Iranian Journal of Fuzzy Systems

Volume 20, Number 5, (2023), pp. 33-45

Original paper

University of Sistan and Baluchestan

Identification of cement rotary kiln using type 2 Takagi-Sugeno neuro-fuzzy system considering the effect of different noisy condition

N. Moradkhani[1] and M. Teshnehlab[2]

[1,2] *Electrical Engineering Department, K.N. Toosi University of Technology, Tehran, Iran*

moradkhani@ee.kntu.ac.ir, teshnehlab@eetd.kntu.ac.ir

Abstract

A cement rotary kiln is the main part of the cement production process, which has always attracted many researchers attention. However, this complex nonlinear system has not been modeled efficiently, which can make an appropriate performance, especially in noisy condition. In this work, the type 2 Takagi-Sugeno neuro-fuzzy system (T2TSNFS) is used to identify the cement rotary kiln, and the gradient descent (GD) algorithm is applied for tuning the parameters of antecedent and consequent parts of fuzzy rules. In addition, the optimal inputs of the system are selected by the genetic algorithm (GA) to achieve less complexity in the fuzzy system. The data relating to the Saveh White Cement (SWC) factory is used in the simulations. The results demonstrate that the proposed identifier has an appropriate performance in noisy conditions. Furthermore, in this work, T2TSNFS is evaluated in noisy conditions, which had not been worked out before in related research works. Also, T2TSNFS and type 1 Takagi-Sugeno neuro-fuzzy system (T1TSNFS) are compared. The simulations show that T2TSNFS has more proper performance when the standard deviation of noise increases.

Keywords: Cement Rotary Kiln, identification, type 2 fuzzy system, feature selection, noisy condition.

1 Introduction

It is a vital issue to reach the model of natural systems in almost all sciences for analyzing system behavior. Especially in engineering fields, a system model is employed, e.g., in the optimization, control, diagnosis, and fault detection cases [21]. On the other hand, intelligent techniques like fuzzy systems have been employed successfully in many applications, such as system identification among several existing methods. For example, using the fuzzy sets theory [36] in industrial control problems [11, 12, 37] and the combination of fuzzy control with neural networks [3, 26] have had good results, particularly for complex systems.

A cement rotary kiln is a complex cylindrical device that is the main part of cement industry equipment, consuming fuel to get pre-heated to a high temperature, which is necessary to produce clinker. It rotates around its axis, and the raw meal dust sticks adhesively to its walls, thus becoming burned and baked. A schematic representation of a cement production unit is shown in Figure 1 [31]. Many effective parts of the cement production process, such as baking the mixture of input materials, occur in the kiln [2], and thus the kilns operation affects the whole plant and it is necessary to obtain an efficient model for it. However, the rotary kiln is a non-linear and time-variant system, which is very complex. We can see a few influential works on the kilns' modeling during 1970-2003 [4, 5, 15, 16, 17, 19, 25, 28, 32] for example, in [17], a model based on computational fluid dynamics (CFD) has been presented. Some other new ideas have been developed for the rotary kilns in recent years, like the research work that has compared the BOX-Jenkins method with the usual linear techniques such as ARMAX and Output-Error (OE) [24]. However, along this procedure, applying artificial intelligence and expert systems are paid attention to widely, and are used for rotary kiln modeling and controls. In one of the models presented, neural networks such as multi-layer perceptron (MLP) have been used [7, 23].

Corresponding Author: N. Moradkhani
Received: May 2022; Revised: October 2022; Accepted: May 2023.
https://doi.org/10.22111/IJFS.2023.7653

Another research work contains a predictor and simulator model for the rotary kiln by the locally linear neuro-fuzzy method [29]. In addition, an adaptive neuro-fuzzy inference system (ANFIS) has been proposed, which uses particular selected inputs to identify each output [9]. The other one is a hierarchical wavelet fuzzy inference system (HWFIS) [31], which applies eight input variables for each output, and compares the mean square error (MSE) of HWFIS with nine other models. To improve the control performance of calcination zone temperature in a lime rotary kiln, a predictive control method based on a support vector machine (SVM) and improved particle swarm optimization (PSO) algorithm are proposed [38]. An artificial neural network for controlling the motor speeds of belt weighers and a separator of a cement grinding unit intended for the production of a three-component cement of various grades was developed and researched too [20]. Another research focused on rotary kiln modeling is a Takagi-Sugeno neuro-fuzzy system that consists of an input selection by a Genetic Algorithm [18]. The operational data from a cement factory are gathered, and the relationships between variables of the plant are analyzed by using an Artificial neural network [1]. Also, a study was issued recently, its objective was to identify and model a rotary cement kiln based on production history data by using an artificial neural network MLP algorithm [6].

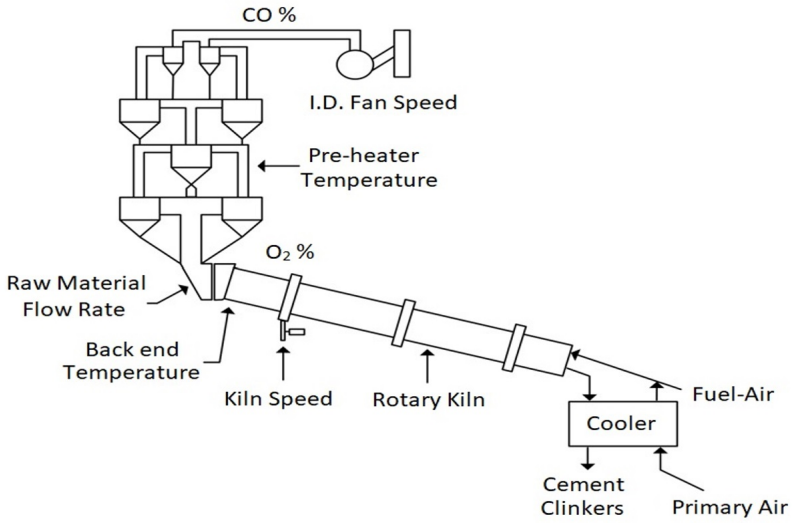

Figure 1: Schematic representation of cement production and rotary kiln operation [10].

