Luis E. Zárate
zarate@pucminas.br
Departamento de Ciência da Computação
Pontificia Universidade Católica de Minas Gerais
Av. Dom José Gaspar 500, Coração Eucarístico
30535-610 Belo Horizonte, MG, Brazil

Predictive Model for the Cold Rolling Process through Sensitivity Factors via Neural Networks

The mathematical modeling of the rolling process involves several parameters that may lead to non-linear equations of difficult analytical solution. Such is the case of Alexander's model (Alexander 1972), considered one of the most complete in the rolling theory. This model requires excessive computational time, which prevents its application in on-line control and supervision systems.
In this paper, the representation of the cold rolling process through Neural Networks trained with data obtained by Alexander's model is presented. This representation is based in sensitivity factors obtained by differentiating a neural network previously trained. The representation allows to obtain equations of the process for different operation points with low computational time. On the other hand, the representation based in sensitivity factors has predictive characteristics that can be used in predictive control techniques. Through predictive model, it is possible to eliminate the time delay in the feedback loop introduced by measurements of the outgoing thickness, normally with X-ray sensors. The predictive model can work as a virtual sensor implemented via software. An example of the application to a single stand rolling mill is presented.
Keywords*: Rolling mills, steel industry, neural networks, predictive model*

Introduction

The technology of rolling mill automation has advanced rapidly in the last decade and, due to modern computer control systems, reached a high level of sophistication. The technology necessarily embraces a broad spectrum of interests, ranging from fundamental process analysis to the solution of special control-theoretic problems. The objective of all industry is to reach a larger productivity and a better quality of the industrialized products. Nowadays, the metallurgical industry is one of the most required, when considered the current market demand.

Many physical models have non-linear characteristics and analytical complexities that avoid its application in on-line control and supervision systems. The first due to the classic control theory that demands linear models, and the other, because of complex models that request numerical solution with critical computational effort. An example is Alexander's mathematical model (Alexander 1972) considered one of the most complete in the rolling theory. The model involves parameters with complex equations of difficult analytical solution. That model is known by requesting significant computational effort for numerical solution.

It is necessary new methods to represent physical processes. The objective of this paper is to propose a new method to represent the rolling process by means of sensitivity factors. These factors are obtained by differentiating a neural network previously trained with smaller computational effort. This method can be used in on-line control and supervision systems.

There is a growing interest in the application of the Artificial Intelligence (AI) in industrial automation systems. Inside of AI, Neural Networks have been the focus of a great attention during the last years, due to its capacity in solving non-linear problems for learning.

On the other hand, an Automatic Gauge Control (AGC) system for rolling mills uses the output thickness as the feedback variable. Usually the thickness sensor is placed too far from the roll-gap, and this causes a time delay in the feedback control loop, which also depends on the strip speed.

The performance of classical techniques for control systems design usually depends on the existence of a good linear model of the plant dynamics in order to achieve an acceptable design. The lack of a good model is still more compelling in the case of a predictive control, in which a plant model is used to estimate, during the sampling, the process variables that are not available for measurement.

A modern alternative to overcome those problems is the use of ANN to face model uncertainties, time delays, and non-linear plants. This paper introduces a new strategy that permits to obtain a predictive model (Eq. 21) whose parameters are obtained directly from the weights of a trained ANN (Eq. 17).

In this work, an application of neural networks to represent the cold rolling process, based on Alexander's model (1972), is presented. The representation uses the sensitivity factors to determinate the variations of its main parameters. The sensitivity equations will be obtained by differentiating a neural network previously trained. The representation is valid for different operation points and it makes possible the application in on-line control systems.

In the next sections, the representation and basic structure of the artificial neural network are discussed and the procedure to obtain the sensitivity expressions through a neural network are presented. In the last section, the application, the results and the conclusions of this paper are presented.

Paper accepted October, 2005. Technical Editor: Anselmo Eduardo Diniz.

Nomenclature

y = Average yielded tensile stress (N/mm^2)
g = Gap (mm)
h_i = Strip input thickness (mm)
h_o = Strip output thickness (mm)
M = Stiffness Rolling Mill modulus (N/mm)
P = Rolling load (N/mm)
P_B = Rolling load x Strip width (N)
Tq = Rolling Torque (N-mm/mm)
R = Roll radius (mm)
t_f = Front tension stress (N/mm^2)
t_b = Back tension stress (N/mm^2)
W or = Strip width (mm.)
E = Young modulus of the strip material
$U_i,\ i = 0,...,N$ are the net inputs (real values original data sets) and $U_0 = 1$ is a polarization entry

$f_i^a(.)$ $i=0,...,N$ *are the normalization entry functions and* $f_0^a(.)=1$

X_i, $i=0,...,N$ *are the normalized inputs* $X_0=U_0$

W_{ij}^h $i=1,...,L$ e $j=0,...,N$ *is the weight corresponding to the neuron i and input j*

$net_j^h=\sum_{i=0}^{N} W_{ji}^h X_i$ $j=1,...,L$ *product of weights times inputs*

$f_j^h(net_j^h)$ $j=0,...,L$ with $f_0^h(net_0^h)=1$ *is the sigmoid function of the hidden layer.*

I_j, $j=0,...,L$ *are the corresponding values of the sigmoid function* $I_0=1$

W_{ij}^o $i=1,...,M$ e $j=0,...,L$ *is the weight of the neuron i and input j for the hidden layer.*

