Hybrid Learning of Interval Type-2 Fuzzy Systems Based on Orthogonal Least Squares and Back Propagation for Manufacturing Applications

Gerardo Mendez, Angeles Hernandez

Abstract:

This paper presents a novel learning methodology based on the hybrid algorithm for interval type-2 (IT2) fuzzy logic systems (FLS). Since in the literature only back-propagation method has been proposed for tuning of both antecedent and consequent parameters of type-2 fuzzy logic systems, a hybrid learning algorithm has been developed. The hybrid method uses recursive orthogonal least-squares method for tuning of consequent parameters as well as the back-propagation method for tuning of antecedent parameters. The systems were tested for three types of inputs: a) interval singleton b) interval type-1 (T1) non-singleton, c) interval type-2 non-singleton. The experimental results of the application of the hybrid interval type-2 fuzzy logic systems for scale breaker entry temperature prediction in a real hot strip mill were carried out for three different types of coils. They proved the feasibility of the systems developed here for scale breaker entry temperature prediction. Comparison with type-1 fuzzy logic systems shows that the hybrid learning interval type-2 fuzzy logic systems improve performance in scale breaker entry temperature prediction under the tested condition.

Keywords: *type-2 fuzzy inference systems, type-2 neuro-fuzzy systems, hybrid learning, uncertain rule-based fuzzy logic systems.*

1. Introduction

Interval type-2 (IT2) fuzzy logic systems (FLS) constitute an emerging technology. In [1] both one-pass and back-propagation (BP) methods are presented as IT2 Mamdani FLS learning methods but only BP is presented for IT2 Takagi-Sugeno-Kang (TSK) FLS systems. The one-pass method generates a set of IF-THEN rules by using the given training data one time, and combines the rules to construct the final FLS. When BP method is used in both IT2 Mamdani and TSK FLS, none of antecedent and consequent parameters of the IT2 FLS are fixed at starting of training process; they are tuned using exclusively BP method. In [1] recursive least-squares (RLS) and recursive orthogonal least-squares (OLS) algorithms are not presented as IT2 FLS learning methods.

The aim of this work is to present and discuss the hybrid learning algorithm for antecedent and consequent parameters tuning during training process for IT2 Mamdani FLS. In the forward pass, the FLS output is calculated and the consequent parameters are tuned using OLS [2] method. In the backward pass, the error propagates backward, and the antecedent parameters are tuned using the BP method. One of the proposed hybrid algorithms elsewhere [3, 4] is based on RLS, since it is a benchmark algorithm for parameter estimation or systems identification. It has been shown [3, 4] that hybrid algorithms improve convergence over the BP method. Convergence of the proposed methods has been practically tested.

Since in the literature, only the back propagation (BP) learning method for IT2 FLS has been proposed, in this work a hybrid learning algorithm for IT2 FLS (OLS-BP) is developed and implemented for temperature prediction. This motivated by the success of the hybrid learning method in type-1 (T1) FLS (ANFIS) [5] over BP only method. Convergence has been practically tested for particular conditions; it is not the purpose of this work the generalization of the algorithm developed here, but only to show preliminary comparative results and feasibility of application. Mathematical proof is still to be done in general for hybrid learning algorithms.

The IT2 FLS were trained using two main learning mechanisms: the back-propagation (BP) method for tuning of both antecedent and consequent parameters and the hybrid training method using recursive orthogonal least-squared (OLS) method for tuning of consequent parameters as well as the BP method for tuning of antecedent parameters. In this work, the former is referred to as IT2 FLS (BP), and the latter as hybrid IT2 FLS (OLS-BP).

IT2 FLS is an emerging technology [1] that accounts for random and systematic components [6] of industrial measurements. Non-linearity of the processes is handled by FLS as identifiers and universal approximators of nonlinear dynamic systems [7, 8, 9, 10]. Such characteristics give IT2 FLS a great potential to model and control industrial processes.

A second but very important purpose of the present work is to estimate the temperature of the incoming bar head-end at scale breaker (SB) entry in a real hot strip mill (HSM) by a hybrid IT2 FLS (OLS-BP) with learning capabilities. Several IT2 FLS were designed and developed for head-end SB entry temperature estimation, and preliminary experimental results are presented and analyzed. Such experiments show it is feasible to apply the hybrid IT2 FLS (OLS-BP) for HSM entry temperature estimation. Although preliminary, validation was carried out with adaptation, since it is the ultimate goal, the main reason for the application of these techniques is their adaptation capabilities. It is important to notice that the experiments run here were carried out for three different types of products separately, while in practice, the same model should be running for all product types sequentially as they are rolled.

In temperature prediction, the inputs of the hybrid IT2 (OLS-BP) fuzzy models, used to predict the SB entry temperatures, are the surface temperature of the transfer bar at the roughing mill (RM) exit (x_1) and the time required by the transfer bar head to reach the SB entry zone (x_2). Currently, the surface temperature is measured using a pyrometer located at the RM exit side. Scale grows at the transfer bar surface producing a noisy temperature measurement. The measurement is also affected by environment water steam as well as pyrometer location, calibration, resolution and repeatability. The head end transfer bar traveling time is estimated by the finishing mill set-up (FSU) model using the finishing mill estimated thread speed. Such estimation has an error associated with the inherent FSU model uncertainty. Although temperature prediction (y) is a critical issue in a HSM the problem has not been fully addressed by fuzzy logic control systems [11-13].

This paper is organized as follows. Section 2 gives the fundamentals of the OLS parameter estimation algorithm. In section 3, the hybrid learning algorithm developed for temperature prediction is presented. A brief introduction of the HSM process fundamentals is given in Section 4. Section 5 deals with the application of the hybrid IT2 FLS for HSM temperature prediction and the experimental results are presented in Section 6. Section 7 summarizes the conclusions.

2. Fundamental Principles of the Orthogonal Least Squares

As mentioned, a hybrid learning IT2 FLS (OLP-BP) is used for SB entry temperature prediction in an HSM. The hybrid learning algorithm is based on BP and OLS learning methods. In this section, a brief presentation of the basic principles of the OLS method is presented. Since IT2 and BP are very well established methodologies the reader is refer to [1], [14] and [15] respectively.

Suppose that, as in [2], a particular system has one input $u(k)$ and one output $y(k)$ with an additive noise $e(k)$ measured during a certain number t of time periods of T, then it is possible to describe it's dynamic behavior using the next differences model:

$$y(k) = \sum_{j=1}^{n} a_j y(k-j) + \sum_{j=0}^{n} b_j u(k-j) + e(k) \tag{1}$$

where $k = 1,2,3...t;\ a_j,\ b_j \in R$ and $n =$ system order. This can be written in more compact form:

$$y(k) = \mathbf{p}^T \mathbf{z}(k) + e(k) \tag{2}$$

where:

$$\mathbf{p}^T = [b_0, a_1, b_1, ..., a_n, b_n] \tag{3}$$

is the parameter estimation matrix of sizeand:

$$\mathbf{z}^T(k) = [u(k), y(k-1), u(k-1), ..., y(k-n), u(k-n)] \tag{4}$$

is the measurements vector of size $2n+1$.

