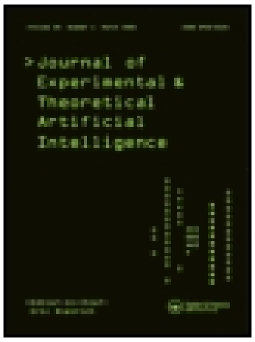

Journal of Experimental & Theoretical Artificial Intelligence

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/teta20

Enhanced type-2 Wang-Mendel Approach

Prashant K. Gupta & Javier Andreu-Perez

To cite this article: Prashant K. Gupta & Javier Andreu-Perez (2022): Enhanced type-2 Wang-Mendel Approach, Journal of Experimental & Theoretical Artificial Intelligence, DOI: 10.1080/0952813X.2022.2135614

To link to this article: https://doi.org/10.1080/0952813X.2022.2135614

Published online: 12 Nov 2022.

JOURNAL OF EXPERIMENTAL & THEORETICAL ARTIFICIAL INTELLIGENCE
https://doi.org/10.1080/0952813X.2022.2135614

ARTICLE

Enhanced type-2 Wang-Mendel Approach

Prashant K. Gupta[a,c] and Javier Andreu-Perez[b,c,d]

[a]Intelligent Information Systems Group, German Research Center for Artificial Intelligence (DFKI) GmbH, Saarbrucken, Germany; [b]Centre for Computational Intelligence, School of Computer Science and Electronic Engineering, University of Essex, Colchester, UK; [c]Institute for Advancing Artificial Intelligence, Colchester, UK; [d]Simbad[2], University of Jaen, Jaen, Spain

ABSTRACT

The Wang-Mendel Approach (WMA) focuses on combining the numerical as well as linguistic information for achieving greater explainability for inference models. The standard WMA models the linguistic information using type-1 (T1) fuzzy sets (FSs), which have a reduced capability to model the semantics of linguistic information. Therefore, we propose a novel Enhanced WMA, which models the linguistic information using the type-2 (T2) FSs. Further, our Enhanced T2 FS-based WMA can be modified to reflect the use of interval type-2 (IT2) FSs, for modelling linguistic uncertainty. IT2 FSs are suitable when better uncertainty handling capabilities are required compared to T1 FSs, however, at a computational cost lesser than the T2 FSs. Performance of Enhanced WMA is demonstrated through a real-world crop-yield prediction problem in smart agriculture and an additional exemplar application on users' satisfaction ratings. Further, we have compared our approach with the performance obtained from the T1 FS-based WMA and the original estimations given in the original data. We found that our Enhanced WMA provides more precise estimates than the other two with 95% confidence level. To the best of our knowledge, this is the first proposal of a T2 FSs method for enhancing the modeling of linguistic uncertainty in the WMA.

ARTICLE HISTORY
Received 8 November 2021
Accepted 9 October 2022

KEYWORDS
Explainable artificial intelligence; Interval type-2 fuzzy sets; Wang-Mendel Approach

Introduction

The recent surge in the computational intelligence provided by sub-symbolic approaches has been mainly due to unprecedented hardware innovations. However, more often than not, they are classified as `black-boxes' (Van Harmelen & Teije, 2019), as they hide the inner details from human beings. Consequently, to have an enriched user experience, there is an increasing trend towards the development of Fuzzy Sets (FSs)[1] based explainable artificial intelligence (XAI) systems (Gupta & Andreu-Perez, 2022; Gupta et al. 2021; Hagras 2018; Restović 2020), which can generate user understandable recommendations. These FSs-based explainable systems derive their description strength by defining the problem domain variables in linguistic terms.

The FSs-based explainable fuzzy systems, however, suffer from two challenges. Firstly, the fuzzy membership functions (MFs) for the linguistic variables are generally designed by an expert. The sub-symbolic systems, on the other hand, use numeric data, which can be easily collected in the various day-to-day industrial processes, in the form of numerical data pairs amongst the system input-output variables. The second limitation is that the MFs are modelled using type-1 (T1) FSs, which have a reduced capability to model the semantics of linguistic information[2] (Zadeh, 1965).

CONTACT Javier Andreu-Perez j.andreu-perez@essex.ac.uk Centre for Computational Intelligence, School of Computer Science and Electronic Engineering, University of Essex, Wivenhoe Park, Colchester CO4 3SQ, UK

We found that the solution to the first limitation is easily provided by an existing data processing approach, called the Wang Mendel Approach (WMA), proposed in Wang and Mendel (1992). The WMA generates soft IF-THEN fuzzy rule-based representations by constructing these rules from numeric data values. It focuses on combining the numerical as well as linguistic information. However, the WMA approach fails to solve the second limitation as it also represents the semantics of linguistic information using T1 FSs.

Thus, the we propose an Enhanced WMA, where the semantics of the linguistic information are modelled using the type-2(T2) FSs. Further, the interval type-2(IT2) FSs are a viable tool for modelling linguistic uncertainty in scenarios that require greater uncertainty modelling than the T1 FSs. But such systems refrain from expenditure on the computational complexity of T2 FSs. Therefore, we have also shown that by minimum modifications in the Enhanced WMA, we can easily use the IT2 FSs for modelling linguistic uncertainty in the Enhanced WMA. Hence, the WMA based on T2 or IT2 FSs is called by the same name viz., Enhanced WMA.

Further, we have also demonstrated the use of Enhanced WMA for crop yield prediction in smart agriculture, using a real-life agritech dataset. Smart agriculture is an active and popular application area in which explainable systems, which can generate linguistic output, are hard to find. The main driving force behind the smart agriculture is the industry 4.0 (Ustundag & Cevikcan, 2017), which has been the reason for automation in various industrial or real-life processes in recent years. This, coupled with the Internet of Things (IoT), is responsible for the use of sensors for collecting data/information about various industrial (or day-to-day) measurements. IoT enables the collection of information in a fast, efficient and easy manner (Rekha et al., 2017). This information can then be utilised for decision making in various application areas, a popular one amongst them being smart agriculture. Numerous works have been published that touch upon various aspects of smart agriculture like smart irrigation (Krishnan et al., 2020), quantifying yield prediction for seasonal workers planning (Amaruchkul, 2021), etc.

We have also compared the crop yield predictions obtained using the Enhanced WMA to those obtained with the existing T1 FS based WMA as well as the ones given in the original dataset. We have found that our Enhanced WMA achieves better precision than the other two using the confidence interval testing at 95%. Therefore, the contributions of the work can be summarised as follows:

- Proposing a novel Enhanced WMA based on T2 FSs.
- Showing how by minimum modifications, the T2 FSs-based Enhanced WMA, can be adapted for use with IT2 FSs.
- Illustrating the application of proposed Enhanced WMA for crop yield prediction in smart agriculture.
- Comparing the results of Enhanced WMA to the existing T1 FS based WMA to establish the supremacy of our proposed approach.

The rest of the paper is organised as follows: in Section 2, we present the related literature, in Section 3 we present the mathematical concepts which are required for understanding the work presented in this paper, in Section 4 we present the details of our Enhanced WMA, in Section 5 we demonstrate the working of our Enhanced WMA using a data vector from a stream of data values pertaining to the real-life agritech dataset of smart agriculture, in Section 6, we present the crop yield prediction results obtained from the applicability of our Enhanced WMA to the test dataset and compare its performance to the existing T1 FS-based WMA as well as the original dataset, along with discussions on important findings obtained from this comparison. Finally, we conclude the present work as well as discuss its future scope in the Section 7.

Related work

In this section, we present the literary works that form the base of the present work. Regarding the methodology, WMA (Wang & Mendel, 1992) is a useful method to learn fuzzy rule-based systems directly from data without prior knowledge. It is an efficient method known for its simplicity, easy implementation, not requiring an iterative learning process, quick to converge, and good performance. This method's crux is to divide the input data into fuzzy sets that are later grouped into candidate rules, which importance degree is assessed and finally grouped into the final Rule Base. However, WM method has been criticised for having a significant dependence on the input data, (Díaz-Pacheco & Reyes-Garcia, 2021; Gou et al., 2015; Hao & Mendel, 2013; Yang et al., 2010), returning poor performance if the data contains high uncertainty. To solve this particular issue, WMA has been fitted with fuzzy clustering, or meta-heuristic optimisation (Gou et al. (2015); Hao et al. (Hao & Mendel, 2013; Yang et al., 2010) at the expense of the simplicity of the original method. We suggest enhancing WM with a Type-2 fuzzy treatment to solve this issue, rendering a new WM method that can cope with higher uncertainty without compromising its original simplicity.

Navarro-Almanza et al., (2022) propose the use of a fuzzy linguistic interpretable model through a neuro-fuzzy system for extracting rules. In Yang et al., (2022), the authors have designed a rule-based system and efficient rule-based modelling and inference procedures. In Jara et al., (2022), authors have proposed a new method to design a fuzzy rule inference method which is fast as well as works when the collection has a large number of fuzzy rules. They have analysed rules in the neighbourhood of a given example instead of going through the complete collection in a sequential manner. In Serrano-Guerrero et al., (2021), authors have presented an extensive study on the use of fuzzy logic for opinion mining. Lucchese et al., (2021) apply the fuzzy inference system (FIS) and artificial neural network (ANN) to the application of landslide susceptibility mapping. The obtained results are compared to determine the potential of both the techniques as well as extract physical relationships from data. Alonso Moral et al. (2021) advocate the use of fuzzy if-then rules for the design of more interpretable systems. Zhai et al., (2021) present the application of WMA for constructing rules from the raw data, in the continuous production process with dynamic and nonlinear characteristics. (Wang et al., (2021) have presented an offline fuzzy logic-based method and online operations for measuring the boiler efficiency. The work of Čubranić-Dobrodolac et al., (2021) proposes the use of a FIS for predicting the driver's propensity to commit a road accident. In Gao et al., (2021), an improved belief rule-based (BRB) system is proposed which is efficient in terms of generating a lesser number of rules than the existing BRB methodology, as well as the learning of the BRB, is made faster by use of selection and reduction strategy. Authors have presented a fuzzy-based framework for the expert recommender system for module advising in Alhabashneh, (2021). However, none of these works says anything about the use of higher order FSs for modeling the semantics of linguistic information in a better manner.

Regarding the application domain, agriculture yield prediction is an important factor in smart agriculture. For instance, Amaruchkul, (2021) develops a stochastic logistic model to determine the allocation of workers and estimate the value of image-based or remote sensing AI-powered yield prediction. As a complementary method, we introduce an XAI fuzzy logic-based approach to provide further interpretative insight into this automated prediction. Fuzzy systems have been applied to deal with information from weather stations in the IoT for agriculture. For example, in Krishnan et al., (2020) the authors have proposed a smart irrigation system that helps farmers water their agricultural fields using the mobile network. Here, the fuzzy logic-controller is used to compute input parameters (e.g. soil moisture, temperature and humidity) and produce motor status outputs. The developed system can switch off the motor during rains to ensure power savings. The authors have compared their proposed system, drip irrigation and manual flooding, to show that water and power conservation is obtained through the proposed smart irrigation system. Nevertheless, this work primarily focuses on the machine control aspects.

Mathematical preliminaries

In this Section, we discuss the basics of T1 FSs, T2 FSs, IT2 FSs and the T1 FS based WMA. All these basics are required for understanding the work presented in this paper.

Type-1 fuzzy sets (T1 Fss)

The concept of FSs was proposed by Prof. Zadeh (1965) through his remarkable work (Zadeh, 1965), as a generalisation of their mathematical counterparts called the crisp sets (or sets). According to Prof. Zadeh, the FSs have a greater capability to model the real life scenarios pertaining to the categorisation problems. By categorisation we mean here the act of dividing the objects into groups. For example, consider a universe of discourse as a set of positive integers, denoted as: $U = \{x|x > 0 \cap x \in Z^+\}$. Let's define a set A on U such that A is a collection of all the integers greater than 10 viz., $A = \{x|x > 10 \cap x \in Z^+\}$. Now clearly, the integers $\{1, 2., 10\}$ do not belong in A, and integers $\{11, 12., \infty\}$, belong to A. However, if we define another set B as a collection of integers *muchlargerthan*10, then will the integers $11, 12, 13, 14$ belong to B or not? When we ask different people their opinions about whether these numbers belong to the set B or not, different people have different opinions. Such situations are encountered quite frequently in real life like a collection of *Tall* men, *Fair* women, etc. Thus, in the example discussed above, the set A is often called the crisp set and the set B is called the FS.

To put it more mathematically, we can associate a degree of membership (or belongingness) to every set element. The degree of membership can also be called the MF value. It is denoted as $\mu(x)$, and referred to as the degree of membership of x into set A. Thus, for a crisp set, A, for every element $x, x \in U$, from Universe of Discourse, the MF for every x is either 0 or 1, the former being an indicator of absence from the set and the latter corresponding to the belongingness into the set.

For the case of the FSs, the MF is not an exclusive number as 0 or 1. Rather it is a precise number between and including both 0 and 1. Thus, for the given FS B, the MF for any element x, taken form Universe of Discourse, $0 \leq \mu(x) \leq 1$. It is pertinent to mention that scales other than 0 to 1 can also be used. Also, the closer the value of $\mu(x)$ to 1, the greater the degree of belongingness of x to FS B. Thus, every element of the FS is twin valued viz., the element and its degree of membership (or MF), denoted in Equation (1) as:

$$B = \{(x, \mu(x))|x \in U, 0 \leq \mu(x) \leq 1\} \tag{1}$$

The FSs defined by Equation (1) have a precise MF and thus have reduced capability to model the linguistic or data uncertainty. Therefore, Prof. Zadeh later on defined higher order FSs, of which one special category is called the T2 FSs, thus leading to the former being called as the T1 FSs. We will discuss the details of T2 FSs in Section 3.2.

