Iranian Journal of Fuzzy Systems

Volume 22, Number 6, (2025), pp. 41-65

Original paper

University of Sistan and Baluchestan

A fuzzy clustering-based genetic TSK system with fuzzy IF part parameters and nonlinear THEN part

S. Askari[1] and S. Zarei[2]

[1] *Mechanical Engineering Department, Amirkabir University of Technology (Tehran Polytechnic), Tehran, Iran*
[2] *Department of Statistics, Faculty of Basic Science, University of Kurdistan, Sanandaj, Iran*

s_askari@aut.ac.ir, sh.zarei@uok.ac.ir

Abstract

Traditional TSK system employs standard membership functions with adjustable crisp parameters in the IF part and linear combinations of input variables in the THEN part, which leads to loss of data fuzziness. Moreover, linear THEN part of this system restricts the number of learning parameters. On the other hand, parameters of the system are tuned to meet a predefined error threshold, but the system does not necessarily display actual structure of data, which impairs its interpretability. This work proposes modifications to the TSK system to tackle these issues and improve its performance. Extended TSK System (ETSK) is presented by replacing crisp parameters of membership functions with fuzzy numbers and adopting arbitrary functions in THEN part of the system. In addition, a Clustering-based TSK system (CTSK) is proposed, which is established from fuzzy clustering algorithms and represents structures and patterns of the data. CTSK has no IF part parameters to tune, which makes the system much faster, but its THEN part is the same as ETSK. Finally, CETSK system is presented as a combination of CTSK and ETSK to study the possibility of further improvements of the systems. The IF parts of TSK, ETSK, and CETSK are tuned by genetic algorithms and their THEN parts are adjusted by the least-squares estimate method. Performance of the systems are studied using 22 datasets in terms of accuracy and runtime. On average, CTSK is 10.82% and 4.84% more accurate than TSK and CETSK; and 91.58% and 92.21% faster than these systems, respectively. CETSK is 6.27% more accurate than TSK, but 6.91% slower because fuzzy clustering and genetic algorithms are jointly employed to tune its IF part whereas only genetic algorithms is used in TSK for this purpose.

Keywords: Fuzzy systems, fuzzy clustering, genetic algorithms, gradient descent algorithm, particle swarm optimization, least-squares estimate method, TSK system, fuzzy numbers.

1 Introduction

1.1 General overview

Upon the advent of fuzzy sets, fuzzy systems were introduced to handle fuzziness of data and processes. A fuzzy system is composed of several fuzzy IF-THEN rules, which are determined through structure identification phase of the system construction. Each fuzzy rule consists of an IF part and a THEN part, which generally describes some parts of the process in a linguistic or qualitative way. However, linguistic terms should be treated quantitatively in the inference phase of the system for decision making, for which fuzzy operators such as t-norm, s-norm, and implication are used, which are termed as fuzzification, inference, and defuzzification. General structure of a fuzzy inference system is shown in Figure 1. Functionality of the modules of this system are as follows.

- **System Inputs:** The input data are fed into the system through this module.

Corresponding Author: S. Zarei
Received: May 2025; Revised: September 2025; Accepted: October 2025.
https://doi.org/10.22111/ijfs.2025.51929.9165

- **Database:** Contains membership functions and fuzzy operators.
- **Rule-Base:** Includes fuzzy IF-THEN rules.
- **Fuzzification:** Fuzzifies the input data using membership functions and fuzzy operators of the Database.
- **Inference Engine:** Calculates fuzzy outputs of the system using the Rule-Base and Database.
- **Defuzzification:** Defuzzifies the fuzzy outputs into crisp values using the Database.
- **System Outputs:** Provides outputs of the system for decision-making.

Takagi-Sugeno-Kang (TSK) system is the most popular fuzzy system since its output is crisp and does not require defuzzification. IF part of the system is composed of fuzzy sets described by membership functions and its THEN part is a linear combination of input variables. Parameters of the IF part and THEN part are tuned based on input/output data such that errors of the systems meet predefined thresholds. Adaptive Network-based Fuzzy Inference System (ANFIS) is a combination of TSK and artificial neural networks. The IF part of the ANFIS is tuned by gradient descent algorithm and its THEN part is adjusted by the least-squares estimate method. There are two types of fuzzy sets in terms of degree of fuzziness. Type-1 fuzzy sets are commonly used in the IF part of fuzzy systems, which results in Type-1 fuzzy systems. On the other hand, Type-2 fuzzy sets, which capture a higher level of fuzziness and uncertainty, but significantly increase complexity, are utilized in Type-2 fuzzy systems. Although Type-2 fuzzy systems handle uncertainty more precisely, they demand considerably higher computational costs compared to the Type-1 fuzzy systems.

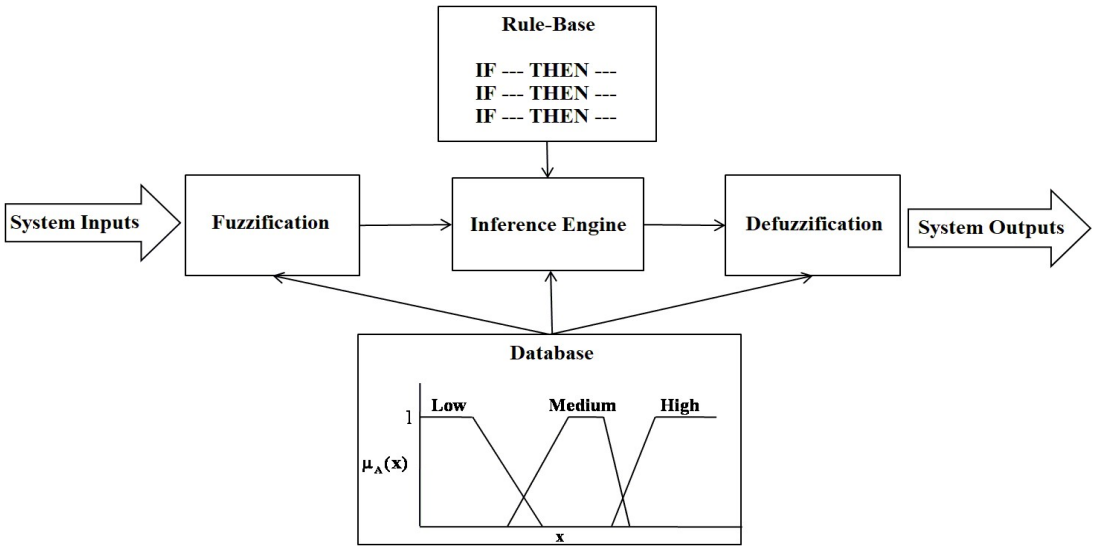


Figure 1: General structure of fuzzy inference systems.

1.2 Related works

Artificial neural networks display high learning performance, but suffer from low interpretability [23]. On the other hand, fuzzy systems handle uncertainties and ambiguities of the data using interpretable linguistic fuzzy IF-THEN rules. Neuro-fuzzy systems are devised by combining these two models to take their advantages simultaneously. Applying TSK-based neuro-fuzzy system to medical datasets demonstrates better performance and interpretability over other models [24]. The primary mathematical framework of the TSK system is proposed as a set of the membership functions of the input space in the IF part and linear functions of the input variables in the THEN part. The system is applied to water cleaning and steel production processes, which proves successful establishment, but lacks well-organized and systematic approach for structure and parameter identification [34].

One of the distinguished properties of fuzzy sets is qualitative and linguistic description of phenomena. A fuzzy system is proposed for qualitative modeling based on linguistic approximation, which is in fact an early form of TSK. Clusters of output data are projected onto the input space to identify membership functions of the IF part. Although this

system shows promising results, the identification procedure is not well-devised and straightforward [33]. A structure identification approach is proposed aiming at deciding on the variables to be included in the IF part and parameters of the membership functions associated to those variables. An unbiasedness criterion is adopted for variable selection [21] and parameters of the membership functions are tuned by successive adjustments. THEN part parameters of the system are also estimated by the unbiasedness criterion approach [32]. The proposed identification methods do not seem to be robust and are applied to only one test case namely, prediction of water flow rate. These methods have now been replaced by computationally efficient and less complex techniques.

A genetic fuzzy system is proposed to predict dynamical behavior of natural gas distribution networks. The system is mainly based on the prototype form of TSK, but the IF part parameter identification is carried out by genetic algorithms and THEN part parameters are computed by the least-squares estimate method. The system is applied to a natural gas distribution network with many end-user nodes and results of the simulations demonstrate satisfactory and efficient performance of the system [7]. A neuro-fuzzy system is proposed based on TSK by granulation of the input space, which is especially designed for regression problems. The system is applied to several benchmark datasets that shows the proposed model outperforms its counterparts [31]. It is well-known that the traditional TSK system is not suitable for processes where inputs and outputs are monotonically related. Monotonic relation-constrained TSK is proposed for classification to maintain both precision and interpretability. Since the monotonicity constraint causes convexity of the membership functions, a regularization strategy is adopted to cope with this problem. This system exhibits better accuracy and performance compared to other systems [16].

In the original TSK system, all attributes of the input data are treated equally, however, some features may be more or less important in special applications. Fuzzy biclustering neuro-fuzzy system is proposed to tackle such problems. This system maintains interpretability of TSK and demonstrates better accuracy and generalization when applied to datasets such as concrete compressive strength, airfoil shape, etc. [30]. TSK system is basically established from Type-1 fuzzy sets, which are capable of handling limited amount of uncertainties and fuzziness. Type-2 fuzzy sets are presented for more complex and ambiguous processes. Type-1 membership functions are simple and two-dimensional, however, Type-2 membership functions are complex and three-dimensional. Although Type-2 fuzzy sets and systems simulate fuzziness more accurately, they suffer from some serious drawbacks such as three-dimensional nature, low linguistic interpretability, complicated inference procedure, and high complexity and computational costs [36]. A trapezoidal Type-2 TSK is proposed where symmetric trapezoidal functions are adopted for both lower and upper membership functions. Central parameters of the membership functions are calculated by partitioning the input data. Consequent module of generalized Type-2 TSK is used in the THEN part of the system. The proposed system displays better performance than Type-2 and adaptive Type-1 fuzzy inference systems [18].

Generalized fuzzy hyperbolic models are less complex than typical fuzzy systems, but they cannot handle fuzziness as effectively as the Type-2 fuzzy systems. A hybrid system is proposed by combining the two systems to incorporate the merits of both systems into a single fuzzy system. The system is employed as a nonlinear controller for closed loop control systems. Analysis of the system performance shows higher accuracy, lower complexity, lower number of parameters, and robustness to noise and external disturbances. A hierarchical fuzzy model is presented to design TSK systems with low complexity suitable for high-dimensional data. The IF part parameters of the system are tuned by the gradient boosting approach and those of the THEN part are adjusted by fuzzy-weighted regularization technique. The system is applied to several high-dimensional datasets and its effectiveness is confirmed [22]. Combination of fuzzy systems with genetic algorithms as an efficient tool for parameter identification, results in genetic fuzzy systems [14, 17]. These systems suffer from the same problem as the original TSK system since genetic algorithms do not analyze and identify data structure and only fit the system to the output data, which impairs its interpretability. Besides genetic algorithms, a number of other techniques are adopted for parameter identification of fuzzy models and systems. A combination of fuzzy partitioning and the least-squares estimate proves effective in stock market predictions [6]. Neuro-fuzzy systems display more of the black-box nature of neural networks than the transparency of fuzzy systems. To tackle this problem, knowledge-based parameter identification is proposed for the TSK system to reveal more aspects of data structures and patterns [35]. It is shown that dynamic fuzzy Q-learning method determines both structure and parameters of fuzzy systems. The system established by this method is used as a controller for mobile robots. Experiments show the superior functionality of this system over other similar controllers [20]. A combination of series-parallel and parallel configurations is used jointly for parameter identification of TSK-based neuro-fuzzy systems. The system displays better performance compared to neural networks when applied to nonlinear dynamic problems [11]. Particle swarm optimization is a population-based optimization technique similar to the genetic algorithms. This algorithm is adopted for the IF part parameter identification of TSK designed to predict natural gas consumption at end-user nodes of gas distribution networks. Experiments show that the proposed system outperforms the TSK system tuned by gradient descent and genetic algorithms [9].

Applying the zero-order TSK system to classify imbalanced data shows better performance and interpretability since

it is capable of handling informative objects in the borderlines and overlapping regions of different classes [40]. Nonlinear mapping capability of this system provides linguistic-based interpretability of imbalanced data with fewer fuzzy rules, lower complexity, and faster running speed [12]. The TSK system provides promising solutions for classification of human-machine communication signals. Fuzziness of the system effectively tackles uncertainties of the brain dynamics and non-stationary nature of its signals [39]. Comparative studies show that the TSK fuzzy classifier outperforms its counterparts in terms of accuracy when applied to several datasets [27]. Incomplete multi-view data are classified using TSK to tackle the scarcity of labeled data by applying transductive learning approach to the system, which proves superiority of TSK over other state-of-the-art techniques [41]. Gaussian TSK system tuned by incremental learning method is employed as a fuzzy classifier and is applied to high-dimensional data. The algorithm demonstrates better generalization aspects and shorter runtime [37], where variable selection and rule extraction are the pivotal ingredients in these applications. Among a wide range of TSK applications, control systems play an important role such as fuzzy control systems for mechatronics [26], inverter controller for grid systems [25], and recurrent adaptive fuzzy controller for nonlinear systems [15]. The system is also widely used for prediction problems such as time series and stock markets [5]. Structural characteristics, advantages, and disadvantages of TSK variants proposed in the aforementioned works are summarized in Table 1. In this article, the following modifications are introduced to the formal structure of TSK

Table 1: A summary of variants of the TSK system and their characteristics.