A combination of fuzzy systems with neural networks leads to an impressive ability to model complex plants. So a Takagi-Sugeno neuro-fuzzy system (TSNFS) is used in this work. But on the other hand, type 2 fuzzy systems play an essential role in a noisy and uncertain situation. The basis of these systems is type 2 fuzzy sets including uncertainty which, with more degree of freedom, can make more accurate models. Thus the proposed identifier of this work is named type 2 Takagi-Sugeno neuro-fuzzy system (T2TSNFS).

The data applied in the simulations is related to the Saveh White Cement (SWC) factory. After a primary data preparation, an input selection is made with the genetic algorithm (GA) to prevent increasing computations and achieve a simple model. Then we start training T2TSNFS. Now, the identifier should be tested. Since noise is an undeniable fact in industrial systems, and the recent works have not focused on this critical subject seriously, a noisy condition is applied in the testing process of T2TSNFS.

This paper is organized as follows. In Section 2, the structure and learning algorithm of T2TSNFS is presented. Data preprocessing is given in Section 3. In Section 4, the simulations and results are presented. Finally, the conclusions are given in Section 5.

2 Type 2 Takagi-Sugeno neuro-fuzzy system

2.1 Structure

The main job in a fuzzy inference system is organized in its fuzzy rules. These rules in a type 2 Takagi-Sugeno fuzzy inference system are presented generally in the following form:

$$R^i : \text{If } x_1 \text{ is } \tilde{A}_1^i, \cdots, x_n \text{ is } \tilde{A}_n^i, \text{ then } f^i = F^i(X, \Gamma^i). \tag{1}$$

Where $x_1, \cdots, x_n$ are the input variables, $\tilde{A}_1^i (i = 1, \cdots, M, j = 1, \cdots, n)$ is the type 2 fuzzy set belonging to the j-th input in i-th fuzzy rule, f^i is the function of the consequent part of the i-th fuzzy rule, and F^i is a function of X and Γ^i, which are defined as following:

$$X = [x_1 \; x_2 \; \cdots \; x_n], \tag{2}$$

$$= \begin{bmatrix} a_0^1 & \cdots & a_0^M \\ a_1^1 & \cdots & a_1^M \\ \vdots & \ddots & \vdots \\ a_n^1 & \cdots & a_n^M \end{bmatrix}_{(n+1)\times M} \triangleq \Gamma. \tag{3}$$

Where a_0^i and $a_j^i (i = 1, \cdots, M, j = 1, \cdots, n)$ are assigned by real values, and since we use a first order Takagi-Sugeno fuzzy inference system, f^i is given by:

$$f^i = a_0^i + \sum_{j=1}^{n} a_j^i x_j, \tag{4}$$

$\mu_{\tilde{A}_j^i}(x_j)$ is the type 2 membership function (MF) of the j-th input in the i-th fuzzy rule, as following:

$$\mu_{\tilde{A}_j^i}(x_j) = \left[\underline{\mu}_{\tilde{A}_j^i}(x_j), \overline{\mu}_{\tilde{A}_j^i}(x_j) \right]. \tag{5}$$

Where $\underline{\mu}_{\tilde{A}_j^i}(x_j)$ is the lower MF and $\overline{\mu}_{\tilde{A}_j^i}(x_j)$ is the upper MF. Thus, the lower and the upper firing strength of this rule are calculated as:

$$\underline{w}^i = \underline{\mu}_{\tilde{A}_1^i}(x_1) * \cdots * \underline{\mu}_{\tilde{A}_n^i}(x_n) = \prod_{j=1}^{n} \underline{\mu}_{\tilde{A}_j^i}(x_j), \tag{6}$$

$$\overline{w}^i = \overline{\mu}_{\tilde{A}_1^i}(x_1) * \cdots * \overline{\mu}_{\tilde{A}_n^i}(x_n) = \prod_{j=1}^{n} \overline{\mu}_{\tilde{A}_j^i}(x_j). \tag{7}$$

Here, $*$ is the t-norm product operator. Thus the normalized firing strength of the i-th rule is given as:

$$W^i = \frac{\underline{w}^i + \overline{w}^i}{\Sigma_{i=1}^{M} \left(\underline{w}^i + \overline{w}^i \right)}. \tag{8}$$

Finally, the output of the Takagi-Sugeno fuzzy inference system is obtained by:

$$\hat{y} = \sum_{i=1}^{M} f^i W^i, \tag{9}$$

where f^i and W^i are determined using Eqs. 4 and 8, respectively.

The structure of the multi input-single output (MISO) T2TSNFS with n inputs and K MFs for each input is given in Figure 2. In the first layer, the input signals are distributed. In the second layer, for each input signal, the lower and upper membership degree belonging to the corresponding type 2 fuzzy set $\left(\mu_{\tilde{A}_j^i}(x_j) \right)$ are calculated. Realizing the inference engine happens in the third layer by determining the lower and upper firing strength of each rule. In the fourth layer, the amount of function in the consequent part is set. The product of fourth layers' output and firing strength for each rule is applied in the fifth layer. Finally, the type reduction and defuzzification is done in the sixth and seventh layers using the Nie-Tan method [22].

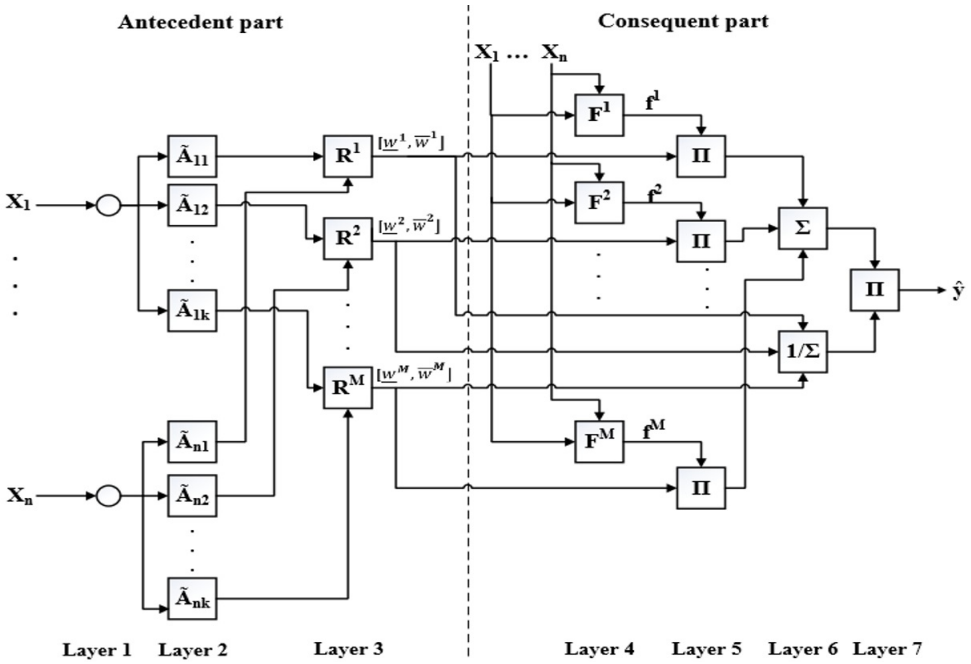

Figure 2: Structure of T2TSNFS.

