Interval Type-2 Recursive Fuzzy C-Means Clustering Algorithm in the TS Fuzzy Model Identification

Tanmoy Dam
Electrical Engineering Department
Indian Institute of Technology Kharagpur
Kharagpur-721302, India
Email: damtanmoy10@gmail.com

Alok Kanti Deb
Electrical Engineering Department
Indian Institute of Technology Kharagpur
Kharagpur-721302, India
Email: alokkanti@ee.iitkgp.ernet.in

***Abstract*—This paper presents an iterative Takagi Sugeno Fuzzy Model (TSFM) identification. Interval Type-2 Recursive Fuzzy C-Means (IT2RFCM) clustering algorithm has been used to classify the data space to obtain premise variable parameters and Weighted Recursive Least Square (WRLS) technique has been used to determine consequence parameters of each linear model. IT2RFCM clustering algorithm has been obtained from type-1 Fuzzy C-Means clustering algorithm by introducing fuzziness parameters. The effectiveness of the proposed IT2RFCM algorithm has been validated on Mackey-Glass time series data.**

I. Introduction

Takagi-Sugeno Fuzzy Model (TSFM) is extensively used in various real-time applications, especially in model-based control and model-based fault diagnosis [1]. The qualitative approach based TSFM is a systematic process to generating fuzzy rules from a given input-output data set [2], [3]. The clustering algorithm has been used to determine the partitions which is related to rule in TSFM domain. However, Type-1 fuzzy clustering based TSFM can not replicate the real plant model with high accuracy. Therefore, a type-2 fuzzy set can be introduced in fuzzy clustering algorithm that makes TSFM to replicate the original model with high accuracy [4], [5]. Mendel extended and developed a new fuzzy set theory from type-1 fuzzy set, known as type-2 fuzzy set theory [6], [7].

The main tasks to design TSFM are how to determine the number of rules and model parameters. Rules are part of structure identification and variables of each rule are part of parameter identification. Two approaches have been used to establish rule-base in a given data set i.e. off-line (batch processing) and on-line (recursive). In an off-line method, all the data samples are available at a time for learning the process. An on-line learning process is entirely different from off-line i.e. learning of rules have been done by an incremental learning process.

There are many off-line clustering algorithm in TSFM identification domain [8]–[10]. But, in the recent years, on-line clustering algorithm has generated interest among researchers for constructing the rules from streaming data. Evolving Takagi-Sugeno model (eTS) is most renowned method existing in online nonlinear fuzzy model identification domain [11]. The further development is done by incorporating informative potential nature in this method. The informative potential helps to capture the new data in rule base. The autonomous modeling skill is found in eTS, known as eTS+ [12]. In eTS+, the user defined parameters such as zone influence, density, utility criteria are required for ensemble learning based fuzzy model. Evolving Participatory Learning (ePL) is an upgraded version of eTS+. ePL has been formulated on participatory learning clustering algorithm [13]. Later, modification of ePL, known as ePL+, is also found [14]. Beside, a conceptually similar approach is found in TSFM domain, known as, Dynamic Evolving Neural Fuzzy Inference System (DENFIS) [3]. The distance based evolving clustering algorithm has been used for identifying the structure and WRLS is adopted to identifying the consequence parameters. A recursive version of vector quantization clustering algorithm has been used to construct Flexible Fuzzy Inference Systems (FLEXFIS) [15].

The self organization neural network is another recursive online method that is generated by error criteria and generalization ability of the network. The self-organizing fuzzy neural network [16] uses nearest heuristic to define width of gaussian membership functions. Self-adaptive fuzzy inference network uses learning-induced partition mechanism to classify the data space. Similar approach is also found in Self-Constructing Fuzzy Neural Network (SCFNN) [17], Generalized Adaptive Neural Fuzzy Inference System (GANFIS) [18]. The other evolving methods are presented in online learning identification domain [19]–[21].

The Fuzzy C-Means (FCM) [22] is most popular clustering algorithm which has been used in TSFM for both cases: offline and online. In this paper, the modification of recursive FCM (type-1)(rFCM) [23] [24] clustering algorithm has been done by interval type-2 fuzzy logic. The modified algorithm is represented as interval type-2 recursive FCM (IT2RFCM) in this study.

The paper is organized as follows: General structure of TSFM is described in section II. Section III describes type-2 fuzzy set theory and its extension to interval type-2 fuzzy set theory. Detailed descriptions of fuzzy c-means clustering algorithm are presented in section IV. IT2RFCM algorithm is presented in Section V. In section VI, the effectiveness of IT2RFCM algorithm based TS fuzzy model has been demonstrated in the validation of a complex nonlinear model. Section VII contains concluding remarks.

978-1-4799-7560-0/15 $31.00 © 2015 IEEE
DOI 10.1109/SSCI.2015.14

II. TS Fuzzy Model

A multi-input multi-output (MIMO) dynamical system can be decoupled and represented as sum of multi-input single-output (MISO) systems. Here, MISO plant is assumed to be modeled by TS fuzzy logic reasoning and its implication. TSFM is represented by a set of "IF-THEN" rules and i-th rule can be represented as,

$$\begin{aligned} P^i:\ & IF\, x_1(k)\, is\, A_1{}^i\, and\, x_2(k)\, is\, A_2{}^i\, and\ \ldots \\ & and\, x_n(k)\, is\, A_n{}^i \\ & THEN\, y^i(k) = p_0{}^i + \sum_{j=1}^{n} x_j(k) * p_j{}^i \\ & = 1 * p_0{}^i + x_1(k) * p_1{}^i + x_2(k) * p_2{}^i + \ldots + x_n(k) * p_n{}^i \\ & = [1\ \mathbf{x}^T(k)]\theta^i \\ & = \bar{\mathbf{x}}(k)\theta^i \end{aligned} \tag{1}$$

where, $\theta^i = [p_0{}^i, p_1{}^i, ..., p_n{}^i]^T \in \mathbb{R}^{(n+1)}$ is the coefficients of consequence parameter of i-th rule, $\mathbf{x}(k) = [x_1(k), x_2(k), ..., x_n(k)]^T \in \mathbb{R}^n$ is the premise variable and $y^i(k)$ is consequence variable generated by i-rule due to k-th sample.