$net_j^o=\sum_{i=0}^{L} W_{ji}^o I_i$ $j=1,...,M$ *product of the weights times entries for the hidden layer.*

$f_j^o(net_j^h)$ $j=1,...,M$ *is the value of the sigmoid function for the exit layer*

Y_j, $j=1,...,M$ *are the normalized outputs of the net, obtained from the sigmoid function*

$f_i^b(.)$ $i=1,...,M$ *are the an-normalization functions of the outputs*

Z_i, $i=1,...,M$ *net outputs (real values of the original data sets)*

$e\max_k, e\min_k$ $k=1,..,N$ *higher and lower value of the inputs*

$s\max_k, s\min_k$ $k=1,..,M$ *higher and lower value of the outputs*

LimiteInf and LimiteSup are the minimum and maximum values of the original data sets respectively.

Greek Symbols

μ = *Friction coefficient*

Cold Rolling Process

Into classical theories, there are several models to calculate roll load and roll torque necessaries for the cold rolling process. These models are non-linear functions of several parameters, Eqs. (1) and (2). Any change in the input thickness (h_i), in the output thickness (h_o), in the back tension (t_b), in the front tension (t_f), in the yield stress (y) and/or friction coefficient (μ), will cause alterations on the rolling load (*P*), on the rolling torque (*Tq*) and, consequently on the outgoing thickness.

$$P=f(h_i,h_o,t_b,t_f,\mu,y,E,R) \quad (1)$$

$$T_q=f(h_i,h_o,t_b,t_f,\mu,y,E,R) \quad (2)$$

where (*E*) is the Young modulus of the strip material and (*R*) is the roll radius.

In Lianis and Ford (1956), Baptista (1986) and Denti (1994) sensitivity expressions (Eqs. (3) and (4)) were obtained by the expressions (1) and (2), using the Bland and Ford (1948), the Roberts (1965) and the Orowan (1944) models respectively.

$$\Delta P=\frac{\partial P}{\partial h_i}\Delta h_i+\frac{\partial P}{\partial h_o}\Delta h_o+\frac{\partial P}{\partial t_b}\Delta t_b+\frac{\partial P}{\partial t_f}\Delta t_f+\frac{\partial P}{\partial \mu}\Delta\mu+\frac{\partial P}{\partial \bar{y}}\Delta\bar{y} \quad (3)$$

$$\Delta Tq=\frac{\partial T_q}{\partial h_i}\Delta h_i+\frac{\partial T_q}{\partial h_o}\Delta h_o+\frac{\partial T_q}{\partial t_b}\Delta t_b+\frac{\partial T_q}{\partial t_f}\Delta t_f+\frac{\partial T_q}{\partial \mu}\Delta\mu+\frac{\partial T_q}{\partial \bar{y}}\Delta\bar{y} \quad (4)$$

where ($\bar{y}$) is the average yield stress of the material.

The sensitivity coefficients in the expressions (3) and (4) are factors that govern the process and these involve expressions of difficult analytical solution. In Denti (1994), a numerical solution for each coefficient based in curve families was presented. For new operation points new numerical solutions and new curves are necessary. That doesn't stimulate its application in on-line systems.

Artificial Neural Networks Applied to the Cold Rolling Process

In the last years, Artificial Neural Networks (ANN) are being proposed as powerful computational tools due to the low time of processing that can be reached when the net is in operation. These times can be 460 µs. approximately (for a neural network of 6 inputs, 2 outputs and a layer hidden with 13 neurons, in a computer pentium 166 MHz., Zárate et al. 1998 b). Nowadays, ANN's are receiving great attention in metallurgical processes, as can be seen in Andersen, et. al. (1992); Smart, (1992); Gunasekera (1998); Zárate et. al. (1998 a, b); Zárate (1998); Schlang (2001); Zárate and Bittencout (2001); Kim (2002); Zárate and Bittencout (2002); Gálvez and Zárate (2003); Yang (2004) and Son (2004).

In Gunasekera et. al. (1998), it was described a neural network model for the cold rolling process. In this work it was proposed a method to obtain a near-optimal neural network structure based on second-order derivative information obtained directly of the process. In Shlang et. al. (2001) it was proposed a hybrid neuro/analytical process model that is dependent of the considered mill, and which permit to calculate the setup for the mill's actuators. In Yang et. al. (2004) a neural network model, to predict roll load, was presented and this model was implemented to on-line roll-gap control.

The ANN used in this work is a multi-layers net which approaches of the cognitive models that try to describe the operation of the human brain. The type of learning of that net is known as supervised learning based on the method "backpropagation". That neural network uses two or more layers with processing neurons. The entry layer receives the external entries, while the output layer is responsible by the generation of the output of the ANN. If there is a third layer, this receives the name of "hidden layer". The definition of the net structure, as the number of hidden layers and the number of neurons in those layers, is still a problem without solution, although there are some approaches. In the case of the number of neurons for the hidden layer is suggested as 2N+1 neurons, where N is the number of inputs of the net (Kovács 1996).

The neural network considered has two layers and 6 inputs where the number of neurons in the hidden layer was chosen as 13 (2N+1). The number of neurons in the output layer is chosen as 2, corresponding to the number of outputs of the net.