The model (2) can be expressed for t input-output data pairs as:

$$\mathbf{Y}^T(t) = \mathbf{P}^T \mathbf{Z}^T(t) + \mathbf{E}^T(t) \tag{5}$$

where the output vector of size t, is:

$$\mathbf{Y}^T(t) = [y(1), y(2), ..., y(t)] \tag{6}$$

the measurements matrix of size $(2n+1)\times t$, is

$$\mathbf{Z}^T(t) = \begin{bmatrix} u(1), u(2), ..., u(t) \\ y(0), y(1), ..., y(t-1) \\ u(0), u(1), ..., u(t-1) \\ .... \\ .... \\ y(1-n), y(2-n), ..., y(t-n) \\ u(1-n), u(2-n), ..., u(t-n) \end{bmatrix} \tag{7}$$

and the noise vector of size t is:

$$\mathbf{E}^T(t) = [e(1), e(2), ..., e(t)] \tag{8}$$

For the estimation of $\mathbf{P}$, it is required to minimize the next criteria:

$$\mathbf{J} = (\mathbf{Y}(t) - \mathbf{Z}(t)\mathbf{P}(t))^T \mathbf{I} (\mathbf{Y}(t) - \mathbf{Z}((t))\mathbf{P}(t)) \tag{9}$$

The symmetric and positive matrix $\mathbf{C}(t+1)$ of size $(2n+1)\times(2n+1)$ is defined as:

$$\mathbf{C}(t+1) = [\,\mathbf{Z}^T(t+1)\mathbf{Z}(t+1]^{-1} \tag{10}$$

which works as a covariance attenuation matrix of the identification process.

On the other hand, the linear equation system

$$\mathbf{A}x = b \tag{11}$$

where $\mathbf{A}$ is a matrix of size $m\times n$, x is a vector of size n, b is a vector of size m, and $m>n$, does not have an exact solution, and can be written as:

$$\mathbf{A}x - b = e \tag{12}$$

where e, a vector of size m, is the error of any solution of (11). If:

$$\mathbf{A}^T\mathbf{A} = \mathbf{F}^T\mathbf{F} \tag{13}$$

where $\mathbf{F}$ is any upper or lower triangular matrix of size n, then (11) can be written as:

$$\mathbf{F}x = (\mathbf{F}^T)^{-1}\mathbf{A}^T b \tag{14}$$

A least-squares solution can be found using (14).

Now, considering the orthogonal transformation or rotational matrix defined as

$$\mathbf{T}^T = \mathbf{T}^{-1} \tag{15}$$

Rewriting (12) as

$$[\mathbf{A} : \mathbf{b}] \begin{bmatrix} \mathbf{x} \\ -1 \end{bmatrix} = \mathbf{e} \tag{16}$$

where $\mathbf{D} = [\mathbf{A} : \mathbf{x}]$ is a matrix of size $m \times (n+1)$, and $\mathbf{x}' = \begin{bmatrix} \mathbf{x} \\ -1 \end{bmatrix}$ is a vector of size $(n+1)$. Now applying the orthogonal transformation matrix $\mathbf{T}$ to (6), we can obtain:

$$\mathbf{TDx}' = \mathbf{Te} \tag{17}$$

If $\mathbf{D}' = \mathbf{TD} = \mathbf{F}$ is a triangular matrix, and $\mathbf{b}' = \mathbf{Te} = (\mathbf{F}^T)^{-1}\mathbf{A}^T\mathbf{b}$ then (17) and (14) are equivalent. $\mathbf{F}$, the resulting upper or lower triangular matrix of size n is the squared-root of (13).

It is possible to apply the orthogonal transformation solution to equations system for parameters identification of discrete models.

The least-squares solution of (5) can be expressed as

$$[\mathbf{Z}^T(t)\,\mathbf{Z}(t)]\mathbf{P}^T = \mathbf{Z}^T(t)\,\mathbf{Y}(t) \tag{18}$$

which can be obtained through the orthogonal transformations algorithm. This equation can be reduced to an equivalent triangular system:

$$\mathbf{F}(t)\mathbf{P} = \mathbf{q}(t) \tag{19}$$

where $\mathbf{F}(t)$ is the square root of $\mathbf{Z}^T(t)\,\mathbf{Z}(t)$ and $\mathbf{q}(t)$ is vector of size of $2n+1$. This method can be used on-line with $\mathbf{F}(0) = 0$ and $\mathbf{q}(0) = 0$ as the initial conditions.

For each period of time, the previous algorithm [2] reduces to zero one row of the compound vector $[\mathbf{z}^T(t)\ \mathbf{y}^T(t)]$ of size $2n+2$. The parameters of $\mathbf{P}(t)$ can easily be calculated by the REDCO routine given in [2].

3. The Hybrid Learning Methodology for IT2 FLS (OLS-BP)

3.1 Limitations of Hybrid Learning for IT2 FLS

In [1], only BP is proposed as learning algorithm for IT2 FLS. During backward pass antecedent and consequent parameters are estimated as shown in Table 1.

Table 1. One pass in learning procedure for IT2 FLS.

	Forward Pass	Backward Pass
Antecedent Parameters	Fixed	BP
Consequent Parameters	Fixed	BP

In the hybrid algorithm developed here, recursive OLS is used during forward pass for consequent parameters tuning, and BP method during backward pass for antecedent parameters tuning as shown in Table 2. This hybrid learning method is an extension of the ANFIS training method proposed for T1 FLS [5, 16].

Table 2. Two passes in the hybrid learning for IT2 FLS.

	Forward Pass	Backward Pass
Antecedent Parameters	Fixed	BP
Consequent Parameters	OLS	Fixed

According to [1], there are three points that prevent the use of OLS for consequent parameters estimation in T2 FLS:

1. The starting point for the OLS method to designing an interval singleton FLS is a T1 FBF expansion. No such FBF expansion exists for a general singleton T2 FLS. It looks like a least-square method can be used to tune the parameters in $\boldsymbol{y}^T_l$ (matrix transpose of M left-points $\boldsymbol{y}^l_l$ of consequent centroids) and $\boldsymbol{y}^T_r$ (matrix transpose of M right-points $\boldsymbol{y}^l_r$ of consequent centroids). Unfortunately, this is incorrect. The problem is that, in order to know the FBF $\boldsymbol{p}_l(x)$ and $\boldsymbol{p}_r(x)$, each $\boldsymbol{y}^l_l$ and $\boldsymbol{y}^l_r$ (the M left-points and right-points of interval consequent centroids) must be known first. Since initially there are no numerical values for those elements, it is impossible to do this; hence the FBF $\boldsymbol{p}_l(x)$ and $\boldsymbol{p}_r(x)$ cannot be calculated. This situation does not occur for T1 FBF expansion. A T2 FLS output $y(\boldsymbol{x})$ is expressed by:

$$y(\mathbf{x}) = \frac{1}{2}\left[\mathbf{y}^T_l\,\mathbf{p}_l(\mathbf{x}) + \mathbf{y}^T_r\,\mathbf{p}_r(\mathbf{x})\right] \tag{20}$$