2.2 Learning algorithm strategy

As it is obvious in Section 2.1, there are some parameters in the fuzzy rules that should be tuned. In this work, we use the reliable gradient descent (GD) algorithm to determine the final amount of these parameters. Mean Square Error (MSE) is used as the cost function, defined by:

$$J = \frac{1}{N}\sum_{p=1}^{N} E_p. \tag{10}$$

Where N is the total number of data samples, and E_p is given as:

$$E_p = \frac{1}{2}\left(y_p - \hat{y_p}\right)^2, \tag{11}$$

where y_p and $\hat{y_p}$ are the targets and the output of fuzzy identifier, respectively. For simplicity, we use E instead of E_p in the following equations.

We chose Gaussian MF for the antecedent part of each rule. Thus, according to Eq. 5 we have the lower and the upper MFs as following:

$$\underline{\mu}_{\tilde{A}_j^i}(x_j) = e^{-\frac{\left(x_j - c_m^j\right)^2}{\left(\underline{\sigma}_m^j\right)^2}}, \tag{12}$$

$$\overline{\mu}_{\tilde{A}_j^i}(x_j) = e^{-\frac{\left(x_j - c_m^j\right)^2}{\left(\overline{\sigma}_m^j\right)^2}}. \tag{13}$$

Where $\underline{\sigma}_m^j$ and $\overline{\sigma}_m^j$ are the lower and the upper standard deviation (STD) and c_m^j is the mean belonging to the m-th MF of the j-th input, respectively. Also $m = 1, \cdots, K$ and $j = 1, \cdots, n$.

Since we consider the STD has a lower and an upper bound, there will be three parameters that belong to MFs should be learned; including $\underline{\sigma}_m^j$, $\overline{\sigma}_m^j$ and c_m^j.

According to the GD algorithm, by applying the momentum factor [27] to improve the convergence speed, $\underline{\sigma}_m^j$ is updated by:

$$\underline{\sigma}_m^j\left(k+1\right)=\underline{\sigma}_m^j(k)-\eta_\sigma(k)\left(\frac{\partial J}{\partial\underline{\sigma}_m^j}\right)_k+\gamma_\sigma\left(\underline{\sigma}_m^j(k)-\underline{\sigma}_m^j(k-1)\right). \tag{14}$$

Where $\eta_\sigma(k)$ is the adaptive learning rate updating by the Bold Driver method [33, 34], γ_σ is the momentum factor, and k indicates the learning step. Also $\frac{\partial J}{\partial\underline{\sigma}_m^j}$ in 14 is obtained according to 10 by:

$$\frac{\partial J}{\partial\underline{\sigma}_m^j}=\frac{1}{N}\sum_{p=1}^{N}\left(\frac{\partial E}{\partial\underline{\sigma}_m^j}\right)_p. \tag{15}$$

Where the derivatives are computed by applying chain rule, so we have:

$$\frac{\partial E}{\partial\underline{\sigma}_m^j}=-(y-\hat{y})\frac{(x_j-c_m^j)^2}{(\underline{\sigma}_m^j)^3}\frac{1}{\Sigma_{i=1}^{M}\left(\underline{w}^i+\overline{w}^i\right)}\times\sum_{q\in Q_\sigma}\underline{w}^q\left(f^q-\hat{y}\right). \tag{16}$$

Where Q_σ is the set consisting of all rule indices containing $\underline{\sigma}_m^j$ (e.g. in a system with $n=5$ and $K=2$, there are 32 rules; $\underline{\sigma}_1^1$ exists in the rules 1 to 16, thus $Q_\sigma=1,2,,16$) and f^q and $\overline{w}^q$ are given by Eqs. 4 and 6, respectively. The same process should be followed to reach the related terms for updating $\overline{\sigma}_m^j$ as in Eqs. 17 to 19:

$$\overline{\sigma}_m^j\left(k+1\right)=\overline{\sigma}_m^j(k)-\eta_\sigma(k)\left(\frac{\partial J}{\partial\overline{\sigma}_m^j}\right)_k+\gamma_\sigma\left(\overline{\sigma}_m^j(k)-\overline{\sigma}_m^j(k-1)\right), \tag{17}$$

$$\frac{\partial J}{\partial\overline{\sigma}_m^j}=\frac{1}{N}\sum_{p=1}^{N}\left(\frac{\partial E}{\partial\overline{\sigma}_m^j}\right)_p, \tag{18}$$

$$\frac{\partial E}{\partial\overline{\sigma}_m^j}=-(y-\hat{y})\frac{(x_j-c_m^j)^2}{(\overline{\sigma}_m^j)^3}\frac{1}{\Sigma_{i=1}^{M}\left(\underline{w}^i+\overline{w}^i\right)}\times\sum_{q\in Q_\sigma}\overline{w}^q\left(f^q-\hat{y}\right). \tag{19}$$

Also updating c_m^j is possible by:

$$c_m^j\left(k+1\right)=c_m^j(k)-\eta_c(k)\left(\frac{\partial J}{\partial c_m^j}\right)_k+\gamma_c\left(c_m^j(k)-c_m^j(k-1)\right). \tag{20}$$

Where $\eta_c(k)$ is the adaptive learning rate, γ_c is the momentum factor, and $\frac{\partial J}{\partial c_m^j}$ is calculated by:

$$\frac{\partial J}{\partial c_m^j}=\frac{1}{N}\sum_{p=1}^{N}\left(\frac{\partial E}{\partial c_m^j}\right)_p, \tag{21}$$