In the Eq. (5) *f* is the non-linear sigmod function chosen in this work as the axon transfer function for being the most consistent with the biophysics of the biological neuron.

$$f=\frac{1}{1+\exp^{-\sum Entries\text{x Weigths}}}=\frac{1}{1+\exp^{-\sum net}} \quad (5)$$

The supervised learning of the ANN uses a train set for the input and output of the net. For each data set, the weights of the net are adjusted to minimize the error in the exits of the net. The error is minimized using the gradient technique with a factor of convergence called "learning rate".

For the training process, it was used the Alexander's model as generator of the database. The inputs of the ANN are: the input thickness (h_i), the output thickness (h_o), the friction coefficient

(μ), the front tension (t_f), the back tension (t_b) and the average yield stress ($\bar{y}$). As outputs were considered the rolling load (P) and the rolling torque (Tq), Eq. (6).

$$(h_i, h_o, \mu, t_b, t_f, \bar{y}) \xrightarrow{ANN} (P, Tq) \tag{6}$$

Generally, the largest care to get a trained neural network lies on collecting and pre-processing neural network input data. The pre-processing operation consists in the data normalization in such a way that the inputs and outputs values will be within the range of 0 to 1.

The following procedure was adopted to normalize the input data before using it in the ANN structure:

a) In order to improve convergence of the ANN training process, the normalization interval [0, 1] was reduced to [0.2, 0.8], because in the sigmod function, Eq. (5), the values [0, 1] aren't reached: $f \to 0$ for $net \to -\infty$ and $f \to 1$ for $net \to +\infty$.

b) The data was normalized through the following formula:

$$f^{a}(Lo) = Ln = (Lo - L\,min) / (L\,max - L\,min) \tag{7a}$$

$$f^{b}(Ln) = Lo = Ln * L\,max + (1 - Ln) * L\,min \tag{7b}$$

where Ln is the normalized value, Lo is the value to normalize, Lmin and Lmax are minimum and maximum variable values, respectively.

c) Lmin and Lmax were computed as follows:

$$Lmin = (4 \text{ x } LimiteInf. - LimiteSup) / 3 \tag{8a}$$

$$Lmax = (LimiteInf. - 0.8 \text{ x } Lmin) / 0.2 \tag{8b}$$

The Eqs. (8a) and (8b) are obtained substituting in the Eq. (7 a) Ln = 0.2 and *Lo* = *LimiteInf*; and *Ln* = 0.8 and *Lo* = *LimiteSup*. Where *LimiteInf* and *LimiteSup* are the minimum and maximum values of the original data sets respectively.

Sensitivity Equations via ANN

In this work, an ANN multi-layer with a hidden layer is used. The ANN has N: inputs, M: outputs and L: neurons in the hidden layer. The differentiation of the neural network is generic and it depends only on N, M, L and on the weights of the hidden and exit layers obtained during the process of training. Figure 1 shows the structure of the neural network considered in this work.

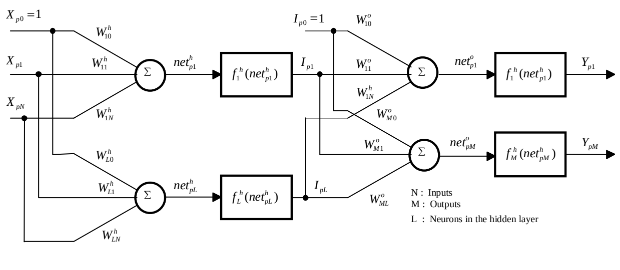

Figure 1. Structure of the Neural Network considered.

The steps to obtain the expressions of the sensitivity factors of the net are as follow:

$$\begin{aligned} Z_1 &= f_1^b(Y_1) \\ Z_2 &= f_2^b(Y_2) \\ &\vdots \\ Z_M &= f_M^b(Y_M) \end{aligned} \tag{9}$$

With an appropriate manipulation of the variables, the Eq. (10) that correlates the inputs with the normalized output of the net can be obtained:

$$\begin{aligned} Z_1 &= f_1^b\Big(f_1^o\Big(\sum_{j=0}^{L} W_{1j}^o f_j^h \Big(\sum_{i=0}^{N} W_{ji}^h f_i^a (U_i)\Big)\Big)\Big) \\ Z_2 &= f_2^b\Big(f_2^o\Big(\sum_{j=0}^{L} W_{2j}^o f_j^h \Big(\sum_{i=0}^{N} W_{ji}^h f_i^a (U_i)\Big)\Big)\Big) \\ &\vdots \\ Z_M &= f_M^b\Big(f_M^o\Big(\sum_{j=0}^{L} W_{Mj}^o f_j^h \Big(\sum_{i=0}^{N} W_{ji}^h f_i^a (U_i)\Big)\Big)\Big) \end{aligned} \tag{10}$$

By substituting the corresponding values for the functions $f^a(.)$, $f^b(.)$, $f^o(.)$, $f^h(.)$, Eq. (11) is obtained:

$$\begin{aligned} Z_1 &= \frac{1}{1+\exp^{-V_1}}\left[\text{smax}_1 - \text{smin}_1\right] + \text{smin}_1 \\ Z_2 &= \frac{1}{1+\exp^{-V_2}}\left[\text{smax}_2 - \text{smin}_2\right] + \text{smin}_2 \\ &\vdots \\ Z_M &= \frac{1}{1+\exp^{-V_M}}\left[\text{smax}_M - \text{smin}_M\right] + \text{smin}_M \end{aligned} \tag{11}$$

where:

$$V_k = \sum_{j=0}^{L} W_{kj}^o f_j^h \Big(\sum_{i=0}^{N} W_{ji}^h f_i^a (U_i)\Big) \quad \text{for } k = 1, \ldots, \text{M} \tag{12}$$

