Information Sciences 179 (2009) 2146–2157

Contents lists available at ScienceDirect

Information Sciences

journal homepage: www.elsevier.com/locate/ins

Hybrid learning for interval type-2 fuzzy logic systems based on orthogonal least-squares and back-propagation methods

Gerardo M. Méndez [a,*], M. de los Angeles Hernandez [b]

[a] Department of Electrical and Electronic Engineering, Instituto Tecnológico de Nuevo León, Calle Septima 822, Col. La Herradura, 67140 Cd. Guadalupe, N.L., Mexico
[b] Department of Economics and Administration Sciences, Instituto Tecnológico de Nuevo León, Cd. Guadalupe, N.L., Mexico

ARTICLE INFO

Article history:
Received 9 November 2007
Received in revised form 13 June 2008
Accepted 10 August 2008

Keywords:
Interval type-2 fuzzy inference systems
Interval type-2 neuro-fuzzy systems
Hybrid learning
Uncertain rule-based fuzzy logic systems

ABSTRACT

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

1. Introduction

Interval type-2 fuzzy logic systems (IT2 FLSs) constitute an emerging technology [1–4]. The processes of financial systems [5–7], hot strip mills [8,9], autonomous mobile robots [10], intelligent controllers [11–13], and plant monitoring and diagnostics [14–16] are characterized by high uncertainty, nonlinearity, and time-varying behavior [17]. Type-2 (T2) fuzzy sets let us model the effects of uncertainties and minimize them by optimizing the parameters of an IT2 fuzzy set during a learning process [18–20]. In [1], both one-pass and back-propagation (BP) methods are presented as IT2 Mamdani FLS learning methods but only the BP method 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 once, and combines the rules to construct the final FLS. When the BP method is used in both IT2 Mamdani and TSK FLS methods, none of the antecedent and consequent parameters of the IT2 FLS are fixed at the start of the training process; they are tuned using exclusively the BP method. The recursive least-squares (RLS) and recursive orthogonal least-squares (OLS) algorithms are not presented as IT2 FLS learning methods in [1].

The aim of this publication is to present and discuss a hybrid learning algorithm based on OLS–BP methods for the tuning of antecedent and consequent parameters during the training process of IT2 Mamdani fuzzy logic systems. In the forward pass, the FLS output is calculated and the consequent parameters are tuned using the OLS method. In the backward pass, the error propagates backward, and the antecedent parameters are tuned using the BP method. One of the hybrid algorithms

* Corresponding author. Tel.: +52 8183172547.
E-mail address: gmm_paper@yahoo.com.mx (G.M. Méndez).

0020-0255/$ - see front matter © 2008 Elsevier Inc. All rights reserved.
doi:10.1016/j.ins.2008.08.008

proposed elsewhere [5,8,9] is based on an RLS method, and since then it has been a benchmark algorithm for parameter estimation and systems identification. It has been shown [5,8,9] that hybrid algorithms provide improved convergence compared with the BP method. The convergence of our proposed methods has been tested practically.

Since only the back-propagation learning method has been proposed in the literature for IT2 FLSs, we have developed a hybrid learning algorithm (using both OLS and BP methods) for IT2 FLSs and implemented it for temperature prediction. This was motivated by the success of the hybrid learning method in type-1 (T1) FLSs (ANFISs) [21] compared with the BP-only method. The convergence has been tested practically for particular conditions; it is not the purpose of this paper to generalize the algorithm developed here, but only to show some preliminary comparative results and the feasibility of application of the algorithm. A mathematical proof for hybrid learning algorithms in general is still to be given.

The IT2 FLSs were trained using two main learning mechanisms: the back-propagation method for tuning both the antecedent and the consequent parameters, and a hybrid training method using a recursive orthogonal least-squares method for tuning the consequent parameters and the BP method for tuning the antecedent parameters. In this paper, the former mechanism is referred to as "IT2 FLS (BP)", and the latter as "hybrid IT2 FLS (OLS–BP)".

The use of IT2 FLSs can account for both random and systematic components [22] of industrial measurements. The nonlinearity of the processes is handled by the FLS by use of identifiers and universal approximators of nonlinear dynamic systems [23–26]. Such characteristics give IT2 FLSs great potential to model and control industrial processes.

In order to test the hybrid adaptation method, IT2 FLS (OLS–BP) forecasters were implemented and used to estimate the temperature of the head end of the incoming bar on entry to the scale breaker in a real hot strip mill. Several IT2 FLSs were designed and developed for estimation of the transfer bar head-end temperature, and preliminary experimental results are presented and analyzed here. Such experiments show that it is feasible to apply the hybrid IT2 FLS (OLS–BP) method for estimation of the entry temperature in a hot strip mill. Although the results were preliminary, validation was carried out with adaptation, since adaptation was the ultimate goal; the main reason for the application of these techniques is their adaptation capabilities. It is important to note that the experiments here were carried out for three different types of product separately, whereas in practice the same model should be run for all product types sequentially as they are rolled.

In this application, the inputs to the hybrid IT2 (OLS–BP) fuzzy models used to predict the scale breaker entry temperature (y) were the surface temperature of the transfer bar at the exit from the roughing mill (x_1) and the time required by the transfer bar head to reach the scale breaker entry zone (x_2). Currently, the surface temperature is measured using a pyrometer located on the exit side of the roughing mill. Scale grows on the transfer bar surface, producing a noisy temperature measurement. The measurement is also affected by water vapor in the environment, as well as by the location, calibration, resolution, and repeatability of the pyrometer. The travel time of the head end of the transfer bar is estimated from a model of the finishing-mill setup using the estimated thread speed of the finishing mill. Such estimation has an error associated with the inherent uncertainty in the model of the finishing-mill setup. Although prediction of the temperature y is a critical issue in a hot strip mill, the problem has not been fully addressed by fuzzy logic control systems [27–29].

