Tabriz Journal of Electrical Engineering (TJEE), vol. 54, no. 3, Autumn 2024 Serial no. 109
DOI: 10.22034/tjee.2024.58036.4692

# ISY: Improved Sugeno-Yasukawa Fuzzy Modelling Approach Using a Novel Clustering and Project Method for Input Partitioning

Sajad Haghzad Klidbary*

[1]Department Electrical and Computer Engineering, University of Zanjan, Zanjan, Iran.
E-mail addresses: s.haghzad@znu.ac.ir
*Corresponding author

Received:18/08/2023, Revised: 27/11/2023, Accepted:09/01/2024.

**Abstract**
Soft computing algorithm such as fuzzy logic, neural networks, and evolutionary algorithms are widely used in many fields. Fuzzy logic, in particular, has gained significant popularity due to its significant ability in modelling. So far, various methods of fuzzy modelling have been presented; each of these methods has its advantages and disadvantages. While all methods start from the input, Sugeno-Yasukawa (SY) differs by initiating the analysis from the output. The popularity of the SY method can be attributed to its effective rule extraction algorithm, which employs a clustering process to determine input membership functions. In this paper, we propose a cluster search algorithm and a new fuzzy partitioning method that enhance the mapping of the output space to the input space by distributing Gaussian functions for each data point within a cluster and calculating their membership values. With this proposed new clustering search method, the performance of the SY method is improved. Through simulations, the proposed method has improved the mean square errors (MSE) criterion by 0.001, and improved the accuracy criterion by 1.5 percent.

**Keywords**
Fuzzy inference system (FIS), fuzzy modelling, Sugeno-Yasukawa (SY), membership function approximation, clustering.

## 1. Introduction

Nowadays, developing computing systems based on artificial intelligence is a primary objective in this field. Among these methods, fuzzy logic stands out as a prominent technique that enables achieving high accuracy [1]. Following the introduction of fuzzy sets, various concepts and solutions have been proposed, including fuzzy algorithms [2], fuzzy decision-making [3], fuzzy ordering [4] and an approach for analyzing complex systems [5]. Fuzzy inference is an approximate simulation of expert decision-making through human-like reasoning. Therefore, fuzzy systems use fuzzy logic and try to make decisions similar to an expert. This fuzzy logic is compatible with fuzzy data and is used to model uncertainty.

Fuzzy systems, similar to humans, take advantage of intuitive knowledge and general understanding and use it for modelling. There are two justifications for the development of fuzzy systems theories, as follows: firstly, the real world is too complex to be easily described by precise relationships. Secondly, as human knowledge and experience have grown over time, it is imperative to incorporate this acquired knowledge into engineering systems.

Classical logic relies on the law of Modus Ponens as the basis for reasoning and inference. Modus Ponens operates on if-then rules, where if the antecedent part is observed, the consequent part is concluded. However, due to the inherent approximations involved in human reasoning and decision-making, a more generalized form of Modus Ponens, known as the Generalized Modus Ponens, is necessary.

Several algorithms utilizing fuzzy techniques have been proposed for fuzzy modelling. Fig. 1 depicts the general structure of a fuzzy system, comprising different components with specific functions. In a real system, many rules are compiled to solve a problem. These rules are stored in a section of the system referred to as the rules database. The fuzzy inference engine combines the "if-then" rules in the rule base by mapping fuzzy sets defined in the input space to fuzzy sets defined in the output space.

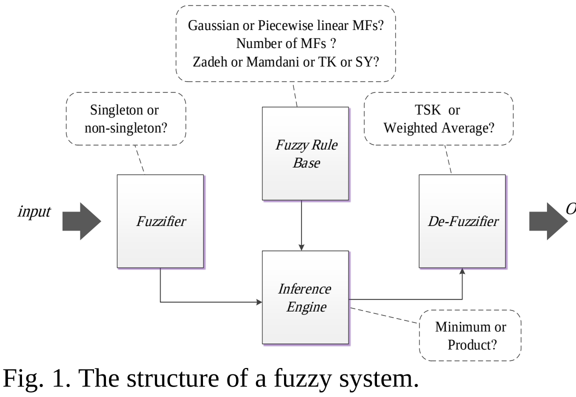



Fig. 1. The structure of a fuzzy system.

Among the most important types of modelling methods, the Mamdani method [6], the TS method [7] were mentioned, and the universal approximation of these methods has been shown in several papers [8-10]. ALM [11] and SY [12] can be mentioned as other methods. In most fuzzy modelling methods, first the parameters of the "if-part", such as the number of partitions performed on the input variables, the shape of the membership functions and the shape of the inference rules are determined. Then, by performing optimization methods, the parameters of the “then-part” are determined. One of the disadvantages of the above modelling is the high number of inference rules. Some of these rules are found with many calculations and do not play a significant role in the modelling result. Therefore, in previous methods such as TS, the algorithm started from the inputs, that is, choosing the appropriate variables for the input and then partitioning them. However, SY's point of view is that a person first observes the phenomenon and then examines the variables affecting it. Therefore, the first task is to find the similarity between these data. For this purpose, it is necessary to perform clustering on the observed outputs. Therefore, this method is considered the main priority of the paper.