Where $\frac{\partial E}{\partial c_m^j}$, by applying chain rule, is given by:

$$\frac{\partial E}{\partial c_m^j}=-(y-\hat{y})(x_j-c_m^j)\frac{1}{\Sigma_{i=1}^{M}\left(\underline{w}^i+\overline{w}^i\right)}\times\left\{\frac{1}{(\overline{\sigma}_m^j)^2}\sum_{q\in Q_c}\underline{w}^q\left(f^q-\hat{y}\right)\right\}. \tag{22}$$

Here, Q_c is the set consisting of all rule indices containing c_m^j. Also in the consequent part of the rules, Γ should be updated by:

$$\Gamma(k+1)=\Gamma(k)-\eta(k)\left(\frac{\partial J}{\partial\Gamma}\right)_k+\gamma\left(\Gamma(k)-\Gamma(k-1)\right). \tag{23}$$

$\frac{\partial J}{\partial\Gamma}$ in Eq. 23 is given according to Eq. 10, as follows:

$$\frac{\partial J}{\partial\Gamma}=\frac{1}{N}\sum_{p=1}^{N}\left(\frac{\partial E}{\partial\Gamma}\right)_p. \tag{24}$$

In which by applying chain rule, we have:

$$\frac{\partial E}{\partial \Gamma} = -(y - \hat{y}) \begin{pmatrix} 1 \\ X^T \end{pmatrix} W. \tag{25}$$

Where W is given as:

$$W = \begin{bmatrix} W^1 & W^2 & \cdots & W^M \end{bmatrix}. \tag{26}$$

3 Data preprocessing

Another issue that we should note, to get the accurate results in an identification process, is to use the real dataset. The data we use in the simulations is collected from the SWC factory, which includes high-frequency noise, offset, and maybe sudden variations, so it is necessary to prepare valid data for applying to the model. We will explain how the data is preprocessed in the rest of this section.

3.1 Input-output variables

In our modeling, the data corresponding to 12 weeks, which consists of nine inputs and outputs, given in Table 1.

Table 1: Input and output variables in cement rotary kiln.

Variable Name	**Abbreviation**	**Symbol**	**Type**
Material feed rate	MAT	x_1	Input
Fuel feed rate	FU	x_2	Input
Kiln speed	KS	x_3	Input
LD fan speed	FA	x_4	Input
Secondary air pressure	AP	x_5	Input
Kiln ampere	KA	y_1	Output
CO content	CO	y_2	Output
Back-end temprature	BE	y_3	Output
Pre-heater temprature	PRE	y_4	Output

3.2 Sampling

Since continuous signals are not usable in computing tools such as computers, a conversion to a discrete type with a proper sampling frequency is required. Here, we determine T_s, the sampling time, as follows:

$$T_s = \frac{\tau_m in}{3}. \tag{27}$$

Where $\tau_m in$ is the smallest time constant of the system that has been calculated to be three minutes [9, 29]. Thus T_s will be 60 seconds by Eq. 27.

3.3 Detecting outlier data

The samples may have different behaviors beyond expectations, called the outlier data. It is possible to obtain a compact clustering for data by getting rid of or replacing them. There are several methods available to determine the outlier data. Here, we use T^2 statistics [35] for this goal. Suppose that X is a matrix containing N samples with dimension P, as follows:

$$X = \begin{bmatrix} x_1^T \\ x_2^T \\ \vdots \\ x_N^T \end{bmatrix}_{N \times P}. \tag{28}$$

First, a clustering is applied to this dataset by the k-means algorithm to reach k clusters named $C_1, \cdots, C_k$ with centers $c_1, \cdots, c_k$, respectively. Then T^2 statistics for samples of each cluster is calculated by:

$$T_j^2 = (x_j - c_i)^T S_i^{-1} (x_j - c_i). \tag{29}$$

Where $x_j \in C_i, i = 1, \cdots, k, j = 1, \cdots, n_i$ and n_i is the number of samples in the i-th cluster. Also S_i in Eq. 29 is obtained by:

$$S_i = \frac{1}{n_i - 1} \sum_{j=1}^{n_i} (x_j - c_i)(x_j - c_i)^T. \tag{30}$$

Finally, each sample will be checked by:

$$T_j^2 \leq \alpha^2. \tag{31}$$

Where α determines the number of outlier data (smaller α results in more outliers). The samples satisfying the condition in Eq. 31 are denoted as normal, otherwise outlier. Then the outliers are replaced with the mean value of the previous and next samples.

3.4 Filtering

In a high-frequency range, some sudden changes and peaks with a large amount of energy may occur in data. Since they degrade the models performance, it is necessary to smooth them by passing the data through a proper filter [31]. To reach this goal, a first-order Butterworth low-pass filter with a 10^{-4} Hz bandwidth is used here.

3.5 Applying input-output delay

The pure delay of a system is always a very effective parameter in the identification process. For the rotary kiln of the SWC factory, as it has been shown in Table 2, we use the delay results mentioned in [13] through the Lipschitz method.

Table 2: Inputoutput delays (minute).

Variable Name	**KA**	**CO**	**BE**	**PRE**
MAT	10	15	30	18
FU	25	5	10	4
KS	0	5	40	36
FA	10	0	5	0
AP	30	4	5	0

3.6 Input selection using GA

Since increasing the number of inputs, results in more complexity in fuzzy systems, the best subset of inputs is selected in this work by applying GA to prevent increasing computations and achieving the best results. For this purpose, first, we gathered the more effective inputs for identifying each output (concluding the effective dynamics of kilns' inputs and the corresponding output) by analyzing the research works that have been done later [7, 8, 9]. Then we used GA to select the best ones. Figure 3 demonstrates the block diagram for rotary kiln identification with input selection utilizing GA, and all the parameters required in GA are gathered in Table 3.

Table 3: Parameters used in GA.

Parameter	**Value**
Population size	100
Maximum generation	20
Crossover rate	0.5
Mutation rate	0.2
W_1	10^4
W_2	1

In Figure 3, U(k) is defined as:

$$U(k) = \begin{bmatrix} x_1 & x_2 & x_3 & x_4 & x_5 \end{bmatrix}^T. \tag{32}$$

The results of applying GA show that first two dynamics of each output i.e., $y_i(k-1), y_i(k-2), i = 1, \cdots, 4$ are the most proper inputs for it [18].