Returning to the T1 FSs, there are various ways of representing the T1 FSs and their associated MFs (graphically as well as mathematically). The two most commonly used forms are the Trapezoidal and Triangular MFs. These are shown in Figures 1 and 2.

If Figure 1, it can be seen that the Trapezoidal MF is described by four points viz., $\{a, b, c, d\}$. These four points are used for constructing the MF by joining the points $(a, 0), (b, 1), (c, 1)$ and $(d, 0)$. Further, it is seen that the trapezoidal MF is shown to be normal (Klir & Yuan, 1995), however subnormal MFs are also possible. Thus, given a point x, lying within the FS B, characterised by the trapezoidal MF (shown in Figure 1), its degree of membership is given in the Equation (2):

$$\mu(x) = \begin{cases} 0, \ x < a \\ \frac{x-a}{b-a}, \ a \leq x < c \\ 1, \ b \leq x < c \\ \frac{d-x}{d-c}, \ c \leq x < d \\ 0, \ x \geq d \end{cases} \tag{2}$$

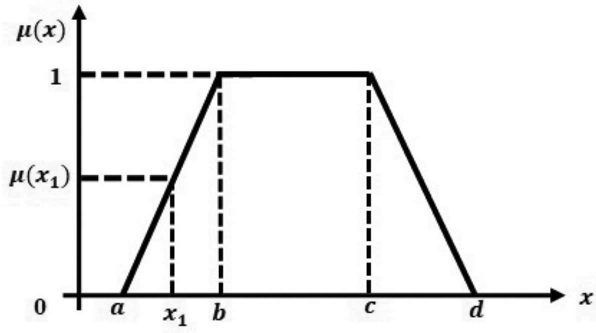

Figure 1. Trapezoidal membership functions of a T1 FS (Klir & Yuan, 1995).

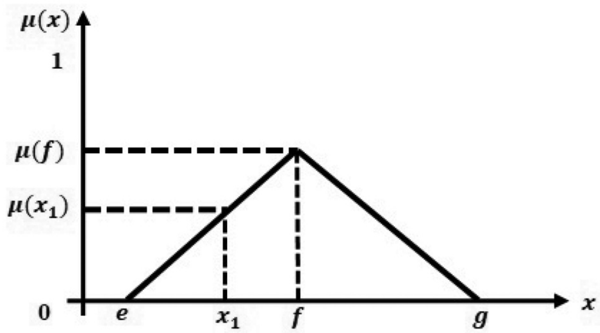

Figure 2. Triangular membership functions of a T1 FS (Klir & Yuan, 1995).

Similarly from Figure 2, it can be seen that the Triangular MF is described by three points viz., $\{e, f, g\}$. Also, the membership value at f of the T1 FS is given as $\mu(f)$. Thus, the triangular MF is obtained by joining the points $(e, 0)$, $(f, \mu(f))$ and $(g, 0)$. It can be seen that the triangular MF shown is subnormal, however normal T1 FS is also possible with $\mu(f) = 1$. So, given a data point x, lying within the FS B, characterised by the triangular MF (shown in Figure 2), its degree of membership is given in the Equation (3):

$$\mu(x) = \begin{cases} 0, \; x < e \\ \frac{\mu(f)(x-e)}{f-e}, \; e \le x < f \\ \frac{\mu(f)(g-x)}{g-f}, \; f \le x < g \\ 0, \; x \ge g \end{cases} \tag{3}$$

In Figures 1 and 2, it can be seen that there lies a data point on the x-axis given as $x = x_1$. Thus, in the trapezoidal (or triangular) MF shown in Figures 1 and 2, the degree of membership of this data point x_1 is given as the y-intercept and it's value is given as $\mu(x_1)$. The value of $\mu(x_1)$ can be found using the Equation (2) or Equation (3), depending on whether the MF shape is trapezoidal or triangular, respectively.

Type-2 fuzzy sets (T2 FSs)

The T1 FSs defined by Equation (1), though an improvement over the crisp sets, still have a reduced capability to model the intra and inter linguistic (data) uncertainty. Intra uncertainty pertains to the different meanings of a linguistic term or word, that a person develops over time. Whereas the inter data uncertainty pertains to the different meanings of a word, that a group of people possess. This is because the MF of the T1 FSs is crisp or precise. This sounds contradictory to the term 'fuzzy'. Therefore, Prof. Zadeh proposed a class of higher order FSs in (Zadeh, 1975), of which a special type of FSs called the T2 FSs. The T2 FSs have a greater capability to model the data uncertainty when compared to the T1 FSs. When one goes from the T1 FSs to the T2 FSs, the degree of membership associated with each set element in Equation (1) begins to be addressed as the primary membership. Further, an additional quantity is appended to each set element viz., the secondary membership, which models the degree of uncertainty about the primary membership. Thus, the T2 FSs are given in Equation (4) as:

$$\tilde{B} = \{(x, \mu(x), \mu_{\tilde{B}}(x, \mu(x)))|x \in U, 0 \leq \mu(x) \leq 1\} \tag{4}$$

In Equation (4), the ˜ sign over the T2 FS $\tilde{B}$ is a denoting of T2 FS. Now, the quantity $\mu(x)$ is called the primary membership and the quantity $\mu_{\tilde{B}}(x, \mu(x))$ is called the secondary membership. A T2 MF is pictorially shown in Figure 3.

In Figure 3, the primary MF is shown as green, and the secondary MF is shown in red. All the primary MFs are contained within the bounding region called the Footprint of uncertainty (FOU) and the secondary MF sits atop the FOU. The spread of the FOU gives an idea of the uncertainty captured by the T2 FS. Further, the FOU is bounded from above and below by a T1 MF called the Upper Membership Function (UMF) and the Lower Membership Function (LMF), respectively. The UMF and LMF are also shown in Figure 3.

Figure 3, though is an exact representation of the T2 FS semantics, however, makes it difficult to visualise and calculate both the primary and secondary MF value for a given data point say x_1. For such calculations, we adapt the Figure 3 to its corresponding 2-D version. It is shown in Figure 4.

In Figure 4, the UMF and LMF are shown to be of trapezoidal and triangular shape respectively. However, both can be trapezoidal (or triangular). In the figure, the depiction of the left hand side is the primary MF and sitting atop it is the secondary MF, as shown on the right side. Consider a data point x_1, lying on the x-axis, as shown in the Fig. Now, the intercept of the data point x_1 on the UMF gives the degree of belongingness to the UMF, whereas its corresponding intercept on the LMF gives the degree of belongingness to the LMF. Therefore, using Equations (2 and 3), the degrees of membership of the data point x_1 can be found in the UMF and LMF, respectively. They are now denoted as shown in Equations (5 and 6) respectively:

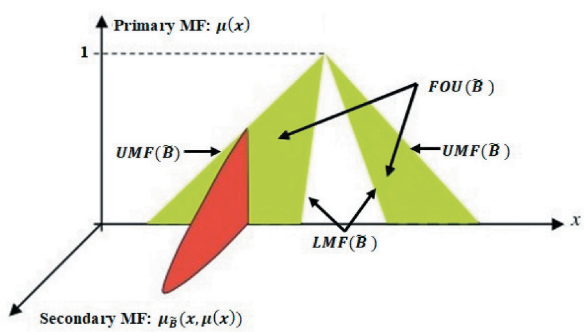

Figure 3. A T2 membership function in 2d: figure on the left is the primary mf and sitting atop it is the secondary mf, shown on right.

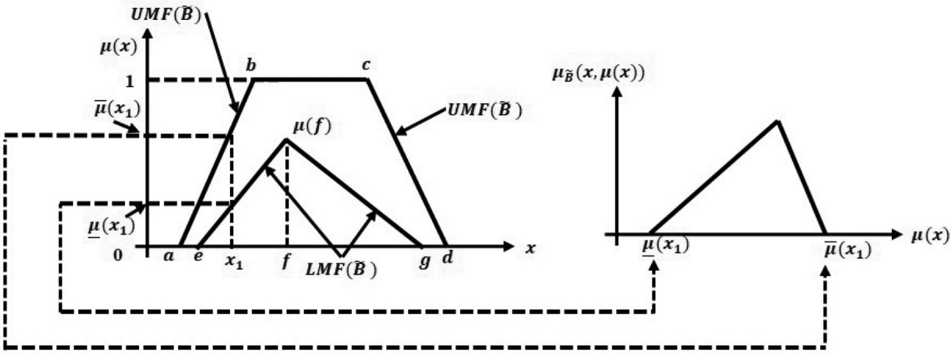

Figure 4. A T2 membership function (Wikipedia Contributors, 2020) public domain.

$$UMF : \bar{\mu}(x_1) = \begin{cases} 0, \ x_1 < a \\ \frac{x_1 - a}{b - a}, \ a \le x_1 < b \\ 1, \ x_1 < b \le x_1 < c \\ \frac{d - x_1}{d - c}, \ c \le x_1 < d \\ 0, \ x_1 \ge d \end{cases} \tag{5}$$

$$LMF : \underline{\mu}(x_1) = \begin{cases} 0, \ x_1 < e \\ \frac{\mu(f)(x_1 - e)}{f - e}, \ e \le x_1 < f \\ \frac{\mu(f)(g - x_1)}{g - f}, \ f \le x_1 < g \\ 0, \ x_1 \ge g \end{cases} \tag{6}$$

A point worth noting is that as we go from Equations (2) to (5), a $^{-}$ symbol appears over the μ. This happens because in Equation. (5), the quantity is now a UMF of a T2 FS and no longer the MF of a T1 FS. A similar analogy exists for the transition from Equations (3) to (6). Therefore, the primary MF value of the data point x_1, is no longer a single precise quantity but rather an interval. It is given in Equation (7) as:

$$\mu(x_1) = \left[\underline{\mu}(x_1), \bar{\mu}(x_1)\right] \tag{7}$$

Considering again the Figure 4, on the right side in this Fig., the primary MF now becomes the horizontal axis and the secondary MF is shown as the vertical axis. It can be seen that for the data point x_1, its primary MF (shown on the left side in the Figure 4) viz., $\left[\underline{\mu}(x_1), \bar{\mu}(x_1)\right]$, becomes the points onto the horizontal axis and the secondary MF takes the form of a function. Again it is pertinent to mention that the Secondary MF for the data point x_1 is shown to be triangular in shape, but in general, it can be of any shape.

Interval type-2 fuzzy sets (IT2 FSs)

The T2 FSs have a greater capability to model the data uncertainty in comparison to the T1 FSs. This greater capability comes with a greater computational cost, both in terms of data representation as well as data processing. However, numerous times the applications (or situations) demand that the data uncertainty should be modelled using the FSs higher than the T1 or using T2 FSs, but the computational complexity should be minimised. Therefore, Prof. Zadeh also conceptualised a special category of the T2 FSs, called the IT2 FSs, in (Zadeh, 1975). In the IT2 FSs, the secondary MF is assumed to be 1 everywhere.

It is pictorially shown in Figure 5 and mathematically given in the form of Equation (8) as:

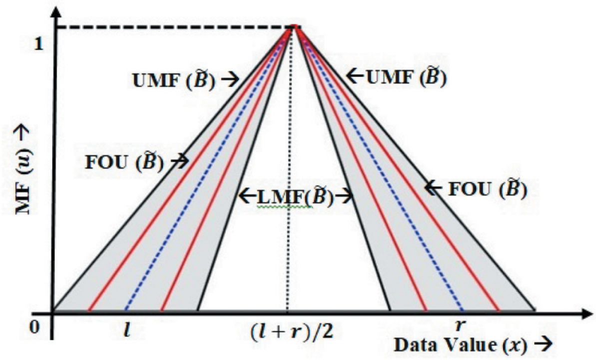

Figure 5. Membership functions of an IT2 FS (Gupta, 2019).

$$\tilde{B} = \{(x, \mu(x), \mu_{\tilde{B}}(x, \mu(x)) = 1)|x \in U, 0 \leq \mu(x) \leq 1\} \tag{8}$$

Thus, for any data point lying within the IT2 FS, its membership function value is completely characterised by its primary MF, which is an interval. It is given by Equation (7). Also, in the Figure 5, it can be seen that a T1 FS is shown inside the FOU of IT2 FS, by a dashed line, whose ends rest on the x-axis at l and r. This T1 FS is called an embedded T1 FS. According to Mendel & Wu, (2010), the FOU of an IT2 FS can be considered as a union of all such embedded T1 FSs.

T1 FS based WMA

The T1 FS-based WMA was proposed in Wang & Mendel, (1992). It provides a five-step methodology for generating fuzzy rules from numerical data values, provided as a collection of input-output pairs.

Consider a collection of numeric data value pairs such that all the x's are inputs and all the y's are outputs. These x's and y's exist in pairs of the form $(x_1^{(j)}, x_2^{(j)}, \ldots, x_n^{(j)}, y_1^{(j)}, y_2^{(j)}, \ldots, y_m^{(j)})$. Here, $x_w^{(j)}, w = 1., n$ and $y_q^{(j)}, q = 1., m$ is the j^{th} instance of the input x_w and the output y_q, respectively.

Step 1 division of input-output data variables into the fuzzy regions

Initially, each input and output variable is associated to a data interval, which aligns closely with the domain of the variable. The domain is the range of values enclosed by an interval and represents the lower as well as upper bound on the value of the variable.