In a general form, Eq. (11) becomes:

$$\begin{aligned} Z_k &= \frac{1}{1+\exp^{-V_k}}\left[\text{smax}_k - \text{smin}_k\right] + \text{smin}_k \\ &\text{for } k = 1, .., \text{M} \end{aligned} \tag{13}$$

The sensitivity factors will be calculated from Eq. (14):

$$\frac{\partial Z}{\partial U} = \begin{bmatrix} \dfrac{\partial Z_1}{\partial U_1} & \dfrac{\partial Z_1}{\partial U_2} & \cdots\cdots & \dfrac{\partial Z_1}{\partial U_N} \\ \cdots\cdots & \cdots\cdots & \cdots\cdots & \cdots\cdots \\ \dfrac{\partial Z_M}{\partial U_1} & \dfrac{\partial Z_M}{\partial U_2} & \cdots\cdots & \dfrac{\partial Z_M}{\partial U_N} \end{bmatrix} \tag{14}$$

where each term of the sensitivity matrix is calculated in the form:

$$\frac{\partial Z_k}{\partial U_i} = [\text{smax}_k - \text{smin}_k] \frac{\exp^{-V_k}}{(1+\exp^{-V_k})^2} \frac{\partial V_k}{\partial U_i} \tag{15}$$

Manipulating the derivative terms of Eq. (15) and taking into account Eq. (12), the following expression is obtained:

$$\frac{\partial V_k}{\partial U_i} = \frac{\partial}{\partial U_i}\Big(W_{k0}^o + \sum_{j=1}^{L} W_{kj}^o f_j^h \Big(\sum_{i=0}^{N} W_{ji}^h f_i^a (U_i)\Big)\Big) \tag{16}$$

by differentiating Eq. (16) and substituting the expression in Eq. (15), the Eq. (17) is obtained, which allows to calculate the sensitivity factors of the net:

$$\frac{\partial Z_k}{\partial U_i}=\begin{bmatrix} R_1W_{11}^o & R_1W_{12}^o & \cdots & R_1W_{1L}^o \\ R_2W_{21}^o & R_2W_{22}^o & \cdots & R_2W_{2L}^o \\ \vdots & \vdots & \vdots & \vdots \\ R_MW_{M1}^o & R_MW_{M2}^o & \cdots & R_MW_{ML}^o \end{bmatrix} \begin{bmatrix} \frac{Q_1W_{11}^h}{emax_1-emin_1} & \frac{Q_1W_{12}^h}{emax_2-emin_2} & \cdots & \frac{Q_1W_{1N}^h}{emax_N-emin_N} \\ \frac{Q_2W_{21}^h}{emax_1-emin_1} & \frac{Q_2W_{22}^h}{emax_2-emin_2} & \cdots & \frac{Q_2W_{2N}^h}{emax_N-emin_N} \\ \vdots & \vdots & \cdots & \vdots \\ \frac{Q_LW_{L1}^h}{emax_1-emin_1} & \frac{Q_LW_{L2}^h}{emax_2-emin_2} & \cdots & \frac{Q_LW_{LN}^h}{emax_N-emin_N} \end{bmatrix} \quad (17)$$

with:

$$Q_k=\frac{\exp^{-(\sum_{i=0}^{N} W_{ki}^h X_i)}}{(1+\exp^{-(\sum_{i=0}^{N} W_{ki}^h X_i)})^2} \qquad k=1,...,\ L$$

$$R_k=(smax_k-smin_k)\frac{\exp^{-V_k}}{(1+\exp^{-V_k})^2}$$

$$k=1,..,\ M$$

and:

$$W^o=\begin{bmatrix} W_{11}^o & W_{12}^o & \cdots & W_{1L}^o \\ W_{21}^o & W_{22}^o & \cdots & W_{2L}^o \\ \vdots & \vdots & \vdots & \vdots \\ W_{M1}^o & W_{M2}^o & \cdots & W_{ML}^o \end{bmatrix}$$

Predictive Model of the Process

During the rolling process, the cylinders are compressed against the strip by a force transmitted by the back rolls. As the rolling mill is not perfectly rigid, the output thickness, can be expressed by the elastic equation of the rolling mill:

$$h_o=g+\frac{P.B}{M} \qquad (18)$$

with (g) being the roll-gap or “gap” and (B) the width of the material being rolled.

Any variation in the gap can be expressed as:

$$\Delta h_o=\Delta g+\frac{\Delta P.B}{M} \qquad (19)$$

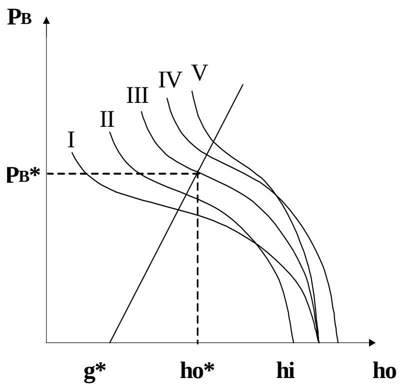

Figure 2. Operation Point of a Rolling Mill.