One of the main drawbacks of this SY is the approximation of trapezoidal functions to the input membership functions, which can result in errors in the output. In this paper, a new fuzzy clustering method is proposed to find an appropriate partitioning of the input data. This method uses a Gaussian function that is assigned to each training data. Finally, by merging all these distributions, a membership function is created for each cluster. Ultimately, by determining the number of memberships in inputs and their belief values, it is possible to formulate rules and create a fuzzy system. Finally, in this paper, by presenting a new clustering method, it can be stated that, the main contribution of the proposed method is no need to fit a trapezoidal function, and thus, calculations based on approximation will not be needed. So the final result is that the speed and accuracy of the conventional SY method are increased.

The structure of the paper is as follows: the second section discusses the types of modelling algorithms, the third section presents the proposed method, the fourth section reviews the analysis and simulations, and the final section includes summaries, conclusions, and suggestions for future work.

## 2. Types of Fuzzy Modelling Methods

In this section, the types of modelling methods are first investigated. To provide a brief explanation of each fuzzy modelling method, a Multi-Input Single-Output (MISO) system is assumed.

### *2.1. Mamdani fuzzy Modelling System*

In order to implement fuzzy systems in engineering, one simple method is to incorporate a fuzzifier at the input, which converts variables with real values into fuzzy sets, and a defuzzifier, which converts fuzzy sets into variables with real values. In this method, the inputs ($X_1$,$X_2$) and output ($Y$) are first fuzzified and the corresponding rules are as follows:

$$If\ X_1\ is\ A_i\ and\ X_2 is\ B_i\ Then\ Y = C_i \tag{1}$$

In this method, the minimum or product operator is used to determine the firing level of each rule, and the maximum operator is used to obtain the final output, which is a fuzzy set. This approach has been applied in various applications [13-16].

### *2.2. TS Fuzzy Modelling System*

TS's inference system is similar to Mamdani's inference, with the main difference being that Sugeno's method uses mathematical functions in the "then-part" instead of fuzzy sets. The rules in Sugeno's inference system are written as follows:

$$If\ X_1\ is\ A_i\ and\ X_2 is\ B_i\ Then\ Y = F_i(X_1, X_2) \tag{2}$$

In this method, the minimum or product operator is used to calculate the firing level of each rule. The final output is a crisp number, which is obtained by taking the weighted sum according to the following formula.

$$\hat{y} = \frac{\sum_{j=1}^{n} w_j \times F_j}{\sum_{j=1}^{n} w_j} \tag{3}$$

Some of the applications of this method include [17-20].

### *2.3. ALM Fuzzy Modelling System*

Instead of relying on numerical values, ALM captures the system's overall behaviour in the form of images. ALM is a fuzzy adaptive learning method that simplifies complex problems by breaking them down into multiple simpler problems. The ALM function is designed for MISO systems, involving decomposing the problem and transforming it into several SISO subsystems. The core of ALM is the Ink Drop Spread (IDS) operator, which models fuzzy interpolation. The IDS operator is used to express data uncertainty modelling. Its effect on the collection of experience points is similar to dropping an ink drop onto any point on a plane. Each Ink Drop Spread unit consists of two main parts: a two-dimensional plane (IDS Plane) representing the relationship between the input and output variables, and the extraction of features related to each plane. The features are called narrow path ($NP_i$) and spread ($SP_i$). The rules for ALM are as follows:

$$R_1: if\ X_1\ is\ B_i\ then\ Y_1 = NP_1 \tag{4}$$
$$R_2: if\ X_2\ is\ A_i\ then\ Y_2 = NP_2 \tag{5}$$

Finally, after extracting all the features, the output relationship is determined by combining the outputs from all the subsystems.

$$\hat{y} = \frac{\sum_{j=1}^{n} \frac{1}{SP_j} \times NP_j}{\sum_{j=1}^{n} \frac{1}{SP_j}} \tag{6}$$

Some of the applications of this method include [21-26].

### *2.4. SY Fuzzy Modelling System*

In SY's method [12], the initial step involves identifying similarities among the data by clustering the observed outputs. This step is crucial in obtaining inference rules and membership function parameters from the data. This distinguishes SY's method from other modelling approaches. In the SY method, the parameters of the "then-part" of the inference system are determined first. Then, these parameters are mapped onto the input space, automatically determining the parameters of the "if-part". The algorithm consists of several steps, including clustering the output space, mapping the resulting clusters to the input space, and constructing inference rules. These steps help refine the modelling process, including determining effective data, the number of clusters in the output space, and the shape of the membership functions. The SY method for identification is shown in Fig. 2 and illustrated as follows:

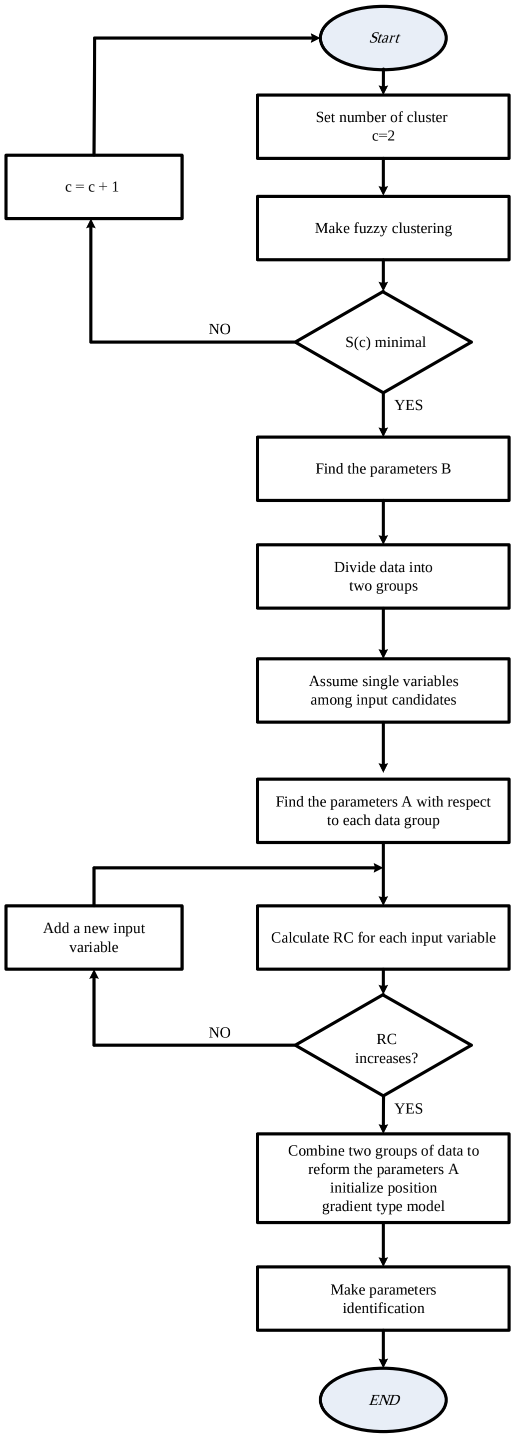



Fig. 2. The algorithm of the SY method for identification [12].

Step 1: The first step in SY's method is to cluster the output data of the system using a fuzzy clustering method such as FCM (Fuzzy C-Means). However, determining the number of clusters is a critical step because it needs to be given in advance to the algorithm. To determine the optimal number of clusters, the following approach can be taken: start by modelling the system with two clusters and measure the modelling error using test data. Increase the number of clusters and repeat the same modelling and error measurement process. Continue increasing the number of clusters and repeating the steps until the first local minimum is reached in the modelling error. At this point, the number of clusters corresponding to the local minimum is considered optimal for the system.

By iteratively adjusting the number of clusters, the method aims to find the most suitable number that minimizes the modelling error and provides the best representation of the system. The relationship used to find the number of suitable clusters is according to the following:

$$SC_m = \sum_{i=1}^{N}\sum_{j=1}^{nc} u_{ij}^m \left(\left\|x_i - v_j\right\|^2 - \left\|v_j - \bar{x}\right\|^2\right) \quad (7)$$

where each parameter is as follows:
$N$: Number of data to be clustered;
$nc$: Number of clusters ($nc \geq 2$);
$x_i$: $i^{th}$ $data$;
$v_j$: Center of $j^{th}$ cluster;
$\bar{x}$: Average of data $x_1, x_2, \dots, x_n$;
$||.||$: Norm;
$u_{ij}$: Grade of $i^{th}$ $data$ belonging to $j^{th}$ cluster;
$m$: Adjustable weight (between 1.5 and usually 3).

Now, the data will be divided into two groups, A and B. A model will be built for each of these categories and the best model will be obtained based on the following relationship:

$$RC = \left[\sum_{i=1}^{k_A}\frac{(y_i^A - y_i^{AB})^2}{k_A} + \sum_{i=1}^{k_B}\frac{(y_i^B - y_i^{BA})^2}{k_B}\right] \quad (8)$$

The lower the value of RC, the greater the effect on the output. In the above relationship, the parameters are as follows:
$k_A$ and $k_B$: the number of data of the groups $A$ and $B$;
$y_A$ and $y_B$: the number of data of the groups $A$ and $B$;
$y_{AB}$: the output data of the intermediate model identified for group $A$ for input data of group $B$;
$y_{BA}$: the output data of the intermediate model identified for group $B$ for input data of group $A$.

Step 2: The second step is to map the identified clusters to the input space and establish inference rules. Assuming that the output data is divided into three clusters, as depicted in Fig. 3, we can assign the following:

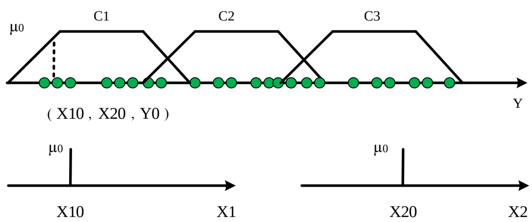



Fig. 3. Example of three clusters for synthetic output data.

For example as shown in Fig. 4, if cluster $C_1$ is mapped to the input space, specifically, $X_1$ and $X_2$. To achieve this, we select individual data belonging to cluster $C_1$ greater than zero, like $Y_0$ with a degree of belonging $\mu_0$. Each corresponding output data has its respective $X_1$ and $X_2$, such as $X_{10}$ and $X_{20}$. In order to map data $Y_0$ onto the input spaces, we need to consider two points, $X_{10}$ and $X_{20}$, on the input space and assign the degree of belonging $\mu_0$ to them. By following the same process for other points in the $C_1$ cluster, we ultimately obtain membership functions. In this algorithm, for simplicity in calculations, these membership functions are assumed to have a trapezoidal shape (as discussed in references [12, 27, 28]).

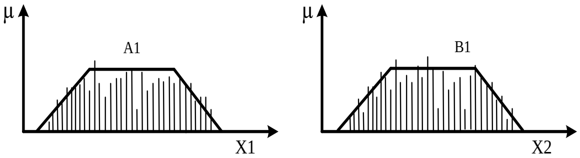



Fig. 4. Approximating the optimal trapezoidal functions associated with the inputs that result in the output of a cluster.