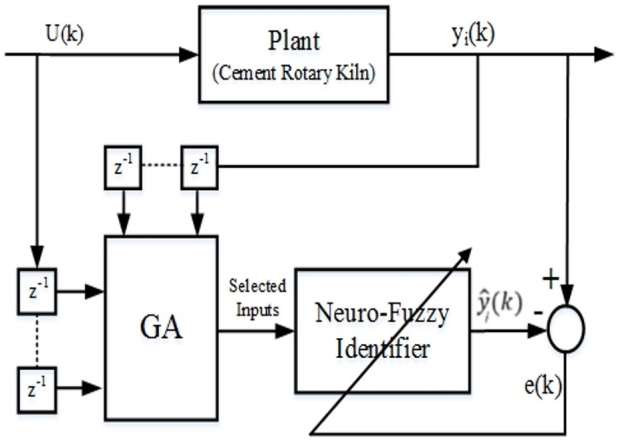

Figure 3: Identification block diagram with input selection by GA [18].

3.7 Normalizing data

We normalize the data in the $[0, 1]$ interval because the inputs and outputs have different ranges that lead to errors in data quantization, and consequently, the plant is not identified well [29, 30].

4 Simulation and results

In this section, the preprocessed data is applied to the T2TSNFS. Also, the type 1 Takagi-Sugeno neuro-fuzzy system (T1TSNFS) is simulated separately to achieve a better comparison. T1TSNFS and T2TSNFS are different in the type of fuzzy sets.

The kiln that is a multi-input multi-output (MIMO) system is supposed as four multi-input single-output (MISO) systems. Each of these dynamic MISO models, is used in the form of schematic representation is shown in Figure 4.

Structure learning of the model to determine the number of inputs is done by GA (mentioned in the previous section). Also, two MFs as the optimal number of MFs for each input have been achieved by trial and error. Thus there are four fuzzy rules with 20 learning parameters for T2TSNFS that are tuned by the GD method. Also the prepared data is divided into three parts consisting of 50% for train, 20% for validation, and 30% for testing. In Figure 4, $n(k)$ is a zero mean Gaussian noise added to the model to evaluate it in different noisy situations depending on the noise STD. In these simulations, we consider two cases. In case I, σ_{Train} is fixed, and σ_{Test} takes different values. In detail, $\sigma_{Train} = 0.7$ for output 1 (KA), $\sigma_{Train} = 0.01$ for output 2 (CO), $\sigma_{Train} = 2$ for output 3 (BE), and $\sigma_{Train} = 1.5$ for output 4 (PRE). σ_{Train} and σ_{Test} denote the STD of noise in the train and the test data, respectively (the amount of σ_{Train} and σ_{Test} are selected based on each outputs' range and STD). In case II, σ_{Test} is fixed, and different values are assigned to σ_{Train}. In detail, $\sigma_{Test} = 0.7$ for output 1 (KA), $\sigma_{Test} = 0.01$ for output 2 (CO), $\sigma_{Test} = 2$ for output 3 (BE) and $\sigma_{Test} = 1.5$ for output 4 (PRE).

According to the results of case I, as it is shown in Table 4, we conclude in the same σ_{Train} if the amount of σ_{Test} increases, the MSE value will grow up for both T1TSNFS and T2TSNFS, which means the operation of the identifier is logical. The more important result occurred in more significant amounts of σ_{Test}. It shows that T2TSNFS performs better than T1TSNFS. In other words, in the situation where more considerable noises may affect the system, T2TSNFS is the more suitable option for identification. The comparison between T1TSNFS and T2TSNFS MSEs, in this case, is shown in Figure 5.

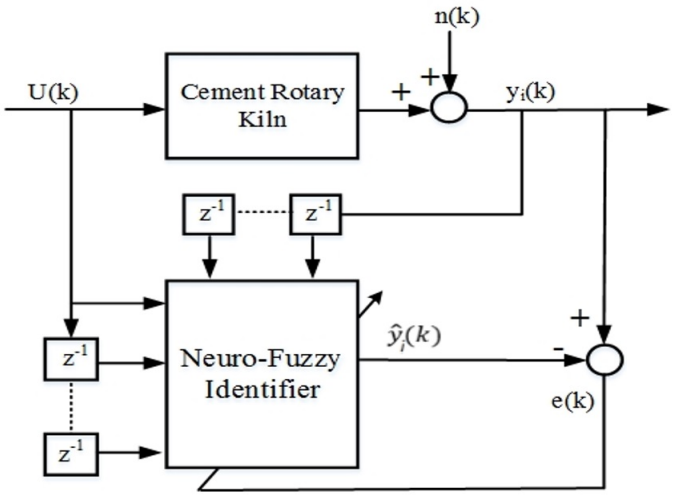

Figure 4: Schematic for dynamic nonlinear fuzzy identification of cement rotary kiln [18].