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. Section 4 gives a brief introduction to the hot strip mill. Section 5 deals with the application of the hybrid IT2 FLS to prediction of the temperature in a hot strip mill, and the experimental results are presented in Section 6. Section 7 summarizes the conclusions.

2. Principles of the orthogonal least-squares method

As mentioned above, a hybrid learning method, which we call IT2 FLS (OLS–BP), was used for prediction of the scale breaker entry temperature in a hot strip mill. The hybrid learning algorithm is based on OLS and BP learning methods. In this section, a brief presentation of the basic principles of the specific OLS method is given. Since the IT2, BP, and OLS methodologies are very well established, readers are referred to [1,30–32], respectively.

Suppose that, as in [32], 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 length T; it is then possible to describe its dynamic behavior using the following model for t input–output data pairs:

$$\mathbf{Y}^{\mathrm{T}}(t) = \mathbf{Z}^{\mathrm{T}}(t)\mathbf{P}^{\mathrm{T}} + \mathbf{E}(t), \tag{1}$$

where $\mathbf{Y}(t)$ is the output vector of size t, $\mathbf{Z}^{\mathrm{T}}(t)$ is the matrix of measurements of size $t \times (2n+1)$, t is the number of input–output data pairs, n is the model number, and $\mathbf{P}$ is the parameter estimation matrix of size $2n+1$. Online estimation of the parameters of $\mathbf{P}(t+1)$ using the measurements for the time $t+1$ can be performed by solving the triangular system of equations

$$F(t+1)P(t+1) = q(t+1), \tag{2}$$

where the upper triangular matrix $\mathbf{F}(t+1)$, of proper size $2n+1$, is the square root of $\mathbf{Z}^{\mathrm{T}}(t+1)\mathbf{Z}(t+1)$ and $\mathbf{q}(t+1)$ is a vector of size $2n+1$. For each period of time, the above algorithm reduces to zero one row of the compound vector $[\mathbf{z}^{\mathrm{T}}(t+1)y(t+1)]$, of size $2n+2$. The parameters of $\mathbf{P}(t+1)$ can easily be calculated by use of the REDCO routine given in [32].

3. The hybrid learning methodology for IT2 FLS (OLS–BP)

3.1. Limitations of hybrid learning for IT2 FLSs

In [1], only BP was proposed as a learning algorithm for IT2 FLSs. During the backward pass, the antecedent and consequent parameters are estimated as shown in Table 1. In the hybrid algorithm developed here, a recursive OLS method is used during the forward pass for tuning of the consequent parameters, and the BP method is used during the backward pass for tuning of the antecedent parameters, as shown in Table 2. This hybrid learning method is an extension of the ANFIS training method proposed for T1 FLSs [21,33].

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

1. The starting point for the OLS method for designing an interval singleton FLS is a T1 fuzzy basis function (FBF) expansion. No such FBF expansion exists for a general singleton T2 FLS. It looks as if it might be possible to use a least-squares method to tune the parameters in $\mathbf{y}_l^T$ (the transpose of the matrix of M left points y_l^l of centroids of the consequents) and $\mathbf{y}_r^T$ (the transpose of the matrix of M right points y_l^l of centroids of the consequents). Unfortunately, this is incorrect. The problem is that, in order to know the FBFs $\mathbf{p}_l(\mathbf{x})$ and $\mathbf{p}_r(\mathbf{x})$, each y_l^l and y_r^l (the M left points and right points of the centroids of the consequents in the interval) must be known first. Since initially there are no numerical values for those elements, it is impossible to do this; hence the FBFs $\mathbf{p}_l(\mathbf{x})$ and $\mathbf{p}_r(\mathbf{x})$ cannot be calculated. This situation does not occur for a T1 FBF expansion. A T2 FLS output $y(\mathbf{x})$ can be expressed as
$$y(\mathbf{x}) = \frac{1}{2}[\mathbf{y}_l^T\mathbf{p}_l(\mathbf{x}) + \mathbf{y}_r^T\mathbf{p}_r(\mathbf{x})]. \tag{3}$$
2. Although y_l and y_r (the end-points of the T2 FLS center-of-sets type-reduced set Y_{COS}) can be expressed as an interval $[\underline{f}^l, \bar{f}^l]$ in terms of their lower ($\underline{f}^l$) and upper ($\bar{f}^l$) M firing sets, whereas the corresponding M left points y_l^l of the consequent can be expressed as
$$y_l = y_l(\bar{f}^1, \ldots, \bar{f}^L, \underline{f}^{L+1}, \ldots, \underline{f}^M, y_l^1, \ldots, y_l^M) \tag{4a}$$
and the corresponding M right points y_r^l of the consequent as
$$y_r = y_r(\underline{f}^1, \ldots, \underline{f}^R, \bar{f}^{R+1}, \ldots, \bar{f}^M, y_r^1, \ldots, y_r^M), \tag{4b}$$
where L is the index of the rule-ordered FBF expansions at which y_l is a minimum and R is the index at which y_r is a maximum, they are not known in advance [1]. Once the points L and R are known, (4) is very useful for organizing and describing the calculations of y_l and y_r.
3. The next problem has to do with the reordering [1] of y_l^l and y_r^l. The T1 FBF expansions always have an inherent rule ordering associated with them, i.e., the rules $R^1, R^2, \ldots, R^M$ are always established as the first, second, …, Mth FBF. In a T2 FBF expansion, this order is lost and needs to be restored for later use.

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

1. Since the values of y_l^l and y_r^l were initially fixed as initial conditions, it is possible to use an OLS method for estimation of the left-point and right-point parameters of the consequent centroids using the standard deviation of the variable in each calculation.
2. The values of $\mathbf{p}_l(\mathbf{x})$ and $\mathbf{p}_r(\mathbf{x})$ in (3) can be calculated using the initial values of y_l^l and y_r^l, and then they can be used as a base for OLS estimation methods.