System	**Advantages**	**Disadvantages**
Dynamic Evolving Neuro-Fuzzy Inference System (DENFIS) [24]	High interpretability	Low accuracy; crisp parameters; linear THEN part
Self-Organizing Fuzzy Neural Network (SOFNN) [24]	High accuracy	Low interpretability; crisp parameters; linear THEN part
Neuro-Fuzzy Classifier (NEFCLASS) [24]	High interpretability	Low accuracy; crisp parameters; linear THEN part
Monotonic Relation-Constrained TSK FS Classifier (MRC-TSK-FSC) [16]	Better performance for monotonic problems	High complexity; crisp parameters; linear THEN part, high runtime
Fuzzy Biclustering Neuro-Fuzzy System (EuBINES) [30]	High accuracy; better interpretation	High runtime; crisp parameters; linear THEN part
Trapezoidal Type-2 Fuzzy Inference System (TT2FIS) [18]	Better generalization	Crisp parameters; linear THEN part; low interpretability; high complexity
Hierarchical Fuzzy Regression Tree (HFRT) [22]	High accuracy; low complexity	Crisp parameters; linear THEN part; low interpretability
Hierarchical Fuzzy Rule Based Classification System (HFR-BCS) [17]	High accuracy	Crisp parameters; low interpretability
Interpretable Sensitivity Ensemble level TSK Fuzzy System (ISE-TSK-FS) [40]	High interpretability; low sensitivity	Crisp parameters; linear THEN part

and it is shown that they have significant impacts on the performance and interpretability of the system through several test cases and datasets.

1. Crisp parameters of the membership functions are replaced with fuzzy numbers to capture fuzziness of the data more precisely.

2. Linear THEN part of the system is replaced with arbitrary functions to extract more knowledge from the data.

3. Membership functions of the system are replaced with fuzzy clusters to increase interpretability of the system by capturing structures and patterns of the data.

4. A hybrid system is formulated to establish a more comprehensive system.

The paper is organized as follows. Formulation and parameter identification of the ETSK system is presented in Section 2. The CTSK system is formulated and studied in Section 3. The hybrid CETSK system is discussed in

Section 4. Applications of the proposed systems to real world data and problems are provided in Section 5 and the paper is summarized in Section 6.

2 ETSK system

For mathematical formulation of the systems, vectors are denoted by an arrow like $\vec{x}_j$, matrices are indicated by Greek capital letters like Φ, and the rest are scalars.

2.1 Description and formulation

The IF part of the TSK system is composed of membership functions with crisp parameters and its THEN part is a linear combination of input variables. The membership function of the q^{th} fuzzy set of the i^{th} rule is defined on the universe of discourse of the q^{th} input variable (feature) x_q as $\mu\left(\vec{A}_{iq}, x_q\right)$ where $\vec{A}_{iq} = \left\{(a_{iq})_1, (a_{iq})_2, \ldots, (a_{iq})_k, \ldots, (a_{iq})_D\right\}$ is the set of parameters of $\mu\left(\vec{A}_{iq}, x_q\right)$, D is the number of parameters of $\mu\left(\vec{A}_{iq}, x_q\right)$, $q \in [1, r]$ where r is the number of input variables. In the TSK system it is assumed that each member of the universal set of the q^{th} fuzzy set of the i^{th} rule is either a member of $\vec{A}_{iq}$ or not, which is equivalent to a singleton. Therefore, for a given parameter $(a_{iq})_k \in \vec{A}_{iq}$, only one member of the universe of discourse exactly coincides with $(a_{iq})_k$. This article aims to soften this assumption and provide each member of the universe of discourse with a chance of being $(a_{iq})_k$ to some degree from a fuzzy logic perspective. For this purpose, instead of crisp parameter $(a_{iq})_k$, the fuzzy set $\mu\left(\left\{\vec{B}_{iq}\right\}_k, x_q\right)$ is used, which assigns a degree of being $(a_{iq})_k$ to each member of the universe of discourse $\vec{X}_q$, and leads to the following description of the parameters.

$$\vec{A}_{iq} = \left\{\mu\left(\left\{\vec{B}_{iq}\right\}_1, x_q\right)\right), \left(\mu\left(\left\{\vec{B}_{iq}\right\}_2, x_q\right), \ldots, \left(\mu\left(\left\{\vec{B}_{iq}\right\}_k, x_q\right)\right), \ldots, \left(\mu\left(\left\{\vec{B}_{iq}\right\}_D, x_q\right)\right)\right\}.$$

Therefore, parameters of the membership functions are fuzzy numbers [28] and each parameter is represented by a fuzzy set defined on the universe of discourse of its associated input variable as shown in Figure 2, where solid lines indicate membership functions and dash lines indicate their fuzzy parameters. Note that fuzzy parameters are not necessarily triangular and any type of fuzzy numbers such as exponential, bell-shape, trapezoidal, etc., can be used. After tuning the IF part parameters $\mu\left(\left\{\vec{B}_{iq}\right\}_k, x_q\right), (a_{iq})_k$ is calculated by defuzzification of $\mu\left(\left\{\vec{B}_{iq}\right\}_k, x_q\right)$ as follows.

$$(a_{iq})_k = \frac{\int_{\vec{X}_q} \left(\mu\left(\{B_{iq}\}_k, x_q\right)\right)^\theta x_q dx_q}{\int_{\vec{X}_q} \left(\mu\left(\{B_{iq}\}_k, x_q\right)\right)^\theta dx_q}. \tag{1}$$

Where, $\vec{X}_q$ is the universe of discourse of x_q and θ is defuzzification parameter, which is set to 1 in this work. Output of the i^{th} rule of TSK for the j^{th} input data vector $\vec{x}_j = \begin{bmatrix} x_{lj} & x_{2j} & \ldots & x_{rj} \end{bmatrix}'$, is defined as $y_{ij} = p_{i1} + p_{i2}x_{lj} + p_{i3}x_{2j} + \cdots + p_{i(r+1)}x_{rj}$, where the number of the THEN part parameters of each rule is fixed and equals $r+1$, which is determined by the number of input variables r. We aim to remove these limitations and use arbitrary functions and an arbitrary number of parameters for output of each rule y_{ij}. Therefore, y_{ij} is defined as follows.

$$y_{ij} = \sum_{s=1}^{n_i} p_{is} f_{is}\left(\vec{x}_j\right). \tag{2}$$

Where, $f_{is}\left(\vec{x}_j\right) = f_{is}\left(x_{1j}, x_{2j}, \ldots, x_{rj}\right), s \in [1, n_i]$ and n_i is the number of parameters of the i^{th} rule. Note that $f_{ik}\left(\vec{x}_j\right) = 0, \forall k \mid k \in (n_i, n_{\max}]$, where $n_{\max} = \max\{n_i, i \in [1, c]\}$ and c is the number of fuzzy rules. The membership functions with fuzzy parameters and y_{ij} given in (2) generalize the traditional TSK to a more flexible one, the ETSK (Extended TSK) system. Output of ETSK is defined as follows.

$$y_j = \frac{\sum_{i=1}^{c} T_{q=1}^{r}\left(\mu\left(\vec{A}_{iq}, x_{qj}\right), \lambda\right) \sum_{s=1}^{n_i} p_{is} f_{is}\left(\vec{x}_j\right)}{\sum_{i=1}^{c} T_{q=1}^{r}\left(\mu\left(\vec{A}_{iq}, x_{qj}\right), \lambda\right)}. \tag{3}$$

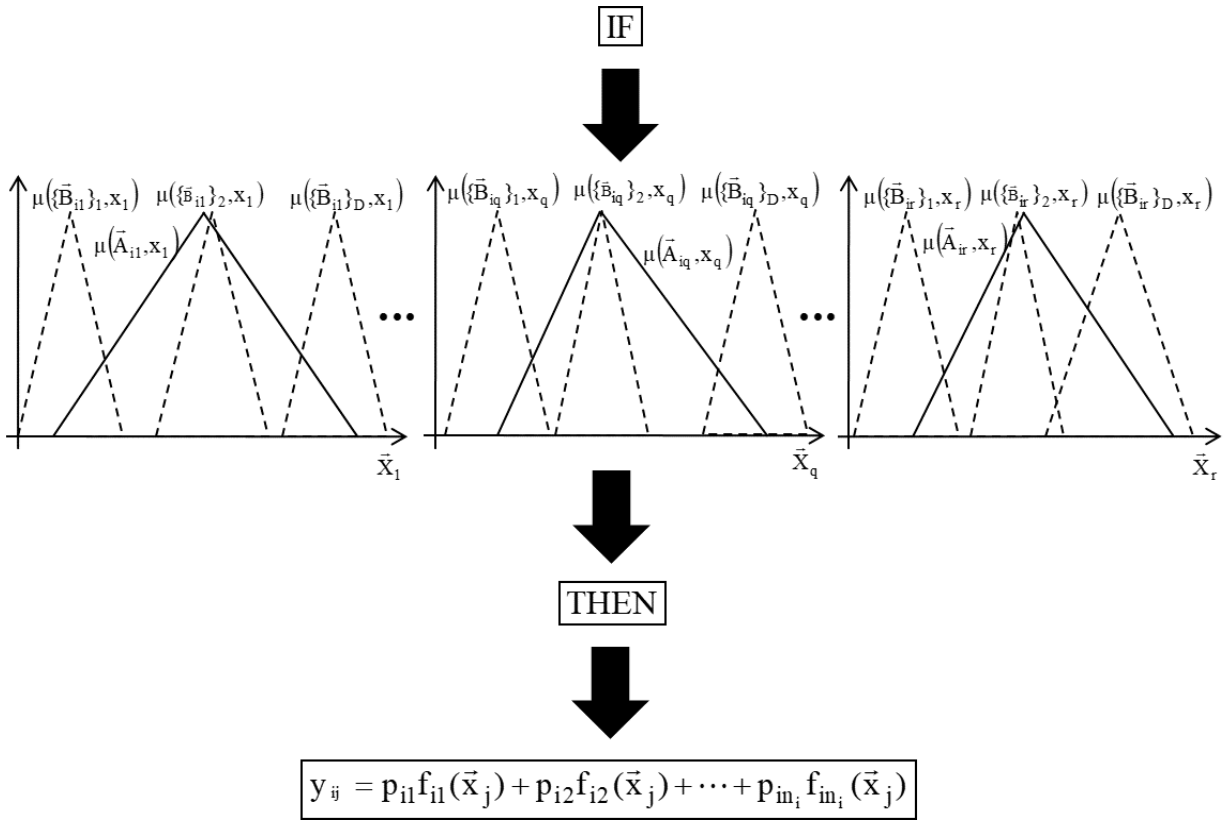


Figure 2: ETSK system with triangular fuzzy parameters and triangular membership functions.

Where, T is the parametric t-norm operator with parameter λ. Equation (3) is rewritten as follows.

$$y_j = \sum_{i=1}^{c} \tau_{ij} \sum_{s=1}^{n_i} p_{is} f_{is}\left(\vec{x}_j\right), \quad \tau_{ij} = \frac{T_{q=1}^{r}\left(\mu\left(\vec{A}_{iq}, x_{qj}\right), \lambda\right)}{\sum_{i=1}^{c} T_{q=1}^{r}\left(\mu\left(\vec{A}_{iq}, x_{qj}\right), \lambda\right)}. \tag{4}$$

Where, τ_{ij} is degree of firing of the i^{th} rule for the input vector $\vec{x}_j$. Parameters of the system including $\lambda, p_{is}, \vec{A}_{iq} i \in [1, c], s \in [1, n_i], q \in [1, r]$ should be calculated to minimize the following cost function.

$$J_1 = \sum_{j=1}^{n} \left(y_j - \sum_{i=1}^{c} \tau_{ij} \sum_{s=1}^{n_i} p_{is} f_{is}\left(\vec{x}_j\right) \right)^2, \quad \tau_{ij} = \frac{T_{q=1}^{r}\left(\mu\left(\vec{A}_{iq}, x_{qj}\right), \lambda\right)}{\sum_{i=1}^{c} T_{q=1}^{r}\left(\mu\left(\vec{A}_{iq}, x_{qj}\right), \lambda\right)}.$$

The IF part and t-norm parameters are tuned using the genetic algorithms and the THEN part parameters are adjusted by the least-squares estimate method. Details of the parameter identification procedure is discussed in the following subsections.

2.2 THEN part identification of the ETSK system

Introducing the training data one by one to the system results in the following set of equations using (4) where n is the number of training samples.

$$\begin{cases} \sum_{i=1}^{c} \tau_{i1} \sum_{s=1}^{n_i} p_{is} f_{is}\left(\vec{x}_1\right) = y_1, \\ \sum_{i=1}^{c} \tau_{i2} \sum_{s=1}^{n_i} p_{is} f_{is}\left(\vec{x}_2\right) = y_2, \\ \vdots \\ \sum_{i=1}^{c} \tau_{ij} \sum_{s=1}^{n_i} p_{is} f_{is}\left(\vec{x}_j\right) = y_j, \\ \vdots \\ \sum_{i=1}^{c} \tau_{in} \sum_{s=1}^{n_i} p_{is} f_{is}\left(\vec{x}_n\right) = y_n. \end{cases} \tag{5}$$

Equation (5) is rewritten as follows.

$$H\vec{P} = \vec{y}, \vec{P} = \left[\begin{array}{cccc} \vec{P}_1 & \vec{P}_2 & \dots & \vec{P}_c \end{array} \right]', \vec{P}_i = \left[\begin{array}{cccc} p_{i1} & p_{i2} & \dots & p_{in_i} \end{array} \right]. \tag{6}$$

Where, $\vec{y} = \left[\begin{array}{cccc} y_1 & y_2 & \dots & y_n \end{array} \right]$ is the output data and H is the following matrix.

$$H = \begin{bmatrix} \tau_{11}\vec{F}_{11} & \tau_{21}\vec{F}_{21} & \dots & \tau_{i1}\vec{F}_{i1} & \dots & \tau_{c1}\vec{F}_{c1} \\ \tau_{12}\vec{F}_{12} & \tau_{22}\vec{F}_{22} & \dots & \tau_{i2}\vec{F}_{i2} & \dots & \tau_{c2}\vec{F}_{c2} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \tau_{1j}\vec{F}_{1j} & \tau_{2j}\vec{F}_{2j} & \dots & \tau_{ij}\vec{F}_{ij} & \dots & \tau_{cj}\vec{F}_{cj} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \tau_{1n}\vec{F}_{1n} & \tau_{2n}\vec{F}_{2n} & \dots & \tau_{in}\vec{F}_{in} & \dots & \tau_{cn}\vec{F}_{cn} \end{bmatrix}, \vec{F}_{ij} = \left[f_{i1}\left(\vec{x}_j\right) f_{i2}\left(\vec{x}_j\right) \dots f_{in_i}\left(\vec{x}_j\right) \right]. \tag{7}$$