2. Although $\boldsymbol{y}_l$ and $\boldsymbol{y}_r$ (the end-points of T2 FLS center-of-sets type-reduced set Y_{COS}) can be expressed as an interval $\left[\underline{f}^l, \overline{f}^l\right]$ in terms of their lower ($\underline{f}^l$) and upper ($\overline{f}^l$) M firing sets, whereas the corresponding M consequent left $\boldsymbol{y}^l_l$ points can be expressed as:

$$y_l = y_l\left(\overline{f}^1,\ldots,\overline{f}^L,\underline{f}^{L+1},\ldots,\underline{f}^M, y^1_l,\ldots,y^M_l\right) \tag{21}$$

and the corresponding M consequent right-points $\boldsymbol{y}^l_r$ as:

$$y_r = y_r\left(\underline{f}^1,\ldots,\underline{f}^R,\overline{f}^{R+1},\ldots,\overline{f}^M, y^1_r,\ldots,y^M_r\right) \tag{22}$$

where L is the index to the rule-ordered FBF expansions at which $\boldsymbol{y}_l$ is a minimum, and R is the index at which $\boldsymbol{y}_r$ is a maximum; they are not known in advance [1]. Once the points L and R are known, (6) is very useful to organize and describe the calculations of $\boldsymbol{y}_l$ and $\boldsymbol{y}_r$.

3. The next problem has to do with the re-ordering [1] of $\boldsymbol{y}^l_l$ and $\boldsymbol{y}^l_r$. The T1 FBF expansions have always had an inherent rule ordering associated with them; i.e. rules R^1, R^2,…, R^M always established as the first, second,…, and Mth FBF. In T2 FBF this order is lost and it is necessary to be restored for later use.

Here, the previous points were overcome using the following approach:

1. Since the values of $\boldsymbol{y}^l_l$ and $\boldsymbol{y}^l_r$ have been initially fixed as initial condition, it is possible to use OLS method for consequent centroids left-point and right-point parameters estimation using the standard deviation of the variable at each calculation.
2. The values of $\boldsymbol{p}_l(x)$ and $\boldsymbol{p}_r(x)$ from (5) can be calculated using the initial values of $\boldsymbol{y}^l_l$ and $\boldsymbol{y}^l_r$, and then use them as the base for OLS estimation methods.
3. The lost rule-ordered FBF expansions can be restored [1] and used for next consequent centroids estimation $\boldsymbol{y}^l_l$ and $\boldsymbol{y}^l_r$.

3.2 The Hybrid Algorithm for IT2 FLS

The membership functions of the developed hybrid training method are Gaussian functions and they are based on initial conditions of the antecedent parameters (x_1 and x_2, the interval mean $m^l_k \in [m^l_{k1}, m^l_{k2}]$ and the standard deviation σ^l_k), consequent parameters (y, y^l_l and y^l_r) and measurement parameters (the interval of standard deviation $\sigma_{Xk} \in [\sigma_{X1}, \sigma_{X2}]$). Antecedent and measurement parameters are tuned using the BP training method, while consequent parameters are tuned using OLS training method. Given N input-output training data pairs, the hybrid training algorithm for E training epochs, should minimize the error function:

$$e^{(t)} = \frac{1}{2}\left[f_{s2}\left(\mathbf{x}^{(t)}\right) - y^{(t)}\right]^2 \tag{23}$$

The hybrid training algorithm is as follows:

1. Initialize all parameters in antecedent and consequent membership functions. Choose the mean values of the T2 gaussian fuzzy numbers to be centered at the measurements and initialize the standard deviation interval end-points of these numbers.
2. Set the counter, ep, of the training epoch to zero; i.e., $ep \equiv 0$.
3. Set the counter, t, of the training data to unity; i.e., $t \equiv 1$.
4. Apply the input $\boldsymbol{x}^{(t)}$ to the IT2 FLS and compute the total firing interval for each rule; i.e. compute $\underline{f}^l$ and $\overline{f}^l$.
5. Compute y_l and y_r using the iterative method described in [1]. Having done this, establish L and R values.
6. Compute the defuzzified output, $f_{s2}(\boldsymbol{x}^{(t)})$.
7. Determine the explicit dependence of y_l and y_r on membership functions. Because L and R obtained in step 5 usually change from one t-iteration to the next, the dependence of y_l and y_r on membership functions will also usually change from one t-iteration to the next.
8. Test each component of $\boldsymbol{x}^{(t)}$ to determine the active branches.
9. Tune the parameters of the active branches of the consequent using OLS algorithm.
10. Tune the parameters of the active branches of the antecedent's membership functions using the steepest descent algorithm.
11. Set $t \equiv t+1$. If $t \equiv N+1$ (N is the input data vector size), go to step 12; otherwise, go to step 4.
12. Set $ep \equiv ep+1$. If $ep \equiv E$ (E is the total number of epochs), STOP; otherwise go to step 3.

Fig. 1 shows a flow chart of the hybrid training algorithm.

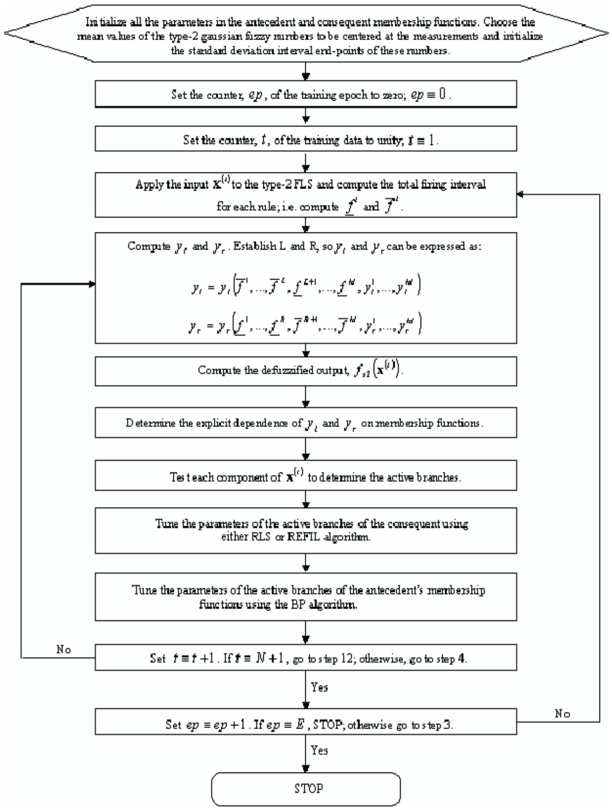

Fig. 1. A flow chart of the hybrid training.