We assume the data intervals corresponding to x_1 be $\left[l_1^{ip}, r_1^{ip}\right]$. Similarly, corresponding to $x_2, \ldots, x_n$ be $\left[l_2^{ip}, r_2^{ip}\right], \ldots, \left[l_n^{ip}, r_n^{ip}\right]$, nd $y_1, \ldots, y_m$ be $[l_1^{op}, r_1^{op}], \ldots, [l_m^{op}, r_m^{op}]$. In these data intervals, the superscript ip and op denote data interval for the input and output, respectively. Now, each of these intervals is individually divided into $2N + 1$ fuzzy regions, N being same or different for any two (or more) variables. It is mentioned here that the semantics of the region are represented using T1 fuzzy MFs. Each of these regions is generally denoted by a linguistic label, like *big*, *small*, *verybig*, etc. Lets say for input $x_w, w = 1., n$, the regions be denoted by linguistic labels as $L_1^{ip_w}, L_2^{ip_w}, \ldots., L_{2N+1}^{ip_w}$. Similarly, for any $y_q, q = 1, \ldots m$, the regions correspond to the linguistic labels: $L_1^{op_q}, L_2^{op_q}, \ldots., L_{2N+1}^{op_q}$.

Step 2 fuzzy rule generation from the numeric data pairs

Now, the task is to generate T1 fuzzy rules from the numeric data values. To exemplify, take any j^{th} data vector containing the variables with values $(x_1^{(j)}, x_2^{(j)}, \ldots, x_n^{(j)}, y_1^{(j)}, y_1^{(j)}, \ldots, y_m^{(j)})$. From the location of these data points of the variables on the information axis, any data point can belong to a maximum of two adjacent linguistic terms with varying degrees of memberships (Klir & Yuan, 1995). Considering that a data point $x_w^{(j)}, w = 1, 2., n$ belongs simultaneously to linguistic terms L_k and L_{k+1}, with degrees of memberships $\mu_{L_k}(x_w^{(j)})$ and $\mu_{L_{k+1}}(x_w^{(j)})$. Therefore, the membership degree of $x_w^{(j)}, w = 1, 2., ., n$ is $max\{\mu_{L_k}(x_w^{(j)}), \mu_{L_{k+1}}(x_w^{(j)})\}$ and the linguistic term is the one in which the variable has the higher degree of membership. In this manner, these degrees of memberships as well as linguistic terms are found for all the input as well as output variables.

Based on these linguistic terms found, the fuzzy IF-THEN rules are generated, which have a form similar to one shown in Equation (9) as:

$$\begin{array}{c} Rule\ j : IFx_1^{(j)}\ is\ L_k^{ip_1}\ and\ x_2^{(j)}\ is\ L_k^{ip_2}, \ldots, x_n^{(j)}\ is\ L_k^{ip_n}, \\ THEN\ y_1^{(j)}\ is\ L_k^{op_1}\ and\ y_2^{(j)}\ is\ L_k^{op_2}, \ldots, y_m^{(j)}\ is\ L_k^{op_m} \end{array} \tag{9}$$

where $x_w^{(j)}, w = 1, 2., n$ ($y_q^{(j)}, q = 1, 2., m$) has highest degree of membership in respective $L_k^{ip_w}$ ($L_k^{op_q}$).

Step 3 degree assignment to each rule

In the previous step, it is possible that the IF-THEN rules generated are conflicting with some of each other, where they have the same IF part but a different THEN part. Hence, a way to resolve this is to calculate the degree of membership of all the rules and retain only those where the degree of membership is highest. The degree of a rule is defined as the product of the memberships of the variables in their respective linguistic terms, given as:

$$\mu_{D_{Rule\ j}} = \prod_{w=1}^{n} \mu_{L_k^{ip_w}}(x_w^{(j)}) \times \prod_{q=1}^{m} \mu_{L_k^{op_q}}(y_q^{(j)}) \tag{10}$$

where $\mu_{L_k^{ip_w}}(x_w^{(j)}), w = 1, 2., n$ is the degree of membership of $x_w^{(j)}$ in the linguistic term $L_k^{ip_w}$. Further, $\mu_{L_k^{op_q}}(y_q^{(j)}), q = 1, 2., m$ is the degree of membership of $y_q^{(j)}$ is the linguistic term $L_k^{op_q}$.

Step 4 creation of a combined fuzzy rule base

Once the conflicting rules have been removed in the previous step, a combined fuzzy rule base is formed. Here, it is possible that some IF-THEN rules for the combination of input data values may not have been formed from the steps above, due to the absence of numeric data values ensuring their construction. Therefore, in this step, generally, an expert is asked to provide the membership degree values of such rules. Further, an expert may also choose to provide a membership degree of a rule obtained following steps above. There, the combined IF-THEN rule base chooses to overcome this ambiguity by retaining the rule with maximum membership degree, out of the one constructed from numeric data value and the one given by the expert.

Step 5 determining a mapping based on the combined fuzzy rule base

Finally centroid defuzzification is used to determine the value of the q^{th} output. To calculate the same, first the degree of membership of the j^{th} rule's inputs is calculated as:

$$\mu_{inps\ Rule\ j}^{(j)} = \prod_{w=1}^{n} \mu_{L_k^{ip_w}}(x_w^{(j)}) \tag{11}$$

where $\mu_{L_k^{ip_w}}(x_w^{(j)}), w = 1, 2., n$ is the membership degree of $x_w^{(j)}$ in the linguistic term $L_k^{ip_w}$.

Using the $\mu_{inps\ Rule\ j}^{(j)}$ from Equation (11), the defuzzified value of q^{th} output $\mu(y_q)$, is given as:

$$\mu(y_q) = \frac{\sum_{j=1}^{K} \mu_{inps\ Rule\ j}^{(j)} \times \mu^{(j)}(y_q)}{\sum_{j=1}^{K} \mu_{inps\ Rule\ j}^{(j)}} \tag{12}$$

where $\mu^{(j)}(y_q)$ is the centroid of the linguistic term corresponding to the q^{th} output variable in the consequent part of j^{th} IF-THEN rule given in Equation (9).

Proposed novel enhanced Wang Mendel Approach (WMA) based on T2 FSs

In this section, we discuss the details of the proposed novel Enhanced WMA based on the T2 FSs. In the proposed Enhanced WMA, the MFs of the linguistic terms are modelled as the T2 FSs. It is an improvement over the existing WMA of (Wang & Mendel, 1992), where the MFs are modelled using the T1 FSs, as we know that T2 FSs have a greater capability to model the semantic uncertainty of the linguistic terms (as discussed in Section 3.2). There have been numerous works on WMA (e.g. Casillas et al. 2000; Guo et al., 2015; Hao & Mendel, 2013; Yang et al., 2010). However, to the best of our knowledge, no work has proposed a design based on T2 FSs so far. Therefore, here we present the Enhanced WMA based on T2 FSs. We will also show that by minimum changes, the T2 FS based design of Enhanced WMA can be converted to the one based on IT2 FSs. Hence, we have chosen to call both the T2 and IT2 FS based design of WMA as the Enhanced WMA. Our Enhanced WMA also consists of a five step procedure as that of (Wang & Mendel, 1992) and as discussed in Section 3.4. Hence the readers of the papers are advised to refer to the respective steps of T1 FS based WMA of Section 3.4, while going through the steps of Enhanced WMA.

Consider, a system with a stream of numeric data values in the form of multiple input variables x_1, x_2, ... , x_n and output variables y_1, y_2, ... , y_m. These inputs and outputs occur in pairs like $(x_1^{(j)}, x_2^{(j)}, \ldots, x_n^{(j)}, y_1^{(j)}, y_2^{(j)}, \ldots, y_m^{(j)})$. Here, $x_w^{(j)}, w = 1., n$ denotes the j^{th} instance of the input x_w and $y_q^{(j)}, q = 1., m$ denotes the j^{th} instance of the output y_q. The Enhanced WMA processes these streams of numeric values to generate output recommendations using the following five step procedure:

Step 1 division of input-output data variables into the fuzzy regions

In this step, we associate with each input and output variable, a data interval. These respective data intervals align closely with the respective domains of input and output variables. Here, the domain means the range of values enclosed by an interval and represents the lower as well as upper bound on the values of the input or output variable. In other words, the respective data intervals align closely with the range of values so that the respective variables (input and output) most probably lie within the respective data interval.

Let these data intervals corresponding to the variables be denoted as: $x_1 \to \left[l_1^{ip}, r_1^{ip}\right]$, $x_2 \to \left[l_2^{ip}, r_2^{ip}\right]$, ... , $x_n \to \left[l_n^{ip}, r_n^{ip}\right]$, $y_1 \to [l_1^{op}, r_1^{op}]$, $y_2 \to [l_2^{op}, r_2^{op}]$, ... , $y_m \to [l_m^{op}, r_m^{op}]$. Thus, in general, the data intervals of any input variable x_w can be denoted as $\left[l_w^{ip}, r_w^{ip}\right], w = 1., n$ and the output variable y_q as $\left[l_q^{op}, r_q^{op}\right], q = 1., m$. Here the superscript *ip* is a depiction of data interval for the input and *op* is a depiction of data interval for the output.

Each of these domains of the input as well as output variables are individually divided into $2N + 1$ fuzzy regions. Here N can be same or different for any two (or more) variables (inputs as well as output). By fuzzy region, we mean that the semantics of the region are represented using T2 FSs. Each of these regions is assigned (generally) a linguistic label. For input x_1, let the regions be denoted as $L_1^{ip_1}, L_2^{ip_1}, \ldots, L_{2N+1}^{ip_1}$. For x_2, let the regions be denoted as $L_1^{ip_2}, L_2^{ip_2}, \ldots, L_{2N+1}^{ip_2}$. Thus, in general for any $w^{th}, w = 1., n$ input $x_w, w = 1., n$, the regions can be denoted by linguistic labels as $L_1^{ip_w}, L_2^{ip_w}, \ldots, L_{2N+1}^{ip_w}$. Similarly, for any $q_{th}, q = 1., m$ output $y_q, q = 1., m$, the regions can be denoted by the linguistic labels as: $L_1^{op_q}, L_2^{op_q}, \ldots, L_{2N+1}^{op_q}$.[3]

Step 2 fuzzy rule generation from the numeric data pairs

In this step, we generate the fuzzy rules using the stream of numeric data values. We pick up individual input output data pairs from the given stream of data values. Then we find out the degree of membership of every numeric data value in the respective linguistic terms (or fuzzy regions). Thus consider any j^{th} data vector containing the values $(x_1^{(j)}, x_2^{(j)}, \ldots, x_n^{(j)}, y_1^{(j)}, y_1^{(j)}, \ldots, y_m^{(j)})$. In this data

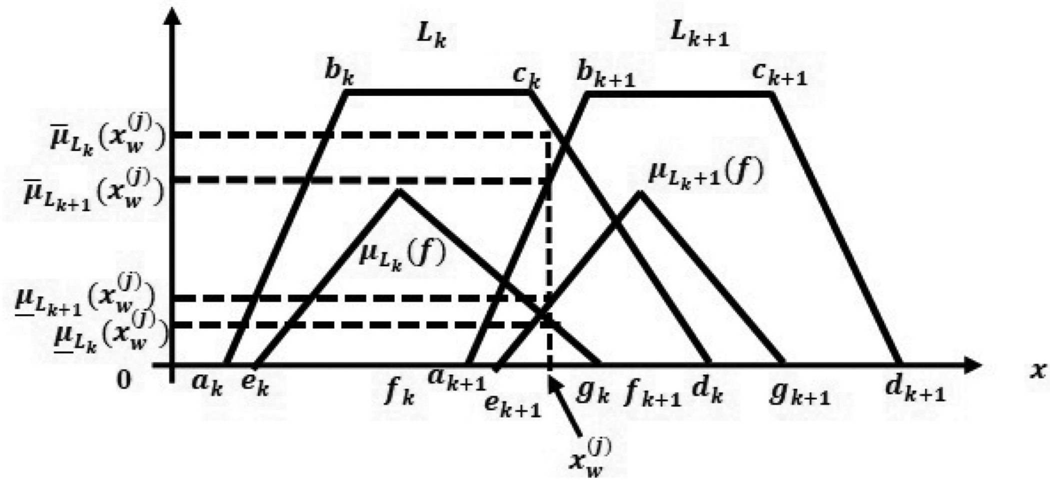

Figure 6. Membership degree of x_w in two adjacent linguistic terms L_k and L_{k+1} represented as T2 FSs.

vector, any input data variable (or the output data variable) can belong to atmost two adjacent linguistic terms, say $L_k, k = 1, 2., 2N$ and $L_{k+1}, k = 1, 2., 2N$. Let's say a data point corresponding to the input variable, $x_w^{(j)}, w = 1, 2., n$, lies onto the x-axis and belongs simultaneously to both L_k and L_{k+1}. This is shown in Figure 6. Here, the L_k is a T2 FS defined in 2-D by the points $\{a_k, b_k, c_k, d_k, e_k, f_k, g_k, \mu_k(f)\}$ and L_{k+1} is defined by the points $\{a_{k+1}, b_{k+1}, c_{k+1}, d_{k+1}, e_{k+1}, f_{k+1}, g_{k+1}, \ \mu_{k+1}(f)\}$, similar to Figure 4 discussed in Section 3.2. The respective secondary MFs will be sitting atop the respective 2-D MFs of the L_k and L_{k+1}, shown in Figure 6. Thus, the degree of belongingness of x_w in L_k and L_{k+1}, respectively is found using the Equation (7), and is given as:

$$\begin{aligned} \mu_{L_k}(x_w^{(j)}) &\rightarrow \left[\underline{\mu}_{L_k}(x_w^{(j)}), \bar{\mu}_{L_k}(x_w^{(j)})\right] \\ \mu_{L_{k+1}}(x_w^{(j)}) &\rightarrow \left[\underline{\mu}_{L_{k+1}}(x_w^{(j)}), \bar{\mu}_{L_{k+1}}(x_w^{(j)})\right] \end{aligned} \tag{13}$$

It is pertinent to mention that in this Equation (13), the quantities L_k and L_{k+1} have been put in the foot of the UMF and the LMF of μ to differentiate between the MF values of the variable $x_w^{(j)}$ in the linguistic terms L_k and L_{k+1}.

Similarly, such intervals can be found for all other $n-1$ input variables as well as the m output variables contained within the j^{th} data vector.