The THEN part parameters $\vec{P}$ should be determined such that the following cost function be minimized, which yields the least-squares estimate of $\vec{P}$.

$$\begin{aligned} &E = (\vec{y} - H\vec{P})'(\vec{y} - H\vec{P}) = \vec{y}'\vec{y} - 2\vec{y}'H\vec{P} + \vec{P}'H'H\vec{P}, \\ &\frac{\partial E}{\partial \vec{P}} = -2H'\vec{y} + 2H'H\vec{P} = 0 \Rightarrow \vec{P} = \left(H'H\right)^{-1} H'\vec{y}. \end{aligned} \tag{8}$$

It is clear that H is the Jacobian matrix of (5). Sometimes the $H'H$ matrix tends to be singular. Hence, the following update equation is preferred instead of (8).

$$\vec{P} = \left(H'H + \omega I\right)^{-1} H'\vec{y}. \tag{9}$$

Where, I is an identity matrix of size $[\sum_{i=1}^{c} n_i \times \sum_{i=1}^{c} n_i]$ and ω is a constant. When $\omega \to 0$, (9) leads to the Gauss-Newton optimization method, which is the same as (8). When $\omega \to \infty$, (9) results in the Gradient Descent optimization algorithm. As an alternative, when $H'H$ is close to singular, the output parameters can be calculated using pseudo-inverse of $H'H$, which leads to $\vec{P} = \left(H'H\right)^{+} H'\vec{y}$. However, the following recursive version of these equation is more popular, which is used in this work [34].

$$\vec{P}_{j+1} = \vec{P}_j + \Phi_{j+1}\vec{h}'_{j+1}\left(\vec{y}_{j+1} - \vec{h}_{j+1}\vec{P}_j\right). \tag{10}$$

$$\Phi_{j+1} = \Phi_j - \frac{\Phi_j \vec{h}'_{j+1}\vec{h}_{j+1}\Phi_j}{1 + \vec{h}_{j+1}\Phi_j\vec{h}'_{j+1}}. \tag{11}$$

Where, $j = 0, 1, \dots, n-1$ and $\vec{h}_{j+1}$ is $(j+1)^{\text{th}}$ row of the H matrix. The starting conditions of these update equations are as follows.

$$\vec{P}_0 = 0, \quad \Phi_0 = \alpha I. \tag{12}$$

Where, α is a sufficiently large number. Updating equations (10)and (11) n times from 0 to $n-1$ results in optimal values of the THEN part parameters of the system $\vec{P}$.

2.3 Tuning the IF part parameters of the ETSK system

The IF part parameters B_{iq} s, t-norm parameter λ, and α, are adjusted using real-coded genetic algorithms [14]. For this purpose, the following string is devised for the i^{th} rule.

$$\vec{S}_i = \left[\begin{array}{cccccccccccccc} \{B_{i1}\}_1 & \{B_{i1}\}_2 & \dots & \{B_{i1}\}_D & \cdots & \{B_{iq}\}_1 & \{B_{iq}\}_2 & \cdots & \{B_{iq}\}_D & \cdots & \{B_{ir}\}_1 & \{B_{ir}\}_2 & \dots & \{B_{ir}\}_D \end{array} \right].$$

These strings are then merged to form string of the entire IF part of the system $\vec{S} = \left[\begin{array}{ccccccc} \vec{S}_1 & \vec{S}_2 & \dots & \vec{S}_c & \lambda & \alpha \end{array} \right]$. It is known that α should be a large number, but there is no clear method for choosing or estimating this starting parameter. Therefore, it is included in the rule-base string to be estimated by the GA. Yager's t-norm is used for all the systems with parametric t-norm in this paper. This t-norm is defined as follows, where x and y are membership grades.

$$T(x, y) = \max\left(0, 1 - \left((1-x)^{\lambda} + (1-y)^{\lambda}\right)^{\frac{1}{\lambda}}\right), \quad \lambda \in [0, \infty). \tag{13}$$

After performing genetic operations, the best string is converted back into the rule parameters from which parameters of the membership functions are calculated using (1). In this work, the inverse of the Root Mean Square Error (RMSE) is used as the Fitness Function (FF).

$$E_{tr}(k) = \sqrt{\frac{1}{n}\sum_{j=1}^{n}\left(y_j - y_j^*\right)^2}. \tag{14}$$

$$FF(k) = \frac{1}{E_{tr}(k)}. \tag{15}$$

Where, y_j and y_j^* are the j^{th} actual and computed outputs for the k^{th} string.

There are three principal operations in the GA including Selection, Crossover, and Mutation. Proportional selection (roulette-wheel selection) is used to choose strings with higher fitness values for crossover operation. Probability of choosing k^{th} string is calculated from the following equation.

$$\Pr(k) = \frac{FF(k)}{\sum_{k=1}^{n_p} FF(k)}. \tag{16}$$

Where, n_P is the number of strings (population). Consider parent strings $\vec{S}_1 = \left[\begin{array}{ccccccc} x_1 & x_2 & x_3 & \dots & x_j & \dots & x_n \end{array}\right]$ and $\vec{S}_2 = \left[\begin{array}{ccccccc} y_1 & y_2 & y_3 & \dots & y_j & \dots & y_n \end{array}\right]$ where n is their length. Flat Crossover is adopted to generate two offspring from the two selected parents, whose alleles are calculated as follows [14].

$$\begin{cases} z_i = R_i x_i + (1 - R_i)\, y_i \\ z_{i+1} = (1 - R_i)\, x_i + R_i y_i \end{cases}, \forall i, i \in [1, n]. \tag{17}$$

Where, R_i is a random number, which is generated separately for each allele. In addition, Random Mutation is employed in which for each allele of each string z_i, a random number R_i is generated. If $R_i > P_M, z_i$ is mutated as follows [14].

$$z_i = a_i + R_i\,(b_i - a_i)\,, \forall i, i \in [1, n] \tag{18}$$

Where, P_M is the mutation probability, a_i and b_i are lower and upper bounds of a predefined interval associated with the i^{th} gene and R_i is again a random number. We choose $a_{iq} = \min\left(\vec{X}_q\right), b_{iq} = \max\left(\vec{X}_q\right)$ for each variable x_q where $\vec{X}_q$ is universe of discourse of this variable. Step-by-step guide to the implementation of the ETSK system is as follows.

Inputs : $X, \vec{y}, c, P_M, t_{\max}$

Create initial population:

$\vec{S}_i = \left[\begin{array}{cccc} \{B_{i1}\}_1 & \{B_{i1}\}_2 & \dots & \{B_{ir}\}_D \end{array}\right] \forall i \in [1, c]$

$S = \left[\begin{array}{cccccc} \vec{S}_1 & \vec{S}_2 & \dots & \vec{S}_c & \lambda & \alpha \end{array}\right]$

for $t = 1$ **to** $t_{\max}$

Calculate $T(i,j)$ $\forall i, j$ (Eq. 13)

Calculate τ_{ij} $\forall i, j$ (Eq. 4)

Calculate H (Eq. 7)

Calculate $\vec{P}$ (Eq. 10-11)

Apply Selection operator on S (Eqs. 14-16)

Apply Crossover operator on S (Eq. 17)

Apply Mutation operator on S (Eq. 18)

Calculate errors of the systems made by each chromosome

Choose the best system

Check termination criterion

If met

Terminate learning process

else

Keep on learning

end

end

Save the system

2.4 Test cases for the ETSK system

2.4.1 Natural gas consumption data

Natural gas daily consumption recorded in a CGS (City Gate Station) is predicted using the ETSK system with seven rules and three independent variables including gas consumption of the two previous days (x_1 and x_2) and temperature of the target day (x_3), which have the maximum impact on the gas consumption. Triangular membership functions with triangular fuzzy parameters and parametric t-norm are used as shown in Figure 3. There are seven rows of membership functions corresponding to the seven fuzzy rules. Each row represents IF part of one of the fuzzy rules and each subplot of the row represents a membership function related to an input variable (x_1, x_2, and x_3). The dash lines on each subplot designate the triangular fuzzy numbers $\mu\left(\left\{\vec{B}_{iq}\right\}_k, x_q\right)$, whose parameters $\left\{\vec{B}_{iq}\right\}_k = [b_1, b_2, b_3]$ are tuned by the genetic algorithms. Since each triangular membership function $\mu\left(\vec{A}_{iq}, x_q\right)$ consists of three parameters $\vec{A}_{iq} = [a_1, a_2, a_3]$, three separate fuzzy numbers are needed. Defuzzification of the first fuzzy number by (1) results in the first parameter a_1, and so on.

The optimal t-norm parameter of the system is identified as $\lambda = 16.81$. Training and testing errors of ETSK are 1.1387 and 1.6248 , respectively, and those of TSK with seven rules, triangular membership functions, parametric tnorm, and linear THEN part are 1.1855 and 1.6506, respectively. Training error of ETSK is less than TSK that demonstrates its higher precision for system identification problems. Testing error of ETSK is also less than TSK that shows better performance of this system in generalization of the trained structure of the data for forecasting. Since structure of the aforementioned ANFIS is very similar to TSK, it is usually considered as a competitive tool to evaluate performance of this system. Thus, these data are also studied using the ANFIS with eight rules and triangular membership functions. Training and testing errors of the ANFIS are 1.6612 and 1.6900, respectively. In this example, 867 data vectors are

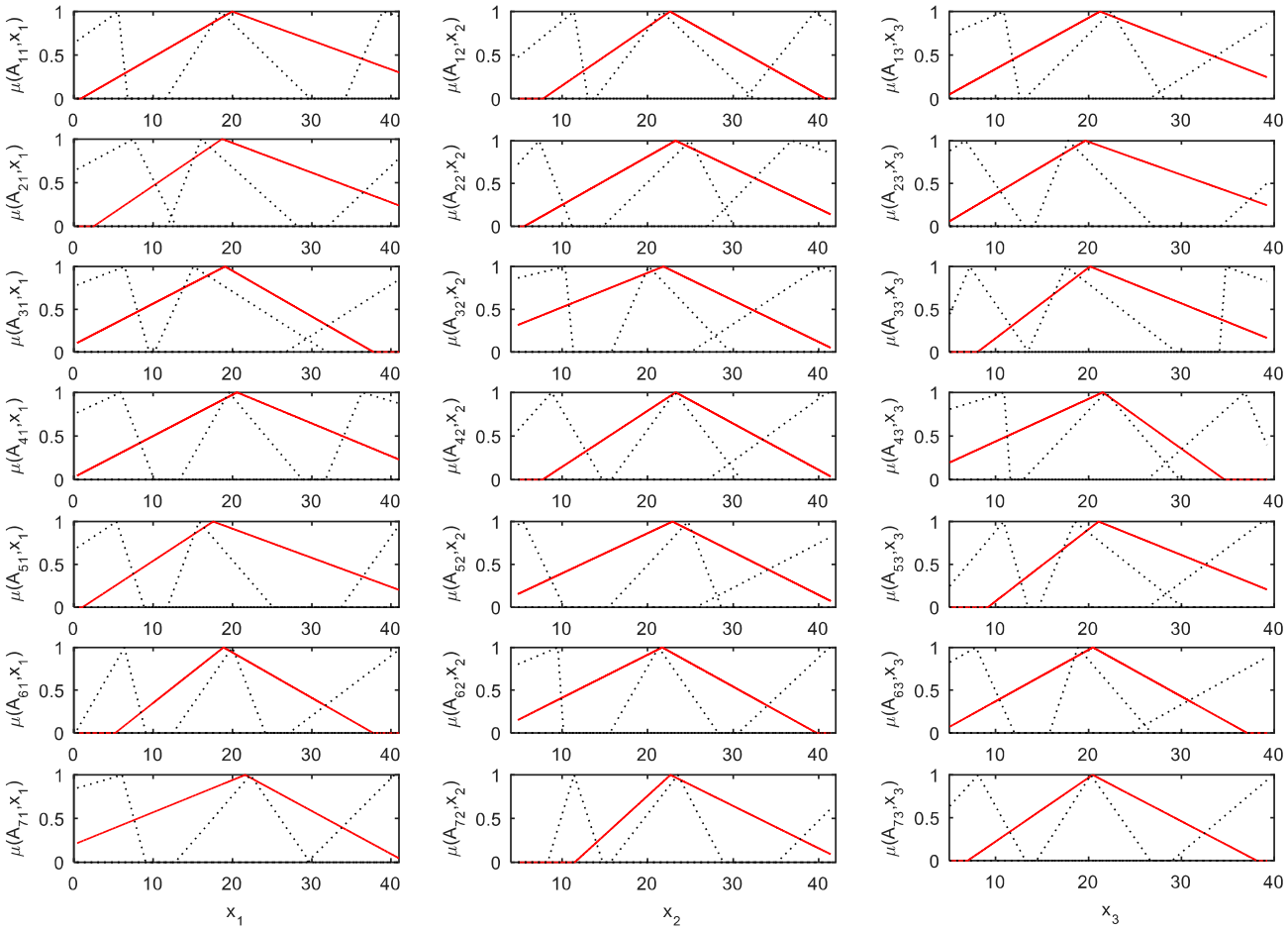


Figure 3: IF part of the ETSK system designed for prediction of natural gas daily consumption.

used for training and 371 data vectors are used for testing. Actual data are compared to the output of ETSK for both training and testing data in Figure 4.