Table 1
Use of one-pass in the BP learning procedure for IT2 FLSs

	Forward pass	Backward pass
Antecedent parameters	Fixed	BP
Consequent parameters	Fixed	BP

Table 2
Use of two passes in a hybrid learning procedure for IT2 FLSs

	Forward pass	Backward pass
Antecedent parameters	Fixed	BP
Consequent parameters	OLS	Fixed

3. The lost rule-ordered FBF expansions can be restored [1] and used for estimation of the next consequent centroids y_l^l and y_r^l.

3.2. The hybrid algorithm for IT2 FLSs

The membership functions of the hybrid training method developed are Gaussian functions and are based on the initial conditions of the antecedent parameters (x_1 and x_2, the interval mean $m_k^l \in [m_{k1}^l, m_{k2}^l]$ and the standard deviation σ_k^l), the consequent parameters (y, y_l^l, and y_r^l) and the measurement parameters (the interval of standard deviations $\sigma_{Xk} \in [\sigma_{X1}, \sigma_{X2}]$). The antecedent and measurement parameters are tuned using the BP training method, while the consequent parameters are tuned using the OLS training method. Given N input–output training data pairs, the hybrid training algorithm for E training epochs should minimize the error function

$$e^{(t)} = \frac{1}{2}[f_{s2}(\mathbf{x}^{(t)}) - y^{(t)}]^2. \tag{5}$$

The hybrid training algorithm is as follows:

1. Initialize all parameters in the antecedent and consequent membership functions. Choose the mean values of the IT2 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 $\mathbf{x}^{(t)}$ to the IT2 FLS and compute the total firing interval for each rule, i.e., compute $\underline{f}^i$ and $\overline{f}^i$.
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 $f_{s2}(\mathbf{x}^{(t)})$.
7. Determine the explicit dependence of y_l and y_r on the membership functions. Because the 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 the membership functions will also usually change from one t-iteration to the next.
8. Test each component of $\mathbf{x}^{(t)}$ to determine the active branches.
9. Tune the parameters of the active branches of the consequent using the OLS algorithm.
10. Tune the parameters of the active branches of the antecedent's membership functions using a steepest-descent algorithm.
11. Set $t \equiv t + 1$. If $t \equiv N + 1$ (where N is the size of the input data vector), go to step 12; otherwise, go to step 4.
12. Set $ep \equiv ep + 1$. If $ep \equiv E$ (where E is the total number of epochs), STOP; otherwise, go to step 3.

4. Hot strip mills

Because of the complexities and uncertainties involved in rolling operations, the development of mathematical theories has been largely restricted to two-dimensional models applicable to heat loss in flat rolling operations.

Fig. 1 depicts a simplified diagram of a hot strip mill, from its initial stage, the entry to the reheat furnace, to the final stage, the coilers. Besides the mechanical, electrical, and electronic equipment, great potential for ensuring good quality is contained in the automation systems and control techniques used. The most critical process in a hot strip mill occurs in the finishing mill. There are several mathematical model-based systems for setting up the finishing mill. There is a model-based setup system [28,34,35] that calculates the working reference values of the finishing mill needed to obtain a given

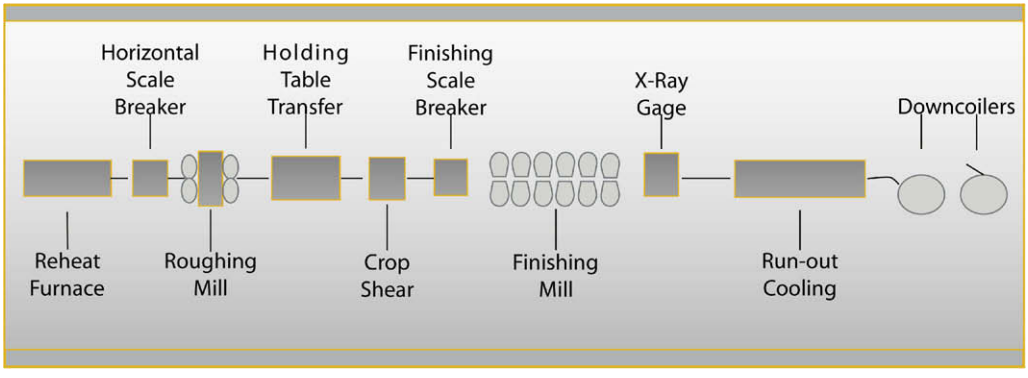

Fig. 1. Schematic view of a typical hot strip mill.

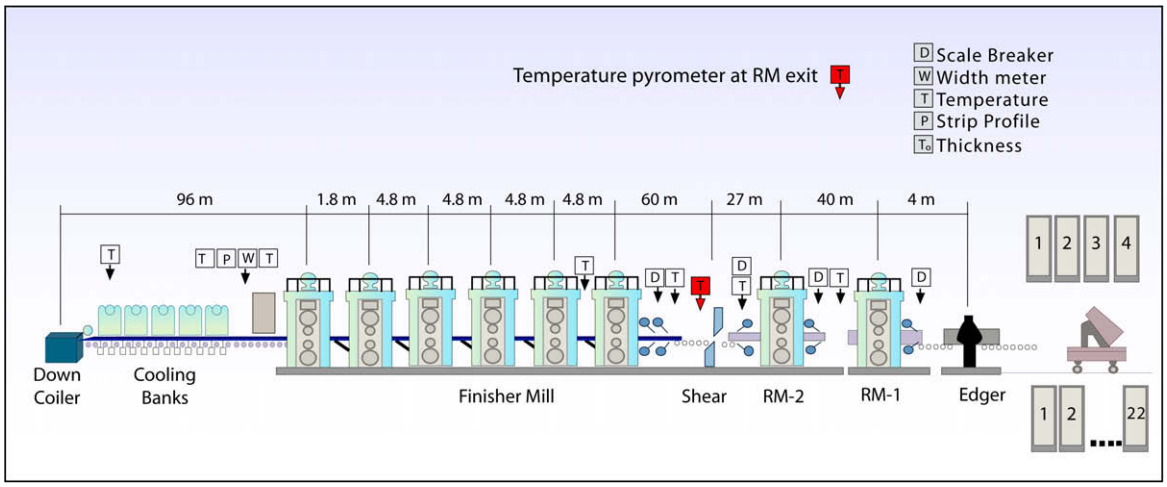

Fig. 2. Detailed schematic view of a hot strip mill.

gauge, width, and temperature at the exit stands of the finishing mill. This system takes as inputs the target gauge at the exit of the finishing mill, the target width and target temperature, the grade of steel, the hardness ratio obtained from the slab chemistry, the load distribution, the gauge offset, the temperature offset, the roll diameters, the load distribution, the transfer bar gauge, the width, and the entry temperature.