4. Hot Strip Mills

In a Hot Strip Mill (HSM), as in any other industrial process, keeping the quality requirements such as thickness and finishing temperature (the latter determines strip mechanical properties) is a major concern. The most critical section of the coil is the head-end. This is due to the uncertainties involved at the head-end of the incoming steel bar, and the varying conditions from bar to bar. Currently, in order to achieve the head-end quality requirements, there are automation systems based on physical modeling, particularly in the Finishing Mill (FM) [11].

The models calculate the rolling variables as temperature, force, and mill stretch in order to Set-Up the FM, i.e. setting the initial FM controller references such as cylinder position, motor speed, strip tension, etc. In order to calculate force and stretch, the models use, in most cases, material deformation curves, one per steel grade or family, and mill stretch curves. The former curves are adapted mainly from force errors via a gain. Besides, there are commonly used additive terms to correct prediction error based on gage and roller gap errors, which generally come from proportional (P) or proportional plus integral (PI) like structures [12]. Since heat bar conservation is crucial, the model calculations have to be performed on-line and at the shortest possible time.

The market is becoming more competitive and worldwide and therefore more demanding [13, 17]. It requires a more stringent control of quality parameters and a more flexible manufacturing, capable of rolling a wider gamma of products in shorter periods of time. Such flexibility requirements, yield higher time varying conditions for the rolling process, thus demanding automation systems with higher adapting capabilities. Most commercial systems compensation techniques (P or PI based) only compensate for error under current conditions, therefore, the first of batch coils frequently present out of specification head-ends [18]. In recent years, research on estimation of process variable in a HSM by Adaptive Neural Networks (ANN) and Fuzzy Logic Systems (FLS) has received particular attention worldwide. ANN and FLS offer the advantages of reliably representing highly non-linear relations, automatically updating the knowledge they contain and providing fast processing times [13, 19].

One of the first works carried out in this field is presented in [17]. An FLS was developed in order to compensate the thickness error based on the force prediction error in two intermediate stands in the FM. The FM Set-Up is performed by the conventional methods. A system which integrates the operator expertise and a FLS to Set-Up gaps and speeds in the FM is presented by Watanabe *et al.* [19]. Several ANN based systems for FM variable estimation, such force, stack temperature, and full set-up have been proposed [11, 18, 19, 20, 21, 22]. In [3, 4], an IT2 FLS with a hybrid learning algorithm for head-end scale breaker (SB) entry temperature prediction is presented, however, the systems are trained to predict mathematical model estimation.

Strip resistance, and therefore, force and gap set-up, highly depends on the strip temperature. Strip temperature of the incoming bar is also essential for speed set-up, since finishing temperature depends on entry temperature. On the other hand, temperature measurement is highly uncertain. SB entry mean and surface temperatures are used by the Finishing Mill Set-Up (FSU) model [12] to preset the FM stand screws and to calculate the transfer bar (TB) thread speed, both required to achieve, respectively, the FM exit target head gage and FM exit target head temperature. However, the bar surface temperature measurement at SB entry is not reliable due to scale formation, therefore it is measured using a pyrometer located at the roughing mill (RM) exit side and later the head-end bar SB entry temperature is estimated and used for FM Set-Up. The measurement at RM exit is affected by noise produced by TB scale growth, environmental water steam, pyrometer location, calibration, resolution and repeatability. The head-end TB traveling time from RM exit to SB entry is estimated by FSU model. This estimation is associated with the inherent FSU model imprecision. Although temperature prediction is a critical issue in a HSM the problem has not been fully addressed by FLS or ANN.

Fig. 2 depicts a simplified diagram of a HSM, from its initial stage, the reheat furnace entry, to the final stage, the coilers.

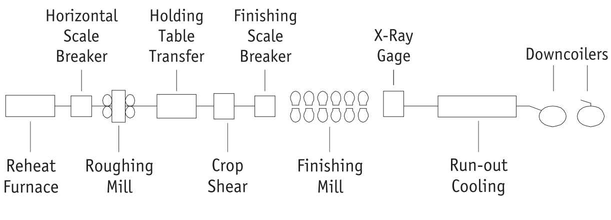

Fig. 2. Schematic view of a typical hot strip mill.

The slab leaves the furnace at about 1300°C and it is transported to the RM by the transfer table. Fig. 1 shows the particular case of a rolling process with two reversible RM´s. The RM gives in several passes the initial thickness reduction to the slab, from about 200mm to about 25.4 mm. The sub-product of the RM is called Transfer Bar (TB).

The TB is taken to the FM where final gage, finishing temperature and final width specifications have to be fulfill. Due the TB length, the transfer stage between RM and FM is about 90m. During the traveling time from RM to FM, scale formation on the TB surface takes place. The scale is washed out by the SB in order to allow proper rolling of the bar.

A great potential to ensure good quality lies in the automation systems and the used of close loop control techniques. The most critical process in a HSM is the FM. The FSU model calculates the finishing mill working references required to fulfill quality specifications at the FM exit stand. The FSU inputs are FM exit target gage, target width, target temperature, steel grade, and hardness ratio from slab chemistry, load distribution, gauge offset, temperature offset, roll diameters, TB gauge, TB width and estimated TB entry temperature. It is very important for the FSU model to have accurate information of the SB entry temperature, since a temperature error would propagate through the entire FM. Due to scale formation, it is not possible to measure the TB temperature at the FM entry, and therefore it has to be estimated from measurements at the RM exit after the last pass.

Because of the complexities and uncertainties involved in the rolling process, the development of mathematical theories for temperature calculations has been largely restricted to one dimensional (assuming infinite length and width) or two-dimensional (assuming only infinite length) models applicable to heat behavior in flat rolling operations.

The physical model estimates the SB entry temperature from the RM exit temperature after the last pass and the TB traveling time to SB entry zone. Therefore the premises to the developed T2 FLS will be the inputs to the physical model: RM exit temperature (x_1) and TB traveling time to SB entry zone (x_2), while the consequent (y) will be the SB entry temperature.

5. TB Surface SB Entry Temperature Prediction

The architecture of the hybrid IT2 FLS for SB entry temperature prediction was established in such a way that parameters are continuously optimized. As mentioned, the antecedents were chosen to be the RM exit surface temperature (x_1) and the TB head traveling time (x_2). These are the inputs to the physical model used for head-end temperature estimation in most of the HSM industrial sites, and they are considered to be the variables which most influence the SB entry temperature. Each antecedent-input space was divided into five fuzzy sets (searching for the compromise between reasonable good performance and a low number of sets to keep low computational resources demand) thus, having twenty-five rules. The output (consequent, y) is the head-end SB

entry surface temperatures.
Gaussian primary membership-functions with uncertain means were chosen for both, antecedents and consequents. Each rule of the three T2 FLS was characterized by six antecedent membership function parameters (two for left-hand and right-hand bounds of the mean and one for standard deviation, for each of the two antecedent Gaussian membership functions) and two consequent parameters (one for left-hand and one for right-hand end points of the centroid of the consequent T2 fuzzy set).