Therefore $C_1$ is obtained by using this simple mapping, the rules in SY's inference system are written as follows:

$$If\ X_1\ is\ A_1\ and\ X_2\ is\ B_1\ the\ Y\ is\ C1 \quad (9)$$

By performing the same procedure for all clusters, we can generate inference rules equivalent to the number of clusters. In the third step, optimization techniques can be employed to enhance the outcomes.

Numerous research efforts have been dedicated to enhancing the performance of this modelling methodology, which we will explore in more detail. In the paper [29], the shortcomings of YS's approach were examined, and solutions were proposed to reduce the number of rules and address unclear details. The paper [30], presented a cluster search algorithm that enhanced the representation of the output space in relation to the input space. In the paper [27], simple solutions were suggested for uncertain aspects of the original paper, including trapezoidal approximation of membership functions, rule generation from sample data points, and selection of significant variables. In the paper [31], modifications were made to the modelling process to identify effective input parameters among a large set of possibilities. The parameter setting phase for the membership function of the input and output parameters was implemented for each step in detecting effective parameters. The paper [32], improved the algorithm by introducing a novel trapezoidal approximation method and incorporating intermediate models into the SY modelling process to achieve the final fuzzy model. In the paper [33], a Genetic Algorithm (GA) based method was utilized for the parameter identification phase, which reportedly yielded superior results compared to the original method and other comparative approaches.

The SY method's rule extraction algorithm has contributed to its popularity. It identifies the partition in the output space and then projects it back to the input space, resulting in a sparse fuzzy rule base. Trapezoid approximation of clustered data is achieved through two steps: determining the convex hull of the original data set, followed by approximating the convex hull using trapezoidal membership functions. Consequently, the key challenges in previous clustering approaches have been approximation and complexity. Based on the aforementioned papers, we will present the proposed algorithm to enhance the performance of the SY method.

## 3. Proposed Method

In this paper, a new fuzzy clustering method is proposed. The flowchart of the proposed method is shown in Fig. 5. In the following, a demonstration of the proposed algorithm is provided.

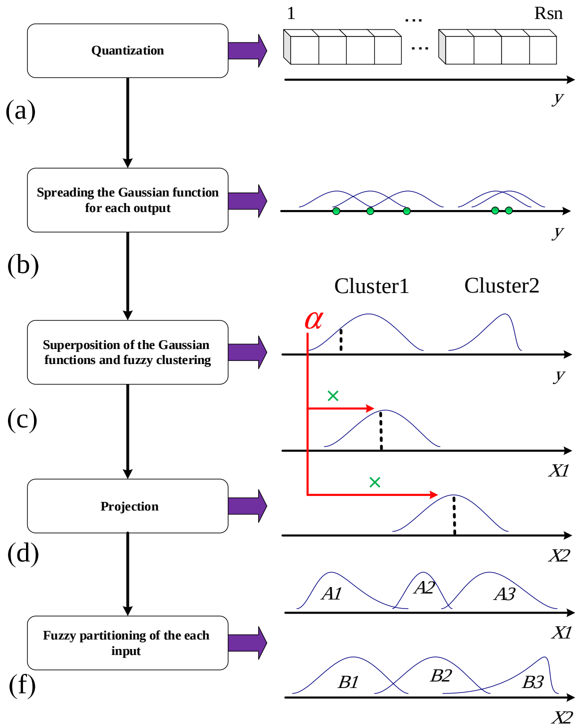



Fig. 5. The flowchart of the proposed method with five steps. (a) Quantization, (b) Spreading the Gaussian function, (c) Fuzzy clustering, (d) Projection, (f) Fuzzy partitioning in input space.

Step 1: First, a one-dimensional grid vector with desired resolution $R_{sn}$ is considered, and the quantization of the output changes is achieved using equation (10).

$$y_q = \left\lfloor \frac{(y_i - y_{min}) \times R_{sn}}{y_{max} - y_{min}} \right\rfloor + 1, y_i \in [y_{min}, y_{max}), \quad (10)$$

$$\overset{Eq.(10)}{\rightleftarrows} \quad Y_q \in \{1, 2, \dots, Rsn\} \quad (11)$$

Where, $y_d$ denotes the quantized output vector, $y_i$ indicates number of output sapmle, $y_{min}$ indicates the minimum value of all output, and $y_{man}$ indicates the maximum value of all output.

Step 2: For each output data in the dataset, Gaussian functions are applied to their centers on the vector according to equations (12) and (13).

$$d(y_T + u) = d(y_T + u) + h(u), -r \leq u \leq r, \quad r = 2 \times \sigma \quad (12)$$

$$h(u) = e^{-\frac{(u - y_T)^2}{2\sigma^2}} \quad (13)$$

$h(u)$ is a Gaussian function, and $\sigma$ represents the standard deviation.
Step 3: The superposition of these Gaussian functions is computed as shown in Fig. 5-b. After that, a fuzzy representation function is obtained. Finally, normalization is performed, and the proposed fuzzy clustering is applied by using the histogram of the vector and a threshold value. The Fig. 6 shows the histogram plot for the output vector.

