2014 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE)
July 6-11, 2014, Beijing, China

Data Driven Fuzzy Membership Function Generation for Increased Understandability

Dumidu Wijayasekara, Milos Manic

Abstract—**Fuzzy Logic Systems (FLS) are a well documented proven method for various applications such as control classification and data mining. The major advantage of FLS is the use of human interpretable linguistic terms and rules. In order to capture the uncertainty inherent to linguistic terms, Fuzzy Membership Functions (MF) are used. Therefore, membership functions are essential for improving the understandability of fuzzy systems. Optimizing FLS for improved accuracy in terms of classification or control can reduce the understandability of fuzzy MFs. Expert knowledge can be used to derive MFs, but it has been shown that this might not be optimal, and acquiring expert knowledge is not trivial. Therefore, this paper presents a data driven method using statistical methods to generate membership functions that describe the data while maintaining the understandability. The presented method calculates key points such as membership function centers, intersections and slopes using data driven statistical methods. Furthermore, the presented method utilizes several understandability metrics to adjust the generated MFs. The presented method was tested on several benchmark datasets and a real-world dataset and was shown to be able to generate MFs that describe the dataset, while maintaining high levels of understandability.**

I. Introduction

FUZZY Logic Systems (FLS) are a well documented and proven method for control, classification, data mining and various other fields [1], [2]. The main advantage of FLS is the utilization human understandable linguistic terms that are capable of capturing uncertainty and vagueness in everyday language [3]. The ability to handle such linguistic terms make FLS attractive as they are highly interpretable and transparent [4]-[7].

However, the understandability of a fuzzy system is heavily dependent on the understandability of its linguistic terms [1], [8]. The linguistic terms are modeled in the data domain via Fuzzy Membership Functions (MFs). Therefore, determination of the linguistic terms, hence the generation of MFs plays a significant role fuzzy system design [1]. The MF should be capable of conveying the knowledge contained in the original data [9], [10].

However, the design of MFs significantly affect the outcome of the FLS [11], [12]. Therefore, many recent work optimize FLS by focusing on the accuracy of the output without considering the understandability of the system [7], [13]-[16]. Such numerical optimization results in highly accurate FLS, however, they pay little attention to the semantical properties of the generated MFs and linguistic terms, thus degrading the understandability of the resulting system [2], [7], [17]. Furthermore, for data mining applications such as linguistic summarization [2], [18], [19] or descriptive rule generation [20], where understandability of data is the goal [19], generating semantically correct MFs is important.

Membership functions derived from expert knowledge are capable of solving understandability issues [12]. However, expert knowledge acquisition can be a difficult task. Experts on the required domain may not be always available, and even when they are available their opinions vary and can be incomplete, varying, and overly precise [4], [17]. Furthermore, due to the large number of dimensions, gathering expert knowledge for highly multi dimensional problems is difficult. Many applications consider pre-defined fuzzy MF [1]. However, this is also sub-optimal as it assumes the data will be distributed in a certain manner and therefore cannot be effectively used to handle specificity of real-world problems [1]. Therefore, the most attractive method of deriving MF is data driven. Furthermore, data driven methods have been shown to accommodate adaptation and self-tuning [4].

Typical MF generation techniques include data histogram based methods, heuristic methods, probability based methods, clustering based methods, neural networks based methods, etc. [21]. However, very little work focuses on the understandability of generated MF [1], [4], [13]-[17]. Furthermore, it has been shown that there is no one optimal way of generating MFs and the optimality depends on the application [21].

Thus, this paper presents a simple, data driven MF generation method that is capable of describing the dataset while maintaining understandability. The presented method utilizes statistical techniques to calculate MF centers, spread, overlap, slope etc. Student's t-test is utilized to identify initial prototypes for MFs. Two different understandability metrics are introduced that measures understandability of MFs. The presented MF generation method utilizes these and several other understandability metrics to generate MFs and thus further increasing the understandability of the MFs. The presented method was tested on several benchmark datasets with known distributions as well as a real-world dataset. The generated MFs were shown to describe the data while maintaining high levels of understandability.

The rest of the paper is organized as follows. Section II describes metrics proposed in the literature for measuring understandability of MFs, and discusses related work. Section III details the presented data driven method for MF generation. Finally, Section IV presents experimental results

Dumidu Wijayasekara is with the Computer Science Dept, University of Idaho, Idaho Falls, ID 83402 USA (phone: 208-533-8127, e-mail: wija2589@vandals.uidaho.edu).

Milos Manic is with the Computer Science Dept, University of Idaho, Idaho Falls, ID 83402 USA (e-mail: misko@ieee.org).

978-1-4799-2072-3/14/$31.00 ©2014 IEEE

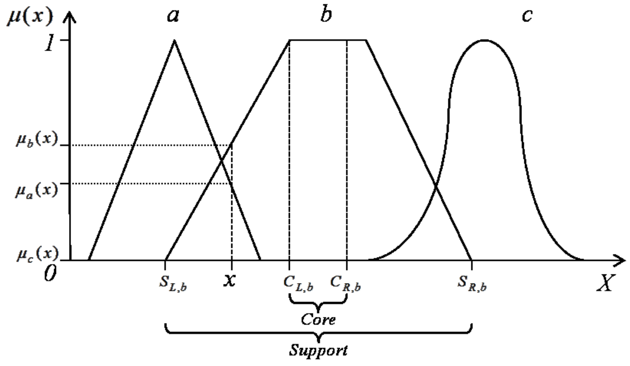


Fig. 1 Typically used Fuzzy Membership Functions (MF)

and Section V concludes the paper.