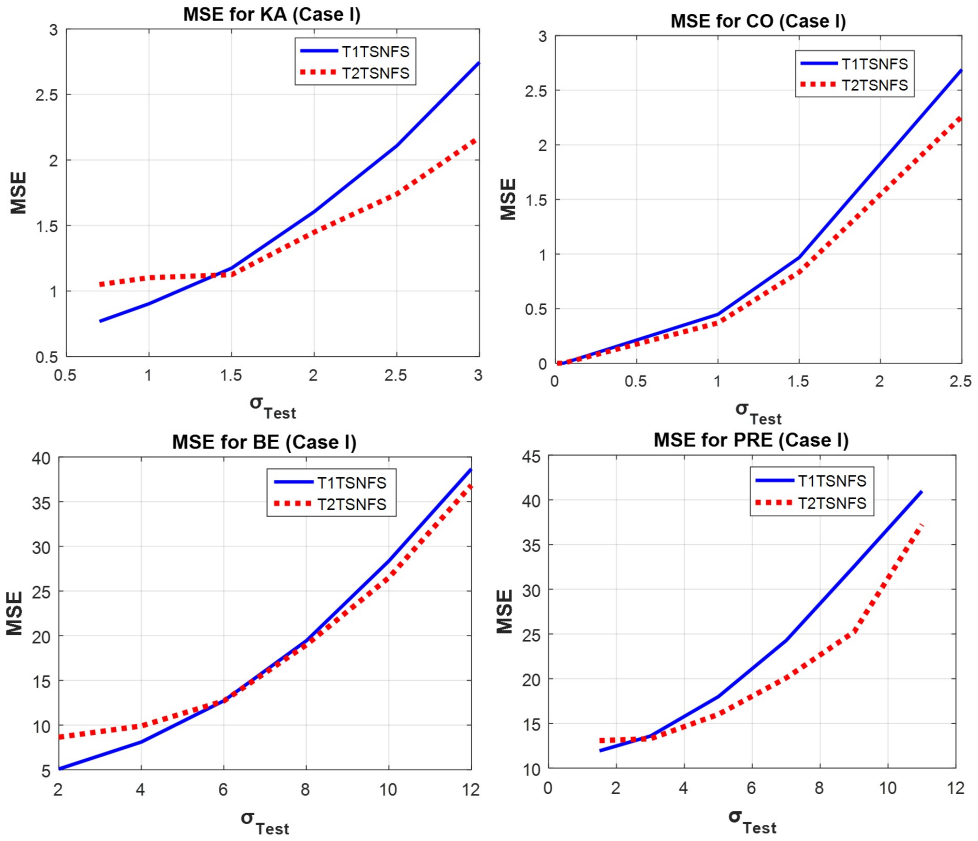

Figure 5: Graphic representation of Comparison between T1TSNFS and T2TSNFS MSEs in noisy condition (Case I: Fixed σ_{Train} and variable σ_{Test})

Moreover, Case II denotes an exciting result; As it is shown in Table 5, by growing σ_{Train} up and setting σ_{Test} fixed, the amount of error decreases for both T1TSNFS and T2TSNFS. It means that when the training data is involved in a situation with more significant noise, both T1TSNFS and T2TSNFS operate better in a situation with lower noise. The actual and identified outputs for KA in case II are shown in Figure 6. (It is necessary to notice, the results that are mentioned in Case II, are related to a fixed σ_{Test}, and we do not compare T1TSNFS and T2TSNFS with each other.

Table 4: MSE for T1TSNFS and T2TSNFS in noisy condition (Case I: Fixed σ_{Train} and variable σ_{Test}).

(a) Output 1: KA ($\sigma_{Train} = 0.7$)

σ_{Test}	MSE	
	T1TSNFS	**T2TSNFS**
0.7	0.76819	1.0495
1	0.90307	1.102
1.5	1.175	1.1245
2	1.6058	1.4493
2.5	2.1089	1.7408
3	2.7462	2.1728

(b) Output 2: CO ($\sigma_{Train} = 0.01$)

σ_{Test}	MSE	
	T1TSNFS	**T2TSNFS**
0.01	2.3417×10^{-4}	6.764×10^{-4}
0.03	5.5446×10^{-4}	10.83×10^{-4}
0.05	12.19×10^{-4}	15.775×10^{-4}
1	0.44848	0.36961
1.5	0.96868	0.83526
2.5	2.6893	2.259

(c) Output 3: BE ($\sigma_{Train} = 2$)

σ_{Test}	MSE	
	T1TSNFS	**T2TSNFS**
2	5.0838	8.6484
4	8.1078	9.9004
6	12.7212	12.7058
8	19.4346	18.9782
10	28.365	26.503
12	38.6691	36.8401

(d) Output 4: PRE ($\sigma_{Train} = 1.5$)

σ_{Test}	MSE	
	T1TSNFS	**T2TSNFS**
1.5	11.9346	13.0809
3	13.5866	13.2868
5	17.9917	16.0115
7	24.2745	20.0891
9	32.5593	25.2379
11	40.9752	37.2347

Table 5: MSE of T1TSNFS and T2TSNFS in noisy condition (Case II: Fixed σ_{Test} and variable σ_{Train}).

(a) Output 1: KA ($\sigma_{Test} = 0.7$)

σ_{Train}	MSE	
	T1TSNFS	**T2TSNFS**
0.7	0.76819	1.0495
1	0.46208	0.80449
1.5	0.1503	0.14388

(b) Output 2: CO ($\sigma_{Test} = 0.01$)

σ_{Train}	MSE	
	T1TSNFS	**T2TSNFS**
0.01	2.3417×10^{-4}	6.764×10^{-4}
0.02	1.7899×10^{-4}	6.3855×10^{-4}
0.05	0.3128×10^{-4}	0.399×10^{-4}

(c) Output 3: BE ($\sigma_{Test} = 2$)

σ_{Train}	MSE	
	T1TSNFS	**T2TSNFS**
2	5.0838	8.6484
3	2.7945	6.2138
4	1.1908	1.1851

(d) Output 4: PRE ($\sigma_{Test} = 1.5$)

σ_{Train}	MSE	
	T1TSNFS	**T2TSNFS**
1.5	11.9346	13.0809
3.5	5.2437	10.7639
5	0.89077	0.7688

The main goal, in this case, is to determine the action of these models when σ_{Train} is varying. Furthermore, an "error condition" has been specified in the training algorithm to reach comparable results. Thus, these MSE values can be significantly lower by changing the "error condition" in a trial and error process.)

5 Conclusion and future work

5.1 Conclusion

In this paper, a type 2 Takagi-Sugeno neuro-fuzzy system (T2TSNFS) was proposed to identify the Saveh White Cement rotary kiln in noisy conditions. The basis of T2TSNFS is a set of fuzzy rules consisting of type 2 fuzzy sets in the antecedent part and a linear function in the consequent part that enables the system to give a better model for nonlinear dynamic plants and handle noisy information effectively. Also, the GD algorithm was applied for updating parameters, and the data preprocessing was done thoroughly. Especially, input selection based on GA, plays a crucial role in the models' simplicity and accurate results.