The Gaussian primary membership function with uncertain means for each antecedent is defined as:

$$\mu_k^l(x_k) = \exp\left[-\frac{1}{2}\left[\frac{x_k - m_{kn}^l}{\sigma_k^l}\right]^2\right] \tag{24}$$

where $m_k^l \in [m_{k1}^l, m_{k2}^l]$ is the uncertain mean, $k=1, 2$ (the number of antecedents), $l=1, 2,..25$; $n=1, 2$ (the lower and upper bounds of the uncertain mean) and σ_k^l is the standard deviation. The means of the antecedent fuzzy sets were initially chosen to be uniformly distributed over the entire input space.

Using initially the calculated mean and standard deviation from measurement of input (x_1) and input (x_2) the values of the antecedent five intervals of uncertainty were established. The initial intervals of uncertainty for input (x_1) were selected as shown in Table 3. Fig. 3 shows the initial membership functions for the antecedent fuzzy sets of input (x_1). The values of the initial intervals of uncertainty for input (x_2) were selected as shown in Table 4. Fig. 4 depicts the initial membership functions for the antecedent fuzzy sets of input (x_2).

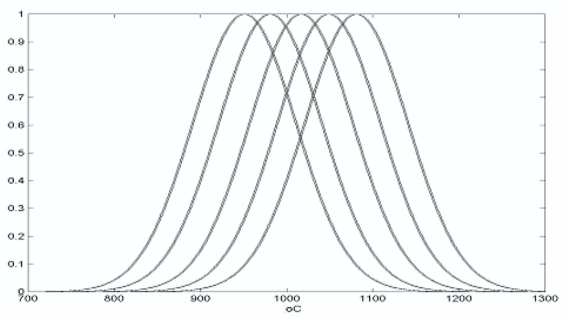

Fig. 3. Membership functions for the antecedent fuzzy sets of x_1 input.

Table 3. Selected intervals of uncertainty for input (x_1).

	m_{11} (°C)	m_{12} (°C)	σ_1 (°C)
1	950	952	60
2	980	982	60
3	1016	1018	60
4	1048	1050	60
5	1080	1082	60

Table 4. Selected intervals of uncertainty for for input (x_2).

	m_{12} (s)	m_{22} (s)	σ_2 (s)
1	32	34	10
2	38	40	10
3	42	44	10
4	48	50	10
5	56	58	10

For the case of IT2 SFLS the resulting T2 FLS uses T1 singleton fuzzification, maximum t-conorm, product t-norm, product implication, and center-of-sets type-reduction, for fuzzy operations see [5]. For IT2 NSFLS-1 the resulting T2 FLS uses T1 non-singleton fuzzification, maximum t-conorm, product t-norm, product implication, and center-of-sets type-reduction. And for the case of IT2 NSFLS-2, the resulting T2 FLS uses T2 non-singleton fuzzification, maximum t-conorm, product t-norm, product implication, and center-of-sets type-reduction.

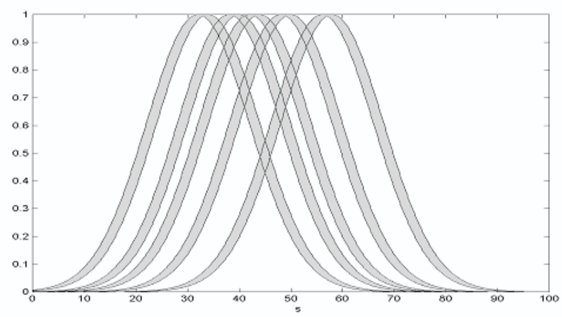

Fig. 4. Membership functions for the antecedent fuzzy sets of x_2 input.

For parameter optimization, the hybrid learning methods, OLS-BP, is used. Experimental results to compare them against BP only method is presented.

The primary membership functions for each input of IT2 NSFLS-1 was:

$$\mu_{X_k}(x_k) = \exp\left[-\frac{1}{2}\left[\frac{x_k - x_k'}{\sigma_{X_k}}\right]^2\right] \tag{25}$$

where: $k = 1,2$ (the number of T1 non-singleton inputs) and $\mu_{Xk}(x_k)$ centered at the measured input $x_k = x'_k$. The standard deviation of RM exit surface temperature measurement σ_{X1} was initially set to 13.0°C and the standard deviation head-end traveling time measurement σ_{X2} was initially set to 2.41s. These values were also selected experimentally.

The primary membership function for each input of IT2 NSFLS-2 was:

$$\mu_{X_k}(x_k) = \exp\left[-\frac{1}{2}\left[\frac{x_k - x_k'}{\sigma_{kn}}\right]^2\right] \tag{26}$$

where: $\sigma_{kn} \in [\sigma_{k1}, \sigma_{k2}]$ $k=1,2$ (the number of T2 non-

singleton inputs), $n = 1, 2$ (the lower and upper bounds of the uncertain mean) and $\mu_{Xk}(x_k)$ centered at the measured input $x_k = x'_k$. The uncertain standard deviation $[\sigma_{11}, \sigma_{12}]$ of RM exit surface temperature measurement was initially set as [11.0, 14.0]°C and the uncertain standard deviation $[\sigma_{21}, \sigma_{22}]$ of head-end traveling time measurement was initially set to [1.41, 3.41]s.
Noisy input-output data pairs of three different coil types with different target gage, target width and steel grade were taken and used as training and validation data, see Table 5, and experiments were carried out for these different coil types.

Table 5. Material type coils.

	Target gage (mm)	Target width (mm)	Steel grade (SAE/AISI)
Coil A	1.95	1104.0	1006
Coil B	5.33	1066.0	1009
Coil C	3.04	939.0	1045

The standard deviation of temperature noise σ_n was initially set to 1.0°C and the standard deviation of time noise σ_n was set to 1.0 s.

The T2 fuzzy rule base consists of a set of IF-THEN rules that represents the model of the system. The interval non-singleton T2 have two inputs $x_1 \in X_1$ and $x_2 \in X_2$ and one output $y \in Y$, which have a corresponding rule base size of $M = 25$ rules of the form:

$$R^l : IF \quad x_1 \quad is \quad \tilde{F}_1^l \quad and \quad x_2 \quad is \quad \tilde{F}_2^l, \qquad THEN \quad y \quad is \quad \tilde{G}^l \tag{27}$$

where $l = 1,2,\ldots 25$. These rules represent a fuzzy relation between the input space $X_1 \times X_2$ and the output space Y, and it is complete, consistent and continuous [30], as shown in Table 6.