II. Understandability of Fuzzy Membership Functions

This section first presents metric proposed in literature to calculate understandability, and then discusses related work.

A. Metrics for Understandability in Membership Functions

Fuzzy Membership Functions (MFs) capture the uncertainty and vagueness of everyday linguistic terms [3]. Therefore, finding optimal partition of the input space, the shape of MFs, the coverage of MFs, and the linguistic term associated with the MFs are significant factors affecting the understandability of FLS [1], [10], [16], [22], [23].

Many metrics have been proposed in literature that measure understandability of MFs [8], [10], [16], [17], [24]. However, since understandability or interpretability depends on multiple factors, and understandability is extremely subjective and domain dependant, metrics for measuring true understandability are difficult to define [8], [10].

Fig. 1 shows three commonly used MFs defined for the input dimension *X*; a) triangular, b) trapezoidal and c) Gaussian. For a given input value *x* the membership degree for each MF $\mu_i(x)$, where *i* is the MF, can be calculated (See Fig. 1). The core or the center of the MF is defined as the set of values whose membership degree is 1: $\forall x \in \mu_i(x) = 1$. Similarly, the footprint or the support of the MF is where the membership degree is greater than 0: $\forall x \in \mu_i(x) > 0$. The core and the support of a MF can be expressed as:

$$C_{L,i} = \min(x : \mu_i(x) = 1) \;\; \forall \; x \in X \tag{1}$$

$$S_{L,i} = \max(x : \mu_i(x) = 0) \;\; \forall \; x \in X \tag{2}$$

$$C_{R,i} = \max(x : \mu_i(x) = 1) \;\; \forall \; x \in X \tag{3}$$

$$S_{R,i} = \min(x : \mu_i(x) = 0) \;\; \forall \; x \in X \tag{4}$$

Equations (1) and (3) describe the left and right boundaries of the core, respectively. Equations (2) and (4) describe the left and right footprint of the MF respectively.

A *Normal* MF is defined as a MF that has a core, i.e. the membership degree is 1 for at least one point in the input dataset, and is considered to be a property for increased understandability [9], [8], [17]. For a given MF *i*, *Normality* can be expressed as:

$$norm_i = \max(\mu_i(x)), \forall x \in X \tag{5}$$

where, $norm_i$ is the normality for i^{th} MF. Thus, if a given MF has a core, then normality will be 1.

It is universally agreed that MFs should be monotonic and convex for increased understandability [3], [9], [8], [23], [24]. In [23] this property is referred to as *Unimodality*, which can be expressed as:

$$\begin{aligned} &\forall x_1, x_2, x_3 \in X : x_1 < x_2 < x_3 \rightarrow \\ &\qquad \mu_i(x_2) \geq \min\{\mu_i(x_1), \mu_i(x_3)\} \end{aligned} \tag{6}$$

Dataset *Coverage* is also considered by many as a simple yet important metric for understandability [8], [17], [24]. This metric states that the range of the input dataset should be covered by at least one MF to a certain degree:

$$\forall x \in X, \mu_i(x) > \beta \tag{7}$$

where, $0 < \beta \leq 1$ and can be preset by the user, and $\forall i \in p$ where *p* is the set of MF for input dimension *X*.

Relatively moderate number of MFs for each dimension is also important for understandability [7], [8], [17]. This characteristic is simplified by authors in [7] as *Partition Granularity* by:

$$part = \frac{1}{p-1} \tag{8}$$

where, *p* is the number of MFs for a given dimension, and it is assumed that $p \geq 2$.

For increased understandability, generated MFs should be sufficiently distinct from each other, with limited overlap [8], [17], [24]. This is typically achieved by a threshold value for intersection points:

$$\forall i, j \in p : \mu_i(x) = \mu_j(x) \;\; \mathit{iff} \; \mu_i(x) < \alpha \tag{9}$$

where, $0 < \alpha < 1$ and can be preset by the user, and *p* is the number of MFs.

The measure *Complementarity* [8] is also closely related to the above measure. Complementarity is a property where for a given input value, the sum of all membership degrees is close to 1, and can be expressed as:

$$\forall x \in X : (1-\delta) < \sum_{i=1}^{p} \mu_i(x) < (1+\delta) \tag{10}$$

where, $\delta \approx 0$ and can be preset by the user

Another measure used to identify the distinctness of MFs is the property of *Separation*. This property states that the cores of adjacent MFs must be separated by at least η :

$$\forall i \in p;\ \left|C_{R,i} - C_{L,j}\right| \geq \eta; \forall j \neq i \in p \tag{11}$$

Symmetry of generated MF is also considered as a relative measure of understandability as it reflects of universal duality and natural relativity of terms [8], [25]. In this paper, the measurement, *Relative Dissymmetry* is proposed to measure symmetry in MFs and is formalized as:

$$diss_i = \left|\sum_{x=S_{L,i}}^{C_{L,i}} \mu_i(x) - \sum_{x=C_{R,i}}^{S_{R,i}} \mu_i(x)\right| \tag{12}$$

where, *diss_i* is the dissymmetry measure for i^{th} MF, and the dissymmetry measure can be normalized for a given MF as:

$$\overline{diss_i} = \frac{diss_i}{\max\left(\sum_{x=S_{L,i}}^{C_{L,i}} \mu_i(x),\ \sum_{x=C_{R,i}}^{S_{R,i}} \mu_i(x)\right)} \tag{13}$$

The dissymmetry measure cannot be calculated for shoulder MFs that describe extremes (minimum and maximum) of data.