Now, the task is to find out that, for each of the input as well as the output data variables, belonging to respective two adjacent linguistic terms, the degree of membership of each of the data variables (input as well as output) is the highest in which one respective linguistic term out of the two adjacent ones. This is accomplished by the following computations. Consider the input data variable $x_w^{(j)}$ from j^{th} data vector. It's secondary MF value in the linguistic terms L_k and L_{k+1} be given as $\mu_{L_k}(x_w^{(j)}, \mu_{L_k}(x_w^{(j)}))$ and $\mu_{L_{k+1}}(x_w^{(j)}, \mu_{L_{k+1}}(x_w^{(j)}))$, respectively which follows directly from Equation (4).

From (Mendel, 2001), it follows that $\mu_{L_k}(x_w^{(j)}, \mu_{L_k}(x_w^{(j)})) \in \left[\underline{\mu}_{L_k}(x_w^{(j)}), \bar{\mu}_{L_k}(x_w^{(j)})\right]$ as well as $\mu_{L_{k+1}}(x_w^{(j)}, \mu_{L_{k+1}}(x_w^{(j)})) \in \left[\underline{\mu}_{L_{k+1}}(x_w^{(j)}), \bar{\mu}_{L_{k+1}}(x_w^{(j)})\right]$. The centre of gravity (COG) associated to the secondary MF can be found (depending on the shape of the secondary MF). Let's say the COG associated to the $\mu_{L_k}(x_w^{(j)}, \mu_{L_k}(x_w^{(j)}))$ as well as $\mu_{L_{k+1}}(x_w^{(j)}, \mu_{L_{k+1}}(x_w^{(j)}))$ be $\mu^{COG}(x_w^{(j)}, \mu_{L_k}(x_w^{(j)}))$ and $\mu^{COG}(x_w^{(j)}, \mu_{L_{k+1}}(x_w^{(j)}))$, respectively. Therefore, if $\mu^{COG}(x_w^{(j)}, \mu_{L_k}(x_w^{(j)})) > \mu^{COG}(x_w^{(j)}, \mu_{L_{k+1}}(x_w^{(j)}))$, then x_w belongs to L_k otherwise it belongs to L_{k+1}.

A special case may arise when the semantics of the linguistic terms are represented using the IT2 FSs. In such cases, we can simplify the calculation of the COG. For an IT2 FS, the COG is given as an average of the LMF and UMF, as shown in Equation (14):

$$\begin{aligned} \mu^{COG}(x_w^{(j)}, \mu_{L_k}(x_w^{(j)})) &= \frac{1}{2}\left[\underline{\mu}_{L_k}(x_w^{(j)}) + \bar{\mu}_{L_k}(x_w^{(j)})\right] \\ \mu^{COG}(x_w^{(j)}, \mu_{L_{k+1}}(x_w^{(j)})) &= \frac{1}{2}\left[\underline{\mu}_{L_{k+1}}(x_w^{(j)}) + \bar{\mu}_{L_{k+1}}(x_w^{(j)})\right] \end{aligned} \tag{14}$$

Thus, in this manner, the belongingness of all the $x_w^{(j)}, w = 1, 2., n$ and $y_q^{(j)}, q = 1, 2., m$ are found into a unique respective linguistic term, given as: $L_k^{ip_w}, w = 1, 2., n, k = 1, 2., 2N + 1$ for is and $L_k^{op_q}, q = 1, 2., m, k = 1, 2., 2N + 1$ for outputs. Finally, in this step, based on the belongingness of the respective input and output variables in the respective linguistic terms, the IF-THEN rules are generated. Therefore, for the j^{th} data pair given as: $(x_1^{(j)}, x_2^{(j)}, \ldots, x_n^{(j)}, y_1^{(j)}, y_1^{(j)}, \ldots, y_m^{(j)})$, based on respective (highest) degree of belongingess to the linguistic terms, the IF-THEN rules are given in Equation (15) as:

$$\begin{gathered} Rule\ j : IF\ x_1^{(j)}\ is\ L_k^{ip_1}\ and\ x_2^{(j)}\ is\ L_k^{ip_2}, \ldots, x_n^{(j)}\ is\ L_k^{ip_n}, \\ THEN\ y_1^{(j)}\ is\ L_k^{op_1}\ and\ y_2^{(j)}\ is\ L_k^{op_2}, \ldots, y_m^{(j)}\ is\ L_k^{op_m} \end{gathered} \tag{15}$$

Step 3 degree assignment to each rule

Now in the third step, we assign a degree to each rule. This is required for conflict resolution between the IF-THEN rules, where conflicting rules are the ones that have the same linguistic terms corresponding to the input variables in the antecedent part of the IF-THEN rule but different linguistic terms corresponding to output variables in the consequent part of the IF-THEN rule. Such rules arise mainly because there is a large number of data pairs in the incoming stream of data values. As pointed out in Wang & Mendel, (1992), one way to resolve such conflicting rules is to accept the rule with the highest degree from each of the conflicting sets. For this, we need to calculate the degree of a rule.

Consider a j^{th} IF-THEN rule as shown in Equation (15). The membership value of the degree of this rule is defined as the product of the membership values of the input and output variables in the respective linguistic terms, which themselves are essentially in the form of intervals generated similar to Equation (7). Thus, using the Equation (7) and the *Rule j* given in Equation (15), the degree of the *Rule j* is an interval, and given as:

$$\begin{aligned} \mu_{D_{Rule\,j}} &= \left[\underline{\mu}_{D_{Rule\,j}}, \bar{\mu}_{D_{Rule\,j}}\right] = \prod_{w=1}^{n} \mu_{L_k^{ipw}}(x_w^{(j)}) \times \prod_{q=1}^{m} \mu_{L_k^{opq}}(y_q^{(j)}) \\ &= \prod_{w=1}^{n}\left[\underline{\mu}_{L_k^{ipw}}(x_w^{(j)}), \bar{\mu}_{L_k^{ipw}}(x_w^{(j)})\right] \times \prod_{q=1}^{m}\left[\underline{\mu}_{L_k^{opq}}(y_q^{(j)}), \bar{\mu}_{L_k^{opq}}(y_q^{(j)})\right] \end{aligned} \tag{16}$$

Also, in the Equation (16), secondary MF value needs to be treated differently. However, in the case of IT2 FSs, the secondary MF is 1 everywhere and therefore separate treatment is not required. Further, it can be seen that the $\mu_{D_{Rule\,j}}$ obtained from Equation (16), is an interval. For the T2 FS representation, we calculate the COG of the $\mu_{D_{Rule\,j}}$, which corresponds to the degree of the *Rule j*. In case the linguistic terms are represented using the IT2 FSs, the COG of the $\mu_{D_{Rule\,j}}$ gives the degree of a *Rule j*. Thus, for the case of IT2 FSs, the degree of *Rule j* is given as:

$$D_{Rule\,j} = \mu_{D_{Rule\,j}}^{COG} = \frac{1}{2}\left[\underline{\mu}_{D_{Rule\,j}} + \bar{\mu}_{D_{Rule\,j}}\right] \tag{17}$$

Sometimes an expert's opinion may also be available about the rules in the form of some apriori information, generally expressed as the membership degree of a rule, say $\mu_{exp}(D_{Rule\ j})$. Therefore, the membership value of the degree of a rule, after incorporating the expert's opinion, causes Equation (16) to become Equation (18):

$$\begin{aligned}\mu_{D_{Rule\ jnew}} &= \left[\underline{\mu}_{D_{Rule\ j\,new}}, \bar{\mu}_{D_{Rule\ j\,new}}\right] = \prod_{w=1}^{n} \mu_{L_k^{ipw}}(x_w^{(j)}) \times \prod_{q=1}^{m} \mu_{L_k^{opq}}(y_q^{(j)}) \times \mu_{exp}(D_{Rule\ j}) \\ &= \prod_{w=1}^{n} \left[\underline{\mu}_{L_k^{ipw}}(x_w^{(j)}), \bar{\mu}_{L_k^{ipw}}(x_w^{(j)})\right] \times \prod_{q=1}^{m} \left[\underline{\mu}_{L_k^{opq}}(y_q^{(j)}), \bar{\mu}_{L_k^{opq}}(y_q^{(j)})\right] \times \left[\underline{\mu}_{exp}(D_{Rule\ j}), \bar{\mu}_{exp}(D_{Rule\ j})\right]\end{aligned} \tag{18}$$

Again, in Equation (18), secondary MF value needs to be treated differently for T2 FSs but not in the case of IT2 FSs. Further, for the T2 FS representation, we calculate the COG of the $\mu_{D_{Rule\ j\,new}}$, which corresponds to the new degree of the *Rule j new*. In case the linguistic terms are represented using the IT2 FSs, the COG of the $\mu_{D_{Rule\ j\,new}}$ gives the degree of a *Rule j* as:

$$D_{Rule\ j\,new} = \mu^{COG}_{D_{Rule\ j\,new}} = \frac{1}{2}\left[\underline{\mu}_{D_{Rule\ j\,new}} + \bar{\mu}_{D_{Rule\ j\,new}}\right] \tag{19}$$

Thus, using Equation (17) or Equation (19), in the absence or presence of an expert, respectively, the conflicting rules are removed and only those rules are accepted which have the highest degree.

Step 4 creation of a combined fuzzy rule base

In this step, a combined fuzzy rule base is constructed based on the IF-THEN fuzzy rules generated from the numeric data pairs (as in Step 3) or the IF-THEN fuzzy rules provided by an expert. In case the IF-THEN fuzzy rule is provided by an expert, he/she provides a membership degree of each rule.

In (Wang & Mendel, 1992), the authors formalised this step by proposing a grid-like structure for the fuzzy rule base. The fuzzy rule base grid is n-dimensional structure, n being the number of input variables. The number of boxes in the grid is equal to the product of the number of linguistic terms corresponding to each input variable in the system. As we discussed in Step 2 that each input variable's domain is partitioned into $2N + 1$ linguistic variables. Therefore, the number of boxes in the grid will be $(2N + 1)^n$. Thus, each box in the grid is representative of a combination of linguistic values corresponding to each of the input data variables coming from some j^{th} IF-THEN rule, as given in Equation (15). Thus, the box in the grid is filled in with the linguistic value of q^{th} output variable available from the j^{th} IF-THEN rule given in Equation (15). Therefore, it can be seen that the number of such fuzzy IF-THEN rule bases constructed are equal to the number of output variables viz., m, one corresponding to each output variable (The number of output variables was considered as m in Step 1).

As a rule from (Wang & Mendel, 1992), if there's more than one rule in one box of the fuzzy rule base, the one with the maximum degree is always chosen.

Step 5 determining a mapping based on the combined fuzzy rule base

Finally, using this step, we find out the value of a q^{th} output variable, $y_q, q = 1, 2., m$ using the defuzzification procedure for the given input data vector, which can also be said as the data vector for which we want to predict the value of the q^{th} output variable, $y_q, q = 1, 2., m$. Consider a j^{th} data vector involving the input variables (however without the output variable as this needs to be predicted), given as: $(x_1^{(j)}, x_2^{(j)}, \ldots, x_n^{(j)})$. Let's say this data vector fires the j^{th} IF-THEN rule given in Equation (15). Thus, the degree of the rule corresponding to all the n input variables is found using Equation (16) (or Equation (18) as the case be). It is given as:

$$\mu^{(j)}_{inps\ Rule\ j} = \left[\underline{\mu}^{(j)}_{inps\ Rule\ j}, \bar{\mu}^{(j)}_{inps\ Rule\ j}\right] = \prod_{w=1}^{n} \mu_{L_k^{ipw}}(x_w^{(j)}) = \prod_{w=1}^{n} \left[\underline{\mu}_{L_k^{ipw}}(x_w^{(j)}), \bar{\mu}_{L_k^{ipw}}(x_w^{(j)})\right] \tag{20}$$

Similarly, the degree of all the rules in the fuzzy rule base, corresponding to all the n input variables (only) is calculated similarly to Equation (20). It is pertinent to mention that only a few of the rules from the rule base will provide non-zero membership degree for the Equation (20). So, we will consider only such rules for further processing. Also, if the linguistic terms' semantics are represented using the T2 FSs, they require separate treatment, but not in the case of IT2 FSs.

Let's say from the combined fuzzy rule base, a total of K IF-THEN rules had non-zero membership degree when fired by the j^{th} data vector involving the input variables. Therefore, using the degrees of all the rules corresponding to inputs only, calculated similar to Equation (20), the membership degree of the q^{th} output variable $y_q, q = 1, 2., m$ is found as:

$$\begin{aligned}\mu(y_q) = \left[\underline{\mu}(y_q), \bar{\mu}(y_q)\right] &= \frac{\sum_{j=1}^{K} \mu^{(j)}_{inps\ Rule\ j} \times \mu^{(j)}(y_q)}{\sum_{j=1}^{K} \mu^{(j)}_{inps\ Rule\ j}} \\ &= \frac{\sum_{j=1}^{K} \left[\underline{\mu}^{(j)}_{inps\ Rule\ j}, \bar{\mu}^{(j)}_{inps\ Rule\ j}\right] \times \left[\underline{\mu}^{(j)}(y_q), \bar{\mu}^{(j)}(y_q)\right]}{\sum_{j=1}^{K} \left[\underline{\mu}^{(j)}_{inps\ Rule\ j}, \bar{\mu}^{(j)}_{inps\ Rule\ j}\right]}\end{aligned} \tag{21}$$

Here, the quantity $\mu^{(j)}(y_q)$ is the centroid of the linguistic term corresponding to the q^{th} output variable in the consequent part of j^{th} IF-THEN rule given in Equation (15). The $\mu^{(j)}(y_q)$ is generally an interval, $\left[\underline{\mu}^{(j)}(y_q), \bar{\mu}^{(j)}(y_q)\right]$, determined using the algorithms like Enhanced Karnik Mendel (EKM) for IT2 FSs.