Figure 2, shows variations in the operation point due to the variation in several parameters. Curve (I) corresponds to an increment in the back and/or in the front tensions of stress. Curve (II) represents an decrement in the input thickness; curve (III) corresponds to the operation point; curve (IV) represents an increment in the output thickness and curve (V) shows possible increment in the friction coefficient and in the average yielded tensile stress or a decrement in the back and/or in the front tensions.

Combining the expressions (3), (4) and (19) the variations in the output thickness of the rolling mill can be obtained:

$$\Delta h_o=\frac{B}{M-B\frac{\partial P}{\partial h_o}}\{\frac{M}{W}\Delta g+\frac{\partial P}{\partial h_i}\Delta h_i+\frac{\partial P}{\partial t_b}\Delta t_b+\frac{\partial P}{\partial t_f}\Delta t_f+\frac{\partial P}{\partial \mu}\Delta\mu+\frac{\partial P}{\partial y}\Delta\bar{y}\} \qquad (20)$$

where: $\Delta h_o=h_o^*-\hat{h}_o$

$$\hat{h}_o=h_o^*-\frac{B}{M-B\frac{\partial P}{\partial h_o}}\{\frac{M}{W}\Delta g+\frac{\partial P}{\partial h_i}\Delta h_i+\frac{\partial P}{\partial t_b}\Delta t_b+\frac{\partial P}{\partial t_f}\Delta t_f+\frac{\partial P}{\partial \mu}\Delta\mu+\frac{\partial P}{\partial y}\Delta\bar{y}\} \qquad (21)$$

Equation (20) uses the partial derivates to calculate the output thickness. The sensitivity factors are obtained by differentiating an ANN and they represent the sensitivities of the rolling load in relationship with the operational parameters (Zárate 1998; Zárate and Helman 1999). It can be observed that Eq. (20) corresponds to the linear terms of the multivariable Taylor’s series expansion of a function (Zárate et al. 1998 b,c).

Equation (21) permits to calculate a predicted value of the output thickness ($\hat{h}_o$), that can be used instead of the current value of h_o, which can not be measured without the interference of a time delay. The value is calculated from the corresponding parameter values for each operating point. In Gálvez and Zárate (2003), it was used a predictive model based in the ANN for thickness control deduced in this paper.

Data for Obtaining the Training Sets

To obtain the training sets, intervals of variation of the parameters were suggested by Bryant (1973), see Table 1.

Table 1. Variations in the rolling parameters.

Limit	h_i $(mm.)$	h_o $(mm.)$	t_b $\frac{N}{mm^2}$	t_f $\frac{N}{mm^2}$	μ	$\bar{y}$ $\frac{N}{mm^2}$
Minimum	4.60	3.492	3.030	62.458	0.096	383.869
Maximum	5.40	3.708	5.619	115.923	0.144	534.940
Increment	0.4	0.108	1.294	26.762	0.024	75.530

Data for the material and rolling mill:
E = 200054 Young modulus of the strip material (N/mm^2)
ν = 0.330 Poisson’s ratio of the strip
B = 900.0 width of strip (mm.)
R = 292.1 roll radius (mm.)
M = 3922640 rigidity rolling mill modulus (N/mm.)
g = roll-gap (mm.)

The values of rolling load and rolling torque used for training were obtained through Alexander's model (1972). Table 2 shows the inferior and superior limits for the load and torque, obtained by simulation

Table 2. Variation of Load and Torque.

	P (N/*mm*.)	Tq (N-*mm*/*mm*)
Minimum *(LimitInf)*	7849.69	3638.35
Maximum *(LimitSup)*	23481.90	185001.51

Training Process

For the training process, aleatory values between -1.0 to +1.0 were considered for net weights. Considering 729 training sets, obtained by the combination of the variables (Table 1) with a learning rate equal to 0.008, the error was of 0.033 after 540,000 training steps (Pentium IV 3.0 GHz).

The final weights for the hidden and output layers with its polarization weight are:

$$W^h_{LxN} = \begin{bmatrix} -2.3454 & 0.7340 & -1.6838 & -0.0022 & 0.6619 & -2.6015 \\ 3.9053 & 3.0809 & -5.1249 & -2.8544 & -5.3234 & -5.6773 \\ -8.4711 & 5.1592 & 6.1539 & -1.5900 & -4.3587 & 3.4279 \\ -11.759 & 0.2278 & -0.9995 & -0.3439 & 0.1404 & 6.6675 \\ -8.4249 & -9.2360 & -5.5525 & 1.1392 & 2.4098 & -5.1437 \\ -1.3987 & 1.6646 & 9.0437 & 0.0068 & 5.0952 & -7.2585 \\ 3.4417 & 0.4525 & 8.6002 & -6.6928 & 5.2036 & 1.8066 \\ -2.0856 & 0.6305 & -0.6184 & 0.0003 & 0.8595 & -1.8555 \\ -4.1919 & 0.5245 & -4.9311 & 7.3975 & 1.3536 & 6.6874 \\ 10.5998 & 3.1858 & -4.8274 & 1.8625 & -3.3113 & -2.8462 \\ 3.8174 & 13.3982 & -3.9207 & -1.5036 & -1.7811 & 0.6947 \\ 3.9009 & -2.6003 & 0.5489 & 2.1262 & -5.7297 & -5.7073 \\ 8.5120 & -6.7926 & -2.3504 & 2.8321 & 4.0079 & -3.0902 \end{bmatrix}$$