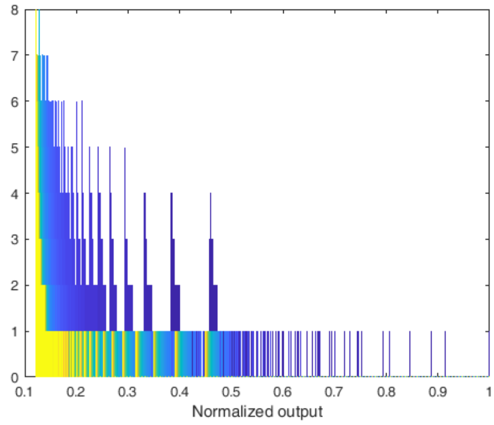



Fig. 6. Histogram bar chart of the elements in the output vector.

Step 4: After clustering, a coefficient corresponding to each cluster is obtained for every output data. The output data is projected to the input space according to the obtained coefficients. The equations are as follows:

$$C_i(X_T+u) = C_i(X_T+u) + g(u), -r \leq u \leq r, \qquad r = 2 \times \sigma \tag{14}$$

$$g(u) = \alpha \times e^{-\frac{(u-X_T)^2}{2\sigma^2}} \tag{15}$$

In the above relationship, $i$ represents the number of features (system dimensions), $g(u)$ is a Gaussian function, $\alpha$ is the membership degree obtained from fuzzy clustering, and $\sigma$ represents the standard deviation. The standard deviation is chosen large when the number of training data is small to cover more space.
Step 5: With the coefficients in the previous step (as shown in Fig. 5-d), partitioning is applied to the input variables based on their values, and a Gaussian function is assigned to each data point based on its degree of belief as shown in Fig. 5-f. The larger these coefficients, the stronger the data's association with the respective cluster, resulting in a greater impact.
It is important to note that instead of the Gaussian function, any function with a degree of belief of one at the center and a decrease in the membership degree with an increase in distance from the center can be considered. Among the advantages of the proposed method, it should be highlighted that it does not use trapezoidal approximation and has the sigma control parameter in the Gaussian function. In the following section, we will demonstrate and simulate the proposed method in the modelling and classification space. The proposed method involves simple calculations, but it exhibits high efficiency.
The proposed algorithm has $S(N) = O((D+1) \times R_{sn})$ space complexity, and $T(N) = O(N \times 2 \times r \times (D+1))$ time complexity. $D$ is system dimensional, and $N$ is the number of samples.

## 4. Simulation Results

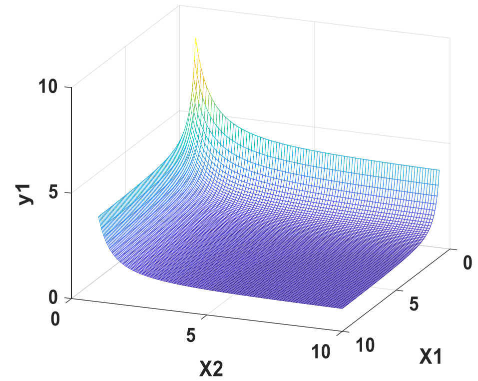
To further investigate the performance of the proposed algorithm was evaluated and compared with the Conventional SY, MLP and ANFIS, and the results are shown in Table I. The simulations are done by MATLAB 2023a environment and Fuzzy Logic Toolbox with Core i5 processor, 2.4 GHz, and 8 GB RAM (a personal computer). The simulation has been done in three different scenarios (Modelling, classification, and parameter sensitivity).

### *4.1. Modelling Non-Linear Systems*

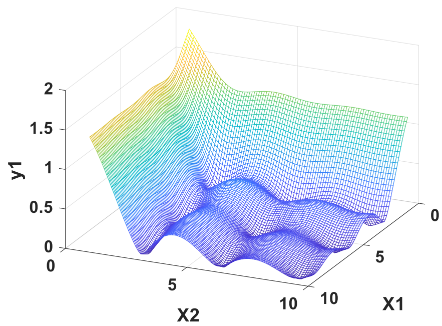
The system modelling in this section focuses on two functions, $y_1$ and $y_2$, which are two non-linear systems with two inputs and a single output Fig. 7 shows two functions $y_1$ and $y_2$. The equations are defined as follows:

$$y_1 = (1 + x_1^{-2} + x_2^{-1.5})^2 \tag{16}$$

$$y_2 = \sqrt[2]{2\left(\frac{sinx_1}{x_1}\right)^2 + 3\left(\frac{sinx_2}{x_2}\right)^2} \tag{17}$$




(a)




(b)

Fig. 7. (a) A function $y_1$ is defined in (16), (b) a function $y_2$ is defined in (17).

To assess the accuracy of the algorithm and evaluate the modelling error, three metrics, Fraction of Variance Unexplained (FVU), Pearson Correlation Coefficient (PCC) and MSE (Mean Square Error) were utilized as used in [34]. The equations for these metrics are as follows:

$$FVU = \frac{\sum_{i=1}^{k}(y^e(x_i) - y(x_i))^2}{\sum_{i=1}^{k}(y(x_i) - \bar{y})^2}, \quad \bar{y} = \left(\frac{1}{k}\right)\sum_{i=1}^{k} y(x_i), \tag{18}$$

Table I: Comparison of SY, proposed SY, MLP and ANFIS based on FVU, PCC, and MSE metrics.