The efficiency of the model is shown through simulations. A comparison between the results for T2TSNFS and T1TSNFS (that has not been noted in the recent studies) showed that it is better to use T1TSNFS in the situation with lower noise, and it is not necessary to involve complex models. Because the T1TSNFS has a smaller MSE de-

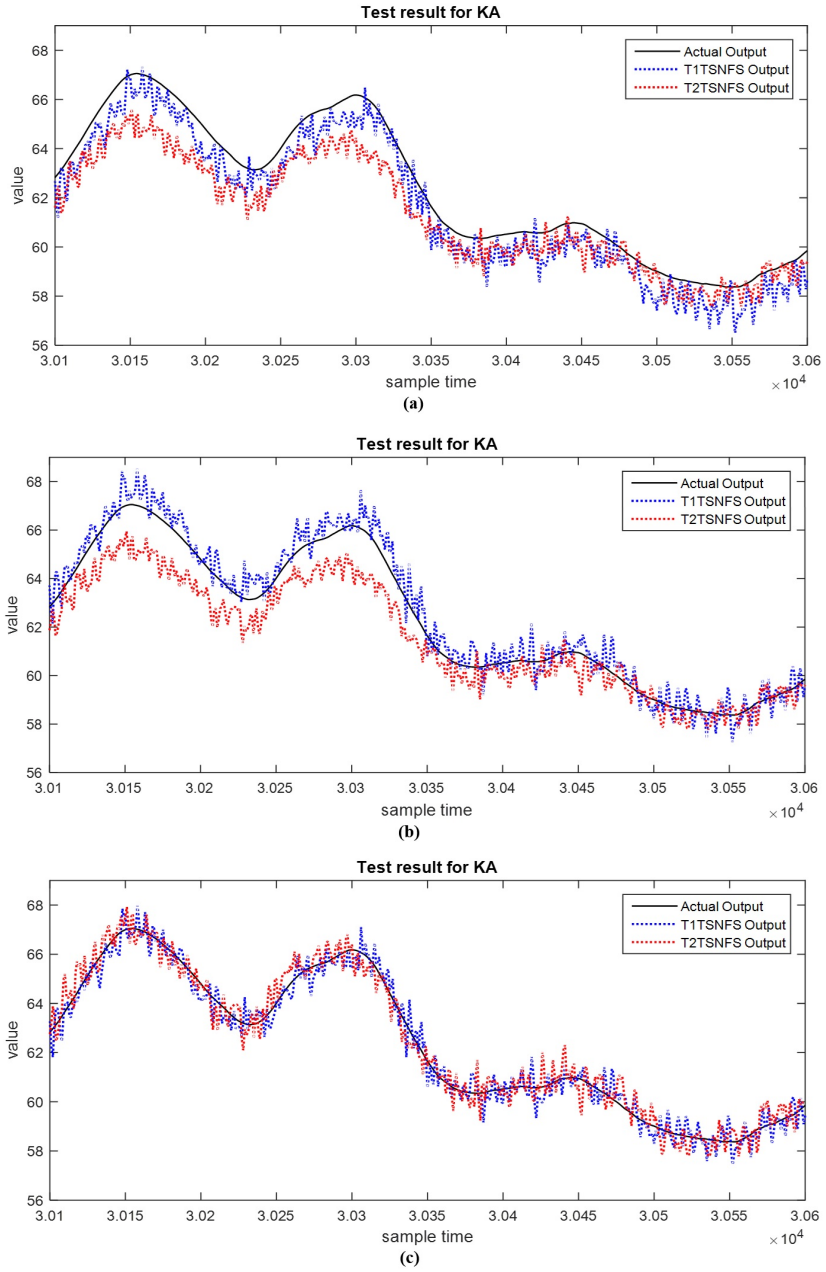

Figure 6: Plot of the actual and identified output for KA in case II (a) σ_{Train}=0.7 (b) σ_{Train}=1 (c) σ_{Train}=1.5.

spite its simple structure, so time and cost will be saved significantly. But when the noise on the test data has a larger STD, T2TSNFS has smaller MSE values and performs better despite the small number of rules and learning parameters.

In fact, type 2 fuzzy systems properties makes it possible. Type 2 fuzzy rules can use an interval to adjust their parameters in which the model finally will find the optimal values. This property is not seen in type 1 fuzzy systems, or the same models have been applied recently.

Also, when T2TSNFS and T1TSNFS were evaluated in the noisy condition, they had relatively successful results; For more significant amounts of noise on test data, the error increases for both models. So we conclude the error is tracking the noise, and the models know the noise well. It means the operation of the identifier is logical.

Besides, we conclude both T2TSNFS and T1TSNFS work better (they had less error) when their parameters are tuned in a more robust noisy condition (more considerable amount of σ_{Train}. Because, a more vigorous training process

causes a more powerful and efficient model, and if we want to use these models as an identifier in a noisy situation, it is better to involve the training data with a relatively considerable amount of noise.

5.2 Future work

In our future work, we will consider the changes, such as using the recurrent structure in the proposed model and having more effective comparisons between different models.

Acknowledgement

The authors wish to express their appreciation for several excellent suggestions for improvements in this paper made by the referees.

References

[1] H. Aghdasinia, S. S. Hosseini, J. Hamedi, *Improvement of a cement rotary kiln performance using artificial neural network*, Journal of Ambient Intelligence and Humanized Computing, **12** (2021), 7765-7776.

[2] M. Akalp, A. L. Dominguez, R. Longchamp, *Supervisory fuzzy control of a rotary cement kiln*, The Proceedings IEEE Electrotechnical Conference, **2** (1994), 754-757.

[3] A. Y. Alanis, E. N. Sanchez, A. G. Loukianov, E. A. Hernandex, *Discrete-time recurrent high order neural networks for nonlinear identification*, Journal of the Franklin Institute, **347**(7) (2010), 1253-1265.

[4] P. V. Barr, J. K. Brimacombe, A. P. Watkinson, *A heat transfer model for the rotary kiln: Part I, Pilot kiln trials*, Metallurgical Transactions B, **20**(B) (1989), 391-402.

[5] A. A. Boateng, P. V. Barr, *A thermal model for the rotary kiln including heat transfer within the bed*, International Journal of Heat and Mass Transfer, **39** (1996), 2131-2147.

[6] M. Z. Doghmane, M. Kidouch, S. Eladj, A. Ouali, *Identification and modeling of a rotary kiln in cement plant based on ANN (MLP)*, Hatti, Artificial Intelligence and Heuristics for Smart Energy Efficiency in Smart Cities. IC-AIRES. Lecture Notes in Networks and Systems, **361**, Springer, Cham, 2021.

[7] M. Fallahpoor, A. Fatehi, B. N. Araabi, M. Azizi, *Identification of cement rotary kiln using neural networks*, First Joint Congress on Fuzzy and Intelligent Systems Ferdowsi University of Mashhad, Iran, (2007), 29-31.