The quantity $\mu(y_q)$ obtained from Equation (21) is an interval. The semantics of this quantity are T2 FSs, if all the quantities on the R.H.S of Equation (21) are represented using T2 FSs.

To find out a single precise quantity corresponding to the $\mu^{(j)}(y_q)$, its COG is computed, which may have value depending on the secondary MF of T2 FS. In case the semantics of the quantities on the R.H.S of Equation (21) are represented using IT2 FSs, the COG value of the $\mu^{(j)}(y_q)$ is given in a straightforward manner as:

$$\mu^{COG}(y_q) = \frac{1}{2}\left[\underline{\mu}(y_q) + \bar{\mu}(y_q)\right] \tag{22}$$

We observed that a special case may occur while processing the non-zero membership degrees, as obtained similar to the Equation (20). Some intervals tend to provide 0.0, the value of the left or right end of the interval. These cause further problems while performing the interval divisions in Equation (21). Therefore, for such scenarios, we propose that the COG be first found for all the $\mu^{(j)}_{inps\,Rule\,j}$, obtained by use of Equation (20). In case the semantics of the linguistic terms are IT2 FSs, the COG can be found as:

$$\mu^{COG(j)}_{inps\ Rule\ j} = \frac{1}{2}\left[\underline{\mu}^{(j)}_{inps\ Rule\ j} + \bar{\mu}^{(j)}_{inps\ Rule\ j}\right] \tag{23}$$

Thus, using the Equation (23), the the membership degree of the q^{th} output variable $y_q, q = 1, 2., m$ is now found as:

$$\mu(y_q) = \left[\underline{\mu}(y_q), \bar{\mu}(y_q)\right] = \frac{\sum_{j=1}^{K} \mu^{COG(j)}_{inps\ Rule\ j} \times \mu^{(j)}(y_q)}{\sum_{j=1}^{K} \mu^{COG(j)}_{inps\ Rule\ j}} \tag{24}$$

Again, to find out a single precise quantity corresponding to the $\mu^{(j)}(y_q)$, its COG is computed, which may have value depending on the secondary MF of T2 FS. In case the semantics of the quantities on the R.H.S of Equation (24) are represented using IT2 FSs, the COG value of the $\mu^{(j)}(y_q)$ is given in a straightforward manner using Equation (22).

A real-world example application of enhanced WMA with agritech data

In this section, we demonstrate the use of the proposed Enhanced WMA from Section 4, for crop yield prediction in smart agriculture. We have used a publicly available agritech dataset. It consists of numerous variables. We found some of the variables redundant and therefore filtered out the relevant 11 input variables. The values of these input variables are together used for crop yield prediction. We have associated five linguistic terms to each of these 11 input data variables and the output variable of crop yield. We discuss all the details in the following two subsections viz., Subsection 5.1 and 5.2. In Subsection 5.1, we give the problem description and show the representation of linguistic variable semantics using IT2 FSs for input as well as output variables. In Subsection 5.2 we demonstrate the applicability of the proposed WMA for crop yield prediction using the word semantics represented in Subsection 5.1.

Problem description

Crop yield prediction in smart agriculture is a real-life application. There have been numerous works on smart agriculture (as discussed in Section 2), which motivated us to work on this research problem. We used a publicly available data set in which numerous variables were processed to predict the crop yield, out of which we filtered out 11 most relevant input variables. Therefore, in the line of the proposed WMA from Section 4, our system is an 11 input 1 output case. The values corresponding to each of these input and output variables were a stream of numbers.

Step-through numeric example for crop-yield prediction

Now we demonstrate the use of Enhanced WMA (from Section 4) for crop yield prediction in Smart agriculture.

Step 1 division of input-output data variables into the fuzzy regions

Now using Step 1 of the Enhanced WMA from Section 4, we assumed that each of these 12 variables (11 input and one output) has five associated linguistic terms. Put in other words, we chose $N = 2$, so that $2N + 1 = 5$. The variables and their associated linguistic terms are shown in Table 1.

In Wang & Mendel, (1992), the semantics for each of the associated variables were represented using uniformly distributed T1 FSs. Say, if the information representation scale for a variable extended from L to R, which had an associated $2N + 1$ linguistic terms, then each linguistic term

Table 1. Input-output variables and associated linguistic terms.

Variables	Associated Linguistic Terms
Latitude (*LA*)	Very Low (*LAE*), Low (*LAL*), Medium (*LAM*), High (*LAH*), Very High (*LAV*)
Longitude (*LO*)	Very Low (*LOE*), Low (*LOL*), Medium (*LOM*), High (*LOH*), Very High (*LOV*)
Apparent Temperature Max (*AM*)	Very Cold (*AME*), Cold (*AMC*), Moderate (*AMM*), Hot (*AMH*), Very Hot (*AMV*)
Apparent Temperature Min (*AI*)	Very Cold (*AIE*), Cold (*AIC*), Moderate (*AIM*), Hot (*AIH*), Very Hot (*AIV*),
Dew Point (*DP*)	Very Low (*DPE*), Low (*DPL*), Moderate (*DPM*), High (*DPH*), Very High (*DPV*)
Humidity (*HH*)	Very Low (*HHE*), Low (*HHL*), Moderate (*HHM*), High (*HHH*), Very High (*HHV*)
Pressure (*PE*)	Very Low (*PEE*), Low (*PEL*), Medium (*PEM*), High (*PEH*), Very High (*PEV*)
Temperature Max (*TM*)	Very Cold (*TME*), Cold (*TMC*), Moderate (*TMM*), Hot (*TMH*), Very Hot (*TMV*)
Temperature Min (*TI*)	Very Cold (*TIE*), Cold (*TIC*), Moderate (*TIM*), Hot (*TIH*), Very Hot (*TIV*)
Wind Bearing (*WB*)	Very Small (*WBE*), Small (*WBS*), Medium (*WBM*), Large (*WBL*), Very Large (*WBV*)
NDVI (*ND*)	Very Less (*NDE*), Less (*NDL*), Moderate (*NDM*), Large (*NDH*), Very Large (*NDV*)
Crop yield (*CY*)	Very Low (*CYE*), Low (*CYL*), Medium (*CYM*), High (*CYH*), Very High (*CYV*)

was represented as a triangular MF, whose centre occurred at $\frac{j\times(R-L)}{2N}, j = 1, 2., 2N-1$, (except the shoulder MFs) and the remaining two ends at the centres of the adjacent linguistic terms. However, in our case, we cannot use these uniform triangular MFs, directly as we choose to model the word semantics using IT2 FSs.

In Mendel & Wu, (2010) (Ch 5, Example 5.1), authors proposed a way to generate interval about a precise data value, say λ, by assuming a displacement of δ on its both sides so that the resulting interval is given as $[\lambda-\delta, \lambda+\delta]$. Thus inspired by this methodology, we propose to convert the uniform T1 FSs to IT2 FSs.

Considering again that the information representation scale of a variable, with associated $2N+1$ linguistic terms, extends from L to R. The centre of each of the linguistic terms occurs $\frac{j\times(R-L)}{2N}, j = 1, 2., 2N-1$, except the left shoulder and right shoulder MF. Therefore, for the MFs (except the shoulder ones), we perform the following computations to transform the T1 MFs to IT2 MFs. Our newly formed IT2 FS will have a trapezoidal UMF and triangular LMF. The UMF will be described by the points a, b, c, d and LMF by e, f, g, μ_f, (collectively called the FOU parameters) as shown in Figure 4 (Please see Section 3.2). To calculate the points describing the UMF and LMF, we assume a displacement of $\delta = 1.0$. Now the calculation procedure for the j^{th} linguistic term is given in Equation (25) as:

$$\begin{aligned} a &= \frac{(j-1)\times(R-L)}{2N},\ b = \frac{j\times(R-L)}{2N} - \delta c = \frac{j\times(R-L)}{2N} + \delta, \\ d &= \frac{(j+1)\times(R-L)}{2N} e = \frac{j\times(R-L)}{2N} - \delta,\ f = \frac{j\times(R-L)}{2N} \\ g &= \frac{j\times(R-L)}{2N} + \delta, \mu_f = 0.5 \end{aligned} \tag{25}$$

Here, $\frac{j\times(R-L)}{2N}$, $\frac{(j-1)\times(R-L)}{2N}$ and $\frac{(j+1)\times(R-L)}{2N}$ is the centre of the j^{th}, $(j-1)^{th}$ and $(j+1)^{th}$ T1 MF, respectively as represented in Wang & Mendel, (1992). It is pertinent to mention that we have assumed the secondary MF to be 1 everywhere. Thus, this transformation from T1 MF to IT2 MF is shown in Figure 7. Also, for the Left and right Shoulder MFs, the FOU parameters are given as:

$$\begin{aligned} &\textit{Left Shoulder}: \\ &a = 0.0,\ b = 0.0,\ c = 0.0,\ d = \frac{R-L}{2N} \\ &e = 0.0,\ f = 0.0,\ g = \frac{R-L}{2N} - \delta, \mu_f = 1.0 \\ &\textit{RightShoulder}: \\ &a = \frac{(2N-1)\times(R-L)}{2N},\ b = 10.0,\ c = 10.0,\ d = 10.0 \\ &e = \frac{(2N-1)\times(R-L)}{2N} + \delta,\ f = 10.0, g = 10.0,\ \mu_f = 1.0 \end{aligned} \tag{26}$$

In our present case, we have assumed the information representation scale as 0 to 10 for all the variables. Therefore $l_w^{ip} = 0$ and $r_w^{ip} = 10, w = 1, 2., 11$ and $l_1^{op} = 0$ and $r_1^{op} = 10$. The number of linguistic terms corresponding to each variable is $2N+1 = 5$. Let's denote these linguistic terms for input variables as: $L_1^{ip_w}, L_2^{ip_w}, L_3^{ip_w}, L_4^{ip_w}, L_5^{ip_w}, w = 1, 2., 11$ and for output variable as: $L_1^{op_1}, L_2^{op_1}, L_3^{op_1}, L_4^{op_1}, L_5^{op_1}$. Of these, $L_1^{ip_w}, w = 1, 2., 11$ and $L_1^{op_1}$ are the left shoulder MFs, as well as $L_5^{ip_w}, w = 1, 2., 11$ and $L_5^{op_1}$ are the right shoulder MFs. The other linguistic terms viz., $L_2^{ip_w}, L_3^{ip_w}, L_4^{ip_w}, w = 1, 2., 11, L_2^{op_1}, L_3^{op_1}, L_4^{op_1}$ are all interior MFs (non shoulder). Thus, using Equations 25 and 26) and $\delta = 1.0$, we get the FOU parameters as:

$$\begin{aligned}
&Non-shoulder(Interior): L_2^{ip_w}, w=1,2.,11, L_2^{op_1}\\
&a=0.0,\ b=1.5,\ c=3.5,\ d=5.0\\
&e=1.5,\ f=2.5, g=3.5,\ \mu_f=0.5\\
&Non-shoulder(Interior): L_3^{ip_w}, w=1,2.,11, L_3^{op_1}\\
&a=2.5,\ b=4.0,\ c=6.0,\ d=7.5\\
&e=4.0,\ f=5.0,\ g=6.0,\ \mu_f=0.5\\
&Non-shoulder(Interior): L_4^{ip_w}, w=1,2.,11, L_4^{op_1}\\
&a=5.0,\ b=6.5,\ c=8.5,\ d=10\\
&e=6.5,\ f=7.5,\ g=8.5,\ \mu_f=0.5\\
&LeftShoulder: L_1^{ip_w}, w=1,2.,11, L_1^{op_1}\\
&a=0.0, b=0.0, c=0.0, d=2.5\\
&e=0.0, f=0.0, g=1.5, \mu_f=1.0\\
&RightShoulder: L_5^{ip_w}, w=1,2.,11, L_5^{op_1}\\
&a=7.5,\ b=10.0,\ c=10.0,\ d=10.0\\
&e=8.5,\ f=10.0,\ g=10.0,\ \mu_f=1.0
\end{aligned} \tag{27}$$

The FOU plots for the linguistic terms corresponding to all the variables are shown in Figure 8 and the corresponding FOU data in Table 2.

Step 2 fuzzy rule generation from the numeric data pairs

In this step we generate fuzzy IF-THEN rules based on the stream of numeric data values. To illustrate this step, consider a numeric data pair corresponding to the input-output variables given as: {*LA*,*LO*, *AM*, *AI*, *DP*, *HH*, *PE*, *TM*, *TI*, *WB*, *ND*, *CY*}= {9.95, 4.58, 5.11, 7.91, 7.12, 5.65, 6.61, 7.32, 7.59, 2.37, 5.27, 4.26}. Consider the illustration using data value $LO = 4.58$. Thus, based on the location of this data point on the *x*-axis, this data value can belong to maximum of two linguistic terms viz., *LOL* and *LOM*. The degree of membership of $LO = 4.58$ in *LOL* is $[0, 0.28]$ and in *LOM* is $[0.29, 1.0]$. As the semantics of the linguistic terms are represented using the IT2 FSs, therefore, using Equation (14), we get $\mu_{LOL}^{COG}(4.58) = 0.14$ and $\mu_{LOM}^{COG}(4.58) = 0.65$. Therefore, $LO = 4.58$ belongs to linguistic term *LOM*.