Type-2 Gaussian membership grade has been used for premise variable and upper membership grade $\bar{A}^i_j(x_j(k))$ is defined as,

$$\begin{aligned} &\bar{A}^i_j(x_j(k)) = \exp\left(-\frac{(x_j(k)-\bar{v}^i_j(k))^2}{\bar{\sigma}^i_j(k)^2}\right) \in [0,1], \\ &for\, i = 1,2,...C \text{ and } j = 1,2,...,n \end{aligned} \tag{2}$$

where, $\bar{v}^i_j(k)$ is the mean and $\bar{\sigma}^i_j(k)$ is width of the upper membership grade. Similarly, lower membership grade of premise variable is defined as,

$$\begin{aligned} &\underline{A}_j{}^i(x_j(k)) = \exp\left(-\frac{(x_j(k)-\underline{v}_j{}^i(k))^2}{\underline{\sigma}_j{}^i(k)^2}\right) \in [0,1], \\ &for\, i = 1,2,...C \text{ and } j = 1,2,...,n \end{aligned} \tag{3}$$

where, $\underline{v}_j{}^i(k)$ is the mean and $\underline{\sigma}_j{}^i(k)$ is width of the lower membership grade.

The estimated output for k-th sample can be obtained by weighted average of all activated rules,

$$\hat{y}^i(k) = \frac{\sum_{i=1}^{C} w^i(k) * y^i(k)}{\sum_{i=1}^{C} w^i(k)}. \tag{4}$$

$$\begin{aligned} &\bar{w}^i(k) = \min\{\bar{A}^i_j(x_j(k))\}. \\ &\underline{w}^i(k) = \min\{\underline{A}_j{}^i(x_j(k))\}. \text{ for } k = 1,2,...,N \\ &w^i(k) = \tfrac{\bar{w}^i(k)+\underline{w}^i(k)}{2} \end{aligned} \tag{5}$$

where, $\bar{A}^i_j(x_j(k))$ and $\underline{A}_j{}^i(x_j(k))$ is upper and lower membership grade of j-th component of k-th premise variable respectively. The overall truth value $(w^i(k))$ of each rule is obtained by mean of upper and lower membership grade.

III. Type-2 Fuzzy Logic System

In the section, type-2 fuzzy set theory is briefly discussed.

A. Type-2 Fuzzy Set (T2FS)

A type-2 fuzzy set [25] in universal set Z is defined as $\tilde{A}$, and its membership value $\mu_{\tilde{A}}(z,u)$ is defined as,

$$\tilde{A} = \int_{z\in Z} \mu_{\tilde{A}}(z)/z = \int_{z\in Z} \left[\int_{u\in J_z} f_z(u)/u\right]/z, \quad J_z \in [0,1] \tag{6}$$

where, $f_z(u)$ is the secondary membership function and J_z is the primary membership function of z, the domain of secondary membership function.

Definition 1. *At each vertical slice z, say $z = z'$, the vertical slice of secondary membership function is $\mu_{\tilde{A}}(z', u)$ in the domain of universal set Z. It is defined as,*

$$\begin{aligned} &\mu_{\tilde{A}}(z = z', u) = \int_{u\in J_{z'}} f_{z'}(u)/u, \ J_{z'} \in [0,1] \\ &0 \le f_{z'}(u) \le 1 \end{aligned} \tag{7}$$

The type-2 fuzzy set [7] is called interval type-2 fuzzy set when the secondary membership grade value is 1 i.e. $f_{z'}(u) = 1, J_{z'} \in [0,1]$.

Definition 2. *An interval type-2 fuzzy $\tilde{A}$, is expressed by its membership function $\mu_{\tilde{A}}(z,u) = 1$, i.e.*

$$\tilde{A} = \{((z,u),1)\,|\forall z \in Z, \forall u \in J_z \in [0,1]\} \tag{8}$$

The footprint of uncertainty (FOU) of $\tilde{A}$ is expressed as union of all the primary membership function,

$$FOU(\tilde{A}) = \cup_{\forall z\in Z} J_z = \{(z,u)\,|u \in J_z \in [0,1]\} \tag{9}$$

The upper membership function (UMF) $\bar{\mu}_{\tilde{A}}(z)$, and lower membership function (LMF) $\underline{\mu}_{\tilde{A}}(z)$, value is associated with FOU [6] [26].

$$\begin{aligned} \bar{\mu}_{\tilde{A}}(z) &= \overline{FOU(\tilde{A})} \\ \underline{\mu}_{\tilde{A}}(z) &= \underline{FOU(\tilde{A})} \end{aligned} \tag{10}$$

IV. Fuzzy C-Means Clustering Algorithm

In the section, offline type-1 FCM clustering algorithm is briefly discussed for explaining proposed IT2RFCM algorithm. Let, the data set $D = \{(\mathbf{x}(1), y(1)), (\mathbf{x}(2), y(2)), ..., (\mathbf{x}(N), y(N))\}$. where, input vector, $\mathbf{x}(k) \in \mathbb{R}^n$. A set of input vector is represented as a matrix form,

$$X = [\ \{\mathbf{x}(1), y(1)\} \ \ \ldots \ \ \{\mathbf{x}(N), y(N)\}\ ]_{(n+1)\times N} \tag{11}$$