A threshold can also be set for the membership degree of a MF inside the core of another MF [8], [16]. This property can be formalized as:

$$\forall x \in [C_{L,i}, C_{R,i}],\ \mu_j(x) < \gamma\ \forall j \neq i \in p \tag{14}$$

where, $0 \leq \gamma \leq 1$ and can be preset by the user. Typically γ is set to zero [16], meaning membership degree of other MFs is zero within the core of a given MF.

Finally, in this paper, a measurement is introduced that measures the compliance of a MF to the data explained by it. This measure is called *Compliance* and can be measured by utilizing a normalized histogram [26] containing *n* bins. The normalized degree of compliance for i^{th} MF $\overline{comp_i}$ can be formalized as:

$$\overline{comp_i} = 1 - \frac{uncomp_i}{\sum_{x=S_{L,i}}^{S_{R,i}} \widehat{h}(x)} \tag{15}$$

where, $\widehat{h}(x)$ is the maximum possible value of the histogram for the input value *x*, which in the case of the normalized histogram is 1. And,

$$uncomp_i = \sum_{x=S_{L,i}}^{S_{R,i}} \left|h(x) - \mu_i(x)\right| \tag{16}$$

where, *h(x)* is the height of the normalized histogram for value *x*.

The normalized degree of compliance is 1 when the MF fully complies with the data and decreases as the MF does not comply with the data.

B. Related Work

Several authors have proposed methods for generating understandable MFs in recent years [8]. Some have used several of the above mentioned understandability (interpretability) measures. Typically the proposed methods used can be separated into three main categories: cluster based methods, evolutionary methods, and combined methods with pre-set MFs.

A constrained Fuzzy C-Means (FCM) based method is used in [3] generating more understandable MFs. Similarly, a modified FCM based method is used in [23]. The authors utilize clustering to generate MFs and then combine similar MFs for increased understandability in [4] and [27]. Clustering and cluster distances are used to derive understandable MFs in [24]. However, clustering based methods has several disadvantages as well as advantages [3].

In [1] the authors use hedge algebra based semantics to assign linguistic terms to information granules and utilize simulated annealing to optimize MFs. Symmetrical MFs are generated using evolutionary algorithms in [25]. In [16] and [17] evolutionary algorithm based approaches are proposed that make use of understandability metrics.

The authors use preset MFs, and tune these using lateral movements in [28]. In [14] and [15] the authors propose an algorithm that utilizes pre-shaped MFs along with Fuzzy C-Means (FCM) clustering to generate transparent MFs. Similar method that also utilizes evolutionary algorithms is proposed in [13] and [29].

Pre-set MFs require the use of experts and may be sub-optimal as mentioned in Section I. The primary drawback of clustering and evolutionary algorithm based methods is the increased computational complexity.

In contrast, the presented method utilizes a deterministic, statistical approach to identify the optimal parameters for MFs. Furthermore, the presented method utilizes several understandability metrics to derive and fine-tune these parameters. This ensures that the understandability of generated MFs is maintained, while describing the data distribution.

III. Data Driven Method For Generating Understandable Membership Functions

The presented data driven, statistics based MF generation method is a 5 step process: **Step1:** Generate initial prototypes, **Setp2:** Refine generated prototypes, **Step3:** Generate initial MFs, **Step4:** Remove unwanted MFs, **Step5:** Refine generated MFs. Each step is focused on increasing the understandability of the MFs while maintaining the ability to describe the data properly. Prior to the MF generation process the dataset is normalized between 0 and 1. Detailed descriptions of each step are given below.

Step1: In the first step initial prototypes for MFs are generated. The first prototypes are the sample mean $\overline{X_{data}}$, minimum, min_{data} and maximum, max_{data} values of the dataset. The Students' t-test for unequal sample size and unequal variance [30] was used to identify portions of the data that are significantly different from the mean, to generate secondary prototypes. The Students' t-test for two mean values $\overline{x_1}$ and $\overline{x_2}$ can be expressed as:

$$t(\overline{x_1},\overline{x_2}) = \frac{\overline{x_1}-\overline{x_2}}{\sqrt{\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}}} \tag{17}$$

where, s and n are standard deviation and sample size of each sample, respectively. If the t value is greater than the *critical t* value then the null hypothesis is rejected, meaning the two sample means are significantly different from each other.