The primary membership function for each consequent is a Gaussian function with uncertain means, as defined in (8). Since the center-of-sets type-reducer replaces each consequent set $C_{\tilde{G}_l}$ by its centroid, then y^l_l and y^l_r are the consequent parameters.

Initially, only the input-output data training pairs $(x^{(1)}:y^{(1)}), (x^{(2)}:y^{(2)}),\ldots, (x^{(N)}:y^{(N)})$ are available and there is no data information about the consequents, hence the initial values for the centroid parameters y^l_l and y^l_r may be determined according to the linguistic rules from human experts or be chosen arbitrarily in the output space [16]. In this work the initial values of parameters y^l_l and y^l_r are such that the corresponding membership functions uniformly cover the output space. Table 6 also shows the initial values of consequent centroids of the twenty-five rules.

Table 6. One pass initial value of antecedent and consequent parameters of IT2 FLS.

l	m_{11}	m_{12}	σ_1	m_{21}	m_{22}	σ_2	y^l_l	y^l_r
1	950	952	60	32	34	10	938	940
2	950	952	60	38	40	10	933	935
3	950	952	60	42	44	10	928	930
4	950	952	60	48	50	10	924	926
5	950	952	60	56	58	10	920	922
6	980	982	60	32	34	10	958	960
7	980	982	60	38	40	10	954	956
8	980	982	60	42	44	10	950	952
9	980	982	60	48	50	10	946	948
10	980	982	60	56	58	10	942	944
11	1016	1018	60	32	34	10	978	980
12	1016	1018	60	38	40	10	974	976
13	1016	1018	60	42	44	10	970	972
14	1016	1018	60	48	50	10	966	970
15	1016	1018	60	56	58	10	962	964
16	1048	1050	60	32	34	10	998	1000
17	1048	1050	60	38	40	10	994	996
18	1048	1050	60	42	44	10	990	992
19	1048	1050	60	48	50	10	986	988
20	1048	1050	60	56	58	10	982	984
21	1080	1082	60	32	34	10	1020	1022
22	1080	1082	60	38	40	10	1016	1018
23	1080	1082	60	42	44	10	1012	1014
24	1080	1082	60	48	50	10	1008	1010
25	1080	1082	60	56	58	10	1002	1004

6. Experimental Results

Three different designs: IT2 SFLS, IT2 NSFLS-1 and IT2 NSFLS-2 were trained to predict the SB entry temperature and then tested. Three different sets of data for the three different coil types mentioned and shown in Table 5 were taken from a real-life mill. Each of these data sets was split into two sets, training and validation sets, taking every other data point. Experiments were run for each product type set independently. For each input-output data pairs, for each product type set, the twenty-five rules of the nine FLS were tuned.

The performance evaluation for each of the learning methods is based on the root mean-squared error (RMSE) criteria:

$$RMSE = \sqrt{\frac{1}{n}\sum_{k=1}^{n}\left[\mathbf{Y}(k) - \mathbf{f}_{s2}\left(\mathbf{x}^{(k)}\right)\right]^2} \tag{28}$$

where $\mathbf{Y}(k)$ is the output validation data vector i.e. the actual SB entry temperature measurements vector for system evaluation, different to the training data vector but from the same coil type, and $\mathbf{f}_{s2}(\boldsymbol{x}^{(k)})$ is the temperature vector predicted by the tested T2 FLS.

Figures 5, 6 and 7 show the RMSE of the three used IT2 FLS systems for type A coils after fifteen epoch computations. The behavior of type B and C coils is similar and not shown here for the sake of briefness.

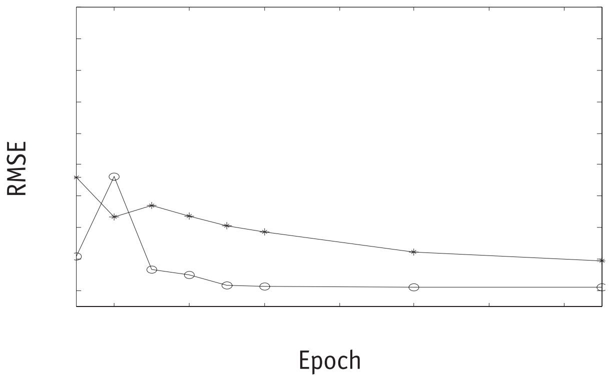

Fig. 5. IT2 SFLS (*) RMSE (BP); (o) RMSE Hybrid (OLS-BP).

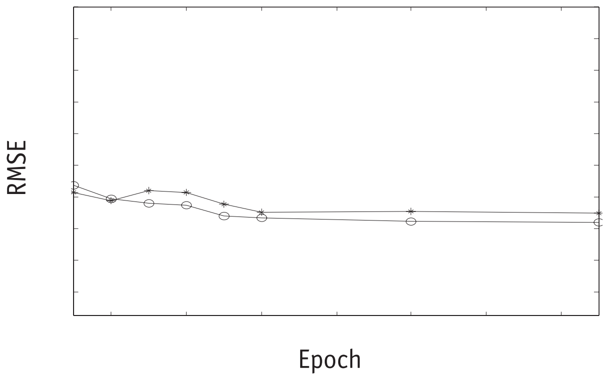

Fig. 6. IT2 NSFLS-1 (*) RMSE (BP); (o) RMSE Hybrid (OLS BP).

In the Figures, the horizontal axis represent the number of training epochs, while in the vertical axis, the RMSE of the validation run with the test set after the corresponding number of training epochs in the horizontal axis is shown, as outlined in [23]. Therefore, the initial value is the validation RMSE after one epoch training. Fifteen epochs were chosen for display purposes since convergence has already taken place for all the systems and for all the experiment. As mentioned, the results presented here are preliminary with the purpose of showing convergence and feasibility of T2 FLS in SB temperature prediction in HSM. However, validation was carried-out allowing tuning (unlike the results of [23]), since the on-line adaptation capabilities mainly motivated rolling variables estimation by either FLS or ANN, and such would be the ultimate applications of these systems.

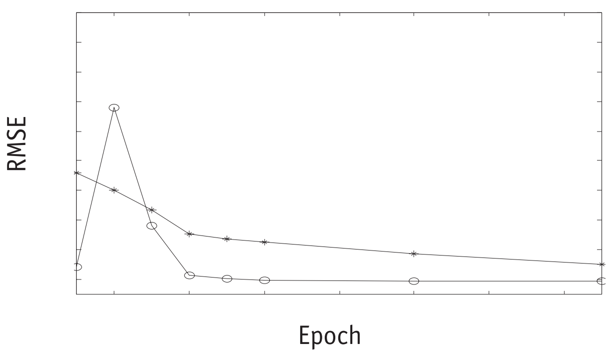

Fig. 7. IT2 NSFLS-2 (*) RMSE (BP); (o) RMSE Hybrid (OLS-BP).

