giacomin, C. N., 2010. "Mathematical Simulation of Plate Rolling at Usiminas: a Tool for Process Enhancement". In: CONFERENCIA DE LAMINACIÓN, 18, 2010, Rosario. IAS, 2010. p. 113-123.

10. RESPONSIBILITY NOTICE