| Algorithm | Function | FVU | PCC | MSE |
|---|---|---|---|---|
| Conventional SY | $y_1$ | $0.0121 \pm 0.0341$ | $0.971 \pm 0.0124$ | $0.0104 \pm 0.0134$ |
| | $y_2$ | $0.0145 \pm 0.0293$ | $0.972 \pm 0.0176$ | $0.0135 \pm 0.0156$ |
| Proposed SY | $y_1$ | $0.0116 \pm 0.0294$ | $0.983 \pm 0.0108$ | $0.0092 \pm 0.0117$ |
| | $y_2$ | $0.0138 \pm 0.0213$ | $0.976 \pm 0.0142$ | $0.0126 \pm 0.0196$ |
| MLP | $y_1$ | $0.0185 \pm 0.0734$ | $0.981 \pm 0.0374$ | $0.0114 \pm 0.0146$ |
| | $y_2$ | $0.0124 \pm 0.0346$ | $0.982 \pm 0.0174$ | $0.0141 \pm 0.0165$ |
| ANFIS | $y_1$ | $0.0079 \pm 0.0028$ | $0.996 \pm 0.0143$ | $0.0089 \pm 0.0138$ |
| | $y_2$ | $0.1117 \pm 0.0339$ | $0.945 \pm 0.0143$ | $0.0127 \pm 0.0186$ |

$$PCC = \frac{\sum_{i=1}^{k}(y_i - \bar{y}) \times (y_i^e - \overline{y^e})}{\sqrt{\sum_{i=1}^{k}(y_i - \bar{y})^2 \times \sum_{i=1}^{k}\left(y_i^e - \overline{y^e}\right)^2}} \tag{19}$$

$$MSE = \frac{1}{k}\sum_{i=1}^{k}(y(x_i) - y^e(x_i))^2 \tag{20}$$

Where, $\hat{y}$ is the model output, $\bar{y}$, and $(\overline{y^e})$ represent the average of output vectors, and $k$ represents the number of test data. The lower the FVU is, the higher the accuracy of the model, and the closer the correlation criterion is to one, the higher the accuracy of the output. FVU is directly related to MSE (Mean Square Error) and both metrics have the same direction of movement. The results are presented in Table I. Fig. 8 shows the MSE values of function $y_1$, for all ten runs of 10-fold cross-validation lower than the SY MSE value.

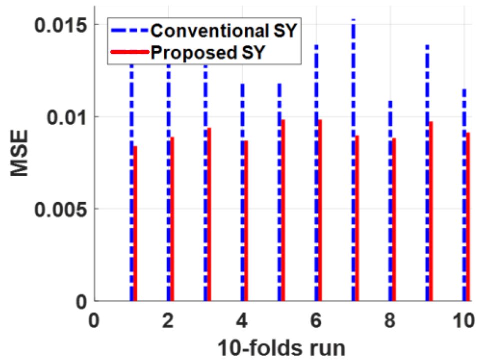



Fig. 8. Comparison of SY and proposed SY performance for a $y_1$ nonlinear system.

### *4.2. Breast Cancer Survival Prediction*

Classification of objects in different classes has always been interesting in different applications [35]. The purpose of this case study was to investigate the applicability of the proposed SY method for a real-world problem. The problem under investigation was the prediction of survival for breast cancer patients. A classification problem was performed on the Australian National Territory and surrounding regions' Breast Cancer Treatment Dataset (ABCTD) (1997-2009) [33]. The dataset consisted of 814 patients and included 12 prognosis factors. Additionally, information on the number of survival days was collected for each patient. The objective of the classification problem was to predict whether a patient would survive or not, using a five-year survival threshold. To investigate the performance of the proposed algorithm, we review and compare the model with algorithms such as conventional SY, proposed SY, MLP, ANFIS, AH and KTLNN (k-two-layer nearest neighbor) [36], and the results are shown in Table II.

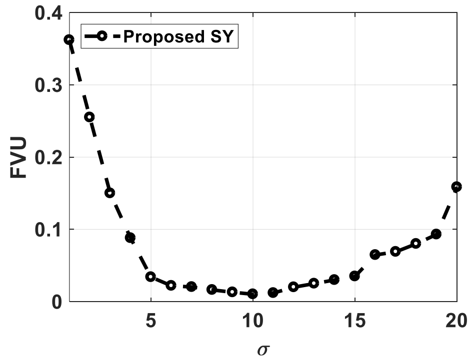
Table II: Accuracy values for five algorithms in ABCTD.

| # | Method | Accuracy |
|---|---|---|
| 1 | Conventional SY | $53.75 \pm 4.00$ |
| 2 | Proposed SY | $65.57 \pm 1.74$ |
| 3 | MLP | $60.47 \pm 3.14$ |
| 4 | ANFIS | $60.20 \pm 3.45$ |
| 5 | AH [33] | $63.57 \pm 2.30$ |
| 6 | KTLNN [36] | $63.97 \pm 1.86$ |

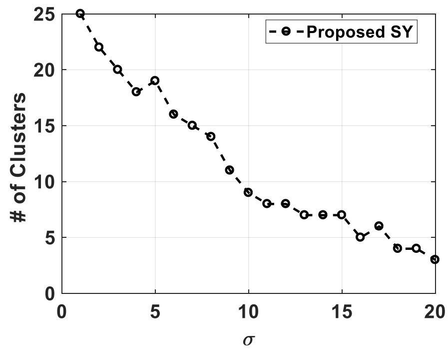
Based on the preceding table, the proposed algorithm not only has a better accuracy than the other four algorithms, but it also has less variance in its accuracy. In the next section, the sensitivity of the accuracy to the algorithm's parameters is investigated.

### *4.3. Parameter sensitivity and noise robustness*

As mentioned, one of the controllable parameters in the proposed SY is the standard deviation ($\sigma$). Fig. 9-a shows the comparison of proposed SY based on FVU and Fig. 9-b shows changes in number of clusters with respect to different values of the $\sigma$ in function$y_1$.