Table 7. Final values of antecedent and consequent parameters at epoch fifteen of hybrid IT2 NSFLS-1 (OLS-BP).

l	m_{11}	m_{12}	σ_1	m_{21}	m_{22}	σ_2	y_l^l	y_r^l
1	947.4	951.7	57.0	32.2	33.2	12.3	830.5	986.3
2	949.8	952.4	59.0	37.5	39.5	11.2	803.4	943.1
3	950.4	951.5	59.9	40.4	42.4	12.9	987.2	1036.2
4	950.4	952.0	60.5	44.8	46.4	12.6	878.2	955.0
5	950.6	951.7	60.6	55.2	55.9	16.0	948.2	970.1
6	979.0	980.7	58.0	26.5	34.4	11.1	917.1	975.5
7	979.7	983.3	61.4	36.3	43.8	9.6	605.3	1088.7
8	978.5	982.3	59.1	41.1	45.4	12.0	909.5	952.6
9	979.0	982.9	59.6	44.2	45.1	14.7	858.8	911.0
10	980.0	981.4	59.7	52.4	56.9	18.0	839.8	1098.4
11	1015.3	1017.7	59.7	35.0	40.7	15.6	913.3	959.7
12	1016.0	1019.0	62.1	38.8	41.6	17.9	550.4	784.8
13	1016.9	1017.0	60.7	39.5	42.8	14.1	742.5	945.7
14	1016.0	1019.3	60.7	50.1	59.7	39.2	1006.8	1067.5
15	1016.0	1017.6	59.7	54.1	58.4	16.0	988.8	991.3
16	1049.1	1050.8	58.6	31.1	38.5	17.0	1112.9	1363.0
17	1047.5	1050.3	59.7	38.2	39.9	12.8	919.5	1319.7
18	1047.7	1048.8	61.7	42.2	42.2	15.4	782.9	870.6
19	1047.7	1050.1	59.9	45.6	49.5	11.6	966.7	972.9
20	1048.0	1050.0	60.0	57.5	62.3	0.6	901.9	922.9
21	1080.3	1082.0	59.9	31.6	34.1	14.4	1029.5	1342.1
22	1078.6	1080.6	62.0	35.6	36.9	14.3	1004.5	1277.6
23	1079.8	1081.8	59.7	33.5	37.9	14.4	1208.8	1263.8
24	1079.2	1081.3	61.7	43.9	48.5	11.6	891.5	980.4
25	1080.5	1082.1	59.7	57.2	57.9	13.1	931.7	992.9

Table 7 shows the final values of the adapted parameters for the hybrid IT2 NSFLS-2 (OLS-BP) system.

As it can be seen from Fig. 5 to 7, all the IT2 FLS developed here for SB temperature estimation, converged for the conditions tested, thus proving experimentally their stability in this application and for the conditions tested. Furthermore, the T2 FLS with the hybrid learning algorithm also showed better performance than the BP IT2 FLS in terms of RMSE. These results show the feasibility of the T2 FLS, both for BP only and hybrid learning, for this particular industrial application. The IT2 FLS antecedent membership functions and consequent centroids absorbed the uncertainty introduced by training noisy data: noisy temperature and noisy traveling time measurements since they finally converged to a RMSE value. The noisy data was not shown here since each data point is the average temperature at the head segment of the TB and the measured time for each TB to get from RM to SB entry. They may vary from one TB to the next for different reasons and not only because of the measurement noise. Therefore, in order to show the noise a deeper analysis is required.

The hybrid learning systems OLS-BP have better performance than only BP methods for the conditions tested and in terms of RMSE. There are not works reported in the literature using OLS algorithm for IT2 FLS adaptation. In [3, 4] the systems developed were trained and validated from the physical model estimations (the system was modeling the physical model) rather than real data from the mill, therefore, as expected, the results reported show better RMSE than those reported here, since in the real data measurement noise and natural random variations of the modeled variable is also presented (real SB entry temperature).

Fig. 8 shows the application of singleton ANFIS. This corresponding to the T2 FLS experiment showed in Fig. 5. By comparison of figures 8 and 5, it can be concluded that the T2 FLS have better performance in terms of RMSE than ANFIS, both for, hybrid and BP only tuning. Table 8 shows the RMSE convergence values after 15 epochs of all hybrid learning ANFIS OLS-BP and IT2 FLS for entry temperature estimation, note that the ANFIS were not tested for IT2 no-singleton inputs. As it can be seen in Table 8, IT2 FLS has consistently lower convergence values and hence, better performance than ANFIS. Table 8 third column shows RMSE improvement in percentage when T2 FLS is applied. The improvement ranges from 14% to 59%, furthermore in four cases out of the six tested, the improvement is 27% or above, which is satisfactory.

Table 8. Comparision between RMSE after 15 epoches of ANFIS and hybrid IT2 FLS.

	IT2 FLS	T1 FLS	Difference (%)
Coil A – singleton	5.1	7.5	47
Coil A – T1 non-singleton	5.8	7	20
Coil B – singleton	5.6	7.9	41
Coil B – T1 non-singleton	6.3	7.5	19
Coil C – singleton	9.4	15	59
Coil C – T1 non-singleton	12.9	14.8	14

However, the greater disadvantage of T2 fuzzy is the high computational resources demand. In order to count for uncertainties, two fuzzy sets are introduced instead of one, therefore the memory requirements to keep these functions double. Consequently, the number of operations increases demanding more CPU time, see Figures 3 and 4. Furthermore, since the hybrid learning algorithm developed here requires to reestablish the lost FBF expansion rule ordered after every iteration the number of operations also increases, see the limitation and step 7 of the algorithm given in section 4, see also Fig. 2. In future, work has to be done in order to get improvements on these aspects.

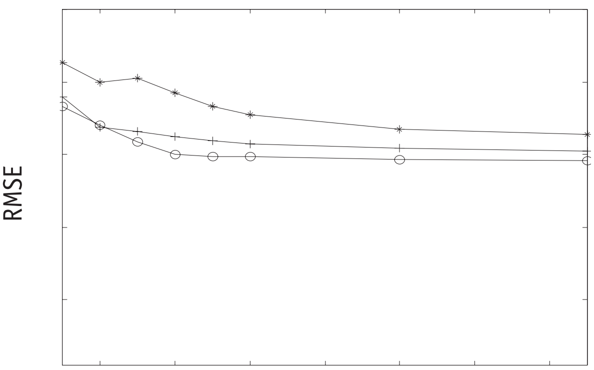

Fig. 8. RMSE for Type A coil and ANFIS for singleton input.

Although more exhaustive experiments with a more complete statistical test of the prediction error are required, as well as exhaustive comparison with the physical model performance, it is believed that the results shown here may motivate further study of IT2 FLS industry application, in general and in particular for rolling variable estimation.