In the clustering algorithm, the main objective is to partition the data space X into C clusters. The fuzzy partition of X is typically represented by fuzzy subsets $B_i (1 \le i \le C)$. These fuzzy subsets are represented by fuzzy membership values $\mu_i(k) \in [0,1]$ that is implicit member of fuzzy partition matrix $U = [\mu_i(k)] \in \mathbb{R}^{C\times N}$. The fuzzy partitions matrix has following properties,

$$\mu_i(k) \in [0,1], \ \ i = 1,2,...,C \text{ and } k = 1,2,...,N \tag{12}$$

$$0 < \sum_{k=1}^{N} \mu_i(k) < N \tag{13}$$

Equation (13) implies that sum of the membership value of samples each cluster is less than the total number of samples, N. The FCM clustering algorithm produces C partitions. The distance between data point to point shaped cluster is expressed as, ${d_i}^2(k) = \left\|X(k) - \mathbf{v}^i\right\|_2^{\ 2} > 0$. A variety of different norms can be used to measure the distance metric such as L_0, L_2, and L_∞ norm [23]. The FCM clustering algorithm can be formulated as an optimization problem,

$$\min_{U,V} \left\{ J_1(D;U,V) = \sum_{i=1}^{C}\sum_{k=1}^{N} {\mu^m}_i(k)\left\|X(k) - \mathbf{v}^i\right\|_2^{\ 2} \right\} \tag{14}$$

$$subject\ to,\ \sum_{i=1}^{C} \mu_i(k) = 1 \tag{15}$$

where, $V = [\mathbf{v}^1, \mathbf{v}^2, ..., \mathbf{v}^C] \in \mathbb{R}^{(n+1)\times C}$ is cluster center matrix, m is the fuzziness parameter. The constrained optimization problem in (14-15) can be converted to unconstrained optimization problem by using N lagrangian multipliers. It can be represented as,

$$L(U,V;X) = J_1(D;U,V) - \sum_{k=1}^{N} \lambda(k)(\sum_{i=1}^{C} \mu_i(k) - 1) \tag{16}$$

If, ${d_i}^2(k) > 0$, the analytical solutions of (16) is obtained as fuzzy partition matrix (U) and cluster center $(\mathbf{v}^i)$.

$$\mu_i(k) = \frac{1}{\sum\limits_{q=1}^{C} \left(\frac{{d^2}_i(k)}{{d^2}_q(k)}\right)^{1/(m-1)}} \tag{17}$$

The cluster center $(\mathbf{v}^i)$ is represented as weighted average sum of the data belonging to the i-th cluster, where the weights are fuzzy membership value and are expressed as follows,

$$\mathbf{v}^i = \frac{\sum\limits_{k=1}^{N} {\mu^m}_i(k)\mathrm{X}(k)}{\sum\limits_{k=1}^{N} {\mu^m}_i(k)} \tag{18}$$

V. Interval Type-2 Recursive Fuzzy C-Means Clustering Algorithm

The recursive algorithm is incorporated into clustering algorithm so that its can predict dynamical system behavior instantly. Recursive version of type-2 FCM clustering algorithm has been obtained by varying fuzziness parameters.

A. Interval type-2 recursive center calculation

We will first introduce the cluster upper and lower center $(\bar{\mathbf{v}}^i(r)\,\mathrm{and}\,\underline{\mathbf{v}}^i(r))$ for constructing IT2RFCM algorithm according to the current observation vector i.e. the weighted mean of the data according to the current membership degrees. The cluster centroid vectors introduces notation $(\bar{\mathbf{v}}^i(r)\,\mathrm{and}\,\underline{\mathbf{v}}^i(r))$, that means centroid at current current sample r which is obtained from the current membership degrees. r stands for the sample index used to calculate cluster centers $(\bar{\mathbf{v}}^i(r)\,\mathrm{and}\,\underline{\mathbf{v}}^i(r))$. In IT2RFCM algorithm, upper and lower initial cluster centers $(\bar{\mathbf{v}}^i(r)\,\mathrm{and}\,\underline{\mathbf{v}}^i(r))$ are obtained by using observed membership function $\mu_i(r)$ value of r-th sample. The upper cluster center for $(r+1)$-th sample is obtained as,

$$\begin{aligned}
\bar{\mathbf{v}}^i(r+1) &= \frac{\sum\limits_{k=1}^{r} {\mu^{\mathrm{m}_2}}_i(k)\mathrm{X}(k) + {\mu^{\mathrm{m}_2}}_i(r+1)\mathrm{X}(r+1)}{\sum\limits_{k=1}^{r} {\mu^{\mathrm{m}_2}}_i(k) + {\mu^{\mathrm{m}_2}}_i(r+1)} \\
&= \frac{\sum\limits_{k=1}^{r} {\mu^{\mathrm{m}_2}}_i(k)\mathrm{X}(k) \Big/ \sum\limits_{k=1}^{r} {\mu^{\mathrm{m}_2}}_i(k)}{\sum\limits_{k=1}^{r+1} {\mu^{\mathrm{m}_2}}_i(k) \Big/ \sum\limits_{k=1}^{r} {\mu^{\mathrm{m}_2}}_i(k)} \\
&\quad + \frac{{\mu^{\mathrm{m}_2}}_i(r+1)\mathrm{X}(r+1)}{\sum\limits_{k=1}^{r} {\mu^{\mathrm{m}_2}}_i(k) + {\mu^{\mathrm{m}_2}}_i(r+1)} \\
&= \frac{\bar{\mathbf{v}}^i(r) \sum\limits_{k=1}^{r} {\mu^{\mathrm{m}_2}}_i(k)}{\sum\limits_{k=1}^{r+1} {\mu^{\mathrm{m}_2}}_i(k)} + \frac{{\mu^{\mathrm{m}_2}}_i(r+1)\mathrm{X}(r+1)}{\sum\limits_{k=1}^{r} {\mu^{\mathrm{m}_2}}_i(k) + {\mu^{\mathrm{m}_2}}_i(r+1)} \\
&= \bar{\mathbf{v}}^i(r) - \frac{\bar{\mathbf{v}}^i(r){\mu^{\mathrm{m}_2}}_i(r+1)\mathrm{X}(r+1)}{\sum\limits_{k=1}^{r+1} {\mu^{\mathrm{m}_2}}_i(k)} \\
&\quad + \frac{{\mu^{\mathrm{m}_2}}_i(r+1)\mathrm{X}(r+1)}{\sum\limits_{k=1}^{r} {\mu^{\mathrm{m}_2}}_i(k) + {\mu^{\mathrm{m}_2}}_i(r+1)} \\
&= \bar{\mathbf{v}}^i(r) + \frac{{\mu^{\mathrm{m}_2}}_i(r+1)(\mathrm{X}(r+1) - \bar{\mathbf{v}}^i(r))}{\sum\limits_{k=1}^{r} {\mu^{\mathrm{m}_2}}_i(k) + {\mu^{\mathrm{m}_2}}_i(r+1)} \\
&= \bar{\mathbf{v}}^i(r) + \Delta\bar{\mathbf{v}}^i(r+1)
\end{aligned} \tag{19}$$