Using the Student's t-test two prototypes to the left and two prototypes to the right of the sample mean are generated. These are formalized as:

$$\begin{aligned} PB_L = \max(x) &: \\ 0 < x < \overline{X_{dara}}&, \\ t(\overline{x_{0,x}},\overline{X_{dara}}) &> t_{Critical} \end{aligned} \tag{18}$$

$$PM_L = \overline{x_{0,PB_L}} \tag{19}$$

Similarly,

$$\begin{aligned} PB_R = \min(x) &: \\ \overline{X_{dara}} < x < &\ 1 \\ t(\overline{x_{x,1}},\overline{X_{dara}}) &> t_{Critical} \end{aligned} \tag{20}$$

$$PM_R = \overline{x_{PB_R,1}} \tag{21}$$

where, PB_L and PM_L are two secondary prototypes to the left of the sample mean and PB_R and PM_R are two prototypes to the right of the sample mean. These prototypes signify portions of the data that are significantly different from the initial prototype (sample mean).

This process is iterated to the left of PB_L and to the right of PB_R using PM_L and PM_R as the initial prototypes, until the newly generated prototypes surpasses the minimum and maximum values.

Step2: In this step the generated initial prototypes are refined. First similar prototypes are combined. Prototypes within ε range of each other are averaged:

$$\forall i \neq j \in k, P_n = {P_i + P_j}\big/_{2} \Rightarrow |P_i - P_j| < \varepsilon \tag{22}$$

where k is the set of generated prototypes, P_i and P_j are prototypes P_n is the new prototype and ε is a preset constant, and $|\ |$ denotes the absolute value. Once this is done, P_i and P_j are removed from k and P_n is added to k.

Secondly, prototypes within ε of the minimum and maximum values are removed, retaining only the minimum and maximum values. This is done because it is generally accepted that data extremes must be prototypes of some MFs [24].

$$P_i = min_{data} \Rightarrow |P_i - min_{data}| < \varepsilon \tag{23}$$

$$P_i = max_{data} \Rightarrow |P_i - max_{data}| < \varepsilon \tag{24}$$

Thirdly, the remaining prototypes are grouped to satisfy the separation property described in (11):

$$\forall i \in k, |P_i - P_{i+1}| < \eta \Rightarrow G_m = \{P_i, P_{i+1}\} \tag{25}$$

$$\forall i \in k, |P_i - P_{i+1}| \geq \eta \Rightarrow G_m = \{P_i\}, G_{m+1}\{P_{i+1}\} \tag{26}$$

where, G is a set of prototypes.

By performing the grouping, portions of the data that are separable, but leads to higher granularity, are combined. During grouping if a group contains more than three prototypes, prototype that is closest to the left or right boundary of the group is deleted. This grouping process is performed until all prototypes satisfy the (25) and (26).

Step3: Once the prototypes are refined, the initial MF are generated. For each of the generated groups G a MF is generated. The cores of the MFs are defined using the minimum and maxim prototypes of a group G:

$$C_{L,i} = \min(P_j) \forall P_j \in G_i; i = 1,2,\ldots,M \tag{27}$$

$$C_{R,i} = \max(P_j) \forall P_j \in G_i; i = 1,2,\ldots,M \tag{28}$$

where, i is the generated MF and M is the set of generated groups in Step2. The support of the i^{th} MF is defined as:

$$S_{L,i} = \max(P_j) \forall P_j \in G_k : P_j < C_{L,i}; k = 1,.,M \tag{29}$$

$$S_{R,i} = \min(P_j) \forall P_j \in G_k : P_j > C_{R,i}; k = 1,.,M \tag{30}$$

Thus, the left support of the i^{th} MF $S_{L,i}$ is the prototype to the immediate left of the left core $C_{L,i}$, and similarly the right support $S_{R,i}$ is the prototype to the immediate right of the right core $C_{R,i}$.

Step4: In this step the some of the generated MF are removed or combined to decrease granularity and increase understanding. Average understandability is defined for the fuzzy system which is used to identify MFs that will be removed or combined. The average understandability, $\overline{AU}$ is defined in terms of normalized dissymmetry (13) $\overline{diss_i}$ and normalized degree of compliance (15) $\overline{comp_i}$, because

TABLE I
PRESET CONSTRAINTS

Description	Symbol	Value
Dataset coverage	β	0.2
Maximum overlap point for two MF	α	0.7
Complementarity	δ	0.15
Core Separation	η	0.15
MF inside the core of another MF	γ	0
Hard threshold for the number of MF	Ω	5
Soft threshold for the number of MF	ω	3

other metrics are already fulfilled:

$$AU = \left(\frac{\sum_{i=1}^{M} \overline{diss_i}}{M} + \frac{\sum_{i=1}^{M} \overline{comp_i}}{M} \right) \div 2 \tag{31}$$

where M is the number of MF in the system, and AU is the averaged understandability.

When a MF is deleted, the prototypes that were used to generate the core of that MF are also deleted. For certain occasions detailed below, MFs are combined. This is done by creating a new MF by defining the core as the leftmost and rightmost prototypes of the MFs that are combined:

$$C_{L,n} = \min(P_k) \forall P_k \in G_i, G_j \tag{32}$$

$$C_{R,n} = \max(P_k) \forall P_k \in G_i, G_j \tag{33}$$

where, i and j are the MFs that are being combined and n is the new MF that is generated.