The errors in the gauge of the transfer bar are absorbed in the first two finishing-mill stands and therefore have little effect on the target exit gauge. It is very important, however, for the model to know the finishing-mill entry temperature accurately. A temperature error will propagate through the entire finishing mill. Owing to scale formation, it is not possible to measure the transfer bar temperature at the entry to the finishing mill, and therefore it has to be estimated from measurements at the roughing-mill exit after the last pass, as shown in Fig. 2.

5. Prediction of scale breaker entry temperature

The architecture of the hybrid IT2 FLS for scale breaker entry temperature prediction was established in such a way that the parameters were continuously optimized. As mentioned earlier, the antecedents were chosen to be the roughing-mill exit surface temperature (x_1) and the travel time of the transfer bar head (x_2). These are the inputs to the physical model used for estimation of the head-end temperature in most hot strip mills used in industrial production, and they are considered to be the variables that have the most influence on the scale breaker entry temperature. Each antecedent–input space was divided into five fuzzy sets (seeking a compromise between reasonably good performance and a low number of sets to keep the demand on computational resources low) and thus had 25 rules. The output (the consequent y) was the head-end surface temperature on entry to the scale breaker.

Gaussian primary membership functions with uncertain means were chosen for both the antecedents and the consequents. Each rule of the three IT2 FLSs was characterized by six antecedent membership function parameters (two for the left-hand and right-hand bounds of the mean and one for the standard deviation, for each of the two antecedent Gaussian membership functions) and two consequent parameters (one for the left-hand and one for the right-hand end point of the centroid of the consequent IT2 fuzzy set).

The primary membership function for each input to the model labeled "IT2 NSFLS-1" was

$$\mu_{x_k}(x_k) = \exp\left[-\frac{1}{2}\left[\frac{x_k - x'_k}{\sigma_{x_k}}\right]^2\right], \tag{6}$$

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

The primary membership function for each input to the model labeled "IT2 NSFLS-2" was

$$\mu_{x_k}(x_k) = \exp\left[-\frac{1}{2}\left[\frac{x_k - x'_k}{\sigma_{kn}}\right]^2\right], \tag{7}$$

where $\sigma_{kn} \in [\sigma_{k1}, \sigma_{k2}]$, $k = 1, 2$ (the number of IT2 non-singleton inputs), $n = 1, 2$ (the lower and upper bounds of the uncertain mean), and $\mu_{Xk}(x_k)$ was centered at the measured input $x_k = x'_k$. The uncertain standard deviation $[\sigma_{11}, \sigma_{12}]$ of the measure-

Table 3
The three types of coil considered

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

ment of the roughing-mill exit surface temperature was initially set to [11.0, 14.0] °C, and the uncertain standard deviation $[\sigma_{21}, \sigma_{22}]$ of the measurement of the head-end travel time was initially set to [1.41, 3.41] s.

Noisy input–output data pairs for three different types of coil with different target gauges, target widths, and grades of steel (see Table 3) were taken and used as training and validation data, and experiments were carried out for these three coil types. The standard deviation of the temperature noise σ_n was initially set to 1.0 °C and the standard deviation of the time noise σ_n was set to 1.0 s. The signal-to-noise ratio of the input variable x_1 was 11.13 dB, that of the input data x_2 was 3.82 dB, and that of the output variable y was 10.68 dB.

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

$$\mu_k^l(x_k) = \exp\left[-\frac{1}{2}\left[\frac{x_k - m_{kn}^l}{\sigma_k^l}\right]^2\right], \tag{8}$$

where $m_k^l \in [m_{k1}^l, m_{k2}^l]$ is the uncertain mean, $k = 1, 2$ (the number of antecedents), $l = 1, 2, \ldots, 25$, $n = 1, 2$ (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 the inputs x_1 and x_2, the values of the antecedent five intervals of uncertainty were established. The initial intervals of uncertainty for the input x_1 were selected as shown in Table 4. Fig. 3 shows the initial membership functions for the antecedent fuzzy sets of the input x_1. The values of the initial intervals of uncertainty for the input x_2 were selected as shown in Table 5. Fig. 4 depicts the initial membership functions for the antecedent fuzzy sets of the input x_2.

In the case of the model labeled "IT2 SFLS", the resulting IT2 FLS used T1 singleton fuzzification, the maximum t-conorm, the product t-norm, product implication, and center-of-sets type-reduction; for details of the fuzzy operations, see [1]. For IT2 NSFLS-1, the resulting T2 FLS used T1 non-singleton fuzzification, the maximum t-conorm, the product t-norm, product implication, and center-of-sets type-reduction. Finally, for IT2 NSFLS-2, the resulting T2 FLS used T2 non-singleton fuzzification, the maximum t-conorm, the product t-norm, product implication, and center-of-sets type-reduction.

The OLS–BP hybrid learning method was used for parameter optimization. Experimental results are presented in the next section for comparison with the results of the BP-only method.