where,

$$\Delta\bar{\mathbf{v}}^i(r+1) = \frac{{\mu^{\mathrm{m}_2}}_i(r+1)[\mathrm{X}(r+1) - \bar{\mathbf{v}}^i(r)]}{\sum\limits_{k=1}^{r} {\mu^{\mathrm{m}_2}}_i(k) + {\mu^{\mathrm{m}_2}}_i(r+1)} \tag{20}$$

where, $\mu_i(r+1)$ is the membership value of $(r+1)$-th sample. m_2 is upper fuzziness parameter. The denominator of (20) for calculating upper cluster center $(r+1)$-th sample value depends on upper membership value of $(r+1)$-th sample. The calculation of current membership value requires all past r-th sample, that is against the recursive approach [23]. The approximation for calculating the current membership value has been done by exponentially weighted constant [23]. The denominator term of (20) can be represented as,

$$\bar{S}^i(r+1) = \xi_1 \bar{S}^i(r) + {\mu^{\mathrm{m}_2}}_i(r+1) \tag{21}$$

where, $\bar{S}^i(r)$ is represented as,

$$\bar{S}^i(r) = \sum_{k=1}^{r} {\mu^{\mathrm{m}_2}}_i(k) \tag{22}$$

$0 < \xi_1 < 1$ is called forgetting factor constant. The upper cluster center can be obtained by using (20) and (21),

$$\Delta\bar{\mathbf{v}}^i(r+1) = \frac{{\mu^{\mathrm{m}_2}}_i(r+1)[\mathrm{X}(r+1) - \bar{\mathbf{v}}^i(r)]}{\bar{S}^i(r+1)} \tag{23}$$

Similarly, the lower cluster center $(\underline{\mathbf{v}}^i(r+1))$ for $(r+1)$-th sample is obtained as,

$$\underline{\mathbf{v}}^i(r+1) = \underline{\mathbf{v}}^i(r) + \Delta\underline{\mathbf{v}}^i(r+1) \tag{24}$$

where,

$$\Delta\underline{\mathbf{v}}^i(r+1) = \frac{{\mu^{\mathrm{m}_1}}_i(r+1)(\mathrm{X}(r+1) - \underline{\mathbf{v}}^i(r))}{\underline{S}^i(r+1)} \tag{25}$$

Lower fuzziness parameter is defined as m_1 and denominator term for lower cluster center is represented as,

$$\underline{S}^i(r+1) = \xi_1 \underline{S}^i(r) + \mu^{\mathrm{m}_1}{}_i(r+1) \tag{26}$$

where, $\underline{S}^i(r)$ is represented as,

$$\underline{S}^i(r) = \sum_{k=1}^{r} \mu^{\mathrm{m}_1}{}_i(k) \tag{27}$$

The membership grade $(r+1)$-th sample is calculated for both the cases,

$$\bar{\mu}_i(r+1) = \max\left\{ \begin{array}{c} \frac{1}{\sum\limits_{q=1}^{C}\left(\frac{\bar{d}_i^2(r+1)}{\bar{d}_q^2(r+1)}\right)^{1/(m_2-1)}}, \\ \frac{1}{\sum\limits_{q=1}^{C}\left(\frac{\underline{d}^2{}_i(r+1)}{\underline{d}^2{}_q(r+1)}\right)^{1/(m_1-1)}} \end{array} \right\} \tag{28}$$

$$\underline{\mu}_i(r+1) = \min\left\{ \begin{array}{c} \frac{1}{\sum\limits_{q=1}^{C}\left(\frac{\bar{d}_i^2(r+1)}{\bar{d}_q^2(r+1)}\right)^{1/(m_2-1)}}, \\ \frac{1}{\sum\limits_{q=1}^{C}\left(\frac{\underline{d}^2{}_i(r+1)}{\underline{d}^2{}_q(r+1)}\right)^{1/(m_1-1)}} \end{array} \right\} \tag{29}$$

where, the distance metric is obtained for both cases,

$$\bar{d}_i(r+1)^2 = (\mathrm{X}(r+1) - \bar{\mathbf{v}}^i(r))^T(\mathrm{X}(r+1) - \bar{\mathbf{v}}^i(r)) \tag{30}$$

$$\underline{d}_i(r+1)^2 = (\mathrm{X}(r+1) - \underline{\mathbf{v}}^i(r))^T(X(r+1) - \underline{\mathbf{v}}^i(r)) \tag{31}$$

B. Interval type-2 recursive fuzzy covariance calculation

Each cluster distribution can be described by its covariance matrix, $\bar{F}_i(r) \in \mathbb{R}^{(n+1)\times(n+1)}$, $i = 1, 2, ..., C$ and r stands for the sample index used to calculate the matrix. The upper $(\bar{F}_i(r))$ and lower $(\underline{F}_i(r))$ fuzzy covariance matrix for r-th sample is obtained as,

$$\bar{F}_i(r) = \frac{\sum\limits_{k=1}^{r} \mu^{m_2}{}_i(k)[X(k) - \bar{\mathbf{v}}^i(r)][X(k) - \bar{\mathbf{v}}^i(r)]^T}{\sum\limits_{k=1}^{r} \mu^{m_2}{}_i(k)} \tag{32}$$

$$\underline{F}_i(r) = \frac{\sum\limits_{k=1}^{r} \mu^{m_1}{}_i(k)[X(k) - \underline{\mathbf{v}}^i(r)][X(k) - \underline{\mathbf{v}}^i(r)]^T}{\sum\limits_{k=1}^{r} \mu^{m_1}{}_i(k)} \tag{33}$$