The MFs generated using the minimum and maximum prototypes of data are not deleted [24]. The remaining MFs are deleted or removed as follows: remove MFs that increases AU while the number of MF is greater than ω. If the number of MFs is still greater than Ω then identify MF that reduces AU the least, and combine it with the closest MF using (32) and (33), until the number of MF is less than or equal to Ω. The constants ω and Ω are preset such that $2 \le \omega \le \Omega$. This ensures that the generated fuzzy system contains at least ω MFs and less than or equal to Ω MFs.

Step5: Finally, the spread of the remaining MFs are adjusted to fulfill criteria (7), (9), (10) and (14). This is achieved using the same method in Step3 using (29) and (30).

IV. EXPERIMENTAL RESULTS

The presented method was tested on several benchmark datasets with known distributions and a real world dataset. In order to verify the understandability of the generated MF, the compliance of each MF to the understandability metrics presented in Section II was used.

In [5] Takagi and Sugeno stated that in order to claim the validity of a generated fuzzy system and the linguistic terms, the MFs must remain same in the presence of noise. Thus, in order to evaluate the validity of the presented system, different levels of noise were introduced to the dataset and MFs was generated. These MFs were then compared to the MFs generated without noise by using the following measure:

TABLE II
SIMILARITY OF GENERATED FUZZY SYSTEMS AT DIFFERENT LEVELS OF NOISE

Dataset	Noise Level (SNR)				
	20dB	14dB	12.5dB	11dB	10dB
Benchmark Uniform	1	0.99	0.97	0.98	0.85
Benchmark Normal	0.99	0.97	0.9	0.87	0.82
Benchmark Bivariate	0.98	0.94	0.91	0.86	0.8
Benchmark Skewed	0.99	0.91	0.85	0.81	0.73
Real-world Zone1	0.9	0.9	0.87	0.8	0.81
Real-world Zone2	0.99	0.93	0.88	0.85	0.8
Real-world Zone3	1	0.92	0.88	0.79	0.79

$$sim_{i,j} = \frac{\sum_x \min(\mu_i(x), \mu_j(x))}{\sum_x \max(\mu_i(x), \mu_j(x))} \tag{34}$$

where, i and j are the MF that are being compared, and i is from the set of original MF and j is from the set of MF created with noisy data. The similarity value, *sim* is 1 when i and j are completely overlapped and 0 when there is no overlap.

The similarity of the fuzzy system was calculated using:

$$sys_{ori,noise} = \frac{\sum_{j=i}^{p_{noise}} \max(sim_{i,j}) \forall i \in p_{ori}}{p_{noise}} \tag{35}$$

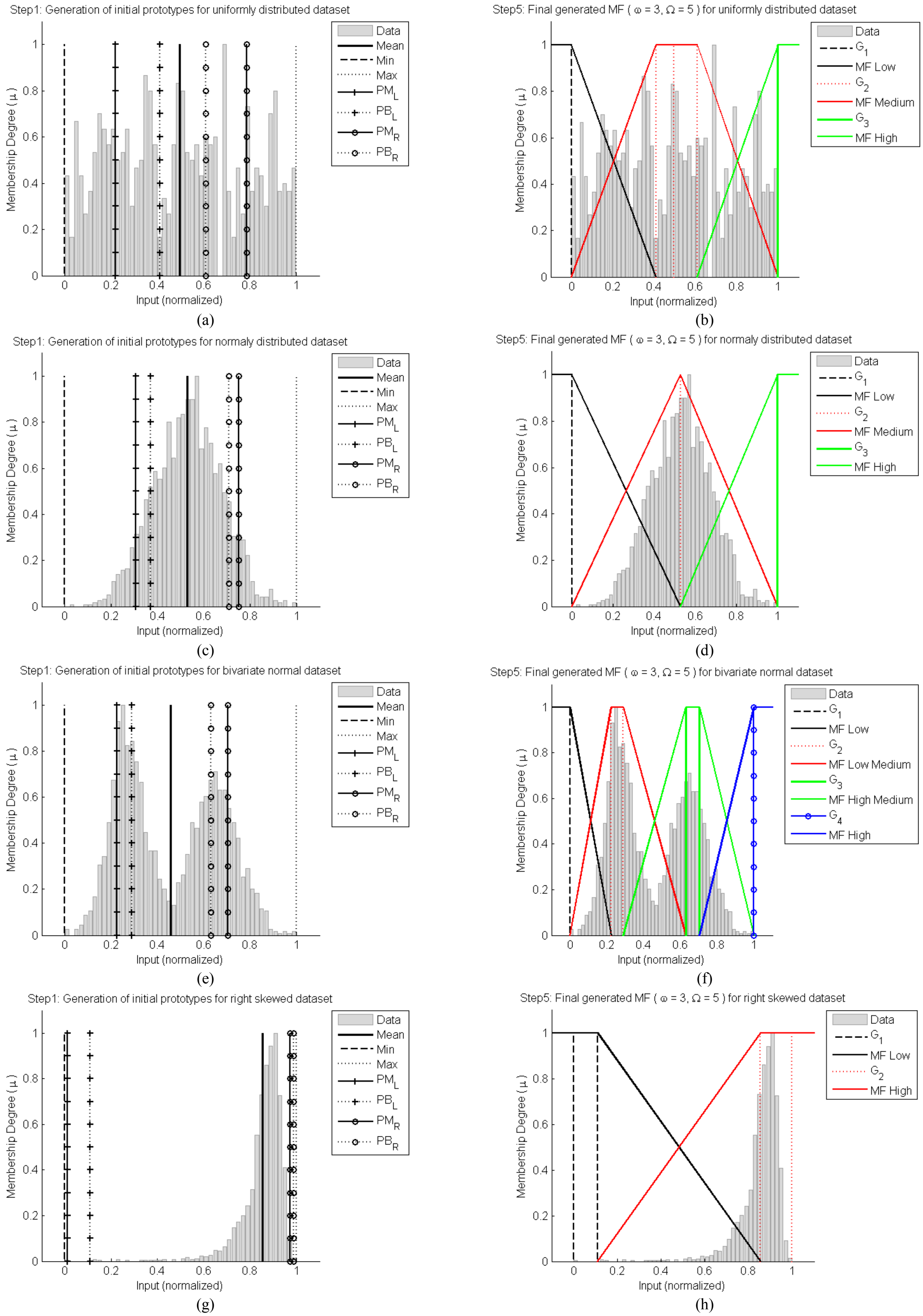
where, *ori* is set of MF generated using the original data and *noise* is the set of MF generated using the noisy data. p_{ori} is the number of MF in *ori* and p_{noise} is the number of MF in *noise*.