Table 4
Intervals of uncertainty selected for the 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

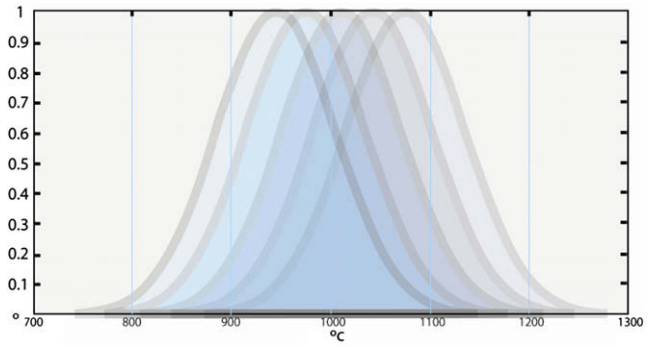

Fig. 3. Membership functions for the antecedent fuzzy sets of the input x_1.

Table 5
Intervals of uncertainty selected for the 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

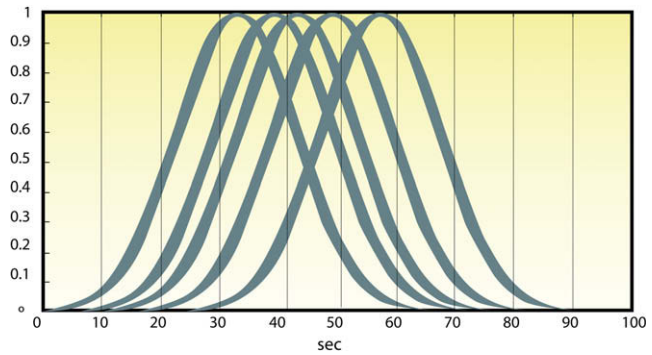

Fig. 4. Membership functions for the antecedent fuzzy sets of the input x_2.

The IT2 fuzzy rule base consisted of a set of IF-THEN rules that represented the model of the system. The interval non-singleton T2 system had two inputs $x_1 \in X_1$ and $x_2 \in X_2$ and one output $y \in Y$, which had a corresponding rule base of size $M = 25$ rules of the form

$$R^l : \text{IF } x_1 \text{ is } \widetilde{F}_1^l \text{ and } x_2 \text{ is } \widetilde{F}_2^l, \text{ THEN } y \text{ is } \widetilde{G}^l, \tag{9}$$

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 [24], as shown in Table 6.

Table 6
Initial values of antecedent and consequent parameters for one pass of the IT2 FLS

l	m_{11}	m_{12}	σ_1	m_{21}	m_{22}	σ_2	y_l^l	y_r^l
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

The primary membership function for each consequent is a Gaussian function with uncertain means, as defined in (6). Since the center-of-sets type-reducer replaces each consequent set $C_{\tilde{G}_l}$ by its centroid, then y_l^l and y_r^l are the consequent parameters.

Initially, only the training pairs $(x^{(1)}:y^{(1)}),(x^{(2)}:y^{(2)}),\ldots,(x^{(N)}:y^{(N)})$ from the input–output data are available and there is no data information about the consequents; hence the initial values of the centroid parameters y_l^l and y_r^l may be determined according to linguistic rules by human experts or be chosen arbitrarily in the output space [24]. In the present work, the initial values of the parameters y_l^l and y_r^l were such that the corresponding membership functions uniformly covered the output space. Table 6 also shows the initial values of the consequent centroids of the 25 rules.

6. Experimental results

The three different models IT2 SFLS, IT2 NSFLS-1, and IT2 NSFLS-2 were trained to predict the scale breaker entry temperature and then tested. Three different sets of data for the three different coil types listed in Table 3 are taken from a real-life mill. Each of these data sets was split into two sets, namely a training and a validation set, taking every other data point. Experiments were run for each coil type independently. The 25 rules of the nine FLSs were tuned for each input–output data pair and for each coil type. Figs. 5–7 show plots of the input–output experimental data for each coil type.

The evaluation of the performance of each of the learning methods was based on the root-mean-square-error (RMSE) criterion

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{k=1}^{n}[\mathbf{Y}(k) - \mathbf{f}_{s2}(\mathbf{x}^{(k)})]^2}, \tag{10}$$

where $\mathbf{Y}(k)$ is the output validation data vector, i.e., the vector of the actual measurements of the scale breaker entry temperature used for system evaluation, different from the training data vector but from the same coil type, and $\mathbf{f}_{s2}(\mathbf{x}^{(k)})$ is the temperature vector predicted by the IT2 FLS tested.

Figs. 8–10 show the RMSE of the three IT2 FLSs used for type A coils after 15 computational epochs. The behavior for type B and C coils was similar and is not shown here for the sake of brevity.

In these figures, the horizontal axis represents the number of training epochs, and the vertical axis represents the RMSE of the validation run with the test set after the corresponding number of training epochs, as outlined in [29]. Therefore, the

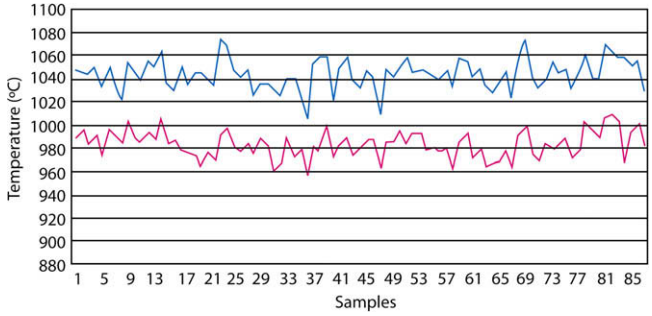

Fig. 5. Input–output training data of coil type A. Blue upper line is temperature of transfer bar at roughing-mill exit zone. Red lower line is temperature of transfer bar at scale breaker. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