(a)




(b)

Fig. 9. (a) The comparison of proposed SY based on FVU with respect to different value of the $\sigma$ in function $y_1$, (b) Changes in the number of clusters compared to changes in $\sigma$ in function $y_1$.

Noise is one of the major challenges in soft computing. The noise resistance of the SY and the proposed SY was compared, and the results are shown in Fig. 10. The percentage of the applied noise is determined according to the following relationship.

$$y_{noisy} = y + \tilde{n}_p\,, \quad \tilde{n}_p = \frac{y\times\tilde{n}\times p}{100}, \tag{21}$$

Where, p represents the applied noise percentage and $\tilde{n}$ is a random number with a uniform distribution in the interval [-1,1], and y is the output data vector. As a result, according to this formula, noise refers to the random displacement of the actual output based on a percentage of the data itself. In Fig. 10, the proposed SY and other algorithms are compared based on PCC with respect to different percentages of the applied noise.

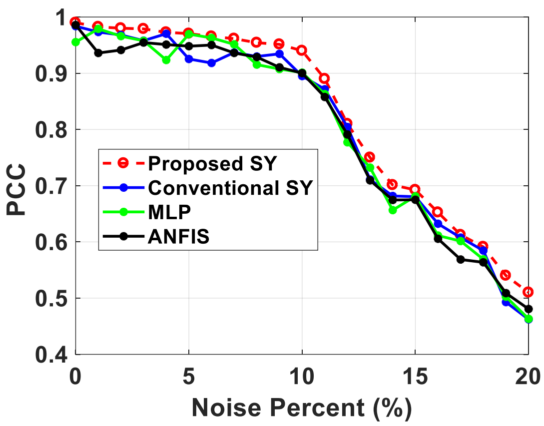



Fig. 10. The comparison of proposed SY, SY, MLP, and ANFIS based on PCC with respect to different percentages of applied noise in function $y_1$.

Finally, considering the simulation results (all figures and tables); it is evident that the proposed SY employs simpler calculations while achieving less error in modelling and higher classification accuracy than SY and other algorithms. This algorithm demonstrates good performance due to the utilization of the smoothing operator, which exhibits inherent similarity to noise. The proposed method identifies outlier data by employing a threshold. Additionally, the proposed method is no need to fit a trapezoidal function, and thus, calculations based on approximation will not be needed. Not using approximation increases the speed.

## 5. Conclusion

Fuzzy modelling methods such as Mamdani, TS, SY, and ALM have garnered attention from researchers in many fields of science and technology. Unlike traditional fuzzy modelling methods, the SY algorithm examines the output first to determine the input parameters. The selection of suitable distributions and appropriate membership functions for input variables is advantageous in enhancing the efficiency of this algorithm. In this paper, the basic performance evaluation method of SY has been discussed by introducing a novel approach to clustering and identifying input parameters. The main objective of our approach was to develop an algorithm that eliminates the need for trapezoidal function approximation. The simulation results confirm the favorable performance of this method. In future work, the proposed algorithm can be further investigated in the type-2 fuzzy set, and the use of the GA (Genetic Algorithm) for finding optimum parameters can be explored.

## 6. References

[1]. Deng, J., & Deng, Y. (2021). Information volume of fuzzy membership function. International Journal of Computers Communications & Control, 16(1).

[2]. Nabijonov, R. (2022). Theories of fuzzy sets and their application in face recognition. Innovation in the modern education system.

[3]. Akram, M., & Naz, S. (2019). A novel decision-making approach under complex Pythagorean fuzzy environment. Mathematical and Computational Applications, 24(3), 73.

[4]. Štěpnička, M., Holčapek, M., & Škorupová, N. (2019, June). Orderings of extensional fuzzy numbers. In 2019 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE) (pp. 1-6). IEEE.

[5]. Tranquillo, J. V. (2019). An introduction to complex systems. Lewisburg: Springer International Publishing.

[6]. Mamdani, E.H. and S. Assilian, *An experiment in linguistic synthesis with a fuzzy logic controller*. International journal of man-machine studies, 1975. 7(1): p. 1-13.

[7]. Takagi, T. and M. Sugeno, *Fuzzy identification of systems and its applications to modeling and control*. IEEE transactions on systems, man, and cybernetics, 1985(1): p. 116-132.

[8]. Ying, H., *Sufficient conditions on uniform approximation of multivariate functions by general Takagi-Sugeno fuzzy systems with linear rule consequent*. IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans, 1998. 28(4): p. 515-520.

[9]. Ying, H., *Sufficient conditions on general fuzzy systems as function approximators*. Automatica, 1994. 30(3): p. 521-525.

[10]. Tanaka, K. and H.O. Wang, *Fuzzy control systems design and analysis: a linear matrix inequality approach*. 2004: John Wiley & Sons.

[11]. Shouraki, S.B., *A novel fuzzy approach to modeling and control and its hardware implementation based on brain functionality and specifications*. 2000.

[12]. Sugeno, M. and T. Yasukawa, *A fuzzy-logic-based approach to qualitative modeling*. IEEE Transactions on fuzzy systems, 1993. 1(1): p. 7.

[13]. Camastra, F., et al., *A fuzzy decision system for genetically modified plant environmental risk assessment using Mamdani inference*. Expert Systems with Applications, 2015. 42(3): p. 1710-1716.