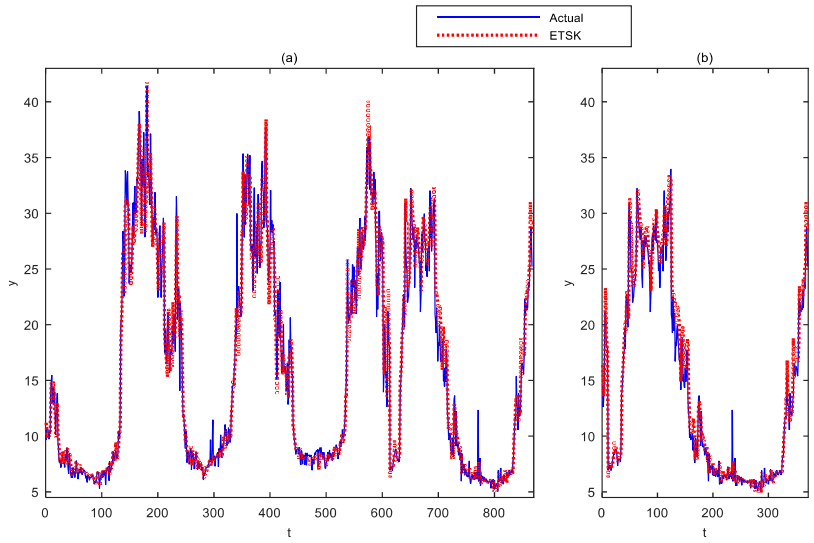


Figure 4: Comparison of the actual data with output of the ETSK system for (a) training data, (b) testing data.

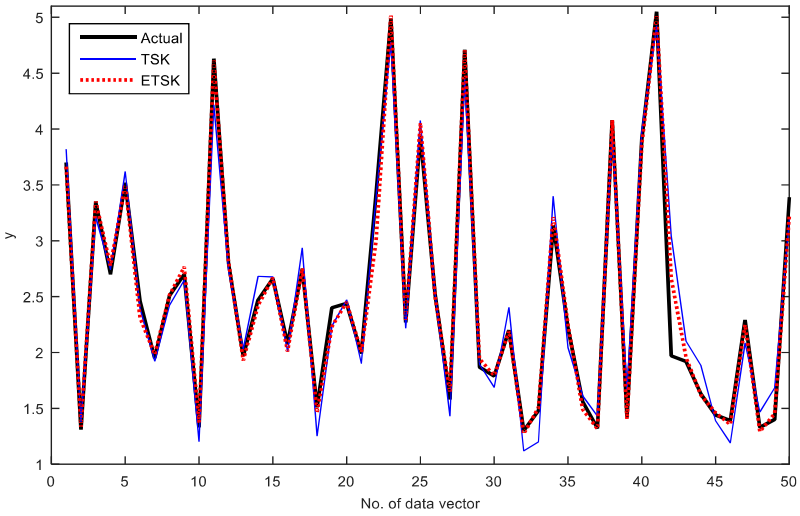


Figure 5: Outputs of the ETSK and TSK systems for identification of the Sugeno-Yasukawa nonlinear system.

2.4.2 Sugeno-Yasukawa nonlinear system

Sugeno-Yasukawa nonlinear system [33] is simulated using ETSK with fuzzy triangular parameters and membership functions and TSK with triangular membership functions. Parametric t-norm is used for both systems and second order polynomial is used for the THEN part of ETSK. These fuzzy systems are employed to identify this nonlinear system from the available data. Errors of the ETSK and TSK are 0.0889 and 0.1562, respectively. Outputs of the systems are

shown in Figure 5. IF part of ETSK is depicted in Figure 6 and its THEN part and t-norm parameters are as follows.

$$y_i = P_{i1} + P_{i2}x_1 + P_{i3}x_2 + P_{i4}x_1^2 + P_{i5}x_2^2 + P_{i6}x_1x_2; \quad \lambda = 1.5389.$$

$$P = \begin{bmatrix} -22.5809 & 12.8988 & 10.9537 & -2.6558 & -2.1558 & 0.8930 \\ -59.5526 & 11.8834 & 30.2828 & 0.2004 & -3.8044 & -2.4150 \\ 5.5513 & -11.7841 & -1.9584 & 3.0160 & 1.7654 & -1.3997 \\ 64.6094 & -37.3767 & -17.6423 & 4.1682 & 2.6215 & 1.5622 \\ -6.9256 & 11.2222 & 5.2728 & -1.9065 & -1.3812 & -0.1597 \end{bmatrix}.$$

This is a system identification problem where the system should learn behavior of the data and only its training error matters. If the traditional TSK system is used, only first three learning parameters of the output polynomial y_i (P_{i1}, P_{i2}, and P_{i3}) can be employed, which may be insufficient to learn the data behavior. However, the proposed ETSK system allows three more parameters (P_{i4}, P_{i5}, and P_{i6}) to assist the system with the learning of the data behavior, which is confirmed by lower error of the system (0.0889 vs. 0.1562) as shown in Figure 5. There is no limit for the number of parameters and terms in the THEN part of the system. However, in applications other than identification, one should be cautious with the number of parameters since high number of parameters results overtraining and overfitting, and increases testing error of the system and decreases its generalization capabilities. In such applications, number of parameters can be increased as far as testing error of the system starts increasing. Three datasets including Box-Jenkins

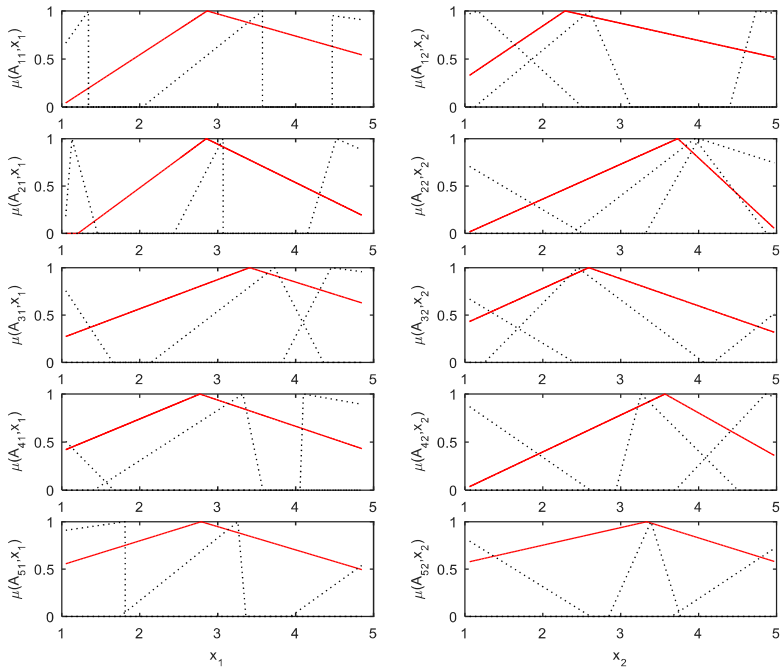


Figure 6: IF part of the ETSK system designed for Sugeno-Yasukawa system.

data, natural gas daily consumption, and TAIEX data are studied using TSK and ETSK with linear THEN part and product and parametric t-norms. They are also modeled using the ANFIS to evaluate their performance with this well-known neuro-fuzzy system. Training and testing errors of these systems are provided in Table 2. Testing error of ETSK is less than those of TSK and ANFIS for both product and parametric t-norms.

3 A clustering-based TSK system (CTSK)

Analysis of TSK, ETSK, and ANFIS shows that training algorithms force the system to tune its parameters such that its error meets a predefined value. This approach ignores patterns and structure of the data and impairs interpretability of the system, which is the basic objective of fuzzy systems. However, an interpretable fuzzy system requires the membership functions to represent the data structure for which they should be placed in dense regions of the data (clusters). In this section, TSK is modified by placing the membership functions on the clusters of the data to incorporate data structure into the system and improve its interpretability.

Table 2: Performance of fuzzy systems applied to various datasets.

	TSK system				ETSK system				ANFIS	
	Product t-norm		Yager t-norm		Product t-norm		Yager t-norm		Product t-norm	
	E_{tr}	E_{ts}	E_{tr}	E_{ts}	E_{tr}	E_{ts}	E_{tr}	E_{ts}	E_{tr}	E_{ts}
Box-Jenkins data	0.1724	0.4146	0.1847	0.3966	0.1764	0.3837	0.1965	0.3614	0.1408	0.4484
Gas Cons. data	1.1824	1.6482	1.1855	1.6706	1.1754	1.6363	1.1387	1.6248	1.6612	1.6900
TAIEX data	25.995	34.586	25.652	35.736	26.041	32.855	25.459	35.148	35.528	35.643

Structure identification is a basic issue in fuzzy systems by which shape of the membership functions and number of fuzzy rules are determined. Determining membership functions by projection of output data clusters onto the input space was one of the earliest structure identification techniques [33]. These membership functions display no standard shape and are approximated by triangular, bell-shape, exponential, or trapezoidal functions. One may also predefine membership functions and then perform parameter identification. Regardless of the type of the training algorithm, shapes and positions of the membership functions are dictated to the system and hence the system is unable to represent actual structure of the data.

Fuzzy clustering partitions the data into similar groups or clusters that disclose hidden structures and patterns of the data [1, 19]. This article presents a Clustering-based TSK system (CTSK), which is directly constructed by applying fuzzy clustering algorithm to the input data from which the IF part of the system is identified. In this approach, the IF part of the system can be directly translated into the input data structure and vice versa. For this purpose, the input data are clustered and $\vec{A}_{iq} = \left\{(a_{iq})_1, (a_{iq})_2, \dots (a_{iq})_k, \dots (a_{iq})_D\right\}$ of the ETSK system is computed from $(a_{iq})_k = v_{qi}, \forall k$ where v_{qi} is the element of the i^{th} cluster center associated with the q^{th} variable. Then, q^{th} element of the input data vector $\vec{x}_j = \begin{bmatrix} x_{1j} & x_{2j} & \dots & x_{qj} & \dots & x_{rj} \end{bmatrix}', x_{qj}$, is fuzzified using v_{qi} from which degree of firing is calculated and inference is performed as in the ETSK system. The most popular fuzzy clustering algorithm namely Fuzzy C-Means (FCM) is employed as the basis of CTSK. The FCM algorithm [3, 8, 10] is fuzzy version of the K-Means algorithm [13]. In this algorithm, cluster centers and membership grades are calculated by minimizing the following cost function [3].

$$J_2 = \sum_{j=1}^{n}\sum_{i=1}^{c} u_{ij}^m \left\|\vec{x}_j - \vec{v}_i\right\|_A^2, \quad \sum_{i=1}^{c} u_{ij} = 1.$$

Where, c is the number of clusters, $\vec{v}_i$ is the center of i^{th} cluster, u_{ij} is membership grade of the j^{th} data vector $\vec{x}_j$ in the i^{th} cluster, $\left\|\vec{x}_j - \vec{v}_i\right\|_A^2 = (\vec{x}_j - \vec{v}_i)' A (\vec{x}_j - \vec{v}_i)$ is the distance between the data vector $\vec{x}_j$ and the cluster center $\vec{v}_i$, $m \in (1, \infty)$ is degree of fuzziness. The constraint $\sum_{i=1}^{c} u_{ij} = 1$ indicates that sum of the membership grades of each data vector in all fuzzy clusters is 1. It is shown that the following cluster centers and membership grades minimize the above cost function [3].

$$\vec{v}_i = \frac{\sum_{j=1}^{n} u_{ij}^m \vec{x}_j}{\sum_{j=1}^{n} u_{ij}^m}, \quad u_{ij} = \left[\sum_{k=1}^{c}\left(\frac{\left\|\vec{x}_j - \vec{v}_i\right\|_A^2}{\left\|\vec{x}_j - \vec{v}_k\right\|_A^2}\right)^{\frac{1}{m-1}}\right]^{-1}. \tag{19}$$

Where, $A_{r\times r}$ is the covariance norm matrix defined as follows.

$$A = \left(\frac{1}{n}\sum_{j=1}^{n}\left(\vec{x}_j - \overline{\vec{v}}\right)\left(\vec{x}_j - \overline{\vec{v}}\right)\right)^{-1}, \quad \overline{\vec{v}} = \frac{1}{n}\sum_{j=1}^{n}\vec{x}_j. \tag{20}$$

3.1 Inference in the CTSK system

In order to identify IF part of CTSK, membership grade of the q^{th} component of the data vector $\vec{x}_j$ in the i^{th} cluster, u_{qij} is calculated from the following equation.

$$u_{qij} = \left[\sum_{k=1}^{c}\left(\frac{\left\|k_{qj} - v_{qi}\right\|^2}{\left\|k_{qj} - v_{qk}\right\|^2}\right)^{\frac{1}{m-1}}\right]^{-1}. \tag{21}$$

Output of CTSK is calculated using (3), which is written as follows by replacing membership function of ETSK with u_{qij}.

$$y_j = \frac{\sum_{i=1}^{c} T_{q=1}^{r}\left(u_{qj}, \lambda\right) \sum_{s=1}^{n_i} p_{is} f_{is}\left(\vec{x}_j\right)}{\sum_{i=1}^{c} T_{q=1}^{r}\left(u_{qji}, \lambda\right)}. \tag{22}$$

Equation (22) is reorganized as follows using degree of firing τ_{ij}.

$$y_j = \sum_{i=1}^{c} \tau_{ij} y_{ij}, \quad \tau_{ij} = \frac{T_{q=1}^{r}\left(u_{qij}, \lambda\right)}{\sum_{i=1}^{c} T_{q=1}^{r}\left(u_{qij}, \lambda\right)}. \tag{23}$$

Therefore, THEN part parameters of CTSK are calculated from (6)-(12) with τ_{ij} and y_{ij} given in (23) and (2), respectively. The product t-norm $\mathrm{T}_{q=1}^{r}\left(u_{qij}, \lambda\right) = \prod_{q=1}^{r} u_{qij}$ is used in CTSK for the sake of simplicity though one can use all types of parametric t-norms found in the literature. It should be noted that there are several extensions of FCM that can be used as a basis for CTSK instead of the FCM itself [2, 4, 29, 38]. Step-by-step guide to the implementation of CTSK is as follows.

Inputs : $X, \vec{y}, c, m, k_{\max}$

for $k = 1$ **to** $k_{\max}$
  Calculate $\vec{v}_i$ **and** u_{ij} $\quad \forall i, j$ (Eqs. 19-20)
  Check convergence criterion.
  If met
    Terminate iterations.
  else
    Keep on iterating
  end
end

Calculate u_{qij} $\quad \forall q, i, j$ (Eq. 21)
Calculate τ_{ij} $\quad \forall i, j$ (Eq. 23)
Calculate H (Eq. 7)
Calculate $\vec{P}$ (Eqs. 10-11)

Save the system.