$$W^o_{LxM} = \begin{bmatrix} -3.1740 & -2.9035 \\ -0.0764 & 0.0828 \\ -0.0657 & -0.0805 \\ 0.1220 & 0.0392 \\ 0.0129 & 0.0412 \\ 0.0697 & -0.0389 \\ 0.0365 & -0.0332 \\ -2.4985 & -2.2298 \\ 0.0431 & -0.0432 \\ -0.0462 & 0.0340 \\ 0.0237 & 0.0011 \\ -0.1453 & 0.1522 \\ -0.0411 & 0.0515 \end{bmatrix} \quad W^h_{Lx1} = \begin{bmatrix} 5.2529 \\ 6.0337 \\ -3.8969 \\ 5.1587 \\ 5.4216 \\ -1.0686 \\ -7.2569 \\ 0.2394 \\ -3.2777 \\ -0.9662 \\ -5.8385 \\ 3.7317 \\ -3.0089 \end{bmatrix} \quad W^o_{Mx1} = \begin{bmatrix} 3.2143 \\ 2.8875 \end{bmatrix}$$

Obtaining of the Sensitivity Factors

The final weights of the ANN can be substituted in the expression (17) to obtain the sensitivity factors for a certain operation point. Observe that for a new operation point a new calculation on the expression (17) is necessary.

The operation points chosen are:

Operation point (A):

$h_i = 5.00$ mm.;

$h_o = 3.60$ mm;

$\mu = 0.12$;

$t_b = 4.3247$ N/mm^2;

$t_f = 89.220$ N/mm^2;

$y = 256.325 + 468.187\varepsilon^{0.4275}$

Operation point (B):

$h_i = 5.25$mm.;

$h_o = 3.60$mm;

$\mu = 0.126$;

$t_b = 4.5404$ N/mm^2;

$t_f = 98.1444$ N/mm^2;

$y = 269.142 + 492.213\varepsilon^{0.4275}$

Tables 3a and 3b show the sensitivity factors for the chosen operation points A and B respectively. The Tables 4a and 4b compare the output thickness values with variations in the parameters (when $\Delta g = 0$), obtained through ANN Eq. (21).

Table 3a. Sensitivity factors for the operation point "A".

$\frac{\partial P}{\partial h_i}$	$\frac{\partial P}{\partial h_o}$	$\frac{\partial P}{\partial \mu}$	$\frac{\partial P}{\partial t_b}$	$\frac{\partial P}{\partial t_f}$	$\frac{\partial P}{\partial \bar{y}}$
498.66	-631.16	7909.25	-86.56	-5.83	56.75

Table 3b. Sensitivity factors for the operation point "B".

$\frac{\partial P}{\partial h_i}$	$\frac{\partial P}{\partial h_o}$	$\frac{\partial P}{\partial \mu}$	$\frac{\partial P}{\partial t_b}$	$\frac{\partial P}{\partial t_f}$	$\frac{\partial P}{\partial \bar{y}}$
609.54	-1207.62	462.74	-53.77	-0.012	29.32

Table 4a. Calculation of the output thickness obtained through of the sensitivity factors.

Var %	h_i	μ	t_b	t_f	$\bar{y}$	h_f *ANN*	h_o *Alex.*	*Erro %*
0.0	5.00	0.12	4.325	89.220	460.106	3.600	3.590	0.28
-1.0	4.95	-	-	-	-	3.585	3.572	0.36
+1.0	5.05	-	-	-	-	3.615	3.623	0.22
-1.0	-	0.119	-	-	-	3.595	3.588	0.19
+1.0	-	0.121	-	-	-	3.605	3.602	0.08
-1.0	-	-	4.276	-	-	3.600	3.599	0.03
+1.0	-	-	4.363	-	-	3.599	3591	0.22
-1.0	-	-	-	88.328	-	3.600	3.597	0.08
+1.0	-	-	-	90.113	-	3.599	3.594	0.14
-1.0	-	-	-	-	455.506	3.584	3.582	0.06
+1.0	-	-	-	-	464.705	3.616	3.606	0.28
-3.0	4.85	-	-	-	-	3.554	3.519	0.99
+3.0	5.15	-	-	-	-	3.646	3.670	0.65
-3.0	-	0.116	-	-	-	3.581	3.580	0.03
+3.0	-	0.124	-	-	-	3.619	3.613	0.17
-3.0	-	-	4.197	-	-	3.600	3.599	0.03
+3.0	-	-	4.452	-	-	3.599	3.591	0.22
-3.0	-	-	-	86.543	-	3.601	3.599	0.03
+3.0	-	-	-	91.898	-	3.599	3.591	0.22
-3.0	-	-	-	-	446.298	3.551	3.550	0.03
+3.0	-	-	-	-	473.266	3.647	3.632	0.41
-5.0	4.75	-	-	-	-	3.524	3.459	1.88
+5.0	5.25	-	-	-	-	3.676	3.723	1.26
-5.0	-	0.114	-	-	-	3.571	3.572	0.03
+5.0	-	0.126	-	-	-	3.629	3.621	0.22
-5.0	-	-	4.109	-	-	3.601	3.598	0.06
+5.0	-	-	4.540	-	-	3.599	3.592	0.19
-5.0	-	-	-	84.758	-	3.602	3.602	0.00
+5.0	-	-	-	93.682	-	3.598	3.589	0.28
-5.0	-	-	-	-	437.099	3.518	3.530	0.34
+5.0	-	-	-	-	483.112	3.681	3.659	0.60
-3.0	4.85	0.116	4.197	86.543	446.298	3.487	3.459	0.81
+3.0	5.15	0.124	4.452	91.898	473.914	3.713	3.733	0.54

Table 4b. Calculation of the output thickness obtained through sensitivity factors.