The upper fuzzy covariance matrix for $(r+1)$-th sample is expressed as follows,

$$\begin{aligned} \bar{F}_i(r+1) &= \frac{\sum\limits_{k=1}^{r} \mu^{m_2}{}_i(k)[X(k)-\bar{\mathbf{v}}^i(r)][X(k)-\bar{\mathbf{v}}^i(r)]^T}{\sum\limits_{k=1}^{r+1} \mu^{m_2}{}_i(k)} \\ &+ \frac{\mu^{m_2}{}_i(r+1)[X(r+1)-\bar{\mathbf{v}}^i(r+1)][X(r+1)-\bar{\mathbf{v}}^i(r+1)]^T}{\sum\limits_{k=1}^{r+1} \mu^{m_2}{}_i(k)} \\ \bar{F}_i(r+1) &= \xi_2 \frac{\bar{S}^i(r)}{\bar{S}^i(r+1)} \bar{F}_i(r) \\ &+ \frac{\mu^{m_2}{}_i(r+1)[X(r+1)-\bar{\mathbf{v}}^i(r+1)][X(r+1)-\bar{\mathbf{v}}^i(r+1)]^T}{\bar{S}^i(r+1)} \end{aligned} \tag{34}$$

Similarly, the lower fuzzy covariance matrix is expressed as,

$$\begin{aligned} \underline{F}_i(r+1) &= \xi_2 \frac{\underline{S}^i(r)}{\underline{S}^i(r+1)} \underline{F}_i(r) \\ &+ \frac{\mu^{m_1}{}_i(r+1)[X(r+1)-\underline{\mathbf{v}}^i(r+1)][X(r+1)-\underline{\mathbf{v}}^i(r+1)]^T}{\underline{S}^i(r+1)} \end{aligned} \tag{35}$$

where, $0 < \xi_2 < 1$ is called forgetting factor constant for upper and lower covariance matrix. The initial value of covariance matrix for both cases are defined as, $\bar{F}_i(0) = \underline{F}_i(0) = I \in \mathbb{R}^{(n+1)\times(n+1)}$.

C. Linear parameters estimation using recursive least square

The fuzzy cluster center and their distribution has been used to define the membership grade of premise variable. Once the premise variable parameters are obtained, WRLS has been used to find the coefficients of consequence parameters of each local linear model. The input membership grade are obtained by using projections of the cluster center onto the input variables. Here, $1, 2, ..., n$-th measured variables are represented as input variables and $(n+1)$ measured variables represent output variable in the data matrix X. The width of gaussian membership grade value can be obtained from cluster distribution i.e. $\bar{\sigma}^i(r+1) = \xi_3 diag(\bar{F}_i(r+1))$ and $\underline{\sigma}^i(r+1) = \xi_3 diag(\underline{F}_i(r+1))$, where, $0 < \xi_3 < 1$ is called overlapping fuzziness factor between the membership grades [23]. The coefficients of consequence parameters of local linear model are obtained by weighted recursive least square (WRLS) [9] approach as follows,

$$\theta^i(k+1) = \theta^i(k) + K(k)[y(k) - \bar{\mathbf{x}}(k)\theta^i(k)] \tag{36}$$

The kalman gain (K) update equation and error covariance matrix is obtained as,

$$K(k) = \frac{Q(k)[\bar{\mathbf{x}}(k+1)]^T}{\frac{1}{w^i(k)} + [\bar{\mathbf{x}}(k+1)]Q(k)[\bar{\mathbf{x}}(k+1)]^T} \tag{37}$$

$$Q(k+1) = [1 - K(k)\bar{\mathrm{x}}(k+1)]\,Q(k) \tag{38}$$

where, $w^i(k)$ is obtained from (5) for k-th samples. The WRLS algorithm is initialized as,

$$\begin{aligned} &Q\,(0) = 10^5 I, \\ &\theta^i(0) = 0; \quad i = 1, 2, ..., C;\ k = 1, 2, ..., r, ..., N. \end{aligned} \tag{39}$$

The forgetting factors (ξ_1, ξ_2) affects tracking the original model values [23]. The general thumb rule for setting the forgetting factor as follows,

$$\xi_t = 1 - (\frac{2}{N}), \quad for\ \ t = 1, 2. \tag{40}$$

Sometimes beyond this thumb rule, the forgetting factor value can also be changed based on identification performances [23] [24]. The high value of Fuzziness factor $(0 < \xi_3 < 1)$ causes high identification error [23]. In the IT2RFCM algorithm, the number of cluster is preassigned that creates big disadvantage for classification problems i.e. it can not be applicable for out-of range data [23]. Since, the range of data space is almost fixed in the process model or control applications such that proposed algorithm can construct the TSFM with predefined number of clusters. The architecture of IT2RFCM clustering algorithm based TSFM is depicted in Fig. 1.