3.2 Test cases for the CTSK system

3.2.1 Natural gas daily consumption data

Product t-norm operator and degree of fuzziness of $m = 2$ are used in this work. Natural gas daily consumption data are used to study performance of the CTSK system. From 1238 available data vectors, 867 are used for training and the remaining 371 samples are reserved for testing the system. The IF part of the system with five fuzzy clusters (fuzzy rules) is shown in Figure 7, where V_{qi} is the component of the i^{th} cluster (fuzzy rule) associated with the q^{th} variable of the input data as described in (19) and (21). THEN part of the system is identified as follows and its output for training and testing data are depicted in Figure 8. Training and testing errors of the system are 1.6681 and 1.6373 , respectively.

$$y_i = P_{i1} + P_{i2}x_1 + P_{i3}x_2 + P_{i4}x_3,$$

$$P = \begin{bmatrix} 6.0897 & -0.0986 & 0.1812 & 0.5839 \\ 1.6302 & -0.0227 & 0.4652 & 0.4200 \\ 14.3700 & -0.3405 & 0.1673 & 0.4571 \\ 4.8220 & -0.1090 & -0.0127 & 0.8110 \\ 9.4842 & -0.2562 & -0.0196 & 0.7768 \end{bmatrix}.$$

Training and testing errors of TSK with five rules, triangular membership functions, product t-norm operator, and linear THEN part for modeling these data are 1.1863 and 1.6469, respectively. Moreover, training and testing errors of the ANFIS with triangular membership functions and eight rules are 1.6612 and 1.6900, respectively. Testing error of CTSK with seven rules is 1.5899 for these data, which is smaller than ETSK discussed above.

3.2.2 Box-Jenkins data

Box-Jenkins data are simulated by CTSK using four clusters. From the 294 available samples, 270 are used for training and the rest are used for testing the system. The IF part of the system is shown in Figure 9. THEN part of CTSK

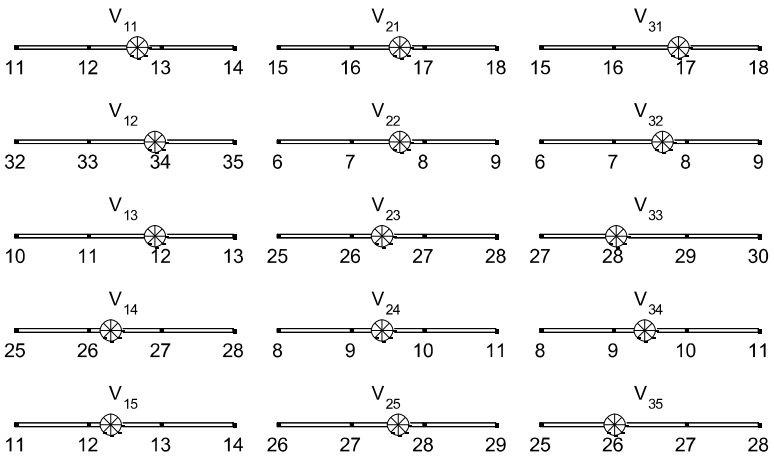


Figure 7: IF part of the CTSK system with five clusters designed for natural gas daily consumption data.

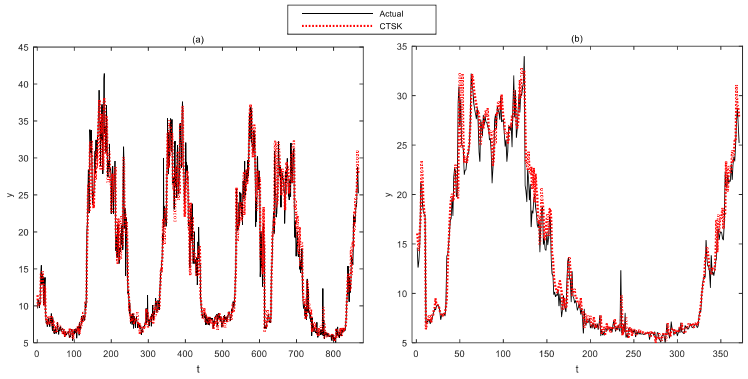


Figure 8: Output of the CTSK system for (a) training data, and (b) testing data.

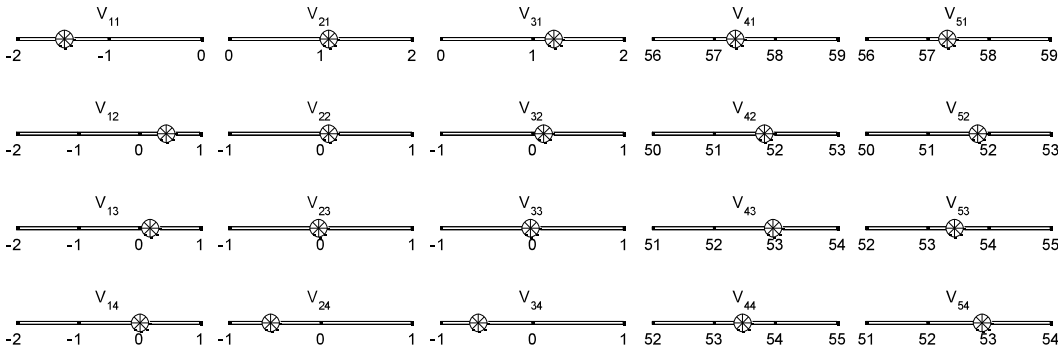


Figure 9: IF part of the CTSK system designed for Box-Jenkins data.

designed for these data is identified as follows.

$$y_i = P_{i1} + P_{i2}x_1 + P_{i3}x_2 + P_{i4}x_3 + P_{i5}x_4 + P_{i6}x_5,$$

$$P = \begin{bmatrix} 2.6902 & -0.0228 & -0.3425 & 0.2303 & -0.6108 & 1.5587 \\ 3.3274 & -0.4325 & 0.0538 & 0.0497 & -0.6336 & 1.5731 \\ 2.6939 & -0.5493 & 0.1286 & 0.0323 & -0.6355 & 1.5842 \\ 2.6141 & 0.1226 & -0.6947 & 0.2268 & -0.6755 & 1.6244 \end{bmatrix}.$$

Training and testing errors of this system are 0.29 and 0.30, respectively. Moreover, these data are studied using the ANFIS with 32 rules and triangular membership functions and TSK with product t-norm, triangular membership functions, linear output, and four fuzzy rules. Outputs of CTSK for training and testing data are compared with the actual data in Figure 10. Training and testing errors of these simulations are as follows.

$$ANFIS:\ E_{tr} = 0.14, E_{ts} = 0.44\ \therefore\ TSK:\ E_{tr} = 0.18, E_{ts} = 0.40.$$

Higher accuracy of CTSK can be attributed to the compatibility of the system with the data structure. Since IF part of the system is identified by the fuzzy clustering algorithm and its THEN part is tuned by the least-squares estimate method, which both are very fast, the system requires very low runtime.

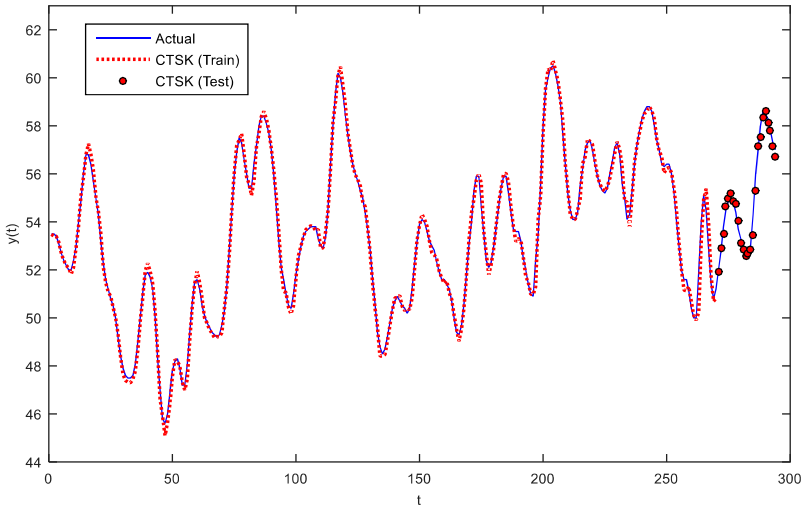


Figure 10: Output of the CTSK system designed for Box-Jenkins data.

4 Combination of the CTSK and ETSK systems

As discussed earlier, CTSK is constructed directly from fuzzy clustering and its fuzzy clusters can be transformed into membership functions. However, membership functions of this rule-base have no standard form and may be non-convex, which are problematic in some applications such as control systems. They usually lead to slow convergence and impair stability of the system. This article proposes CETSK system as a combination of CTSK and ETSK to overcome this problem. FCM is employed for structure identification and central parameter of each membership function is set to the corresponding cluster center v_{qi}, but other parameters are computed by training algorithms. This type of rule-base still represents actual structure of the data to some degrees and provides more learning from the input/output data by tuning other parameters of the membership functions. In order to compare CTSK, CETSK, and TSK, three datasets are studied and results are provided in Table 3. Two types of t-norms including product t-norm and Yager's parametric t-norm are employed. In addition, triangular, bellshape, and exponential membership functions are used in CETSK and TSK with linear THEN parts. Training error of CTSK is higher than CETSK and TSK, but simulates the testing data more accurately due to lower testing error. Therefore, the CTSK system generalizes structure and behavior of the training data to the testing data more precisely, which is essential for forecasting and classification. Similarly, training error of CETSK is higher than TSK, but provides lower testing error. Runtime of CTSK, CETSK, and TSK for two datasets are provided in Table 4. Parametric and non-parametric t-norms and three different membership functions are used. Runtime of CTSK is considerably less than TSK and CETSK because no tuning process is required for the IF part of this system. Runtime of CETSK is higher than TSK because of excessive computation for the clustering of the input data. The main difference between CTSK and CETSK is their IF part structures. IF part of the former is constructed from fuzzy clusters. However, IF part of the latter is established from membership functions whose central parameters are similarly computed by fuzzy clustering, but the rest are identified by optimization algorithms such as GA, PSO, etc., which takes considerably longer time and makes the system much slower than CTSK. For instance, for a dataset with five attributes, CTSK system with four fuzzy rules has twenty IF part parameters each for one fuzzy set as shown in Figure 9. However, in CETSK system with triangular membership functions and the same number of fuzzy

Table 3: Errors of the CTSK, CETSK, and TSK systems for various datasets and membership functions.

System	Membership Function	t-norm	TAIEX data		Gas data		Box-Jenkins data	
			Six rules		Seven rules		Six rules	
			E_{tr}	E_{ts}	E_{tr}	E_{ts}	E_{tr}	E_{ts}
CTSK	-	Product	37.0745	29.7777	1.6746	1.5899	0.3239	0.2936
CETSK	Triangular	Product	26.0808	34.2545	1.1883	1.6488	0.1753	0.3975
TSK	Triangular	Product	26.2977	34.4799	1.1900	1.6674	0.1709	0.4079
CETSK	Triangular	Yager	25.7951	33.9531	1.1626	1.6337	0.1561	0.4240
TSK	Triangular	Yager	25.8110	35.0666	1.1855	1.6506	0.1568	0.4348
CETSK	Bell-shape	Product	26.2618	34.1905	1.1823	1.6330	0.1986	0.2972
TSK	Bell-shape	Product	26.0721	34.9372	1.1736	1.6363	0.1752	0.3747
CETSK	Bell-shape	Yager	31.0853	34.9232	1.1599	1.6378	0.1521	0.4148
TSK	Bell-shape	Yager	26.9758	35.5158	1.1530	1.6358	0.1537	0.4292
CETSK	Exponential	Product	26.2849	32.3951	1.1707	1.6294	0.1931	0.3039
TSK	Exponential	Product	25.8068	35.8649	1.1713	1.6368	0.1743	0.4006
CETSK	Exponential	Yager	27.0693	34.7791	1.1574	1.6353	0.1527	0.4276
TSK	Exponential	Yager	29.8278	35.5284	1.1541	1.6386	0.1564	0.4210

Table 4: Runtime (seconds) of the CTSK, CETSK, and TSK systems for different datasets and membership functions.

System	Membership Function	t-norm	Data	
			TAIEX data	Box-Jenkins data
			Six rules	Six rules
CTSK	-	Product	3.2165	3.1517
CETSK	Triangular	Product	19.4444	30.8247
TSK	Triangular	Product	20.0625	33.0188
CETSK	Triangular	Yager	24.4895	33.3350
TSK	Triangular	Yager	22.3330	36.8565
CETSK	Bell-shape	Product	22.8452	35.1203
TSK	Bell-shape	Product	19.8049	32.9647
CETSK	Bell-shape	Yager	25.1359	38.5711
TSK	Bell-shape	Yager	22.5680	38.7263
CETSK	Exponential	Product	23.0521	34.5999
TSK	Exponential	Product	18.3782	33.1250
CETSK	Exponential	Yager	26.1551	39.7043
TSK	Exponential	Yager	21.1147	34.8489

rules, central parameter of each membership function is the same as the only parameter of its corresponding fuzzy set in CTSK, which is calculated by fuzzy clustering. Each triangular membership function requires two additional side parameters, which are computed via optimization algorithms. Considering the five membership functions in each rule, CETSK employs sixty IF part parameters, which requires significantly longer runtime to tune them (Table 4). It should be noted that THEN parts of both systems are the same and different runtimes originates only from their IF parts. Moreover, larger number of parameters may cause overtraining and overfitting, which impairs generalization ability of the system and leads to higher testing error as observed in (Table 3).

5 Real world applications of the proposed systems

Several applications of the proposed systems to real world data including natural gas metering systems, meteorological data, stock market data, and classification problems are discussed in this section.