Noise from Signal-to-Noise Ratio (SNR) 20dB to SNR 10dB was introduced to the original data to generate the noisy data, where SNR is defined as:

$$SNR_{dB} = 10\log_{10}\left(\frac{A_{signal}}{A_{noise}}\right) \tag{36}$$

where, A_{signal} and A_{noise} are amplitude of the input and amplitude of the noise, respectively.

The preset values as tested for the understandability constraints presented in Section II are shown in Table I. Using these preset values four different benchmark datasets were tested. Each dataset was generated using a random number generator to follow a known distribution and contained 2000 data points. The tested known distributions were: uniform distribution, normal distribution, bivariate normal distribution, and right skewed distribution. Fig. 2 shows the initially generated prototypes and final MF for each benchmark dataset.



(a)

(b)

(c)

(d)

(e)

(f)

(g)

(h)

Fig. 2 Initial prototypes for (a) uniformly distributed data, (c) normally distributed data, (e) bivariate normal data, and (g) right skewed data, along with generated MF for, (b) uniformly distributed data, (d) normally distributed data, (f) bivariate normal data, and (h) right skewed data.

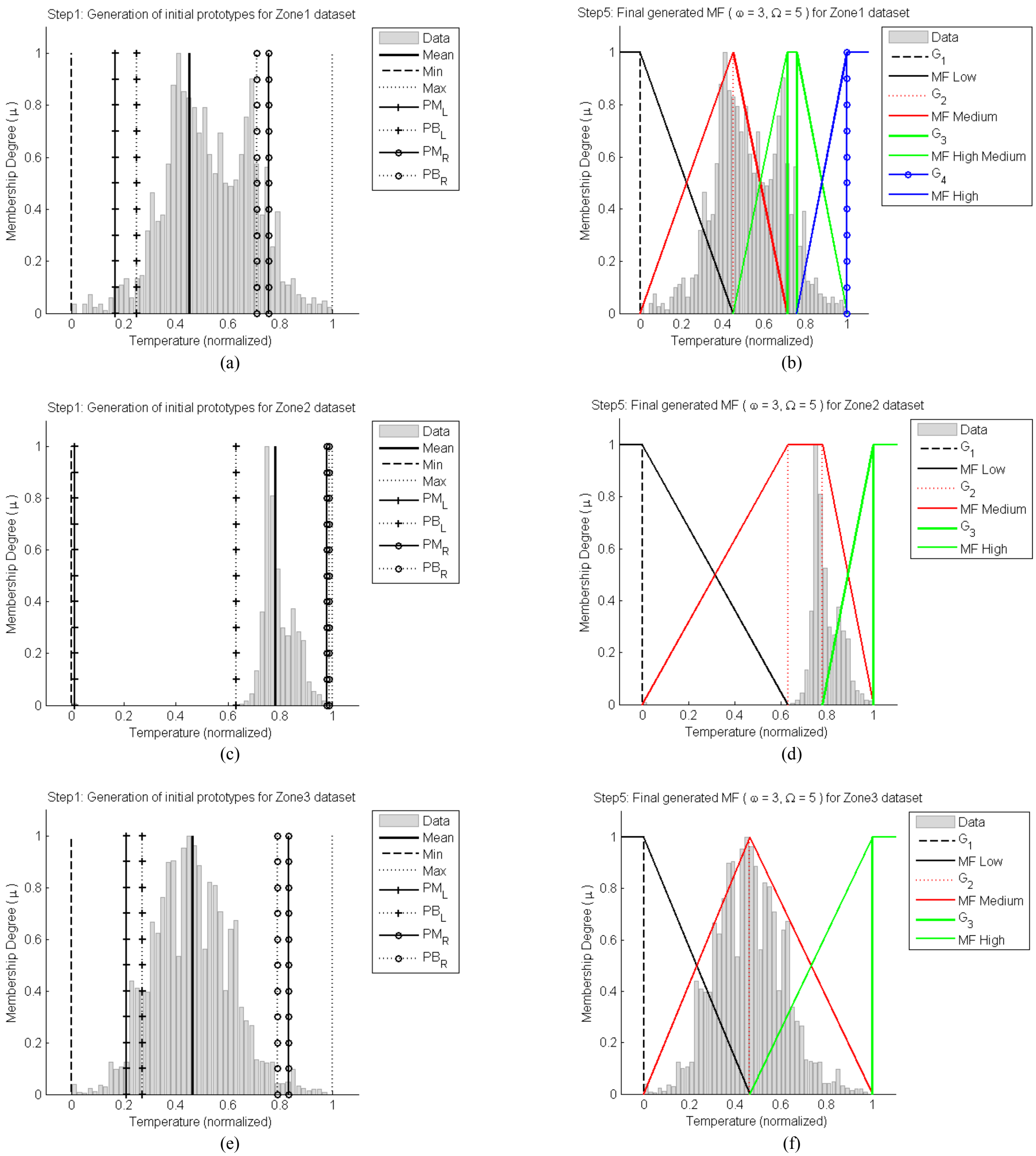


Fig. 3 Initial prototypes for real world zone temperature data - (a), (c) and (e), along with generated MF - (b), (d) and, (f)

For further validation, the presented method was tested on a real-world dataset obtained form an office building containing temperature of 3 zones. These zones were selected for their varying distributions of data. The dataset contained temperature values collected every 30 minutes for a period of a month. The initial prototypes and the generated MFs for the real world dataset are shown in Fig. 3.

For all cases tested, all understandability constraints given in Section II were met. This is because the presented method ensures each of the understandability constraints is met during the MF generation process.

The similarity of the generated fuzzy systems for different noise levels is shown in Table II. Even for high noise levels, the generated fuzzy systems are close to the ones generated for the original dataset, thus confirming that the generated fuzzy systems are valid and describe the data consistently.

V. Conclusion

This paper presented a novel, data driven, statistical method for generating understandable Fuzzy Membership Functions (MFs) that describe data. The presented method uses classical statistical methods to identify initial MF prototypes. Two understandability measures were introduced. These, along with several other understandability metrics were used to generate and fine-tune MFs for increased understandability.