Variation (%)	h_i	μ	t_b	t_f	$\bar{y}$	h_o ANN	h_o Alex	*Erro %*
0.0	5.25	0.126	4.540	98.144	496.165	3.600	3.590	0.28
-3.0	5.092	0.122	4.403	95.202	481.278	3.514	3.452	1.76
+3.0	5.407	0.130	4.383	101.086	511.041	3.686	3.744	1.57

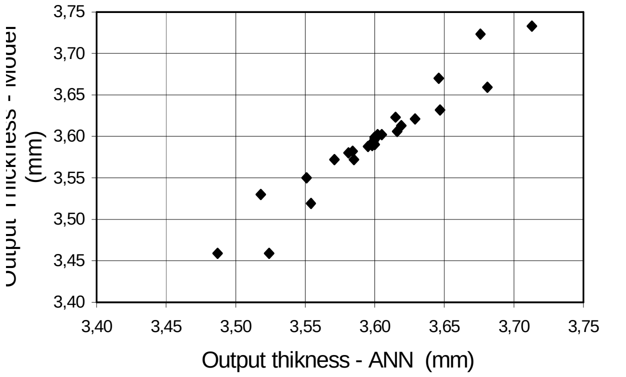

Figure 3. Output thickness obtained by Alexander's Model and through ANN.

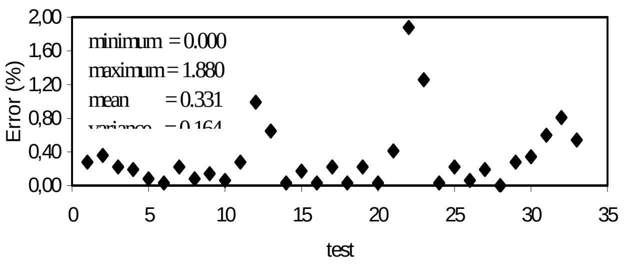

Figure 4. Errors obtained through ANN.

Discussion of the Results

Figure 3 shows the estimated values for output thickness (for data given in Table 4a) obtained through ANN and Alexander's Model. For those values it was possible to calculate the correlation coefficient 0.96 that indicates a satisfactory response of the neural representation for this kind of process. Figure 4 shows minimum and maximum errors reached, 0.00 and 1.88 respectively, and the mean error of 0.331. This last error was calculated for variations of 1%, 3% and 5% in the operational parameters.

In Table 4a, for variations of 1% in each operational parameter, is possible to observe that the output thickness reached a minimum error of 0.03% and a maximum error of 0.36%. For variations of 3% the minimum error was 0.03% and the maximum error 0.99%. For variations of 5% the minimum error was 0.0% while that maximum error increased to 1.88%. While the variation increases, larger became the errors. This happens mainly because the output thickness is calculated through Eq. (21) that uses the sensitivity factors valid to little variations neighboring the operational point correspondent to variations of 0% in nominal parameters, Table 4a. Equation (21) is equivalent to the linearization equation obtained through Taylor's series (Zárate et. al. 1998c).

It is possible to observe that small errors are present in the variations of the back tension, in the front tension and in the average yield stress. This happens because the sensitivity factors for those parameters are the lowest for the operation point chosen. This can be verified in Table 3a, where $|\partial P/\partial t_b| = 86.56$ and $|\partial P/\partial t_f| = -5.83$

Table 4b shows the robustness of the predictive model Eq. (21). For this simulation, each operational parameter had a variation of 3% as recommended in Briant et. al. 1973. The error reached was 1.76%, which was considered acceptable for this kind of process.

Conclusions

In this work, the use of Artificial Neural Networks to represent the rolling process through Alexander's model was presented and discussed. The neural representation uses the sensitivity equations, where the sensitivity factors are calculated through the differentiation of a neural network trained previously.

The neural representation, based on sensitivity factors, allows to calculate the output thickness considering the rolling load, which can be measured directly from the process. This fact can eliminate the thickness sensor, usually X-ray, placed in each stand of a rolling mill.

Observe that Eq. (21) considers as operational parameters those that occur exactly in the cylinder-material deformation area. Thus, the obtained value ($\hat{h}_o$) provides the same value that would be obtained by an X-ray sensor. Therefore, the Eq (21) has predictive characteristics.

The proposed method allows to calculate the sensitivity of the rolling process in any operation points inside the range of data sets for which the net was trained. For operation points out of the training groups and out of the space of generalization of the net, it is necessary a new training process.

The neural representation proposed becomes useful when the sensitivity factors, obtained through complex physical models, have complex analytical expressions with numerical solution with critical computational effort. Through the proposed method, it is possible to apply techniques based on predictive model in on-line control systems.