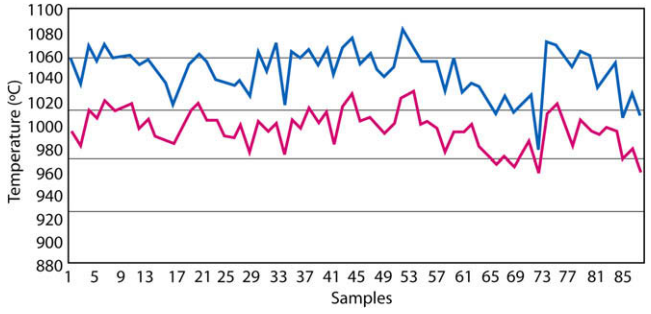

Fig. 6. Input–output training data of coil type B. Blue upper line is temperature of transfer bar at roughing-mill exit zone. Red lower line is temperature of transfer bar at scale breaker. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

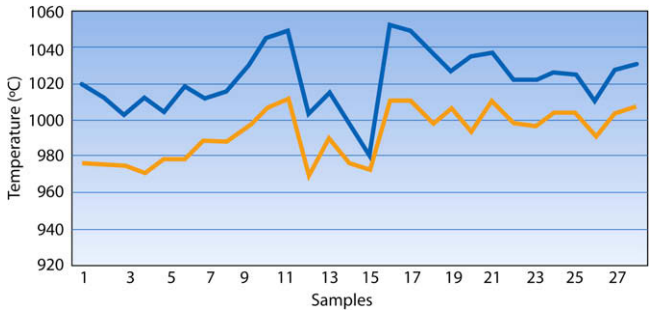

Fig. 7. Input–output training data of coil type C. Blue upper line is temperature of transfer bar at roughing-mill exit zone. Orange lower line is temperature of transfer bar at scale breaker. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

initial value is the RMSE of the validation run after one training epoch. Fifteen epochs were chosen for display purposes because convergence had taken place by the end of that period for all of the systems and experiments considered here. As mentioned earlier, the results presented here are preliminary and are presented with the purpose of showing the convergence and feasibility of T2 FLSs for prediction of the scale breaker temperature in a hot strip mill. However, validation was carried out allowing tuning (unlike the case for the results in [29]), since the online adaptation capabilities motivated us to consider

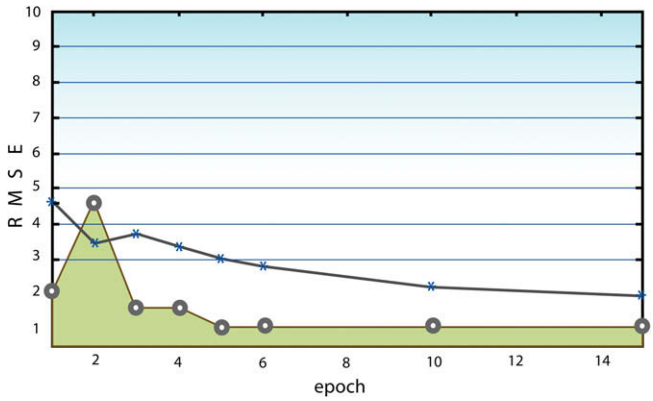

Fig. 8. Results from IT2 SFLS (*), BP training method (○) and OLS–BP hybrid training method.

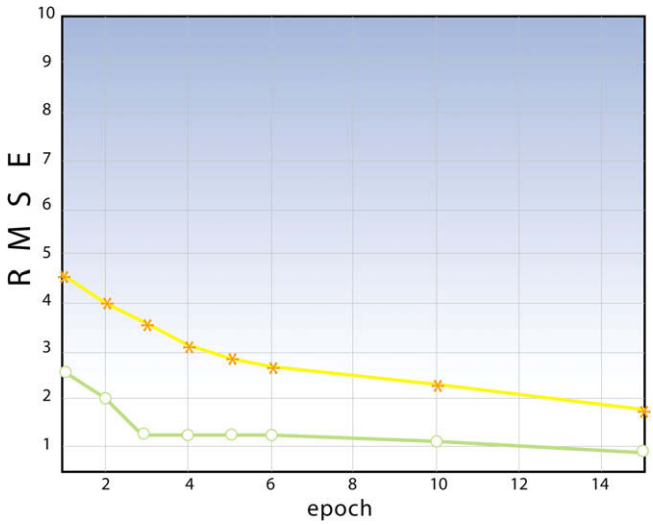

Fig. 9. Results from IT2 NSFLS-1 (*), BP training method (○) and OLS–BP hybrid training method.

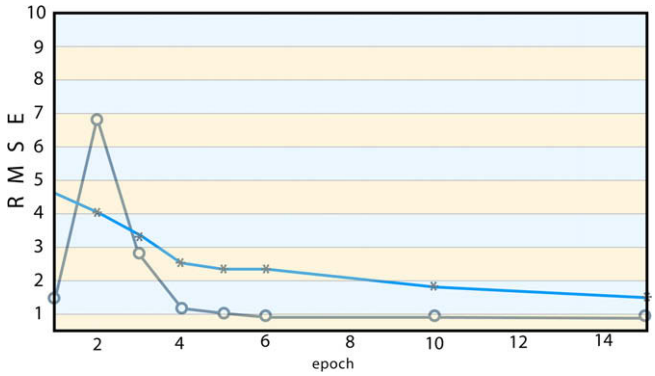

Fig. 10. Results from IT2 NSFLS-2 (*), BP training method (○) and OLS–BP hybrid training method.

estimating the rolling variables by means of either FLSs or artificial neural networks, and this would be the ultimate application of these systems.

Table 7 shows the final values of the adapted parameters for the hybrid system IT2 NSFLS-1. As can be seen from Figs. 8–10, all of the IT2 FLSs developed here for scale breaker temperature estimation converged for the conditions tested, thus proving experimentally their stability in this application under these conditions. Furthermore, the T2 FLS with the hybrid learning algorithm showed better performance than the BP IT2 FLS in terms of RMSE. These results show the feasibility of using a T2 FLS, both with BP only and with hybrid learning, in this particular industrial application. The IT2 FLS antecedent membership functions and consequent centroids absorbed the uncertainty introduced by training with noisy data, i.e., noisy temperature and travel time measurements, since they finally converged to an RMSE value. The noisy data has not been shown here, since each data point consisted of the average temperature at the head segment of the transfer bar and the measured time for each transfer bar to get from the roughing mill to the scale breaker entrance. These values may vary from one transfer bar to the next for various reasons, and not only because of the measurement noise. Therefore, in order to show the noise, a deeper analysis will be required.