The presented method was tested on several benchmark datasets with known distributions as well as real-world datasets. The presented method was shown to produce meaningful MFs that describe the data while maintaining high degree of understandability.

As future work, the presented method will be compared to other methods for understandable MF generation. The presented method will also be expanded further using more advanced statistical methods and will extended to accommodate the generation of Gaussian or Bezier MFs. Furthermore, the presented method can also be extended to facilitate the generation of type-2 MFs. Hybrid of classical methods and the presented method can be used to improve accuracy in classification and control system FLS while maintaining high levels of understandability.

References

[1] C. H. Nguyen, W. Pedrycz, T. L. Duong, T. S. Tran, “A genetic design of linguistic terms for fuzzy rule based classifiers,” in *Int. Journal of Approximate Reasoning*, vol. 54, no. 1, pp. 1-21, Jan. 2013.

[2] A. Wilbik, J. M. Keller, J. C. Bezdek, “Generation of prototypes from sets of linguistic summaries,” in *Proc. IEEE Int. Conf .on Fuzzy Systems*, *FUZZ-IEEE*, pp 1-8, Jun. 2012.

[3] T. W. Liao, A. K. Celmins, R. J. Hammell II, “A fuzzy c-means variant for the generation of fuzzy term sets,” in *Fuzzy Sets and Systems*, vol. 135, no. 2, pp. 241-257, Apr. 2003.

[4] M. Setnes, R. Babuska, H. B. Verbruggen, “Transparent Fuzzy Modelling,” in *Int. Journal of Human-Computer Studies*, vol. 49, no. 2, pp. 159-179, Aug. 1998.

[5] H. Takagi, M. Sugeno, “Fuzzy identification of systems and its applications to modeling and control,” in *IEEE Trans. On Systems, Man. and Cybernetics*, vol. 15, no. 1, pp. 116-132, Jan. 1985.

[6] W. Duch, R. Adamczak, K. Grabczewski, “A new methodology of extraction, optimization and application of crisp and fuzzy logical rules,” in *IEEE Trans. on Neural Networks*, vol. 12, no. 2, pp. 277-306, Mar. 2001.

[7] D. D. Nauck, “Measuring interpretability in rule-based classification systems,” in *Proc. IEEE Int. Conf. on Fuzzy Systems, FUZZ-IEEE*. vol. 1, pp. 196-201, May 2003.

[8] M. J. Gacto, R. Alcalá, F. Herrera, “Interpretability of linguistic fuzzy rule-based systems: an overview of interpretability measures,” in *Information Sciences*, vol. 181, pp. 4340–4360. 2011.

[9] A. L. Medaglia, S. C. Fang, Henry L.W. Nuttle, J. R. Wilson, “An efficient and flexible mechanism for constructing membership functions,” in *European Journal of Operational Research*, vol. 139, no. 1, pp. 84-95, May 2002.