[8] M. Fallahpour, A. Fatehi, B. N. Araabi, M. Azizi, *A neuro-fuzzy controller for rotary cement kilns*, Proceedings of the 17th World Congress The International Federation of Automatic Control, Seoul, Korea, July 6-11, 2008.

[9] A. Fatemi Moghaddam, A. Sharifi, M. Teshnehlab, *Prediction and identification of nonlinear rotary cement kiln system with neuro-fuzzy ANFIS network by using feature selection with genetic algorithm*, Journal of Control, **5**(2) (2011), 22-33.

[10] V. Hoenig, *Evaluation of the energy performance of cement kilns in the context of co-processing (Technical Report)*, European Cement Research Academy GmbH (ecra), 2016, Available on: https://cembureau.eu/media/oyahklgk/12042-ecra-energy-performance-cement-kilns-2017-10-15.pdf.

[11] D. Karaboga Bagis, *Evolutionary algorithm-based fuzzy PD control of spillway gates of dams*, Journal of the Franklin Institute, **344**(8) (2007), 1039-1055.

[12] C. C. Ku, P. H. Huang, W. J. Chang, *Passive fuzzy controller design for nonlinear systems with multiplicative noises*, Journal of the Franklin Institute, **347**(5) (2010), 732-750.

[13] I. Makaremi, A. Fatehi, B. N. Araabi, *Lipschitz numbers: A medium for delay estimation*, Proceedings of the 17th IFAC World Congress, 2008.

[14] I. Makaremi, A. Fatehi, B. N. Araabi, *Abnormal condition detection in cement rotary kiln with system identification methods*, Elsevier Ltd. Journal of Process Control, **19** (2009), 1538-1545.

[15] M. A. Martins, L. S. Oliveira, A. S. Franca, *Modeling and simulation of petroleum coke calcination in rotary kilns*, Fuel, **80** (2001), 1611-1622.

[16] M. A. Martins, L. S. Oliveira, A. S. Franca, *Modeling and simulation of limestone calcination in rotary kilns*, ZKG International, **55**(4) (2002), 76-87.

[17] E. Mastorakos, A. Massias, C. D. Tsakiroglou, D. A. Goussis, V. N. Burganos, *CFD predictions for cement kiln including flame modeling, heat transfer and clinker chemistry*, Applied Mathematical Modeling, **23** (1999), 55-76.

[18] N. Moradkhani, M. Teshnehlab, *Identification of cement rotary kiln in noisy condition using Takagi-Sugeno neuro-fuzzy system*, Journal of AI and Data Mining, **7**(3) (2019), 367-375.

[19] K. S. Mujumdar, V. V. Ranade, *CFD simulations of heat transfer in transverse plane of rotating kilns*, Chemcon, 2002.

[20] E. A. Muravyova, R. R. Mustaev, *Development of an artificial neural network for controlling motor speeds of belt weighers and separator in cement production*, Optical Memory and Neural Networks, **26** (2018), 289-297.

[21] O. Nelles, *Nonlinear system identification*, Springer-Verlag, Berlin, 2001.

[22] M. Nie, W. W. Tan, *Towards an efficient type-reduction method for intervaltype-2 fuzzy logic systems*, IEEE International Conference on Fuzzy Systems, Hong Kong, 2008.

[23] G. Noshirvani, A. Fatehi, B. N. Araabi, M. Shirvani, M. Azizi, *Comparison of rotary cement kiln identified models*, Proceedings of the IEEE International Conference on Control and Automation, (2009), 1290-1295.

[24] R. Noshirvani, M. Shirvani, A. Fatehi, *Linear identification of white rotary cement kiln*, ICINCO, Italy, 2009.

[25] M. Paul, K. S. Mujumdar, V. V. Ranade, *Modeling of rotary cement kilns*, Chemcon, 2002.

[26] M. L. Presti, R. Poluzzi, A. Zanaboni, *Synthesis of fuzzy controller through neural networks*, Fuzzy Sets and Systems, **71**(1) (1996), 46-70.

[27] N. Qian, *On the momentum term in gradient descent learning algorithms*, Neural Networks, **12** (1999), 145-151.

[28] V. V. Ranade, K. S. Mujumdar, *CFD simulations of solid motion in the transverse plane of rotating kilns*, Proceedings of the third International Conference on CFD in the Minerals and Process Industries, 2003.

[29] M. Sadeghian, A. Fatehi, *Identification of nonlinear predictor and simulator models of a cement rotary kiln by locally linear neuro-fuzzy technique*, Proceedings of the Second International Conference on Computer and Electrical Engineering, (2009), 168-173.

[30] Z. Saghian, A. Esfahanipour, B. Karimi, *A novel Kumaraswamy interval type-2 TSK fuzzy logic system for subway passenger demand prediction*, Iranian Journal of Fuzzy Systems, **19**(3) (2022), 69-87.

[31] A. Sharifi, M. Aliyari Shoorehdeli, M. Teshnehlab, *Identification of cement rotary kiln using hierarchical wavelet fuzzy inference system*, Journal of the Franklin Institute, **349** (2012), 162-183.

[32] H. A. Spang, *A dynamic model of a cement kiln*, Automatica, **8** (1972), 309-323.

[33] The Netlib Website, *The bold driver network*, 1994. Available on: http://www.netlib.org/utk/lsi/pcwLSI/text/node 226.htmlfig3opt (accessed 12 October 2022).

[34] The Willamette University website, *Momentum and Learning Rate Adaptation*, 1999. Available on: https://www.willamette.edu/ gorr/classes/cs449/momrate.html (accessed 12 October 2022).

[35] J. D. Williams, W. H. Woodall, J. B. Birch, J. H. Sullivan, *Distribution of hotelling's T2 statistic based on the successive differences estimator*, Journal of Quality Technology, **38**(3) (2006), 217-229.

[36] L. A. Zadeh, *Fuzzy sets*, Information and Control, **8** (1965), 338-353.

[37] L. A. Zadeh, *A rationale for fuzzy control*, World Scientific Publishing Co. (1996), 123-126.

[38] T. Zhongda, L. Shujiang, W. Yanhong, W. Xiangdong, *SVM predictive control for calcination zone temperature in lime rotary kiln with improved PSO algorithm*, Transactions of the Institute of Measurement and Control, **40**(10) (2018), 3134-3146.