5.1 Natural gas metering systems

Natural gas is widely used as a clean energy resource for heating in many industries. Mass flow rate cannot be measured due to its gaseous state, so volumetric flow rate is measured instead, posing challenges because gas temperature and pressure must be measured concurrently since volumetric flow depends on operational pressure and temperature. A typical metering system (Figure 11) measures supply line pressure and temperature with sensors. The gas meter generates digital electronic pulses transmitted via Modbus protocol through RS234 or RS485 ports to a flow computer that calculates gas flow rate proportional to pulse rate. However, raw data cannot be used for billing because mass flow depends on line pressure, temperature, and gas compressibility, which change over time. Therefore, gas flow rate at line conditions is converted to equivalent flow at standard conditions (101.325 kPa and 15.56°C) for billing by the flow computer using the following equation where Vl, Pr, Tr, and $Z(Pr, Tr, \vec{G})$ indicate flow rate, pressure, temperature, and gas compressibility at line conditions, respectively. Note that compressibility is a function of pressure, temperature, and chemical compositions of the gas ($\vec{G}$), which is calculated from appropriate state equations by the flow computer. Moreover, $Vl_s, \mathrm{Pr}_s, Tr_s$, and $Z\left(\mathrm{Pr}_s, Tr_s, \vec{G}\right)$ are flow rate, pressure, temperature, and compressibility at the standard conditions, respectively.

$$Vl_s = Vl \frac{\mathrm{Pr}}{\mathrm{Pr}_s} \frac{\mathrm{Tr}_s}{\mathrm{Tr}} \frac{Z\left(\mathrm{Pr}_s, \mathrm{Tr}_s, \vec{G}\right)}{Z\left(\mathrm{Pr}, \mathrm{Tr}, \vec{G}\right)}. \tag{24}$$

One of the major problems with the gas metering systems is failure of the pulse generator, which causes loss of data (Vl in (24)) and legal issues in custody transfer applications. In this section, lost gas flow is estimated based on available historical data stored in the flow computer of a real metering station. Gas flows continuously through the meter. However, due to memory and power supply limitations, the flow computer is set to interrogate line conditions including pressure, temperature, and pulse rate every thirty seconds and compute standard flow rate. Finally, these data are summed up and stored every hour in the memory alongside average hourly pressure and temperature. A total of 2144 hourly samples are used to test performance of the aforementioned fuzzy systems where 1500 data vectors are used for training and the rest are reserved for testing the systems. CTSK with three fuzzy rules is employed to estimate 644 test data points, which are assumed to be lost due to pulse generator failure. Actual data and CTSK outputs for training and testing data are compared in Figure 12 that indicates fairly good estimation of the lost data.

Performance of different systems in terms of training error E_{tr}, testing error E_{ts}, and runtime t on the lost data estimation are compared in Table 5. Three fuzzy rules and product t-norm are employed for all of the systems and impacts of different membership functions on their performance are also studied that shows exponential membership functions provide the systems with better accuracy and speed followed by the bell-shape ones.

CTSK is the most accurate and fastest one among the systems employed in this case study. In addition, the hybrid CETSK system is more accurate than the traditional TSK system, however it is slower due to additional clustering module employed to identify its IF part fuzzy clusters. In order to statistically validate the performance differences observed in Table 5., a Friedman test with Nemenyi post-hoc analysis is conducted ($\alpha = 0.05$). The test confirmed that CTSK superiority in testing error (325.8177) over all other systems (range: 326.29-461.77) is significant ($p < 0.01$). Moreover, MAE analysis with 95% confidence intervals for CTSK (MAE = 18.7±2.3), demonstrates consistent accuracy beyond MAE metrics.

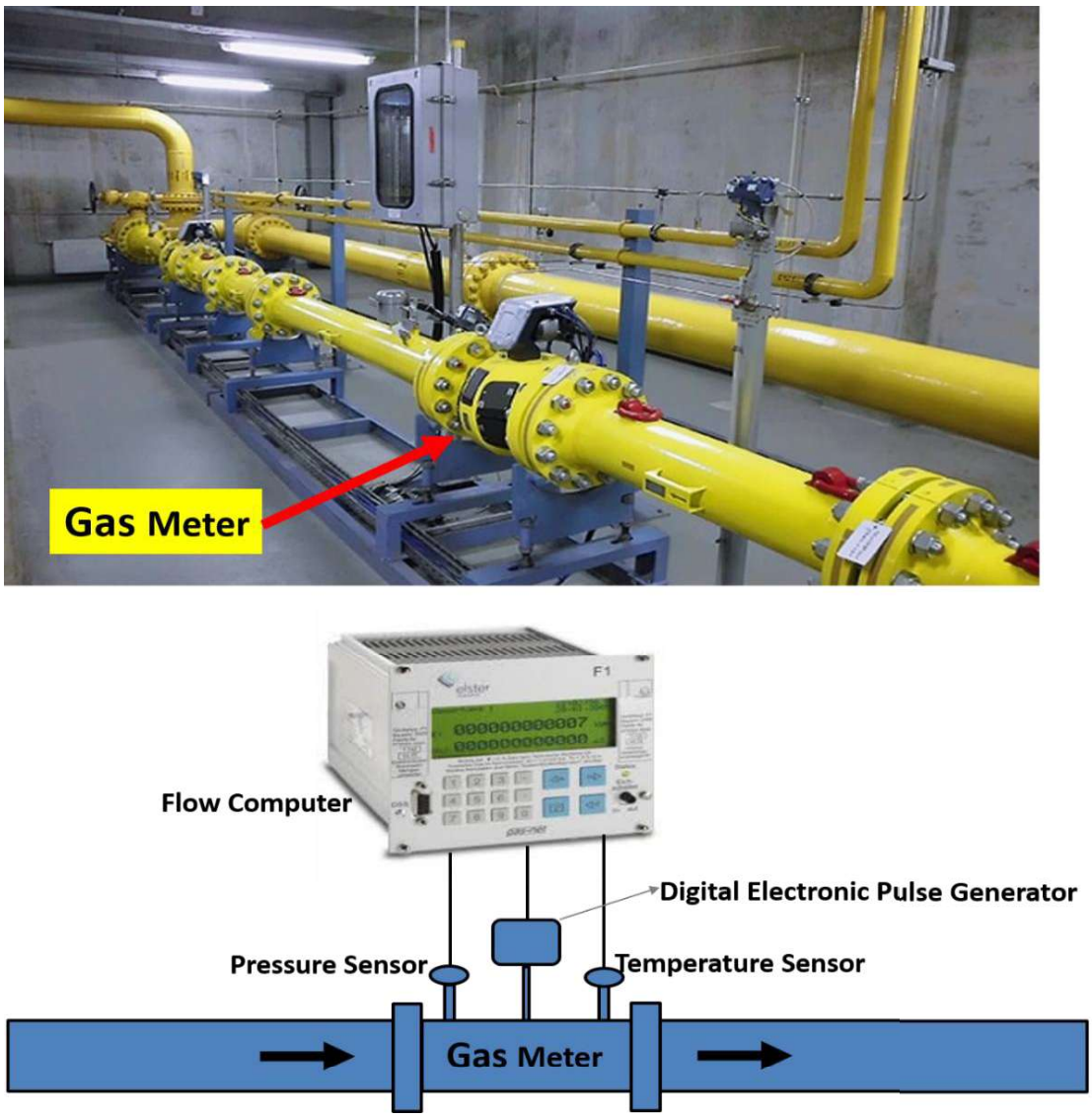


Figure 11: Natural gas metering system for custody transfer purposes.

5.2 Forecasting hourly temperature

Hourly temperature of the Zahedan city is recorded from 2017 to 2019 by the local meteorological organization. There are 10252 available samples that 7200 are used as training data and the rest as testing data. In fuzzy logic perspective, there are three temperatures conceivable, low, moderate, and high. Therefore, three clusters (fuzzy rules) are devised to forecast the 3052 hourly temperatures ahead of the training data. Statistics of the predictions performed by the above-mentioned fuzzy systems are provided in Table 6.

CTSK again outperforms the other systems in terms of both accuracy and runtime. Output of the system for part of the test data is illustrated in Figure 13 that demonstrates good agreement between the actual and predicted data. Statistical significance of CTSK performance with testing error of 3.3725 is verified via Wilcoxon signed-rank tests against all benchmark systems ($p < 0.05$ for all pairwise comparisons). The RMSE values are also calculated, which align with the MAE trends. Confidence intervals of 95% for CTSK predictions are ± 0.35°C, which reinforce its reliability.

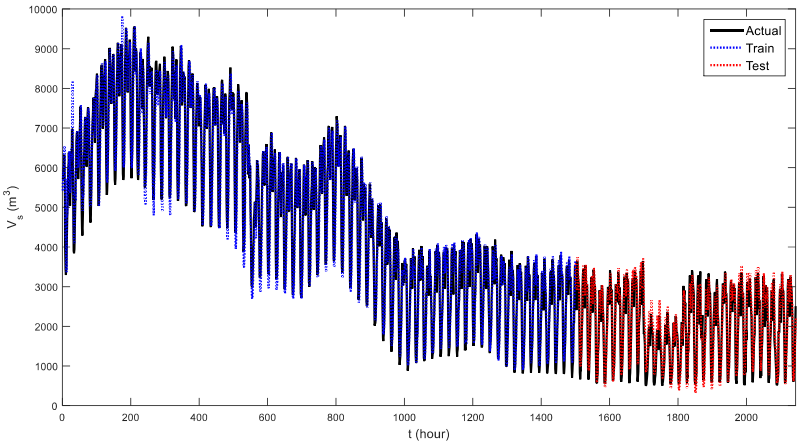


Figure 12: Output of CTSK for training and testing data of natural gas flow rate compared to the actual data.

Table 5: Performance of different systems on lost data estimation of gas metering systems.

System	Membership Function	E_{tr}	E_{ts}	Runtime (sec)
CETSK	Triangular	421.2998	353.0218	182.2588
TSK	Triangular	414.7894	461.7717	162.1489
CETSK	Bell-shape	418.1505	333.9622	153.7828
TSK	Bell-shape	364.8188	377.5074	136.1333
CETSK	Exponential	384.5405	326.2909	142.1986
TSK	Exponential	434.5458	353.6288	131.4795
CTSK	-	420.0414	325.8177	8.8882

5.3 Stock market prediction

TAIEX data from 1999 to 2014 including sixteen datasets are studied using TSK tuned by Genetic Algorithms (GATSK), TSK tuned by Particle Swarm Optimization (PSOTSK), TSK tuned by Gradient Descent algorithm (GDTSK), and the proposed CETSK and CTSK. Product t-norm is employed for all systems and triangular membership functions are used for both TSK and CETSK. Training and testing errors and runtime of the systems are reported in Table 7. The number of fuzzy rules or clusters is chosen such that the testing error becomes minimum. CTSK is the most accurate one followed by CETSK. Moreover, it is also the fastest system because there are no parameters in its IF part to be tuned. Since CETSK employs both FCM and GA algorithms to tune its IF part parameters, it is slower than the TSK system, which uses only the GA algorithm for identification of its IF part. CTSK system is benchmarked against SVM (RBF kernel) and ANN (2-layer MLP), which shows that the system reduces testing errors by 12–18% versus ML benchmarks across all 16 datasets (Friedman test p = 0.003). Moreover, 95% confidence intervals confirm robustness of the system (e.g., for TAIEX 2014, CTSK: 30.95±1.8, SVM: 35.12±2.1).

Table 6: Performance of different fuzzy systems on the prediction of hourly temperature.

System	Membership Function	E_{tr}	E_{ts}	Runtime (sec)
CETSK	Triangular	3.5713	3.3727	508.5128
TSK	Triangular	3.4978	3.4111	470.3696
CETSK	Bell-shape	3.5974	3.3793	435.9672
TSK	Bell-shape	3.5411	3.3995	395.8556
CETSK	Exponential	3.6096	3.3751	470.0741
TSK	Exponential	3.6317	3.4360	401.8267
CTSK	-	3.5613	3.3725	36.0074

Table 7: Statistics of forecasting TAIEX data from 1999 to 2014 by TSK, CETSK, and CTSK systems.