Refererences

Alexander. J.M.. 1972. "On the theory of rolling", Proc. R. Soc. Lond. A. 326. pp. 535-563.

Andersen, K., Cook, G.E. and Barnett, R.J., 1992, "Gas Tungsten Arc Welding Process Control Using Artificial Neural Networks", International Trends in Welding, Science and Technology, ASM, pp. 877-1030.

Baptista. E. 1986, "Uma metodologia para o desenvolvimento de modelos em variável de estados para laminadores de chapas a frio", Anais do 6° Congresso Brasileiro de Automática. Belo Horizonte. Brasil. pp. 585-589.

Bland. D.R. and Ford. H. 1948, "The calculation of roll force and torque in cold strip rolling with tensions", Proc. Inst. Mech. Eng.. Vol. 158. pp.174.

Bryant. G.F.. Edwards. W.J. and McClure. C.H. 1973, "Automation of Tandem mills", London: The Iron and Steel Institute. Cap. 1.pp.1-29.

Denti. J. 1994, "Um método de controle dinâmico de laminadores reversíveis", Doctoral Thesis. EEUFMG. Belo Horizonte. Brasil.

Gálvez, J.M, Zárate, L.E. and Helman H. "A Model-Based Predictive Control Scheme for Steal Rolling Mills using Neural Networks", Revista Brasileira de Ciências Mecânicas, RBCM vol. XXV, no. 1, 2003

Gunasekera, J.S., Zhengjie Jia, Malas J.C., Rabelo, L. "Development of a neural network model for a cold rolling process", Engineering Applications of Artificial Intelligence, Vol. 11, pp. 597-603, 1998.

Kim, D.H., Lee, Y. and Kim B.M. "Applications of ANN for the dimensional accuracy of workpiece in hot rod rolling process", Journal of Materials Processing Technology", Vol. 130-131, pp. 214-218, 2002.

Kovács. Z.L. 1996, "Redes Neurais Artificiais", Edição Acadêmica São Paulo, Cap 5, pp.75-76. São Paulo. Brasil.

Lianis. G. and Ford. H. 1956, "Control equations of multistand cold rolling mills", University of London. pp. 201-216.

Orowan. E. 1944, "The calculation of roll pressure in hot and cold rolling", Proc. Inst. Mech. Eng.. Vol. 150. pp. 140-167.

Roberts. W.L. 1965, "A simplified cold rolling model", Iron and Steel Eng. Vol 42. n° 10. pp. 75-87.

Shlang, M. Lang, B. Poppe, T., Runkler T., Weinzier K. “Current and Future development in neural computation in steel processing”, Control Engineering Practice, Vol. 9, pp. 975-986, 2001.

Smartt, H.B., 1992, “Intelligent Sensing and Control of Arc Welding”, International Trends in Welding, Science and Technology, ASM, pp. 843-851.

Son, J.S., Lee, D.M., Kim, I.S. and Choi, S.K. “A study on genetic algorithm to select architecture of a optimal neural network in the hot rolling process”, Journal of Materials Processing Technology”, Vol. 153-154, pp. 643-648, 2004.

Zárate, L.E., 1998, Doctoral Thesis, "Um Método para Análise da Laminação Tandem a Frio", Universidade Federal de Minas Gerais, Escola de Engenharia Metalurgica e de Minas, Belo Horizonte, MG, Brasil.

Zárate. L.E.; Gálvez. J.M. and Helman H. 1998a. "A Neural Networks Based Controller for Steel Rolling Mills By Using Sensitivity Functions". XII Brazilian Automatic Control Conference. Vol. I. pp.29-34 - Set 14-18. MG. Brasil.

Zárate. L.E.; Song. M. and Helman H. 1998b. "Algoritmo para Linearização de Processos via redes neurais e sua análise de comportamento assintótico". V Simpósio Brasileiro de Redes Neurais. Dez. Belo Horizonte. Brasil. pp.199-204

Zárate. L.E.; Helman H. and Gálvez. J.M. 1998 c. "Um Método para Linearização de Modelos Utilizando Redes Neurais e sua Aplicação em Processos de Laminação". VIII Congreso Latinoamericano de Control Automático". Vol. II. pp.709-714 - Nov. 9-13. Viña del Mar. Chile.

Zárate. L.E. and Helman H. 1999. "Determination of the Thickness control parameters of the rolling process through the sensitivity method. using neural networks". The Second International Conference on Intelligent Processing and Manufacturing of Materials. Jul. 10-15. Big Island. Hawaii.

Zárate L.E. and Bittencout, F.R., 2001. “Analise Qualitativa de Processos Através dos Fatores de sensibilidade via Redes Neurais e sua Aplicação na Laminação em Tandem”. V Congresso Brasileiro de Redes Neurais, V.1 p.421-426, Rio de Janeiro, Rj, Brazil.

Zárate L.E. and Bittencout, F.R., 2002. “Controle da Laminação a frio Baseado em Redes Neurais com Capacidade de Generalização e Lógica Nebulosa via Fatores de Sensibilidade”. XIV CBA, Vol. I, September, Natal, RN, Brazil.

Yang, Y.Y., Linkens, D. and Talamantes-Silva, “Roll load prediction – data collection, analysis and neural network modelling”, Journal of Materials Processing Technology”, Vol. 152, pp. 304-315, 2004.