It is also important to consider in future some extra factors which may influence the temperature behavior and may be difficult to incorporate to the mathematical model, such as: steel chemical composition, TB thickness, and rolling pace among others. This would increase of course the number of fuzzy sets and operations required and therefore resources demand. As mentioned, the experiments run here were carried out for three different types of products, while in practice; the same model should be running for all product types or at least for a wide range of products. Tests on different products have to be performed. Estimation of temperatures for the rest of the rolling process is also to be studied.

7. Conclusions

New applications of IT2 FLS, using a hybrid learning method, are presented. The IT2 FLS antecedent membership functions and consequent centroids absorbed the uncertainty introduced by training noisy data: noisy temperature and noisy traveling time measurements. The uncertainty of the input data and measurements was modeled as stationary additive noise using T1 fuzzy sets and as non-stationary additive noise using T2 fuzzy sets. The only BP and the hybrid OLS-BP methods were tested and have demonstrated the power of hybrid parameters estimation. There is a substantial improvement in performance and stability of the hybrid method over BP. The hybrid OLS-BP achieves the better RMSE performance as can be seen in the experimental results. They show that the developed hybrid algorithm can be applied for modeling entry SB temperature of steel bars. Comparative results show that IT2 FLS over-performs ANFIS, in most cases substantially. The results shown here may motivate a further study in this topic. In future, the algorithm developed here has to be tested more exhaustively, against the physical model and for different products; incorporation of extra factors may also be useful in this matter. On the other hand, optimization of the algorithm would be desirable. Estimation of temperatures for the rest of the rolling process is also to be studied.

AUTHORS
Gerardo Mendez, Angeles Hernandez - Instituro Tecnologico de Nuevo Leon, Av. Eloy Cavazos #2001, CP 67170, Cd. Guadalupe, NL, México. E-mails:
gmm_paper@yahoo.com.mx,
ahernandez@yturria.com.mx.

References

[1] Mendel J. M., *Uncertain Rule-Based Fuzzy Logic Systems: Introduction and New Directions*, Upper Saddle River NJ: Prentice-Hall, 2001.

[2] Aguado A., *Temas de Identificación y Control Adaptable*, La Habana 10200 Cuba, Instituto de Cibernética, Matemáticas y Física, 2000, (In Spanish).

[3] Méndez M., Cavazos A., Leduc L., Soto R., "Hot Strip Mill Temperature Prediction Using Hybrid Learning Interval Singleton Type-2 FLS". In: *IASTED International Conference on Modeling and Simulation*, Palm Springs CA, 2003, pp. 380-385.

[4] Méndez M., Cavazos A., Leduc L., Soto R., "Modelling of a Hot Strip Mill Temperature Using Hybrid Learning for Interval Type-1 and Type-2 Non-Singleton Type-2 FLS". In: *IASTED International Conference on Artificial Intelligence and Applications*, Benalmádena Spain, 2003, pp. 529-533.

[5] Jang J. S. R., Sun C. -T., Mizutani E., *Neuro-Fuzzy and Soft Computing: A Computational Approach to Learning and Machine Intelligence*, Upper Saddle River, NJ: Prentice-Hall, 1997.

[6] Taylor B. N., Kuyatt C. E., *Guidelines for Evaluating and Expressing the Uncertainty of NIST Measurement Results*, NIST Technical Note 1297, aitherburg MD, September 1994.

[7] Wang, X., "Fuzzy Systems are Universal Approximators". In: *Proceedings of the IEEE Conf. on Fuzzy Systems*, San Diego 1992, pp. 1163-1170.

[8] Wang X., *Fuzzy Systems as Nonlinear Mapping. A Course in Fuzzy Systems and Control*, Upper Saddle River, NJ: Prentice Hall PTR, 1997, pp. 118-127.

[9] Wang L. X., „Fuzzy Systems as Nonlinear Dynamic System Identifiers". In: *Proceedings of the 31th IEEE Conference on Decision and Control*, Tucson, Arizona, 1992, pp. 897-902.

[10] Wang L. X., Mendel J.M., "Back-Propagation Fuzzy Systems as Nonlinear Dynamic System Identifiers". In: *Proceedings of the IEEE Conf. On Fuzzy Systems*, San Diego, CA, 1992, pp. 1409-1418.

[11] Martinetz T., Protzel P., Gramchow O., Sorgel G., *Neural Network Control for Rolling Mills*, ELITE Foundation EUFIT 94, Achen, Germany, 1994, pp. 147-152.

[12] *General Electric Models: Users reference*, Roanaoke VA, vol. 1, 1993.

[13] Lee Y., Cho H. S., "Neural Network Approach to the Control of the Plate Width in Hot Plate Mills", *IEEE International Joint Conference on Neural Networks*, vol. 5, Washington DC, 1999, pp. 3391-3396.

[14] Ljung L., *System Identification: Theory for the User*, Prentice Hall Inc., Englewood Cliff NJ, 1987.

[15] Fausett L., *Fundamental of Neural Networks*, Prentice Hall Inc., Englewood Cliff NJ, 1994.

[16] Jang,S. R. J., Sun, C. -T., "Neuro-Fuzzy Modelling and Control". In: *Proceedings of the IEEE*, vol. 3, 1995, pp. 378-406.

[17] Sato N., Kamada N., Naito S., Fukushima T., Fujino M., "Application of Fuzzy Control System to Hot Strip Mill". In: *Proceedings of the IEEE International Conference on Industrial Electronics, Control, Instrumentation*, San Diego CA, 1992, pp. 1202-06.

[18] Bissessur Y., Martin E.B., Morris A.J., Kitson P., "Fault Detection in Hot Steel Rolling Using Neural Networks and Multivariate Statistics", *IEE Proceedings Control Theory and Application*, vol. 147, 2000, pp. 633-640.

[19] Watanabe T., Narazaki H., Kitamura A., Takahashi Y., Hasegawa H., "A New Mill-setup System for Hot Strip Rolling Mill that Integrates a Process Model and Expertise". *IEEE International Conference on Computational Cybernetics and Simulation*, vol. 3, Orlando FL, 1997, pp. 2818-2822.

[20] Schalng M., Poppe T., "Neural Network for Steel Manufacturing", *IEEE Expert*, vol. 11, 1996, pp. 8-9.

[21] Yao X., Tieu A.K., Fang X.D., Frances D., "Neural Network Application to Head & Tail Width Control in Hot Strip Mill", *IEEE International Conference on Neural Networks*, vol. 1. Orlando FL, 1995, pp. 433-437.

[22] Kim Y.S., Yum B.J., Kim M., "Robust Design of Artificial Neural Network for Roll Force Prediction in Hot Strip Mill", *IEEE International Joint Conference on Neural Network*, vol. 4, Washington DC, 2001, pp. 2800-2804.

[23] Liang Q., Mendel J. M., "Interval type-2 fuzzy logic systems: Theory and design, *Trans. Fuzzy Systems*, 2000, pp. 535-550.