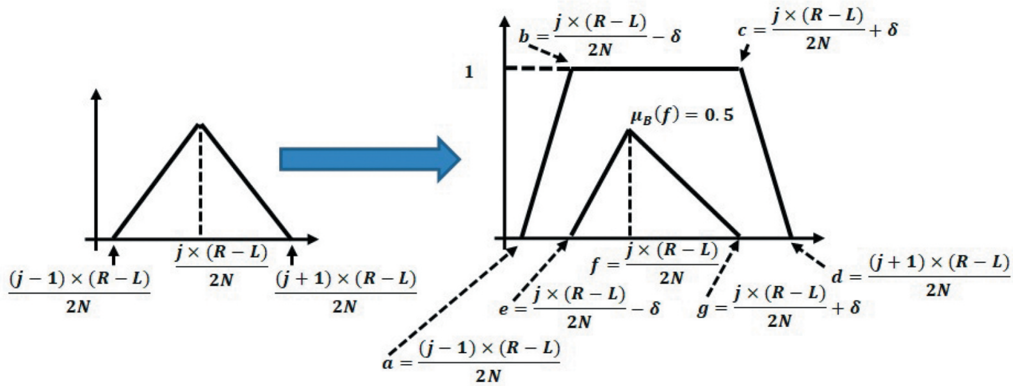

Figure 7. Converting T1 MF to IT2 MF.

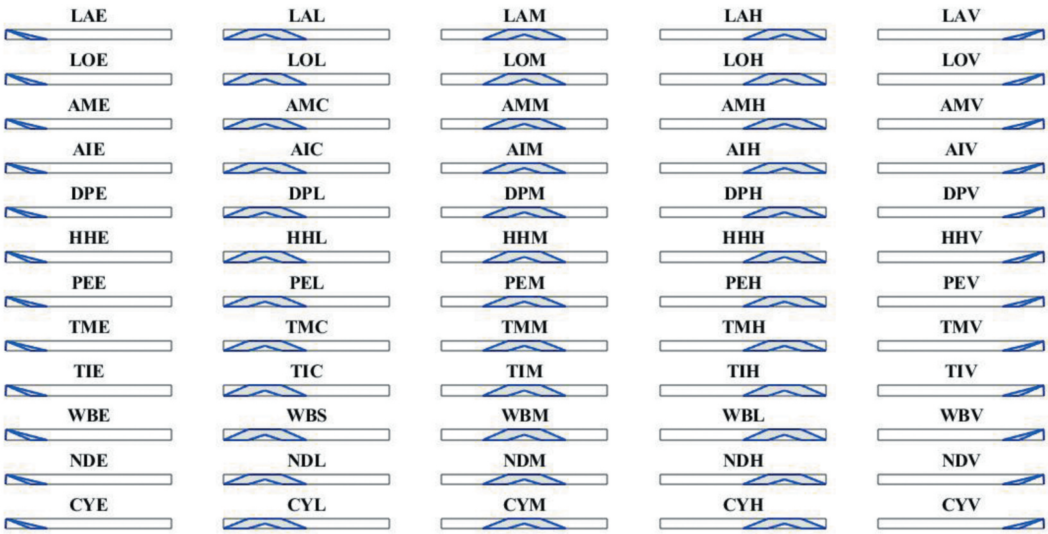

Figure 8. FOU plots for the linguistic terms of variables.

Similarly, the degree of belongingness of $LA = 9.95$, $AM = 5.11$, $AI = 7.91$, $DP = 7.12$, $HH = 5.65$, $PE = 6.61$, $TM = 7.32$, $TI = 7.59$, $WB = 2.37$, $ND = 5.27$, and $CY = 4.26$ is found respectively in *LAV*, *AMM*, *AIH*, *DPH*, *HHM*, *PEH*, *TMH*, *TIH*, *WBS*, *NDM*, and *CYM* with respective degrees $[0.96, 0.98]$, $[0.44, 1.0]$, $[0.3, 1.0]$, $[0.31, 1.0]$, $[0.18, 1.0]$, $[0.05, 1.0]$, $[0.41, 1.0]$, $[0.45, 1.0]$, $[0.43, 1.0]$, $[0.36, 1.0]$, and $[0.13, 1.0]$. Therefore based on this data vector, the IF-THEN rule (similar to Equation (15)) generated is given as:

$$\begin{aligned}&Rule\ j: IF x_1^{(j)}\ is\ LAV\ and\ x_2^{(j)}\ is\ LOM\ and\ x_3^{(j)}\ is\ AMM\ and\ x_4^{(j)}\ is\\&AIH\ and\ x_5^{(j)}\ is\ DPH\ and\ x_6^{(j)}\ is\ HHM\ and\ x_7^{(j)}\ is\ PEH\ and\ x_8^{(j)}\ is\\&TMH\ and\ x_9^{(j)}\ is\ TIH\ and\ x_{10}^{(j)}\ is\ WBS\ and\ x_{11}^{(j)}\ is\ NDM,\\&THEN\ y_1^{(j)}\ CYM\end{aligned} \tag{28}$$

Step 3 degree assignment to each rule

Now we will calculate the degree of the rule (given in Equation (28)), according to Equation (16). We will assume that the expert opinion is uniform and therefore the $\mu_{exp}(D_{Rule\ j}) = 1$, and thus we will only use Equation (16). Therefore, using the data from above step, the degree of the *Rule j* from Equation (28) is given as:

$$\begin{aligned}\mu_{D_{Rule\ j}} &= \left[\underline{\mu}_{D_{Rule\ j}}, \bar{\mu}_{D_{Rule\ j}}\right] = [0.96, 0.98] \times [0.29, 1.0] \times [0.44, 1.0]\times\\&[0.3, 1.0] \times [0.31, 1.0] \times [0.18, 1.0] \times [0.05, 1.0] \times [0.41, 1.0]\times\\&[0.45, 1.0] \times [0.43, 1.0] \times [0.36, 1.0] \times [0.13, 1.0] = [0.0, 0.98]\end{aligned} \tag{29}$$

As the left end of the interval given in Equation (29) has a value 0.0, therefore using Equation (17), we get the COG of the rule in Equation (29) as:

$$D_{Rule\ j} = \mu^{COG}_{D_{Rule\ j}} = \frac{1}{2}\left[\underline{\mu}_{D_{Rule\ j}} + \bar{\mu}_{D_{Rule\ j}}\right] = \frac{1}{2}[0.0 + 0.98] = 0.49 \tag{30}$$

Table 2. FOU data for the linguistic terms of the variables.

		FOU data										
		UMF				LMF				Centroid		
Variables	Associated Linguistic terms*	a	b	c	d	e	f	g	μ_f	c_l	c_r	c_{avg}
	LAE	0.0	0.0	0.0	2.5	0.0	0.0	1.5	1.0	0.5	0.86	0.68
	LAL	0.0	1.5	3.5	5.0	1.5	2.5	3.5	0.5	1.59	3.4	2.5
Latitude (*LA*)	*LAM*	2.5	4.0	6.0	7.5	4.0	5.0	6.0	0.5	4.09	5.90	5.0
	LAH	5.0	6.5	8.5	10.0	6.5	7.5	8.5	0.5	6.59	8.4	7.5
	LAV	7.5	10.0	10.0	10.0	8.5	10.0	10.0	1.0	9.13	9.49	9.3
	LOE	0.0	0.0	0.0	2.5	0.0	0.0	1.5	1.0	0.5	0.86	0.68
	LOL	0.0	1.5	3.5	5.0	1.5	2.5	3.5	0.5	1.59	3.4	2.5
Longitude (*LO*)	*LOM*	2.5	4.0	6.0	7.5	4.0	5.0	6.0	0.5	4.09	5.90	5.0
	LOH	5.0	6.5	8.5	10.0	6.5	7.5	8.5	0.5	6.59	8.4	7.5
	LOV	7.5	10.0	10.0	10.0	8.5	10.0	10.0	1.0	9.13	9.49	9.3
	AME	0.0	0.0	0.0	2.5	0.0	0.0	1.5	1.0	0.5	0.86	0.68
Apparent Temperature Max (*AM*)	*AMC*	0.0	1.5	3.5	5.0	1.5	2.5	3.5	0.5	1.59	3.4	2.5
	AMM	2.5	4.0	6.0	7.5	4.0	5.0	6.0	0.5	4.09	5.90	5.0
	AMH	5.0	6.5	8.5	10.0	6.5	7.5	8.5	0.5	6.59	8.4	7.5
	AMV	7.5	10.0	10.0	10.0	8.5	10.0	10.0	1.0	9.13	9.49	9.3
	AIE	0.0	0.0	0.0	2.5	0.0	0.0	1.5	1.0	0.5	0.86	0.68
Apparent Temperature Min (*AI*)	*AIC*	0.0	1.5	3.5	5.0	1.5	2.5	3.5	0.5	1.59	3.4	2.5
	AIM	2.5	4.0	6.0	7.5	4.0	5.0	6.0	0.5	4.09	5.90	5.0
	AIH	5.0	6.5	8.5	10.0	6.5	7.5	8.5	0.5	6.59	8.4	7.5
	AIV	7.5	10.0	10.0	10.0	8.5	10.0	10.0	1.0	9.13	9.49	9.3
	DPE	0.0	0.0	0.0	2.5	0.0	0.0	1.5	1.0	0.5	0.86	0.68
	DPL	0.0	1.5	3.5	5.0	1.5	2.5	3.5	0.5	1.59	3.4	2.5
Dew Point (*DP*)	*DPM*	2.5	4.0	6.0	7.5	4.0	5.0	6.0	0.5	4.09	5.90	5.0
	DPH	5.0	6.5	8.5	10.0	6.5	7.5	8.5	0.5	6.59	8.4	7.5
	DPV	7.5	10.0	10.0	10.0	8.5	10.0	10.0	1.0	9.13	9.49	9.3
	HHE	0.0	0.0	0.0	2.5	0.0	0.0	1.5	1.0	0.5	0.86	0.68
	HHL	0.0	1.5	3.5	5.0	1.5	2.5	3.5	0.5	1.59	3.4	2.5
Humidity (*HH*)	*HHM*	2.5	4.0	6.0	7.5	4.0	5.0	6.0	0.5	4.09	5.90	5.0
	HHH	5.0	6.5	8.5	10.0	6.5	7.5	8.5	0.5	6.59	8.4	7.5
	HHV	7.5	10.0	10.0	10.0	8.5	10.0	10.0	1.0	9.13	9.49	9.3
	PEE	0.0	0.0	0.0	2.5	0.0	0.0	1.5	1.0	0.5	0.86	0.68
	PEL	0.0	1.5	3.5	5.0	1.5	2.5	3.5	0.5	1.59	3.4	2.5
Pressure (*PE*)	*PEM*	2.5	4.0	6.0	7.5	4.0	5.0	6.0	0.5	4.09	5.90	5.0
	PEH	5.0	6.5	8.5	10.0	6.5	7.5	8.5	0.5	6.59	8.4	7.5
	PEV	7.5	10.0	10.0	10.0	8.5	10.0	10.0	1.0	9.13	9.49	9.3
	TME	0.0	0.0	0.0	2.5	0.0	0.0	1.5	1.0	0.5	0.86	0.68
	TMC	0.0	1.5	3.5	5.0	1.5	2.5	3.5	0.5	1.59	3.4	2.5
Temperature Max (*TM*)	*TMM*	2.5	4.0	6.0	7.5	4.0	5.0	6.0	0.5	4.09	5.90	5.0
	TMH	5.0	6.5	8.5	10.0	6.5	7.5	8.5	0.5	6.59	8.4	7.5
	TMV	7.5	10.0	10.0	10.0	8.5	10.0	10.0	1.0	9.13	9.49	9.3
	TIE	0.0	0.0	0.0	2.5	0.0	0.0	1.5	1.0	0.5	0.86	0.68
	TIC	0.0	1.5	3.5	5.0	1.5	2.5	3.5	0.5	1.59	3.4	2.5
Temperature Min (*TI*)	*TIM*	2.5	4.0	6.0	7.5	4.0	5.0	6.0	0.5	4.09	5.90	5.0
	TIH	5.0	6.5	8.5	10.0	6.5	7.5	8.5	0.5	6.59	8.4	7.5
	TIV	7.5	10.0	10.0	10.0	8.5	10.0	10.0	1.0	9.13	9.49	9.3
	WBE	0.0	0.0	0.0	2.5	0.0	0.0	1.5	1.0	0.5	0.86	0.68
	WBS	0.0	1.5	3.5	5.0	1.5	2.5	3.5	0.5	1.59	3.4	2.5
Wind Bearing (*WB*)	*WBM*	2.5	4.0	6.0	7.5	4.0	5.0	6.0	0.5	4.09	5.90	5.0
	WBL	5.0	6.5	8.5	10.0	6.5	7.5	8.5	0.5	6.59	8.4	7.5
	WBV	7.5	10.0	10.0	10.0	8.5	10.0	10.0	1.0	9.13	9.49	9.3
	NDE	0.0	0.0	0.0	2.5	0.0	0.0	1.5	1.0	0.5	0.86	0.68
	NDL	0.0	1.5	3.5	5.0	1.5	2.5	3.5	0.5	1.59	3.4	2.5
NDVI (*ND*)	*NDM*	2.5	4.0	6.0	7.5	4.0	5.0	6.0	0.5	4.09	5.90	5.0
	NDH	5.0	6.5	8.5	10.0	6.5	7.5	8.5	0.5	6.59	8.4	7.5
	NDV	7.5	10.0	10.0	10.0	8.5	10.0	10.0	1.0	9.13	9.49	9.3
	CYE	0.0	0.0	0.0	2.5	0.0	0.0	1.5	1.0	0.5	0.86	0.68
	CYL	0.0	1.5	3.5	5.0	1.5	2.5	3.5	0.5	1.59	3.4	2.5
Crop Yield (*CY*)	*CYM*	2.5	4.0	6.0	7.5	4.0	5.0	6.0	0.5	4.09	5.90	5.0
	CYH	5.0	6.5	8.5	10.0	6.5	7.5	8.5	0.5	6.59	8.4	7.5
	CYV	7.5	10.0	10.0	10.0	8.5	10.0	10.0	1.0	9.13	9.49	9.3

Note: *For extended forms of the words, please see Table 1.