Algorithm 1: IT2RFCM clustering algorithm in TS fuzzy model identification

Initializations:
Define the number of clusters (C).
Set the control parameters ($\xi_1,\ \xi_2$ and $\xi_3 = 0.25\ to\ 1$)
Fuzziness value for upper (m_1) and lower (m_2)
Define Error Covariance matrix $Q_i(0)$ for first sample using (39).
Define the initial upper ($\bar{F}_i(r)$) and lower ($\underline{F}_i(r)$) cluster covariance matrix by using $I \in \mathbb{R}^{(n+1)\times(n+1)}$
Define upper and lower cluster center $\bar{v}^i(k) = \mathbf{x}(k)$ and $\underline{v}^i(k) = \mathbf{x}(k)$ for $i = k = 1, 2, ..., C$ respectively.
The upper and lower membership grade for premise variable is assigned as, $\bar{A}^i_j(x_j(k)) = \begin{cases} 1, & \text{for } i = k \\ 0, & \text{else} \end{cases}$ $\underline{A}_j{}^i(x_j(k)) = \begin{cases} 1, & \text{for } i = k \\ 0, & \text{else} \end{cases}$
Define upper and lower membership value
$\bar{\mu}_i(k) = diag(C)$,
$\underline{\mu}_i(k) = diag(C)$ for $k = 1, 2, ..., C$
Calculate the initial upper ($\bar{S}^i$) and lower ($\underline{S}^i$) by using (22) and (27).
Repeat
Step 1
Calculate the upper ($\bar{\mu}^i(r+1)$) and lower ($\underline{\mu}_i(r+1)$) membership value by using (28) and (29).
Type reduction of membership value by using simple mean technique. i.e. $\mu_i(r+1) = \frac{\bar{\mu}_i(r+1)+\underline{\mu}_i(r+1)}{2}$
Step 2
Update upper $\bar{S}^i(r+1)$) and lower $\underline{S}^i(r+1)$) by using (21) and (26) respectively.
Step 3
Update cluster centers ($\Delta\bar{\mathbf{v}}^i(r+1)$ and $\Delta\underline{\mathbf{v}}^i(r+1)$) for both cases by using (20) and (25) respectively.
Step 4
Update cluster centers ($\bar{\mathbf{v}}^i(r+1)\ and\ \underline{\mathbf{v}}^i(r+1)$) for both cases by using (19) and (24) respectively.
Step 5
Calculate new fuzzy covariance matrix ($\bar{F}^i(r+1)\ and\ \underline{F}^i(r+1)$) for both cases by using (34) and (35) respectively.
Step 6
Upper and lower membership grade for premise variable is obtained by using (2) and (3) respectively.
Type reduction for membership grade same as in **Step 1**.
Step 7
Local linear model parameters are determined by using (36-38).
Until all samples (N) are over.

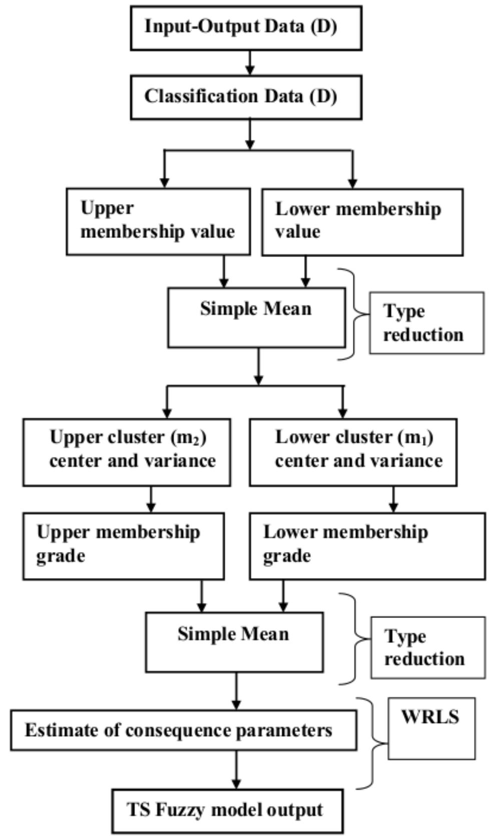

Fig. 1: Architecture of the IT2RFCM based TS Fuzzy Model

VI. Experimental Validation

The proposed method has been tested on Mackey-Glass time series data. Non Dimensional Error Index (NDEI) has been considered as performance index to check the modeling precision. It is evaluated as,

$$\text{RMSE} = \sqrt{\frac{\sum_{k=1}^{N}(y(k) - \hat{y}(k))^2}{N}} \tag{41}$$

$$\text{NDEI} = \frac{\text{RMSE}}{std(Y)} \tag{42}$$

where, $Y = [y(1), y(2), ..., y(N)]^T$, $y(k)$ is model output, $\hat{y}(k)$ is estimated output of TSFM, $std(.)$ is standard deviation function, and N is number of samples.

A. Mackey-Glass Time Series Prediction

The Mackey-Glass (M-G) chaotic time series [23] is represented by delayed differential equation,

$$\frac{dx(t)}{dt} = \frac{0.2(t-\tau)}{1 + x^{10}(t-\tau)} - 0.1x(t) \tag{43}$$

Here, the parameters were set to $x(0) = 1.2$ and $\tau = 17$

for generating the data. The data set (D) has been created as,

$$D = [\mathbf{x}(t), y(t)], t = 1, 2, ..., 5500$$
$$\mathbf{x}(t) = [x(t-18)\ x(t-12)\ x(t-6)\ \ x(t)]^T \quad (44)$$
$$y(t) = x(t+85)$$

The data set (D) is contained with 5500 samples from which 3000 ($t = 200$ to 3200) samples are used to train the model. Model validation has been done by remaining 500 ($t = 5000$ to 5500) samples, same as in [3] [23]. IT2RFCM algorithm has been tested on M-G time series data by varying two different fuzzifiers (m_1, m_2). The fuzziness factor and forgetting factor parameters were chosen based on performing the experiments in terms of NDEI value. The tuning parameters are listed in Table I. The obtained results are listed in Table II, III, IV, V under different rule conditions (3, 10, 58 and 100).