Data	System	No. of rules/clusters	Membership function	t-norm	Training error	Testing error	Runtime (sec)
	GATSK	2	Triangular	Product	52.4886	42.6819	11.2319
	PSOTSK	2	Triangular	Product	52.1994	41.4739	9.5156
TAIEX 1999	GDTSK	2	Triangular	Product	57.8649	40.1629	7.3678
	CETSK	2	Triangular	Product	57.5161	41.9555	11.5788
	CTSK	2	–	Product	58.6463	41.3793	0.3859
	GATSK	2	Triangular	Product	79.7550	65.9595	13.0613
	PSOTSK	2	Triangular	Product	79.4872	64.8974	10.9866
TAIEX 2000	GDTSK	2	Triangular	Product	83.4426	61.6385	8.6147
	CETSK	2	Triangular	Product	83.7338	60.4629	13.4552
	CTSK	2	–	Product	83.7513	61.7984	0.7045
	GATSK	3	Triangular	Product	48.1608	66.9105	17.9071
	PSOTSK	3	Triangular	Product	42.7482	58.2227	14.9715
TAIEX 2001	GDTSK	3	Triangular	Product	44.6297	56.8933	15.5965
	CETSK	3	Triangular	Product	44.3828	57.4626	18.6598
	CTSK	3	–	Product	44.4784	56.2567	2.1828
	GATSK	2	Triangular	Product	43.2023	38.0224	13.2125
	PSOTSK	2	Triangular	Product	43.1298	37.8779	11.2934
TAIEX 2002	GDTSK	2	Triangular	Product	43.7641	36.6060	8.5295
	CETSK	2	Triangular	Product	43.5966	36.9275	13.9623
	CTSK	2	–	Product	43.7833	36.8348	0.3754
	GATSK	2	Triangular	Product	32.1419	26.2111	13.2950
	PSOTSK	2	Triangular	Product	30.2870	25.0883	11.2025
TAIEX 2003	GDTSK	2	Triangular	Product	30.9435	25.7730	8.8623
	CETSK	2	Triangular	Product	30.9035	24.5147	13.7812
	CTSK	2	–	Product	30.8952	24.5656	0.3382
	GATSK	5	Triangular	Product	47.1443	31.8449	21.0120
	PSOTSK	5	Triangular	Product	34.6600	30.3801	26.7313
TAIEX 2004	GDTSK	5	Triangular	Product	38.4409	29.3764	36.1027
	CETSK	5	Triangular	Product	37.8510	30.6006	23.0162
	CTSK	5	–	Product	36.6781	29.8940	3.0950
	GATSK	2	Triangular	Product	21.9689	25.4299	13.2736
	PSOTSK	2	Triangular	Product	20.8280	26.0578	11.0505
TAIEX 2005	GDTSK	2	Triangular	Product	21.7122	25.7213	10.8292
	CETSK	2	Triangular	Product	22.1207	25.2975	13.8092
	CTSK	2	–	Product	22.2619	25.0528	0.3819
	GATSK	4	Triangular	Product	32.1408	54.1577	22.9408
	PSOTSK	4	Triangular	Product	30.0691	24.6076	21.0271
TAIEX 2006	GDTSK	4	Triangular	Product	22.2286	26.0644	8.8289
	CETSK	4	Triangular	Product	31.2577	44.9788	24.5729
	CTSK	4	–	Product	31.1232	23.9283	1.4371
	GATSK	3	Triangular	Product	48.7611	82.3226	17.9330
	PSOTSK	3	Triangular	Product	33.3744	67.6961	14.9341
TAIEX 2007	GDTSK	3	Triangular	Product	36.2549	66.0154	15.0980
	CETSK	3	Triangular	Product	48.2784	69.3356	18.6710
	CTSK	3	–	Product	34.8153	65.6973	0.8158
	GATSK	4	Triangular	Product	73.5533	47.0935	23.5152
	PSOTSK	4	Triangular	Product	53.5951	54.7054	19.1069
TAIEX 2008	GDTSK	4	Triangular	Product	56.1861	63.8176	24.7112
	CETSK	4	Triangular	Product	57.5036	47.0458	26.2566
	CTSK	4	–	Product	56.7466	46.2195	1.2601
	GATSK	2	Triangular	Product	40.2582	32.9366	13.1406
	PSOTSK	2	Triangular	Product	38.6334	32.7775	11.8098
TAIEX 2009	GDTSK	2	Triangular	Product	41.0024	33.4366	8.6760
	CETSK	2	Triangular	Product	40.5388	32.4830	14.4672
	CTSK	2	–	Product	40.7317	32.0383	0.3718
	GATSK	3	Triangular	Product	33.0366	28.5058	16.0722
	PSOTSK	3	Triangular	Product	30.3511	28.4622	15.3538
TAIEX 2010	GDTSK	3	Triangular	Product	33.0244	19.9979	16.0952
	CETSK	3	Triangular	Product	32.9156	22.7786	18.6485
	CTSK	3	–	Product	32.6734	19.4845	1.2680
	GATSK	2	Triangular	Product	40.4843	54.2169	12.4700
	PSOTSK	2	Triangular	Product	38.0892	47.0847	11.6664
TAIEX 2011	GDTSK	2	Triangular	Product	38.9384	39.2210	8.6118
	CETSK	2	Triangular	Product	38.7815	40.3189	13.4598
	CTSK	2	–	Product	38.7877	39.0463	0.3042
	GATSK	4	Triangular	Product	61.7872	54.7856	23.1485
	PSOTSK	4	Triangular	Product	29.1634	30.4982	18.9722
TAIEX 2012	GDTSK	4	Triangular	Product	29.8346	30.3862	24.8113
	CETSK	4	Triangular	Product	30.2462	32.3412	26.1850
	CTSK	4	–	Product	30.1742	29.0207	1.1131
	GATSK	2	Triangular	Product	25.1924	20.7908	12.9519
	PSOTSK	2	Triangular	Product	24.8148	20.7466	10.8175
TAIEX 2013	GDTSK	2	Triangular	Product	25.3307	19.9617	8.4532
	CETSK	2	Triangular	Product	24.8724	20.7825	13.4572
	CTSK	2	–	Product	25.3134	20.0716	0.3919
	GATSK	7	Triangular	Product	28.8569	34.1694	40.7812
	PSOTSK	7	Triangular	Product	23.3298	34.1249	35.1608
TAIEX 2014	GDTSK	7	Triangular	Product	24.4652	33.3333	63.7111
	CETSK	7	Triangular	Product	24.8178	32.0504	46.5101
	CTSK	7	–	Product	24.3961	30.9524	3.2428

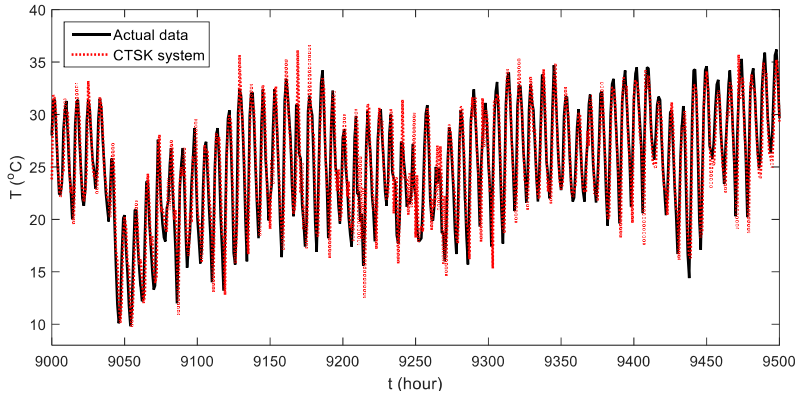


Figure 13: Predicted hourly temperature by CTSK in the time domain [9000,9500] hours.

Table 8: Training and testing errors and runtime of the systems employed for classification of the IRIS data.

System	E_{tr}	E_{ts}	Runtime (sec)
TSK	0.1847	0.3075	14.0832
CETSK	0.1666	0.2660	12.6149
CTSK	0.1720	0.2580	0.6752

5.3.1 Classification problems

IRIS data are a well-known benchmark dataset for classification with 150 samples, three classes, and four features. The TSK, CETSK, and CTSK systems are employed to classify these data where 130 samples are used for training and the rest are reserved for testing the systems. Exponential membership functions are employed for TSK and CETSK. Since there are three classes within the data, three fuzzy rules are used in the systems. Training and testing errors and runtime of the systems are given in Table 8. CTSK is the most accurate and fastest system followed by CETSK. For this problem, CTSK achieves 98.3% accuracy (testing error of 0.258) versus SVM (96.7%) and ANN (95.8%). Moreover, Wilcoxon tests confirm significance of the system performance (p ¡ 0.01). We report F1-scores of 0.981 and 95% confidence intervals for error rates as 0.258±0.04 for CTSK to statistically contextualize its performance beyond runtime advantages. Cluster centres matrix of the IRIS data $V_{4\times3}$ (four features and three clusters) is computed by FCM as follows.

$$V = \begin{bmatrix} 5.0049 & 6.7047 & 5.8914 \\ 3.4151 & 3.0546 & 2.7538 \\ 1.4850 & 5.5686 & 4.3294 \\ 0.2611 & 2.0666 & 1.3700 \end{bmatrix}.$$

The TSK system tuned based on the input/output data, may simulate behavior of the process satisfactorily, but does not necessarily represent actual structures and patterns of the data, which are characterized by the clusters. This undermines the main goal of the fuzzy systems, which is high interpretability compared to the black-box tools such as artificial neural networks. The IF part of the CTSK system is constructed from the fuzzy clusters identified by the FCM algorithm. Therefore, this system represents structure of the data as well as their behavior as observed in the above applications. The IF part of CTSK designed for the IRIS data is shown in Figure 14. Regarding the above matrix of cluster centers, it is observed that heights (location of maximum point) of the membership functions coincide with the cluster centers and one can clearly see the data structure by viewing these membership functions and the IF part configuration of the system.

The main drawback with this system is non-convex shape of its membership functions. They can be replaced with convex functions such as triangular, exponential, trapezoidal, or bell-shape membership functions while keeping their centers unchanged. The modified system, which is called CETSK, still represents structure of the data and maintains a good degree of interpretability as discussed above. The IF part of CETSK for the IRIS data is shown in Figure 15, which is perfectly convex though less representative of the structure of the IRIS data compared to the CTSK system. The IF part of the traditional TSK system designed for these data is shown in Figure 16. It is observed that height of its membership functions does not coincide with the cluster centers given in the above matrix. The IF part parameters of this system are tuned by the GA algorithm such that the system follows output data properly, but the learning

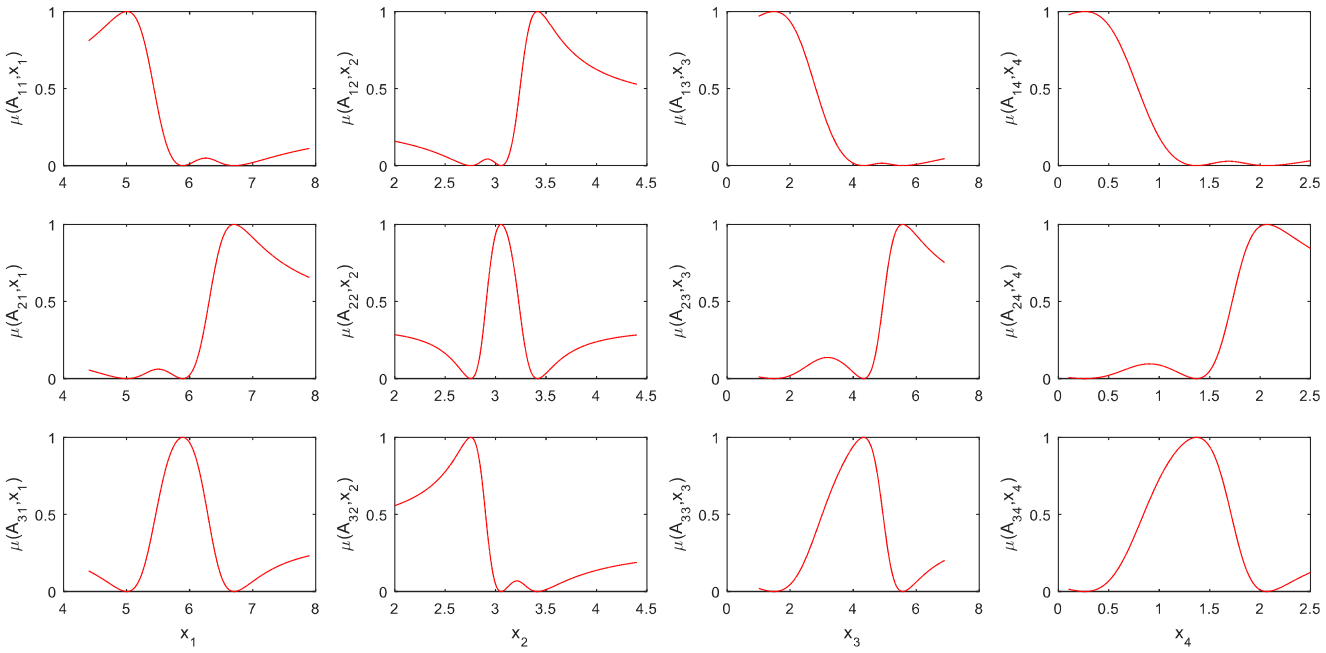


Figure 14: IF part of the CTSK system designed for classification of the IRIS data.

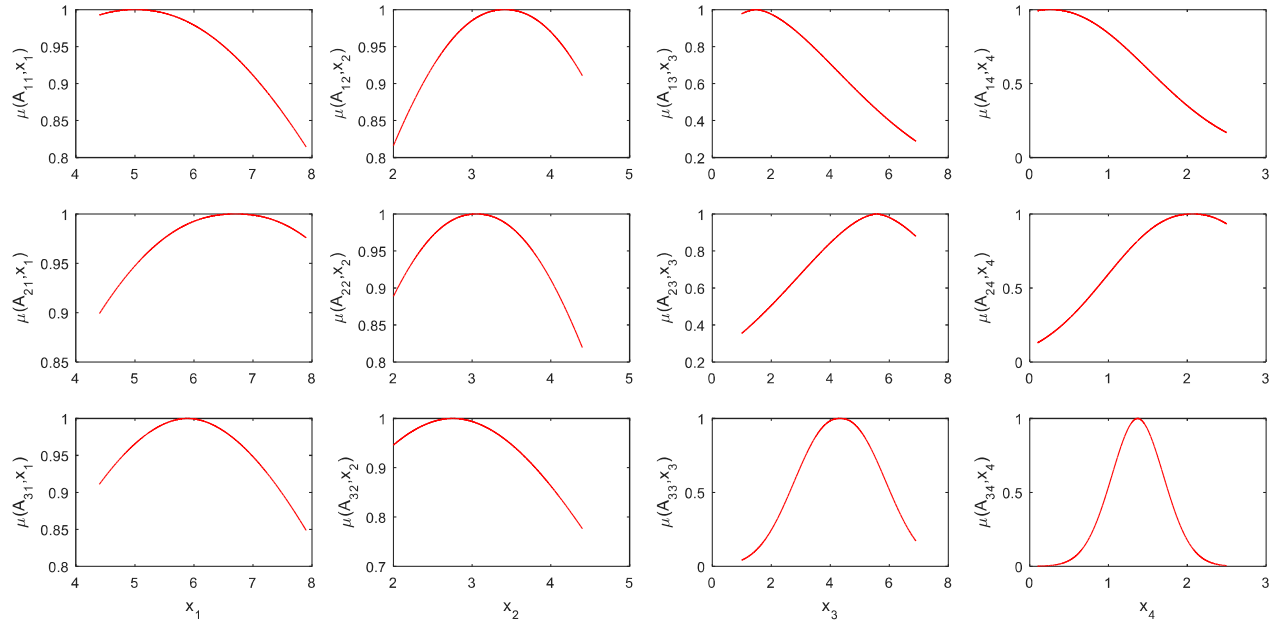


Figure 15: IF part of the CETSK system designed for classification of the IRIS data.

algorithm cannot fit IF part of the system to the data structure. However, the clustering algorithms such as the FCM, can do this perfectly as shown in Figure 14 and Figure 15.