[14]. Khan, D.A. and S. Abbas, *Intelligent Transportation System for Smart-Cities using Fuzzy Logic. Lahore*

*Garrison Univ.* Res. J. Comput. Sci. Inf. Technol, 2018. 2: p. 64-79.

[15]. Rustum, R., et al., *Sustainability ranking of desalination plants using mamdani fuzzy logic inference systems.* Sustainability, 2020. 12(2): p. 631.

[16]. Martinez-Gil, J. and J.M. Chaves-Gonzalez, *Interpretable ontology meta-matching in the biomedical domain using Mamdani fuzzy inference.* Expert Systems with Applications, 2022. 188: p. 116025.

[17]. Georg, S., H. Schulte, and H. Aschemann. *Control-oriented modelling of wind turbines using a Takagi-Sugeno model structure*. in *2012 IEEE International Conference on Fuzzy Systems*. 2012. IEEE.

[18]. Salgado, C.M., et al., *Takagi–Sugeno fuzzy modeling using mixed fuzzy clustering.* IEEE Transactions on Fuzzy Systems, 2016. 25(6): p. 1417-1429.

[19].Elias, L.J., et al., *Stability analysis of Takagi–Sugeno systems using a switched fuzzy Lyapunov function.* Information Sciences, 2021. 543: p. 43-57.

[20]. Chaubey, S. and V. Puig, *Autonomous Vehicle State Estimation and Mapping Using Takagi–Sugeno Modeling Approach.* Sensors, 2022. 22(9): p. 3399.

[21]. Javadian, M., A. Hejazi, and S.H. Klidbary, *Obtaining Fuzzy Membership Function of Clusters With the Memristor Hardware Implementation and On-Chip Learning.* IEEE Transactions on Emerging Topics in Computational Intelligence, 2022. 6(4): p. 1008-1025.

[22]. Javadian, M., et al., *Refining membership degrees obtained from fuzzy C-means by re-fuzzification.* Iranian Journal of Fuzzy Systems, 2020. 17(4): p. 85-104.

[23]. Jokar, E., et al., *Hardware-algorithm co-design of a compressed fuzzy active learning method.* IEEE Transactions on Circuits and Systems I: Regular Papers, 2020. 67(12): p. 4932-4945.

[24]. Klidbary, S.H., S.B. Shouraki, and B. Linares-Barranco, *Digital hardware realization of a novel adaptive ink drop spread operator and its application in modeling and classification and on-chip training.* International Journal of Machine Learning and Cybernetics, 2019. 10: p. 2541-2561.

[25]. Klidbary, S.H., et al. *Outlier robust fuzzy active learning method (ALM)*. in *2017 7th international conference on computer and knowledge engineering (ICCKE)*. 2017. IEEE.

[26]. Merrikh-Bayat, F., S.B. Shouraki, and A. Rohani, *Memristor crossbar-based hardware implementation of the IDS method.* IEEE Transactions on Fuzzy Systems, 2011. 19(6): p. 1083-1096.

[27]. Tikk, D., et al., *Improvements and critique on Sugeno's and Yasukawa's qualitative modeling.* IEEE Transactions on Fuzzy Systems, 2002. 10(5): p. 596-606.

[28]. Hadad, A.H., T.D. Gedeon, and B.S.U. Mendis. *Finding input sub-spaces for polymorphic fuzzy signatures*. in *2009 IEEE International Conference on Fuzzy Systems*. 2009. IEEE.

[29]. Tikk, D., et al. *Implementation details of problems in Sugeno and Yasukawa's qualitative modeling*. in *Research Working Paper RWP-IT-02-2001, School of Information Technology*. 2001.

[30]. Wong, K.W., et al. *Improvement of the cluster searching algorithm in Sugeno and Yasukawa's qualitative modeling approach*. in *Computational Intelligence. Theory and Applications: International Conference, 7th Fuzzy Days Dortmund, Germany, October 1–3, 2001 Proceedings 7*. 2001. Springer.

[31]. Hadad, A.H., et al., *A modification of Sugeno-Yasukawa modeler to improve structure identification phase.* ACSE J, 2006. 6(3): p. 33-40.

[32]. Hadad, A.H., et al. *A modified version of Sugeno-Yasukawa modeler*. in *Advances in Computer Science and Engineering: 13th International CSI Computer Conference, CSICC 2008 Kish Island, Iran, March 9-11, 2008 Revised Selected Papers*. 2009. Springer.

[33]. Hadad, A.H., B.S.U. Mendis, and T.D. Gedeon. *Improvements in Sugeno-Yasukawa modelling algorithm*. in *International Conference on Fuzzy Systems*. 2010. IEEE.

[34]. سجاد حق زاد کلیدبری "ارائه اپراتور جدید جایگزین پخش قطره جوهر در روش یادگیری فعال" مجله مهندسی برق دانشگاه تبریز. ۴۹.۳ ۱۰۵۵-۱۰۶۶(۲۰۱۹).

[35]. کدخدا, اکبرزاده توتونچی, صباحی. (۲۰۲۱). طبقه‌بند همباشی ادراکی مبتنی بر منطق فازی توسعه یافته. مجله مهندسی برق دانشگاه تبریز, ٥٠(٤), ١٧٧٣-١٧٨٤.

[36]. Wang, Y., Z. Pan, and J. Dong, A new two-layer nearest neighbor selection method for kNN classifier. Knowledge-Based Systems, 2023. 235: p. 107604.