Step 4 creation of a combined fuzzy rule base

We process all the data pairs (in the stream of data values) according to Step 3 to generate IF-THEN fuzzy rules. Now, for generating the combined fuzzy rule base we use the Step 4 given in Section 4, and remove the conflicting rules. We will discuss these results in Section 6.

Step 5 determining a mapping based on the combined fuzzy rule base

In this step, we calculate the value of the crop yield using the defuzzification equations in Step 5 of Section 4, pertaining to a data vector. For example, consider the data vector corresponding to the values of the input data variables given as: $\{LA, LO, AM, AI, DP, HH, PE, TM, TI, WB, ND\} = \{9.95, 4.58, 5.11, 7.91, 7.12, 5.65, 6.61, 7.32, 7.59, 2.37, 5.27\}$. Let's say this vector is used to fire the rules in the IT-THEN rule base and of all the rules, two IF-THEN rules gave non-zero firing values. Let these two rules be given as:

$$\begin{aligned}&Rule\ 1: IF x_1^{(j)}\ is\ LAV\ and\ x_2^{(j)}\ is\ LOM\ and\ x_3^{(j)}\ is\ AMM\ and\ x_4^{(j)}\ is\\&AIH\ and\ x_5^{(j)}\ is\ DPH\ and\ x_6^{(j)}\ is\ HHM\ and\ x_7^{(j)}\ is\ PEH\ and\ x_8^{(j)}\ is\\&TMH\ and\ x_9^{(j)}\ is\ TIH\ and\ x_{10}^{(j)}\ is\ WBS\ and\ x_{11}^{(j)} is\ NDM,\\&THEN\ y_1^{(j)}\ is\ CYM\end{aligned} \tag{31}$$

$$\begin{aligned}&Rule\ 2: IF x_1^{(j)}\ is\ LAH\ and\ x_2^{(j)}\ is\ LOL\ and\ x_3^{(j)}\ is\ AMH\ and\ x_4^{(j)}\ is\\&AIV\ and\ x_5^{(j)}\ is\ DPM\ and\ x_6^{(j)}\ is\ HHV\ and\ x_7^{(j)}\ is\ PEM\ and\ x_8^{(j)}\ is\\&TMM\ and\ x_9^{(j)}\ is\ TIV\ and\ x_{10}^{(j)}\ is\ WBE\ and\ x_{11}^{(j)}\ is\ NDL,\\&THEN\ y_1^{(j)}\ is\ CYL\end{aligned} \tag{32}$$

Thus, for the data vector given above, the degrees of belongingness of the input variables in the respective linguistic terms from the antecedent part of the rule from Equation (31) are found as: $[0.96, 0.98]$, $[0.29, 1.0]$, $[0.44, 1.0]$, $[0.3, 1.0]$, $[0.31, 1.0]$, $[0.18, 1.0]$, $[0.05, 1.0]$, $[0.41, 1.0]$, $[0.45, 1.0]$, $[0.43, 1.0]$, $[0.36, 1.0]$. The corresponding degrees of the input variables in the respective linguistic terms from the antecedent part of the rule from Equation (32) are found as: $[0.0, 0.03]$, $[0.0, 0.28]$, $[0.0, 0.07]$, $[0.0, 0.16]$, $[0.0, 0.25]$, $[0.0, 0.43]$, $[0.0, 0.6]$, $[0.0, 0.12]$, $[0.0, 0.03]$, $[0.0, 0.05]$, $[0.0, 0.18]$.

Now the degree of rule from Equation (31) corresponding to the inputs only is given (using Equation (20)) as:

$$\begin{aligned}&\mu_{inps\ Rule\ 1}^{(1)} = \left[\underline{\mu}_{inps\ Rule\ 1}^{(1)}, \bar{\mu}_{inps\ Rule\ 1}^{(1)}\right] = [0.96, 0.98] \times [0.29, 1.0] \times\\&[0.44, 1.0] \times [0.3, 1.0] \times [0.31, 1.0] \times [0.18, 1.0] \times [0.05, 1.0] \times [0.41, 1.0]\\&\times [0.45, 1.0] \times [0.43, 1.0] \times [0.36, 1.0] = [0.0, 0.98]\end{aligned} \tag{33}$$

As we can see from Equation (33), the left end of the interval has a value 0.0. Therefore, using Equation (23), the COG of the $\mu_{inps\ Rule\ 1}^{(1)}$ is found as:

$$\mu_{inps\ Rule\ 1}^{COG(1)} = \frac{1}{2}[0.0 + 0.98] = 0.49 \tag{34}$$

Similarly, the degree of rule from Equation (32) corresponding to the inputs only is given (using Equation (20)) as:

$$\begin{aligned}&\mu_{inps\ Rule\ 2}^{(2)} = \left[\underline{\mu}_{inps\ Rule\ 2}^{(2)}, \bar{\mu}_{inps\ Rule\ 2}^{(2)}\right] = [0.0, 0.03] \times [0.0, 0.28] \times\\&[0.0, 0.07] \times [0.0, 0.16] \times [0.0, 0.25] \times [0.0, 0.43] \times [0.0, 0.6] \times [0.0, 0.12]\\&\times [0.0, 0.03] \times [0.0, 0.05] \times [0.0, 0.18] = [0.0, 0.0]\end{aligned} \tag{35}$$

As again, we can see from Equation (35), the left (as well as right) end of the interval has a value 0.0. Therefore, using Equation (23), the COG of the $\mu^{(2)}_{inpsRule\ 2}$ is found as:

$$\mu^{COG(2)}_{inps\ Rule\ 2} = \frac{1}{2}[0.0 + 0.0] = 0.0 \tag{36}$$

Now consider the linguistic terms in the consequent parts of the Rules 1 and 2 given in Equations (31 and 32), respectively. These linguistic terms are *CYM* and *CYL*. The semantics of these linguistic terms are represented using the IT2 FSs (Please see Figure 8). Thus, the centroids of the *CYM* and *CYL* are found using the EKM algorithm and are respectively given as: $[4.09, 5.9]$ and $[1.59, 3.4]$. Thus, putting the values of the COGs for rule inputs from Equation (34), Equation (36) and these centroid values of the output variable into the Equation (24), we get the defuzzified value of the crop yield as:

$$\begin{aligned}\mu(y_q) &= \left[\underline{\mu}(y_q), \bar{\mu}(y_q)\right] = \frac{(0.49 \times [4.09, 5.9]) + (0.0 \times [1.59, 3.4])}{0.49 + 0.0} \\ &= \frac{[2.0, 2.89] + [0.0, 0.0]}{0.49} = \frac{[2.0, 2.89]}{0.49} = [4.09, 5.9]\end{aligned} \tag{37}$$

Finally, using Equation (22), the value of the crop yield is given as:

$$\mu^{COG}(y_q) = \frac{1}{2}[4.09 + 5.9] = 4.99 \tag{38}$$

Results and discussions

In this section we will present the results obtained by applicability of our Enhanced WMA to the agritech data, user satisfaction rating computation while using a battery operated device using the Enhanced WMA[4] as well as discuss important findings obtained from it.

Results with exemplar application on real-world agritech data

Our original dataset consisted of 359,427s data vectors. Each data vector was a unique combination of the numeric values for these 12 input-output variables (please see Table 1). When we processed these 359,427 data vectors using Step 2 of Section 4, we obtained 359427 number of IF-THEN rules, as each dataset generated one rule.

Now out of these 359,427 rules, many of them can be assumed to form equivalence classes. In an equivalence class, we categorise all those rules where the linguistic terms corresponding to the variables in the antecedent are the same, but the ones corresponding to the consequents are different. Therefore, by application of Step 3 of Section 4, we calculated the degree for each (of the 359,427) rule falling in an equivalence class. Then for constructing the combined IF-THEN rule base using Step 4 of Section 4, we selected only one rule from every equivalence class which had the highest degree. Thus, the combined IF-THEN rule base consisted of now 6279 unique rules.

We took 300 test sample data vectors from the agritech data and predicted the crop yield for these data vectors. The mean value of the crop yield (as predicted) by Enhanced WMA was found to be 2.4, with a standard deviation of 0.19. Thus, with 95% confidence, we predicted the yield value to lie within $[2.39, 2.43]$. When compared to the actual yield values, the mean and standard deviation were respectively 3.65 and 2.26. With 95% confidence, the yield values lie within $[3.38, 3.90]$. Thus, clearly, it can be seen that the data interval around the crop yield values as predicted by Enhanced WMA is narrower when compared to the one around the actual data values. As it follows from (Tan & Tan, 2010) that a narrower interval means higher accuracy, hence our Enhanced WMA has higher accuracy than those obtained with the actual data values.

Further, we can also generate linguistic recommendations for the crop yield values. Given a numeric crop yield value (as shown in Equation (38)), we can calculate the distance of this value from the centroids of the linguistic terms for the crop yield as given in Table 1, and then recommend the linguistic term with the smallest distance.

We also compared the results obtained by our Enhanced WMA to those obtained with the existing T1 FS-based WMA (Wang & Mendel, 1992). When applied to the agritech data, the T1 FS based WMA generated 6556 unique rules against the 6279 generated by our proposed WMA. When we subjected the same 300 agritech data vectors (as used for Enhanced WMA in Section 6.1) for crop yield prediction using the T1 FS based WMA, we found the mean yield value to be 2.58 with a standard deviation of 0.51. With 95% confidence, we predicted the yield value to lie within $[2.52, 2.64]$. Clearly, the band interval spread of the T1 FS based WMA is lesser than that of the original dataset but more than that of the Enhanced WMA. Therefore, our Enhanced WMA achieves better precision and that too using a smaller sized rule base.

Results on enhanced WMA for exemplar real-world application on computing users' satisfaction ratings

In this subsection, we present the use of Enhanced WMA for user satisfaction computation while using a battery operated device, using the dataset available in (Wang & Mendel, 1992), Chapter 3). In the Chapter 3 of this thesis, a novel power management policy has been proposed for the battery operated devices. The users were asked to provide feedback about the perceived system performance using the linguistic terms associated to four system parameters viz., Battery Life, Type of Application, Amount of time spent and Application ratings. Each of these parameters had an associated five linguistic variables. These are given in the Table 3. Using the Enhanced WMA, their FOUs were constructed which are defined by the FOU data as given in Table 4.

The dataset containing 810,000 data vectors was constructed to test the performance of Enhanced WMA against the original T1 FS based one. Each data vector was unique. Hence, processing these data vectors we obtained 810,000 number of IF-THEN rules. Of these rules, many were of the form where the linguistic terms corresponding to the variables in the antecedent are the same, but the ones corresponding to the consequents were different. Therefore, by application of Step 3 of Section 4, the degree for each rule was calculated and combined IF-THEN rule base was constructed using Step 4 of Section 4.

We took 500 test sample data vectors from the dataset and predicted the user satisfaction rating for these data vectors. The mean value of the user satisfaction computer by Enhanced WMA was found to be 6.49, with a standard deviation of 1.42. Thus, with 95% confidence, we predicted the yield value to lie within $[5.07, 7.98]$. When compared to the existing T1 FS based WMA, the mean and standard deviation were respectively 5.36 and 2.23. With 95% confidence, the yield values lie within $[3.11, 7.58]$. Thus, clearly, it can be seen that the data interval around the crop yield values as predicted by Enhanced WMA is narrower when compared to the one around the actual data values. As it follows from (Tan & Tan, 2010) that a narrower interval means higher accuracy, hence our Enhanced WMA has higher accuracy than those obtained with the actual data values. Also, from the 810,000 data vectors, the Enhanced WMA generated much lesser unique rules than the existing T1 FS based WMA.

Table 3. Input-output variables and associated linguistic terms.

Variables	Associated Linguistic Terms
Battery Life	Very Low (*BVL*), Low (*BL*), Medium (*BM*), High (*BH*), Very High (*BEH*)
Type of Application	Absolutely uninteresting (*AU*), Somewhat interesting (*SI*), Fairly Interesting (*FI*), More interesting (*MI*), Absolutely interesting (*AI*)
Amount of Time spent	Very Little (*VL*), Small (*S*), Moderate (*M*), Large (*L*), Very Large (*VLA*)
Application Ratings	Very Slow (*AVS*), Slow (*AS*), Moderate (*AM*), Fast (*AF*), Extremely Fast (*AEF*),

Table 4. FOU data for the linguistic terms of the variables.

Variables	Associated Linguistic terms*	FOU data: UMF a	b	c	d	LMF e	f	g	μ_f	Centroid c_l	c_r	c_{avg}
Battery Life	*BVL*	0.0	0.0	0.0	2.5	0.0	0.0	1.5	1.0	0.5	0.86	0.68
	BL	0.0	1.5	3.5	5.0	1.5	2.5	3.5	0.5	1.59	3.4	2.5
	BM	2.5	4.0	6.0	7.5	4.0	5.0	6.0	0.5	4.09	5.90	5.0
	BH	5.0	6.5	8.5	10.0	6.5	7.5	8.5	0.5	6.59	8.4	7.5
	BEH	7.5	10.0	10.0	10.0	8.5	10.0	10.0	1.0	9.13	9.49	9.3
Type of Application	*AU*	0.0	0.0	0.0	2.5	0.0	0.0	1.5	1.0	0.5	0.86	0.68
	SI	0.0	1.5	3.5	5.0	1.5	2.5	3.5	0.5	1.59	3.4	2.5
	FI	2.5	4.0	6.0	7.5	4.0	5.0	6.0	0.5	4.09	5.90	5.0
	MI	5.0	6.5	8.5	10.0	6.5	7.5	8.5	0.5	6.59	8.4	7.5
	AI	7.5	10.0	10.0	10.0	8.5	10.0	10.0	1.0	9.13	9.49	9.3
Amount of Time Spent	*VL*	0.0	0.0	0.0	2.5	0.0	0.0	1.5	1.0	0.5	0.86	0.68
	S	0.0	1.5	3.5	5.0	1.5	2.5	3.5	0.5	1.59	3.4	2.5
	M	2.5	4.0	6.0	7.5	4.0	5.0	6.0	0.5	4.09	5.90	5.0
	L	5.0	6.5	8.5	10.0	6.5	7.5	8.5	0.5	6.59	8.4	7.5
	VLA	7.5	10.0	10.0	10.0	8.5	10.0	10.0	1.0	9.13	9.49	9.3
Application Ratings	*AVS*	0.0	0.0	0.0	2.5	0.0	0.0	1.5	1.0	0.5	0.86	0.68
	AS	0.0	1.5	3.5	5.0	1.5	2.5	3.5	0.5	1.59	3.4	2.5
	AM	2.5	4.0	6.0	7.5	4.0	5.0	6.0	0.5	4.09	5.90	5.0
	AF	5.0	6.5	8.5	10.0	6.5	7.5	8.5	0.5	6.59	8.4	7.5
	AEF	7.5	10.0	10.0	10.0	8.5	10.0	10.0	1.0	9.13	9.49	9.3

*For extended forms of the words, please see Table 3.