The OLS–BP hybrid learning systems have better performance than methods using only BP for the conditions tested, in terms of RMSE. There are no reports in the literature of the use of OLS algorithms for IT2 FLS adaptation.

Fig. 11 shows the application of a singleton ANFIS. This corresponds to the IT2 FLS experiment shown in Fig. 8. By comparison of Figs. 8 and 11, it can be concluded that the IT2 FLS has better performance than the ANFIS in terms of RMSE, both for hybrid and for BP-only tuning. Table 8 shows the RMSE convergence values after 15 epochs of all runs of the hybrid learn-

Table 7
Final values of antecedent and consequent parameters at epoch 15 using the hybrid system IT2 NSFLS-1

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

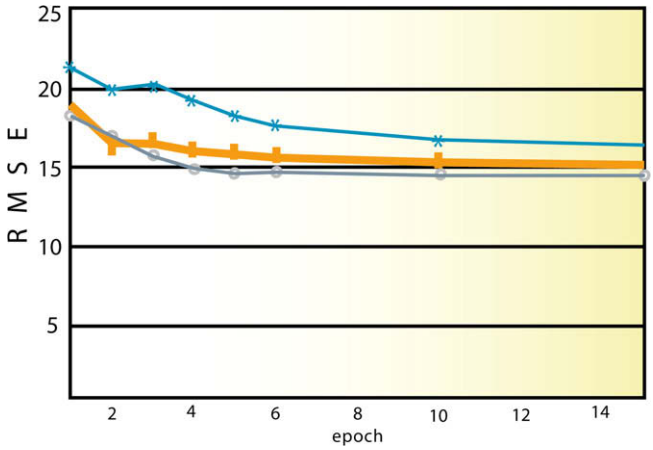

Fig. 11. RMSE for type A coils (∘,+) and ANFIS (*) with singleton input.

Table 8
Comparison between RMSE values after 15 epochs of ANFIS and hybrid IT2 FLS

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

ing (OLS–BP) ANFIS and the IT2 FLS for entry temperature estimation; note that the ANFIS was not tested for IT2 non-singleton inputs. As can be seen in Table 8, the IT2 FLS has consistently lower convergence values and hence better performance than the ANFIS. The third column of Table 8 shows the percentage RMSE improvement when the 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.

However, the great disadvantage of T2 fuzzy logic is the high demand on computational resources. In order to account for uncertainties, two fuzzy sets are introduced instead of one, and therefore the memory requirements for storing these functions are doubled. Consequently, the number of operations increases, demanding more CPU time; see Figs. 3 and 4. Furthermore, since the hybrid learning algorithm developed here requires us to reestablish the lost FBF expansion rule, ordered after every iteration, the number of operations also increases; see the limitations described in Section 3.1 and step 7 of the algorithm given in Section 3.2. In future, work will have to be done to obtain improvements in relation to these aspects.

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

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

7. Conclusions

A new application of IT2 FLSs, using a hybrid learning method, has been presented. The IT2 FLS antecedent membership functions and consequent centroids absorbed the uncertainty introduced by training using noisy data, namely noisy temperature and travel time measurements. The uncertainty of the input data and the measurements was modeled as stationary additive noise using T1 fuzzy sets and as non-stationary additive noise using IT2 fuzzy sets. A method using only BP and a hybrid OLS–BP method were tested, and we have demonstrated the power of hybrid estimation of parameters. Because

of OLS algorithm is much faster than BP algorithm, at first epoch of training the IT2 FLS converges to RMSE values lower than 2. It is a substantial improvement in performance, therefore the hybrid IT2 FLS (OLS–BP) is expected to be used in modeling and predicting the behavior of system's variables where there is the chance of only one epoch of training. These results show that the hybrid algorithm developed here can be applied to the prediction of the temperature of steel bars on entry to the scale breaker of a hot strip mill. Also, they show that the hybrid IT2 FLS (OLS–BP) method outperforms the ANFIS method, in most cases substantially, motivating further studies of this topic. In future, the algorithm developed here will have to be tested more exhaustively in several control applications under uncertain environments.

Acknowledgements

The authors acknowledge the assistance given by Ternium Hylsa, S.A., by providing data from the company's hot strip mill plant No. 1, and that given by Globalove – Graphic Design.