[10] O. Cordón, “A historical review of evolutionary learning methods for Mamdani-type fuzzy rule-based systems: designing interpretable genetic fuzzy systems,” in *Int. Journal of Approximate Reasoning*, vol. 52. pp. 894–913, 2011.

[11] I. Rojas, H. Pomares, J. Ortega, A. Prieto, “Self-organized fuzzy system generation from training examples,” in *IEEE Trans. on Fuzzy Systems*, vol. 8, no. 1, pp. 23-36, Feb. 2000.

[12] M. Mottaghi-Kashtiban, A. Khoei, Kh. Hadidi, “Optimization of rational-powered membership functions using extended Kalman filter,” in *Fuzzy Sets and Systems*, vol. 159, no. 23, pp. 3232-3244, Dec. 2008.

[13] L. Chen, C. L. Philip Chen, “Transparent linguistic interface generation and its application in fuzzy decision trees,” in *Proc. IEEE Int. Conf. on Systems, Man. and Cybernetics*, pp. 1337-1342, 2008.

[14] L. Chen, C. L. Philip Chen, “Pre-shaped fuzzy c-means algorithm (pfcm) for transparent membership function generation,” in *Proc. IEEE Int. Conf. Systems, Man. and Cybernetics*, pp. 789-794, 2007.

[15] L. Chen, C. L. Philip Chen, “Gradient pre-shaped fuzzy C-means algorithm (GradPFCM) for transparent membership function generation,” in *Proc. IEEE Int. Conf .on Fuzzy Systems*, *FUZZ-IEEE*, pp. 428-433, 2008.

[16] A. Riid, E. Rüstern, “Transparent fuzzy systems and modeling with transparency protection,” in *Proc. IFAC Symp. on Artificial Intelligence in Real Time Control*, pp. 229-234, 2000.

[17] J. Valente de Oliveira, “Semantic Constraints for Membership Function Approximation,” in *IEEE Trans. on Systems, Man. and Cybernetics, part A*, vol. 29, no. 1, pp. 128-138, 1999.

[18] C. Martinez-Cruz, D. Sanchez, G. Trivino, “Three main components of experience base in linguistic description of data,” in *Proc. IEEE Int. Conf. on Fuzzy Systems FUZZ-IEEE*, pp. 1-6, Jul. 2013.

[19] D. Wu, J. M. Mendel, J. Joo, “Linguistic Summarization Using IF-THEN Rules,” in *Proc. IEEE Int. Conf. on Fuzzy Systems FUZZ-IEEE*, pp. 1 - 8, Jul. 2010.

[20] D. Wijayasekara, M. Manic, “Visual, Linguistic Data Mining Using Self-Organizing Maps,” in *Proc. of Intl. Joint Conference on Neural Networks, IJCNN*, Jun. 2012.

[21] S. Medasani, J. Kim, R. Krishnapuram, “An overview of membership function generation techniques for pattern recognition,” in *Int. Journal of Approximate Reasoning*, vol. 19, pp. 391-417, 1998.

[22] M. Sugeno, G. T. Kang, “Structure identification of fuzzy model,” in *Fuzzy Sets and Systems*, vol. 28, no. 1, pp 13-33, Oct. 1988.

[23] F. Hoppner, F. Klawonn, “A new approach to fuzzy partitioning,” in *Proc. 20th NAFIPS Int. Conf.*, vol. 3, pp. 1419-1424, Jul. 2001.

[24] G. Castellano, A. M. Fanelli, C. Mencar, “Design of transparent Mamdani fuzzy inference systems,” in *Proc of HIS 2003*, pp. 468-477, 2003.

[25] C. H. Chang; Y. C. Wu, “The genetic algorithm based tuning method for symmetric membership functions of fuzzy logic control systems,” in *Proc. IEEE Conf. on Industrial Automation and Control: Emerging Technologies*, pp. 421-428, May 1995.

[26] L. Chengyu, L. Peng, L. Zhao, J. Yang, L. Liping, “Evaluation method for heart failure using RR sequence normalized histogram,” in *Computing in Cardiology*, pp. 305-308, Sept. 2011.

[27] M. Y. Chen, D. A. Linkens, “Rule-base self-generation and simplification for data-driven fuzzy models,” in *Proc. IEEE Int. Conf. on Fuzzy Systems, FUZZ-IEEE*, vol. 1, pp. 424-427, 2001.

[28] R. Alcalá, Y. Nojima, F. Herrera, H. Ishibuchi, “Multi-objective genetic fuzzy rule selection of single granularity-based fuzzy classification rules and its interaction with the lateral tuning of membership functions,” in *Soft Computing*, vol. 15, pp. 2303–2318, 2011.

[29] L. Chen, W. Pedrycz, C. L. P. Chen, “Computational intelligence techniques for building transparent construction performance models,” in *Proc. IEEE Int. Conf. on Systems, Man. and Cybernetics*, pp. 1166-1171, 2006.

[30] C. A. Markowski, E. P. Markowski, “Conditions for the Effectiveness of a Preliminary Test of Variance,” in *The American Statistician*, vol. 44, no. 4, pp. 322–326, 1990.