Discussion

Now we discuss the important findings obtained on our study of the use and results of the proposed novel Enhanced WMA to real-world data.

It is pertinent to mention that we have proposed a simplified method of generating IT2 MFs for the linguistic terms (or words) in Step 1 of Section 5, which we have subsequently used in the Enhanced WMA. There exist other methods in the literature too for generating the MFs for the words. For example, the works (Choi & Rhee, 2009; Medasani et al., 1998) present various techniques for MF generation like the use of heuristics, neural networks, etc. However, all these techniques may not be computationally efficient.

Further, in the works (Hao & Mendel, 2015; Mendel & Wu, 2010; Wu et al., 2011), novel methods have been proposed to generate IT2 MFs by collecting data from group of people and subjecting it to various data processing steps. These works are suitable for modelling the subjectivity of linguistic terms, however, data collection is a time-consuming task. Further, many subjects don't take the data collection seriously and don't provide genuine feedback.

However, with the advent of new IoT technologies (Rekha et al., 2017), data collection from sensors has become very fast, efficient, and seamless. Therefore, this motivated us to propose a novel way of generating IT2 MFs, without using the computationally expensive methods as well as collecting data from group of users. Still, we feel that the IT2 MFs generated by us are based on the opinions of a single expert. Scope of improvement exists for putting worth a better methodology for generating IT2 MFs, which model the linguistic uncertainty in a much more data-adaptive way.

Conclusions and future scope

In this paper, we have proposed a novel Enhanced WMA that models the semantics of linguistic terms using T2 FSs, improving the existing WMA based on T1 FSs. T2 FSs model the linguistic uncertainty in a better manner than the T1 FSs. We have demonstrated the working of our Enhanced

WMA for crop yield prediction in smart agriculture (through the case study of real-life agritech data). We have also compared the crop yield prediction results obtained by our Enhanced WMA to those obtained with the existing T1 FS-based WMA and the ones given in the original dataset. We found that our Enhanced WMA has a greater precision owing to smaller interval width for confidence level of 95%, when compared to the existing T1 FS-based WMA and the original dataset.

Further, our method gives a very simplified way of generating the T2 FS MFs, without using computationally expensive techniques as well as data collection from a group of people. However, we feel that the T2 MFs are mainly generated using limited expert knowledge.

In the future, we aim to provide a method to generate T2 MFs from a stream of data values that are more in sync with data and model the linguistic uncertainty better.

Notes

1. The concept of FSs was proposed in 1965 by Prof. Zadeh (Zadeh, 1965), as an extension of classical (or crisp) sets. FSs are used to classify objects in real-life scenarios, where no sharp classification boundaries exist. Further, FSs have an inherent capability to model the linguistic uncertainty in a manner similar to the human cognitive process and human beings naturally understand (and express themselves) linguistically. The FSs make use of membership function (MF) or membership degrees to model the linguistic (or data) uncertainty (we will discuss the details in section 3).
2. The T1 FSs proposed in (Zadeh, 1965) had a precise degree of membership and therefore contradicted the very notion a "fuzzy quantity being precise". Therefore, later higher order FSs were proposed by Prof. Zadeh in (Zadeh, 1975), thereby leading to former (the ones in the work (Zadeh, 1965)) being addressed as the Type-1 (T1) FSs. A category of the higher order FSs from (Zadeh, 1975) is called the Type-2 (T2) FSs. The T2 FSs have a greater capability to model the linguistic uncertainty through their membership degrees (Hagras & Wagner, 2012).
3. In (Wang & Mendel, 1992), these regions were named as $SmallN., Small1, Center, Big1 \ldots, BigN$. However, we have chosen to keep our terminology general.
4. For purpose of illustration only as smart agriculture is the main focus of this paper.

Disclosure statement

No potential conflict of interest was reported by the author(s).

ORCID

Prashant K. Gupta http://orcid.org/0000-0003-4220-1954
Javier Andreu-Perez http://orcid.org/0000-0002-7421-4808

References

Alhabashneh, O. (2021). Fuzzy-based adaptive framework for module advising expert system. *Annals of Emerging Technologies in Computing (AETiC)*, *5*(1), 13–27. https://doi.org/10.33166/AETiC.2021.01.002

Alonso Moral, J. M., Castiello, C., Magdalena, L., & Mencar, C. (2021). Designing interpretable fuzzy systems. In *Explainable fuzzy systems* (pp. 119–168). Springer.

Amaruchkul, K. (2021). Planning migrant labor for green sugarcane harvest: A stochastic logistics model with dynamic yield prediction. *Computers & Industrial Engineering*, *154*, 107016. https://doi.org/10.1016/j.cie.2020.107016

Casillas, J., Cordón, O., & Herrera, F. (2000). Improving the wang and mendel's fuzzy rule learning method by inducing cooperation among rules. In *Proceedings of the 8th information processing and management of uncertainty in knowledge-based systems conference, Milan, Italy* (pp. 1682–1688).

Choi, B. -I., & Rhee, F. C. -H. (2009). Interval type-2 fuzzy membership function generation methods for pattern recognition. *Information Sciences*, *179*(13), 2102–2122. https://doi.org/10.1016/j.ins.2008.04.009

Čubranić-Dobrodolac, M., Švadlenka, L., Marković, G. Z., & Dobrodolac, M. (2021). A decision support model for transportation companies to examine driver behavior. *IEEE Transactions on Engineering Management*.

Díaz-Pacheco, A., & Reyes-Garcia, C. A. (2021). A classification-based fuzzy-rules proxy model to assist in the full model selection problem in high volume datasets. *Journal of Experimental & Theoretical Artificial Intelligence*, *34*(5), 815-844.

Gao, F., Zhang, A., Bi, W., & Ma, J. (2021). A greedy belief rule base generation and learning method for classification problem. *Applied Soft Computing*, *98*, 106856.ssssf. https://doi.org/10.1016/j.asoc.2020.106856

Gou, J., Hou, F., Chen, W., Wang, C., & Luo, W. (2015). Improving wang–mendel method performance in fuzzy rules generation using the fuzzy c-means clustering algorithm. *Neurocomputing*, *151*, 1293–1304. https://doi.org/10.1016/j.neucom.2014.10.077

Gupta, P. (2019). Perceptual computing for power management, linguistic optimization and decision making. *PhD Thesis*.

Gupta, P. K., & Andreu-Perez, J. (2022). A gentle introduction and survey on computing with words (cww) methodologies. *Neurocomputing*, *500*, 921–937. https://doi.org/10.1016/j.neucom.2022.05.097

Gupta, P. K., Sharma, D., & Andreu-Perez, J. (2021). Enhanced linguistic computational models and their similarity with yager's computing with words. *Information Sciences*, *574*, 259–278. https://doi.org/10.1016/j.ins.2021.05.038

Hagras, H. (2018). Toward human-understandable, explainable ai. *Computer*, *51*(9), 28–36. https://doi.org/10.1109/MC.2018.3620965

Hagras, H., & Wagner, C. (2012). Towards the wide spread use of type-2 fuzzy logic systems in real world applications. *IEEE Computational Intelligence Magazine*, *7*(3), 14–24. https://doi.org/10.1109/MCI.2012.2200621

Hao, M., and Mendel, J. M. (2013). Extracting if-then rules from numerical data using wang-mendel methods. In *Spe western regional & aapg pacific section meeting 2013 joint technical conference*, *Monterey, California, USA*

Hao, M., & Mendel, J. M. (2015). Encoding words into normal interval type-2 fuzzy sets: Hm approach. *IEEE Transactions on Fuzzy Systems*, *24*(4), 865–879. https://doi.org/10.1109/TFUZZ.2015.2486814

Jara, L., Ariza-Valderrama, R., Fernández-Olivares, J., González, A., & Pérez, R. (2022). Efficient inference models for classification problems with a high number of fuzzy rules. *Applied Soft Computing*, *115*, 108164. https://doi.org/10.1016/j.asoc.2021.108164

Klir, G., & Yuan, B. (1995). Fuzzy sets and fuzzy logic. *Prentice Hall New Jersey*, *4*.

Krishnan, R. S., Julie, E. G., Robinson, Y. H., Raja, S., Kumar, R., Thong, P. H., & Son, L. H. (2020). Fuzzy logic based smart irrigation system using internet of things. *Journal of Cleaner Production*, *252*, 119902.

Lucchese, L. V., de Oliveira, G. G., & Pedrollo, O. C. (2021). Mamdani fuzzy inference systems and artificial neural networks for landslide susceptibility mapping. *Natural Hazards*, *106*(3), 2381–2405. https://doi.org/10.1007/s11069-021-04547-6

Medasani, S., Kim, J., & Krishnapuram, R. (1998). An overview of membership function generation techniques for pattern recognition. *International Journal of Approximate Reasoning*, *19*(3–4), 391–417. https://doi.org/10.1016/S0888-613X(98)10017-8

Mendel, J. M. (2001). The perceptual computer: An architecture for computing with words. In *10th ieee international conference on fuzzy systems, Melbourne, Australia* (Vol. 1, pp. 35–38).

Mendel, J., & Wu, D. (2010). *Perceptual computing: Aiding people in making subjective judgments* (Vol. 13). John Wiley & Sons.

Navarro-Almanza, R., Sanchez, M. A., Castro, J. R., Mendoza, O., & Licea, G. (2022). Interpretable Mamdani neuro-fuzzy model through context awareness and linguistic adaptation. *Expert Systems with Applications*, *189*, 116098. https://doi.org/10.1016/j.eswa.2021.116098

Rekha, P., Rangan, V. P., Ramesh, M. V., & Nibi, K. (2017). High yield groundnut agronomy: An iot based precision farming framework. In *2017 ieee global humanitarian technology conference (ghtc), Toronto, Canada* (pp. 1–5).

Restović, I. (2020). Why not fuzzy logic? In Sandro Skansi (ed.), *Guide to deep learning basics* (pp. 29–40). Springer.

Serrano-Guerrero, J., Romero, F. P., & Olivas, J. A. (2021). Fuzzy logic applied to opinion mining: A review. *Knowledge-Based Systems*, *222*, 107018. https://doi.org/10.1016/j.knosys.2021.107018

Tan, S. H., & Tan, S. B. (2010). The correct interpretation of confidence intervals. *Proceedings of Singapore Healthcare*, *19*(3), 276–278. https://doi.org/10.1177/201010581001900316

Ustundag, A., & Cevikcan, E. (2017). *Industry 4.0: Managing the digital transformation*. Springer.

Van Harmelen, F., & Teije, A. T. (2019). A boxology of design patterns for hybrid learning and reasoning systems Journal of Web Engineering 18 Combined Issue 1, 2 & 3 97–124 .

Wang, L. -X., & Mendel, J. M. (1992). Generating fuzzy rules by learning from examples. *IEEE Transactions on Systems, Man, and Cybernetics*, *22*(6), 1414–1427. https://doi.org/10.1109/21.199466

Wang, Y., Wang, M., Liu, Y., Yin, L., Zhou, X., Xu, J., & Zhang, X. (2021). Fuzzy modeling of boiler efficiency in power plants. *Information Sciences*, *542*, 391–405. https://doi.org/10.1016/j.ins.2020.06.064

Wikipedia contributors. (2020). *Type-2 fuzzy sets and systems — Wikipedia, the free encyclopedia*. Retrieved November 27, 2020, from https://en.wikipedia.org

Wu, D., Mendel, J. M., & Coupland, S. (2011). Enhanced interval approach for encoding words into interval type-2 fuzzy sets and its convergence analysis. *IEEE Transactions on Fuzzy Systems*, *20*(3), 499–513.

Yang, L. -H., Liu, J., Ye, F. -F., Wang, Y. -M., Nugent, C., Wang, H., & Martínez, L. (2022). Highly explainable cumulative belief rule-based system with effective rule-base modeling and inference scheme. *Knowledge-Based Systems*, *240*, 107805. https://doi.org/10.1016/j.knosys.2021.107805

Yang, X., Yuan, J., Yuan, J., & Mao, H. (2010). An improved wm method based on pso for electric load forecasting. *Expert Systems with Applications*, *37*(12), 8036–8041. https://doi.org/10.1016/j.eswa.2010.05.085

Zadeh, L. A. (1965). Fuzzy sets. *Information and Control*, *8*(3), 338–353. https://doi.org/10.1016/S0019-9958(65)90241-X

Zadeh, L. A. (1975). The concept of a linguistic variable and its application to approximate reasoning—i. *Information Sciences*, *8*(3), 199–249. https://doi.org/10.1016/0020-0255(75)90036-5

Zhai, Y., Lv, Z., Zhao, J., Wang, W., & Leung, H. (2021). Data-driven inference modeling based on an on-line Wang-Mendel fuzzy approach. *Information Sciences*, *551*, 113–127. https://doi.org/10.1016/j.ins.2020.10.018