TABLE I: Tuning Parameters List

Rule	ξ_1	ξ_2	ξ_3
3	0.65	0.85	0.25
10	0.99	0.15	0.25
58	0.99	0.40	0.25
100	0.99	0.10	0.25

TABLE II: Results on IT2RFCM based TSFM with 3 rules by varying two different fuzzifiers

m_1	m_2	Rules	NDEI
1.95	2.05	3	0.5172
1.90	2.10	3	0.5070
1.85	**2.15**	3	**0.4855**
1.80	2.20	3	0.5071
1.5	2.5	3	0.5349
1.3	2.7	3	0.5332
1.1	2.9	3	0.5096
2	3	3	0.5282
3	4	3	0.5415

The best NDEI value has been obtained for fuzziness parameters set at $m_1 = 1.85$ and $m_2 = 2.15$ for rules 3, 10, 58. However, for 100 rules, the best NDEI value has been achieved for fuzziness parameters set at $m_1 = 1.1$ and $m_2 = 2.9$. The actual model and IT2RFCM based TSFM with 100 clusters performances are shown in Fig. 2. The performance of IT2RFCM based TSFM and existing in literature model [23] results are listed in Table VI. It is observed that IT2RFCM based TSFM gives reasonable accuracy with less number of rules (or nodes) as compared to other existing model in literature. For setting 3 rules, rFCM [23] based fuzzy model gives better performance compared to IT2RFCM algorithm based TSFM. Gaussian noise with variance of a 0.01 has been added to M-G time series prediction model for checking the robustness capability of IT2RFCM based TSFM. The obtained performance is depicted in Fig. 3. The obtained NDEI value was 0.4731 for noisy M-G time series data while the rules were set to 100 with fuzziness parameters ($m_1 = 1.1$ and $m_2 = 2.9$).

TABLE III: Results on IT2RFCM based TSFM with 10 rules by varying two different fuzzifiers

m_1	m_2	Rules	NDEI
1.95	2.05	10	0.4545
1.90	2.10	10	0.4559
1.85	**2.15**	10	**0.4543**
1.80	2.20	10	0.4562
1.5	2.5	10	0.4745
1.3	2.7	10	0.4637
1.1	2.9	10	0.4760
2	3	10	0.4670
3	4	10	0.4996

TABLE IV: Results on IT2RFCM based TSFM with 58 rules by varying two different fuzzifiers

m_1	m_2	Rules	NDEI
1.95	2.05	58	0.2711
1.90	2.10	58	0.2668
1.85	**2.15**	58	**0.2644**
1.80	2.20	58	0.2666
1.5	2.5	58	0.2679
1.3	2.7	58	0.2917
1.1	2.9	58	0.2848
2	3	58	0.2875
3	4	58	0.5063

TABLE V: Results on IT2RFCM based TSFM with 100 rules by varying two different fuzzifiers

m_1	m_2	Rules	NDEI
1.95	2.05	100	0.1280
1.90	2.10	100	0.1285
1.85	2.15	100	0.1277
1.80	2.20	100	0.1252
1.5	2.5	100	0.1237
1.3	2.7	100	0.1192
1.1	**2.9**	100	**0.1175**
2	3	100	0.3594
3	4	100	0.4291

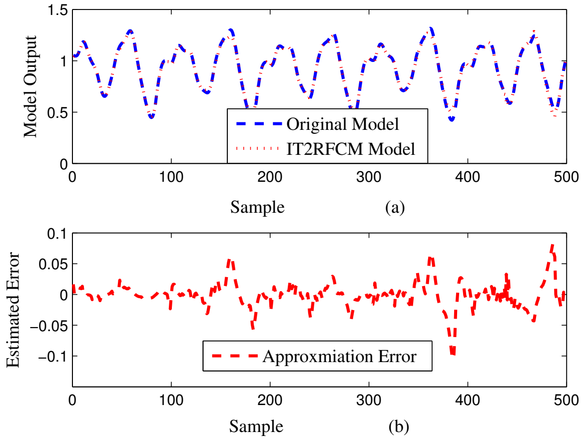

Fig. 2: (a) 85 step ahead prediction of M-G time series data ($t = 101$ to 600 samples). Red line defines IT2RFCM based TSFM. (b) Approximation error between real model and predicted model.

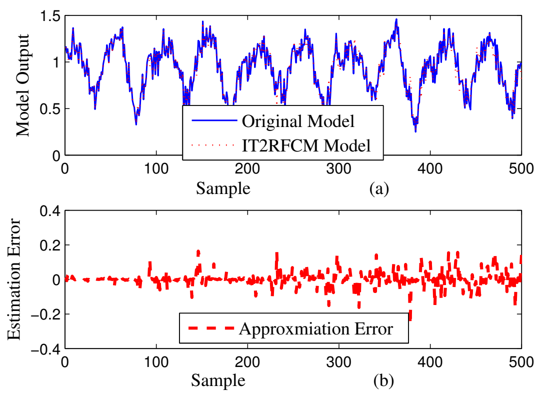

Fig. 3: (a) 85 step ahead prediction of noisy M-G time series data ($t = 101$ to 600 samples). Red line defines IT2RFCM based TSFM. (b) Approximation error between real model and predicted model.

VII. Conclusion

In this paper, a recursive fuzzy c-means clustering algorithm has been modified by using an interval type-2 fuzzy set i.e. IT2RFCM. IT2RFCM clustering algorithm has been used to classify the streaming data to identify the structure of TSFM. Premise variable parameters are also obtained by using IT2RFCM algorithm and WRLS has been used to estimate the coefficients of each local linear model of TSFM. The proposed algorithm has been validated on M-G time series prediction under different fuzzifiers and obtained results show competitive accuracy to other evolving processes. The recursive nature of IT2RFCM algorithm can be performed well in noisy environment. The algorithm parameters are easy to tune for achieving model accuracy, but it does not require complete data at a time to estimate parameters. In the future, adaptive technique can be considered in the proposed IT2RFCM based TSFM to obtain optimal number of clusters.

TABLE VI: Comparison of IT2RFCM based TSFM with existing algorithms in literature on 85 step ahead M-G time's series data.

Methods	Rules	NDEI	Change	Parameters
DENFIS [23]	58	0.276	N.A.	58×13
eTS [23]	113	0.0954	N.A	113×13
rFCM [23]	3	0.4849	N.A.	3×13
rFCM [23]	10	0.4562	N.A.	10×13
rFCM [23]	58	0.3085	N.A.	58×13
rFCM [23]	100	0.1250	N.A.	100×13
IT2RFCM	3	0.4855	-0.6 (%)	3×13
IT2RFCM	10	0.4543	0.2 (%)	10×13
IT2RFCM	58	0.2644	14.2 (%)	58×13
IT2RFCM	100	0.1175	6.4(%)	100×13

References

[1] A. C. Aras and O. Kaynak, "Interval type-2 fuzzy neural system based control with recursive fuzzy c-means clustering," *Int. J. Fuzzy Syst*, vol. 16, no. 3, pp. 317–326, 2014.