References

[1] J.M. Mendel, Uncertain Rule-Based Fuzzy Logic Systems: Introduction and New Directions, Prentice-Hall, Upper Saddle River, NJ, 2001.
[2] J.M. Mendel, Advances in type-2 fuzzy sets and systems, Information Sciences 177 (1) (2007) 84–110.
[3] J.M. Mendel, H. Wu, New results about the centroid of an interval type-2 fuzzy set, including the centroid of a fuzzy granule Mendel, Information Sciences 177 (2) (2007) 360–377.
[4] F. Lui, An efficient centroid type-reduction strategy for general type-2 fuzzy logic system, Information Sciences 178 (9) (2008) 2224–2236.
[5] A. Hernandez, G.M. Mendez, Modeling and prediction of the MXNUSD exchange rate using interval singleton type-2 fuzzy logic systems, in: Proceedings of the IEEE International Conference, Palm Springs, CA, 2007, pp. 380–385.
[6] L. Anastasakis, N. Mort, Prediction of the GSPUSD exchange rate using statistical and neural network models, in: Proceedings of the IASTED International Conference on Artificial Intelligence and Applications, September 2003, pp. 493–498.
[7] A. Lendasse, E. de Boot, V. Wertz, M. Verleysen, Non-linear financial time series forecasting – application to the bel 20 stock market index, European Journal of Economics and Social Systems 14 (2000) 81–91.
[8] M. Mendez, A. Cavazos, L. Leduc, R. Soto, Hot strip mill temperature prediction using hybrid learning interval singleton type-2 FLS, Proceedings of the IASTED International Conference on Modeling and Simulation, Artificial Intelligence and Applications, Palm Springs, 2003, pp. 380–385.
[9] M. Mendez, A. Cavazos, L. Leduc, R. Soto, Modeling of a hot strip mill temperature using hybrid learning for interval type-1 and type-2 non-singleton type-2 fuzzy logic systems, in: Proceedings of the IASTED International Conference on Artificial Intelligence and Applications, Benalmádena, Spain, 2003, pp. 529–533.
[10] H.A. Hagras, A hierarchical type-2 fuzzy logic control architecture for autonomous mobile robots, IEEE Transactions on Fuzzy Systems 12 (4) (2004) 524–539.
[11] R. Sepulveda, O. Castillo, P. Melin, O. Montiel, An efficient computational method to implement type-2 fuzzy logic in control applications, in: Advances in Soft Computing 41: Analysis and Design of Intelligent Systems Using Soft Computing Techniques, IFSA 2007, Cancún, Mexico, June 2007, Springer-Verlag, pp. 45–52.
[12] R. Sepulveda, O. Castillo, P. Melin, A. Rodriguez-Diaz, O. Montiel, Experimental study of intelligent controllers under uncertainty using type-1 and type-2 fuzzy logic, Information Sciences 177 (10) (2007) 2023–2048.
[13] M.Y. Hsiao, T.H. Li, J.Z. Lee, C.H. Chao, S.H. Tsai, Design of interval type-2 fuzzy sliding-mode controller, Information Sciences 178 (6) (2008) 1696–1716.
[14] O. Castillo, P. Melin, A new hybrid approach for plant monitoring and diagnostics using type-2 fuzzy logic and fractal theory, in: Proceedings of the International Conference FUZZ'2003, May 2003, pp. 102–107.
[15] O. Castillo, G. Huesca, F. Valdez, Evolutionary computing for optimizing type-2 fuzzy systems in intelligent control of non-linear dynamic plants, in: Proceedings of the IEEE NAFIPS'05 International Conference, June 2005, pp. 247–251.
[16] P. Melin, O. Castillo, An intelligent hybrid approach for industrial quality control combining neural networks, fuzzy logic and fractal theory, Journal of Information Sciences 177 (23) (2007) 1543–1557.
[17] D. Wu, J.M. Mendel, Uncertainty measures for interval type-2 fuzzy sets, Information Sciences 177 (23) (2007) 5378–5393.
[18] Q. Liang, J.M. Mendel, Interval type-2 fuzzy logic systems: theory and design, Transactions on Fuzzy Systems 8 (October) (2000) 535–550.
[19] R.I. John, Embedded interval valued type-2 fuzzy sets, in: Proceedings of 2002 IEEE International Conference on Fuzzy Systems, vols. 1 and 2, Honolulu, Hawaii, 2002, pp. 1316–1321.
[20] J.M. Mendel, R.I. John, Type-2 fuzzy sets made simple, IEEE Transactions on Fuzzy Systems 10 (2) (2002) 117–127.
[21] J.-S.R. Jang, C.-T. Sun, E. Mizutani, Neuro-Fuzzy and Soft Computing: A Computational Approach to Learning and Machine Intelligence, Prentice-Hall, Upper Saddle River, NJ, 1997.
[22] B.N. Taylor, C.E. Kuyatt, Guidelines for Evaluating and Expressing the Uncertainty of NIST Measurement Results, NIST Technical Note 1297, Gaithersburg, MD, September 1994.
[23] L.-X. Wang, Fuzzy systems are universal approximators, in: Proceedings of the IEEE Conference on Fuzzy Systems, San Diego, CA, 1992, pp. 1163–1170.
[24] L.-X. Wang, Fuzzy Systems as Nonlinear Mapping: A Course in Fuzzy Systems and Control, Prentice-Hall PTR, Upper Saddle River, NJ, 1997. pp. 118–127.
[25] L.-X. Wang, Fuzzy systems as nonlinear dynamic system identifiers, in: Proceedings of the 31st Conference on Decision and Control, Tucson, AZ, 1992, pp. 897–902.
[26] L.-X. Wang, J.M. Mendel, Back-propagation fuzzy systems as nonlinear dynamic system identifiers, in: Proceedings of the IEEE Conference on Fuzzy Systems, San Diego, CA, 1992, pp. 1409–1418.
[27] T. Martinetz, P. Protzel, O. Gramchow, G. Sorgel, Neural network control for rolling mills, in: EUFIT 94, ELITE Foundation, Aachen, Germany, 1994, pp. 147–152.
[28] General Electric, Models Reference Manual 1, Roanoke, VA, 1993.
[29] D.Y. Lee, H.S. Cho, A neural network approach to the control of the plate width in hot plate mills, in: International Joint Conference on Neural Networks, Washington, DC, vol. 5, 1999, pp. 3391–3396.
[30] L. Ljung, System Identification: Theory for the User, Prentice-Hall, Englewood Cliffs, NJ, 1987.
[31] L. Fausett, Fundamentals of Neural Networks, Prentice-Hall, Englewood Cliffs, NJ, 1994.
[32] A. Aguado, Temas de Identificación y Control Adaptable, Instituto de Cibernética, Matemáticas y Física, La Habana 10200, Cuba, 2000.
[33] J.-S.R. Jang, C.-T. Sun, Neuro-fuzzy modelling and control, Proceedings of the IEEE 3 (March) (1995) 378–406.
[34] N. Sato, N. Kamada, S. Naito, T. Fukushima, M. Fujino, 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. 1202–1206.
[35] Y.S. Kim, B.J. Yum, M. Kim, Robust design of artificial neural network for roll force prediction in hot strip mill, in: IEEE International Joint Conference on Neural Networks, Washington, DC, 2001, pp. 2800–2804.