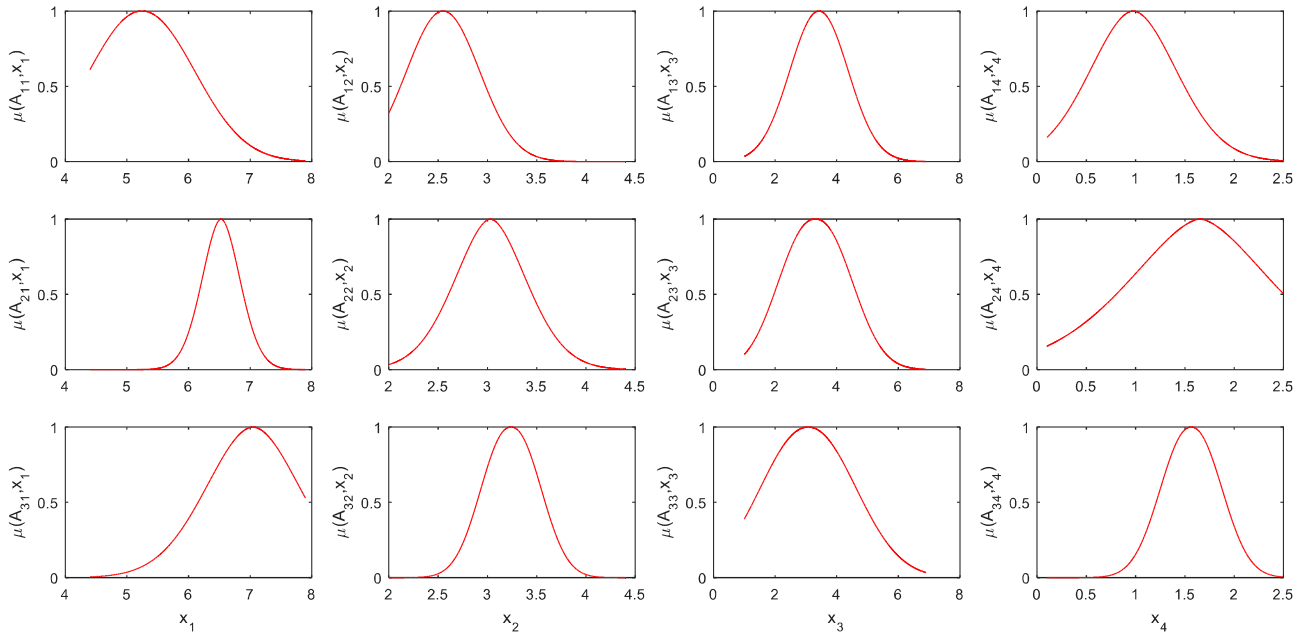


Figure 16: IF part of the TSK system designed for classification of the IRIS data.

6 Conclusions

Traditional TSK system is extended to a more general form as the ETSK system with fuzzy parameters in the IF part and nonlinear functions in the THEN part. Performance of ETSK is examined by forecasting, identification, and classification problems, which demonstrates superiority of the system over TSK. In order to incorporate structures and patterns of data into the fuzzy inference system, the clustering-based fuzzy system CTSK is introduced. CTSK replaces membership functions with the cluster centers by which the input data are fuzzified. Runtime of CTSK is shorter because no IF part parameters are tuned in this system. Moreover, CTSK is more accurate than TSK and ETSK. Finally, CTSK and ETSK are combined as the hybrid CETSK system to utilize merits of both systems.

Acknowledgement

The authors would like to thank the editors and the four anonymous reviewers for their insightful comments on the earlier version of this paper. These suggestions were instrumental in clarifying the content and strengthening the arguments.

References

[1] M. Arefi, *Geometric clustering fuzzy regression based on c-means clustering*, Iranian Journal of Fuzzy Systems, **22** (2025), 87-101. https://doi.org/10.22111/ijfs.2025.51386.9078

[2] S. Askari, *Oil reservoirs classification using fuzzy clustering*, International Journal of Engineering, Transactions C: Aspects, **30** (2017), 1391-1400. https://doi.org/10.5829/idosi.ije.2017.30.09c.12

[3] S. Askari, *Fuzzy c-means clustering algorithm for data with unequal cluster sizes and contaminated with noise and outliers: Review and development*, Expert Systems with Applications, **165** (2021), 113856. https://doi.org/10.1016/j.eswa.2020.113856

[4] S. Askari, *Noise-resistant fuzzy clustering algorithm*, Granular Computing, **6** (2021), 815-828. https://doi.org/10.1007/s41066-020-00230-6

[5] S. Askari, N. Montazerin, *A high-order multivariable fuzzy time series forecasting algorithm based on fuzzy clustering*, Expert Systems with Applications, **42** (2015), 2121-2135. https://doi.org/10.1016/j.eswa.2014.09.036

[6] S. Askari, N. Montazerin, M. H. F. Zarandi, *A clustering based forecasting algorithm for multivariable fuzzy time series using linear combinations of independent variables*, Applied Soft Computing, **35** (2015), 151-160. https://doi.org/10.1016/j.asoc.2015.06.028

[7] S. Askari, N. Montazerin, M. H. F. Zarandi, *Forecasting semi-dynamic response of natural gas networks to nodal gas consumptions using genetic fuzzy systems*, Energy, **83** (2015), 252-266. https://doi.org/10.1016/j.energy.2015.02.020

[8] S. Askari, N. Montazerin, M. H. F. Zarandi, *Generalized possibilistic fuzzy c-means with novel cluster validity indices for clustering noisy data*, Applied Soft Computing, **53** (2017), 262-283. https://doi.org/10.1016/j.asoc.2016.12.049

[9] S. Askari, N. Montazerin, M. H. F. Zarandi, *Modeling energy flow in natural gas networks using time series disaggregation and fuzzy systems tuned by particle swarm optimization*, Applied Soft Computing, **92** (2020), 106332. https://doi.org/10.1016/j.asoc.2020.106332

[10] S. Askari, N. Montazerin, M. H. F. Zarandi, E. Hakimi, *Generalized entropy based possibilistic fuzzy c-means for clustering noisy data and its convergence proof*, Neurocomputing, **219** (2017), 186-202. https://doi.org/10.1016/j.neucom.2016.09.025

[11] A. Banakar, M. F. Azeem, *Parameter identification of TSK neuro-fuzzy models*, Fuzzy Sets and Systems, **179** (2011), 62-82. https://doi.org/10.1016/j.fss.2011.05.003

[12] Z. Bian, J. Zhang, Y. Nojima, F. Chung, S. Wang, *Hybrid-ensemble-based interpretable TSK fuzzy classifier for imbalanced data*, Information Fusion, **98** (2023), 101845. https://doi.org/10.1016/j.inffus.2023.101845

[13] I. D. Borlea, R. E. Precup, F. Dragan, A. B. Borlea, *Centroid update approach to K-means clustering*, Advances in Electrical and Computer Engineering, **17** (2017), 3-10. https://doi.org/10.4316/AECE.2017.04001

[14] O. Cordon, F. Herrera, F. Hoffmann, L. Magdalena, *Genetic fuzzy systems, Evolutionary tuning and learning of fuzzy knowledge bases*, Advances in Fuzzy Systems–Applications and Theory, World Scientific, **19** (2001). https://doi.org/10.1142/4177

[15] A. Dass, S. Srivastava, R. Kumar, *A novel Lyapunov-stability-based recurrent-fuzzy system for the identification and adaptive control of nonlinear systems*, Applied Soft Computing, **137** (2023), 110161. https://doi.org/10.1016/j.asoc.2023.110161

[16] Z. Deng, Y. Cao, Q. Lou, K. Choi, S. Wang, *Monotonic relation-constrained Takagi-Sugeno-Kang fuzzy system*, Information Sciences, **582** (2022), 243-257. https://doi.org/10.1016/j.ins.2021.09.015

[17] M. Fazzolari, B. Gigli, R. Alcalá, F. Marcelloni, F. Herrera, *A study on the application of instance selection techniques in genetic fuzzy rule-based classification systems: Accuracy-complexity trade-off*, Knowledge-Based Systems, **54** (2013), 32-41. https://doi.org/10.1016/j.knosys.2013.07.011

[18] S. Huang, G. Zhao, Z. Weng, S. Ma, *Trapezoidal Type-2 fuzzy inference system with tensor unfolding structure learning method*, Neurocomputing, **473** (2022), 54-67. https://doi.org/10.1016/j.neucom.2021.12.011

[19] M. Javadian, R. Vaziri, S. Haghzad Klidbary, A. Malekzadeh, *Refining membership degrees obtained from fuzzy C-means by re-fuzzification*, Iranian Journal of Fuzzy Systems, **17** (2020), 85-104. https://10.22111/ijfs.2020.5408

[20] E. M. Joo, D. Chang, *Online tuning of fuzzy inference systems using dynamic fuzzy Q-learning*, IEEE Transactions on Systems, Man, and Cybernetics, **34** (2004), 1478-1489. https://doi.org/10.1109/TSMCB.2004.825938

[21] U. Kilic, E. S. Essiz, M. K. Keles, *Binary anarchic society optimization for feature selection*, Romanian Journal of Information Science and Technology, **26** (2023), 351-364. https://doi.org/10.59277/ROMJIST.2023.3-4.08

[22] Z. Mei, T. Zhao, X. Xie, *Hierarchical fuzzy regression tree: A new gradient boosting approach to design a TSK fuzzy model*, Information Sciences, **652** (2024), 119740. https://doi.org/10.1016/j.ins.2023.119740

[23] J. Ning, *Neural network-based pattern recognition in the framework of edge computing*, Romanian Journal of Information Science and Technology, **27** (2024), 106-119. https://doi.org/10.59277/ROMJIST.2024.1.08

[24] H. Ouifak, A. Idri, *On the performance and interpretability of Mamdani and Takagi-Sugeno-Kang based neuro-fuzzy systems for medical diagnosis*, Scientific African, **20** (2023), e01610. https://doi.org/10.1016/j.sciaf.2023.e01610

[25] T. Pidikiti, Shreedevi, B. Gireesha, M. Subbarao, V. B. M. Krishna, *Design and control of Takagi-Sugeno-Kang fuzzy based inverter for power quality improvement in grid-tied PV systems*, Measurement: Sensors, **25** (2023), 100638. https://doi.org/10.1016/j.measen.2022.100638

[26] R. E. Precup, A. T. Nguyen, S. Blazic, *A survey on fuzzy control for mechatronics applications*, International Journal of Systems Science, **55** (2024), 771-813. https://doi.org/10.1080/00207721.2023.2293486

[27] B. Qin, F. Chung, Y. Nojima, H. Ishibuchi, S. Wang, *Fuzzy rule dropout with dynamic compensation for wide learning algorithm of TSK fuzzy classifier*, Applied Soft Computing, **127** (2022), 109410. https://doi.org/10.1016/j.asoc.2022.109410

[28] F. Salas-Molina, J. Reig-Mullor, D. Pla-Santamaria, A. Garcia-Bernabeu, *A multidimensional approach to rank fuzzy numbers based on the concept of magnitude*, Iranian Journal of Fuzzy Systems, **20** (2023), 137-153. https://doi.org/10.22111/ijfs.2023.43939.7738

[29] Y. Shan, S. Li, F. Li, Y. Cui, S. Li, M. Chen, X. He, *Fuzzy self-consistent clustering ensemble*, Applied Soft Computing, **151** (2024), 111151. https://doi.org/10.1016/j.asoc.2023.111151

[30] K. Siminski, *FuBiNFS–fuzzy biclustering neuro-fuzzy system*, Fuzzy Sets and Systems, **438** (2022), 84-106. https://doi.org/10.1016/j.fss.2021.07.009

[31] K. Siminski, *Prototype based granular neuro-fuzzy system for regression task*, Fuzzy Sets and Systems, **449** (2022), 56-78. https://doi.org/10.1016/j.fss.2022.03.001

[32] M. Sugeno, G. Kang, *Structure identification of fuzzy model*, Fuzzy Sets and Systems, **28** (1988), 15-33. https://doi.org/10.1016/0165-0114(88)90113-3

[33] M. Sugeno, T. Yasukawa, *A fuzzy logic based approach to qualitative modeling*, IEEE Transactions on Fuzzy Systems, **1** (1993), 1-24. https://doi.org/10.1109/TFUZZ.1993.390281

[34] T. Takagi, M. Sugeno, *Fuzzy identification of systems and its application to modeling and control*, IEEE Transactions on Systems, Man, and Cybernetics, SMC-15, (1985), 116-132. https://doi.org/10.1109/TSMC.1985.6313399

[35] A. Tewari, M. U. Macdonald, *Knowledge-based parameter identification of TSK fuzzy models*, Applied Soft Computing, **10** (2010), 481-489. https://doi.org/10.1016/j.asoc.2009.08.034

[36] D. Wu, J. M. Mendel, *Aggregation using the linguistic weighted average and interval Type-2 fuzzy sets*, IEEE Transactions on Fuzzy Systems, **15** (2007), 1145-1161. https://doi.org/10.1109/TFUZZ.2007.896325

[37] R. Xie, S. Wang, *A wide interpretable Gaussian Takagi-Sugeno-Kang fuzzy classifier and its incremental learning*, Knowledge-Based Systems, **241** (2022), 108203. https://doi.org/10.1016/j.knosys.2022.108203

[38] H. Yu, L. Jiang, J. Fan, S. Xie, R. Lan, *A feature-weighted suppressed possibilistic fuzzy c-means clustering algorithm and its application on color image segmentation*, Expert Systems with Applications, **241** (2024), 122270. https://doi.org/10.1016/j.eswa.2023.122270

[39] Y. Zhang, W. Ding, *Motor imagery classification via stacking-based Takagi-Sugeno-Kang fuzzy classifier ensemble*, Knowledge-Based Systems, **263** (2023), 110292. https://doi.org/10.1016/j.knosys.2023.110292

[40] Y. Zhang, G. Wang, X. Huang, W. Ding, *TSK fuzzy system fusion at sensitivity-ensemble level for imbalanced data classification*, Information Fusion, **92** (2023), 350-362. https://doi.org/10.1016/j.inffus.2022.12.014

[41] Y. Zhang, G. Wang, T. Zhou, X. Huang, S. Lam, J. Sheng, K. Choi, J. Cai, W. Ding, *Takagi-Sugeno-Kang fuzzy system fusion: A survey at hierarchical, wide and stacked levels*, Information Fusion, **101** (2024), 101977. https://doi.org/10.1016/j.inffus.2023.101977