[2] T. Takagi and M. Sugeno, "Fuzzy identification of systems and its applications to modeling and control," *IEEE Trans. on Systems, Man and Cybernetics*, no. 1, pp. 116–132, 1985.

[3] N. K. Kasabov and Q. Song, "Denfis: dynamic evolving neural-fuzzy inference system and its application for time-series prediction," *IEEE Trans. on Fuzzy Systems*, vol. 10, no. 2, pp. 144–154, 2002.

[4] N. N. Karnik and J. M. Mendel, "Introduction to type-2 fuzzy logic systems," *IEEE World Congress on Fuzzy Systems*, vol. 2, pp. 915–920, 1998.

[5] J. M. Mendel, "Type-2 fuzzy sets and systems: an overview," *IEEE Computational Intelligence Magazine*, vol. 2, no. 1, pp. 20–29, 2007.

[6] J. M. Mendel and R. I. B. John, "Type-2 fuzzy sets made simple," *IEEE Trans. on Fuzzy Systems*, vol. 10, no. 2, pp. 117–127, 2002.

[7] N. N. Karnik and J. M. Mendel, "Operations on type-2 fuzzy sets," *Fuzzy Sets and Systems*, vol. 122, no. 2, pp. 327–348, 2001.

[8] D. Kukolj and E. Levi, "Identification of complex systems based on neural and Takagi-Sugeno fuzzy model," *IEEE Trans. on Systems, Man and Cybernetics-Part B*, vol. 34, no. 1, pp. 272–282, 2004.

[9] T. Dam and A. K. Deb, "TS fuzzy model identification by a novel objective function based fuzzy clustering algorithm," *IEEE Symposium on Computational Intelligence in Ensemble Learning (CIEL)*, pp. 1–7, Dec 2014.

[10] S. L. Chiu, "Fuzzy model identification based on cluster estimation." *Journal of Intelligent and Fuzzy systems*, vol. 2, no. 3, pp. 267–278, 1994.

[11] P. P. Angelov and D. P. Filev, "An approach to online identification of Takagi-Sugeno fuzzy models," *IEEE Trans. on Systems, Man and Cybernetics-Part B*, vol. 34, no. 1, pp. 484–498, 2004.

[12] P. Angelov, D. P. Filev, and N. Kasabov, *Evolving intelligent systems: methodology and applications*, John Wiley & Sons, 2010, vol. 12.

[13] L. Maciel, F. Gomide, and R. Ballini, "Enhanced evolving participatory learning fuzzy modeling: an application for asset returns volatility forecasting," *Evolving Systems*, vol. 5, no. 2, pp. 75–88, 2014.

[14] E. Lima, M. Hell, R. Ballini, and F. Gomide, "Evolving fuzzy modeling using participatory learning," *Evolving Intelligent Systems: Methodology and Applications*, pp. 67–86, 2010.

[15] E. D. Lughofer, "Flexfis: a robust incremental learning approach for evolving takagi–sugeno fuzzy models," *IEEE Trans. on Fuzzy Systems*, vol. 16, no. 6, pp. 1393–1410, 2008.

[16] J. Qiao and H. Wang, "A self-organizing fuzzy neural network and its applications to function approximation and forecast modeling," *Neurocomputing*, vol. 71, no. 4, pp. 564–569, 2008.

[17] F.-J. Lin, C.-H. Lin, and P.-H. Shen, “Self-constructing fuzzy neural network speed controller for permanent-magnet synchronous motor drive,” *IEEE Trans. on Fuzzy Systems*, vol. 9, no. 5, pp. 751–759, 2001.

[18] M. F. Azeem, M. Hanmandlu, and N. Ahmad, “Structure identification of generalized adaptive neuro-fuzzy inference systems,” *IEEE Trans. on Fuzzy Systems*, vol. 11, no. 5, pp. 666–681, 2003.

[19] K. Subramanian, A. K. Das, S. Sundaram, and S. Ramasamy, “A meta-cognitive interval type-2 fuzzy inference system and its projection based learning algorithm,” *Evolving Systems*, vol. 5, no. 4, pp. 219–230, 2014.

[20] K. Subramanian and S. Suresh, “A meta-cognitive sequential learning algorithm for neuro-fuzzy inference system,” *Applied Soft Computing*, vol. 12, no. 11, pp. 3603–3614, 2012.

[21] H.-J. Rong, N. Sundararajan, G.-B. Huang, and G.-S. Zhao, “Extended sequential adaptive fuzzy inference system for classification problems,” *Evolving Systems*, vol. 2, no. 2, pp. 71–82, 2011.

[22] J. C. Bezdek, R. Ehrlich, and W. Full, “FCM: The fuzzy c-means clustering algorithm,” *Computers & Geosciences*, vol. 10, no. 2, pp. 191–203, 1984.

[23] D. Dovžan and I. Škrjanc, “Recursive fuzzy c-means clustering for recursive fuzzy identification of time-varying processes,” *ISA Transactions*, vol. 50, no. 2, pp. 159–169, 2011.

[24] L. Maciel, F. Gomide, and R. Ballini, “Recursive possibilistic fuzzy modeling,” in *IEEE Symposium on Evolving and Autonomous Learning Systems (EALS)*, pp. 9–16, 2014.

[25] B.-I. Choi and F. C.-H. Rhee, “Interval type-2 fuzzy membership function generation methods for pattern recognition,” *Information Sciences*, vol. 179, no. 13, pp. 2102–2122, 2009.

[26] E. Rubio and O. Castillo, “Interval type-2 fuzzy clustering for membership function generation,” *IEEE Workshop on Hybrid Intelligent Models and Applications (HIMA)*, pp. 13–18, 2013.