International Journal of Approximate Reasoning 54 (2013) 273–289

Contents lists available at SciVerse ScienceDirect

International Journal of Approximate Reasoning

journal homepage: www.elsevier.com/locate/ijar

A 2uFunction representation for non-uniform type-2 fuzzy sets: Theory and design

Seyyedeh Fatemeh Molaeezadeh, Mohammad Hassan Moradi *

Department of Biomedical Engineering, Amirkabir University of Technology (Tehran Polytechnic), Tehran 15875-4413, Iran

ARTICLE INFO

Article history:
Received 3 December 2011
Received in revised form 17 August 2012
Accepted 28 August 2012
Available online 11 September 2012

Keywords:
Non-uniform type-2 fuzzy sets
Non-uniform random variables
2uFunction representation

ABSTRACT

The theoretical and computational complexities involved in non-uniform type-2 fuzzy sets (T2 FSs) are main obstacles to apply these sets to modeling high-order uncertainties. To reduce the complexities, this paper introduces a 2uFunction representation for T2 FSs. This representation captures the ideas from probability theory. By using this representation, any non-uniform T2 FS can be represented by a function of two uniform T2 FSs. In addition, any non-uniform T2 fuzzy logic system (FLS) can be indirectly designed by two uniform T2 FLSs. In particular, a 2uFunction-based trapezoid T2 FLS is designed. Then, it is applied to the problem of forecasting Mackey–Glass time series corrupted by two kinds of noise sources: (1) stationary and (2) non-stationary additive noises. Finally, the performance of the proposed FLS is compared by (1) other types of FLS: T1 FLS and uniform T2 FLS, and (2) other studies: ANFIS [54], IT2FNN-1 [54], T2SFLS [3] and Q-T2FLS [35]. Comparative results show that the proposed design has a low prediction error as well as is suitable for online applications.

1. Introduction

Type-2 fuzzy sets (T2 FSs) have the capability to model more uncertainties in rule-based fuzzy logic systems. They have the potential to outperform type-1 fuzzy sets (T1 FSs), because their membership grades are fuzzy. However, the computational and theoretical complexities in inference and type-reduction are major obstacles to the utility of type-2 fuzzy logic systems (T2 FLSs) in the real-world and on-line applications. Nevertheless, the wonderful benefits of these systems make them interesting objects for many researchers [1].

The major objectives of the researches on T2 FSs are to: (1) reduce the complexities, (2) exploit the potential of general T2 FSs (hereafter we call them non-uniform T2 FSs) for practical applications, and (3) develop fuzzy set theory. During two past decades, the researches focused more on simplest T2 FSs, i.e. interval T2 FSs (hereafter we call them uniform T2 FSs).

The researches on uniform T2 FSs can be divided into two main groups. First group aims to reduce the complexities in the inference. For example, Mendel et al. [2,3] used wavy slice representation [4] for defining efficient inference operations. They showed that these operations can be obtained by T1 operations. Coupland and John [5] introduced a geometric inference by the use of the techniques from computational geometry. Second group tries to decrease the complexities in the type-reduction. For example, Karnik and Mendel [6] proposed an iterative type-reduction, which is widely used in uniform T2 FLSs. Greenfield et al. [7] introduced a collapsing method for the type-reduction. By eliminating type-reduction, Wu and Mendel [8] defined a centre based on uncertainty bounds. As well, Coupland and John [5] suggested a geometric centre.

The above researches lead to implement uniform T2 FLSs in hardware and make them more applicable. To date, these systems have been used in various applications such as clustering [9], classification [10–16], person recognition [17–19],

* Corresponding author. Tel.: +98 21 64542399; fax: +98 21 66468186.
E-mail address: mhmoradi@aut.ac.ir (M.H. Moradi)

0888-613X/$ - see front matter
http://dx.doi.org/10.1016/j.ijar.2012.08.009

speech recognition [20,21], forecasting [3,22–25], estimation [26], control [17,27], adaptive noise cancellation [17,23], image segmentation [9], modeling [28], etc.

These applications demonstrated the significant improvements in the performance of uniform T2 FLSs over T1 FLSs. However, uniform T2 FSs cannot model convex uncertainties in primary membership grades. Hence, handling non-uniform and convex uncertainties is an object of growing attention in recent years.

Most of the researches on non-uniform T2 FSs are based on presenting new representations of these sets. These representations are briefly as follows. Zadeh [29] proposed an α *-level set representation* for non-uniform T2 FSs. This representation causes to generalize uniform FSs to non-uniform FSs. Hence, the computations of T2 operations can be easily done by uniform ones. Starczewski [30,31] introduced a *triangular representation* originating from triangular fuzzy truth numbers (triangular FTNs). Coupland et al. [5,32] suggested a *geometric representation* inspiring from computational geometry. This representation reduces the execution time of T2 operations. Lucas et al. [33] presented a *general footprint of uncertainty representation* (GFOU). This representation helps to simplify the understanding of non-uniform T2 FLSs. Liu et al. [34–36] introduced an α *-plane representation*, which is the extended version of alpha-cut for T1 FSs. Wagner et al proposed a *zSlice representation* [37,38] by slicing a non-uniform T2 FS in the third dimension.

Among these representations, the zSlice,α-plane, and triangular representations let us model non-uniform T2 FLSs in terms of uniform T2 FLSs. However, they can only provide a computationally efficient structure for handling the triangular and trapezoid uncertainties. To date, an efficient approach has not been considered for modeling every form of uncertainty such as Gaussian. This is the main motivation for this paper. Another motivation is the challenges involved in computing the set theoretic operations and the type-reduced set. These challenges make non-uniform T2 FSs (except Triangular T2 FSs) have no applications in rule-based fuzzy logic systems.

The existence of resemblances between theories of fuzzy set and probability leads us to use probability topics as the key to dealing with all non-uniform uncertainties. In the probability theory, it is difficult to directly generate non-uniform random numbers (RNs), while uniform RNs are simply generated [39,40]. Hence, in practice, a non-uniform RN is generated by uniform RNs. For example, a Gaussian RN can be obtained by a function of two uniform RNs [41]. Similarly, in fuzzy set theory, it is difficult to design non-uniform T2 FLSs. This is because of their computational and theoretical complexities. As a result, applications rarely employ these systems. However, there are many efficient methods to design uniform T2 FLSs. As well, these systems are extensively used in many applications.

As it can be seen, there are two similar problems in these two theories: the difficulty in generating non-uniform RVs and in analyzing non-uniform T2 FLSs. While the first problem has some solutions such as a function of two uniform RVs, the other is still an open problem. Therefore, this paper proposes a solution similar to the one proposed in probability. The proposed solution is to design a non-uniform T2 FLS by two uniform T2 FLSs. To do this, a new representation is presented for non-uniform T2 FSs called a 2uFunction representation. In brief, our main contributions are as follow:

- A new representation is introduced for non-uniform T2 FSs in such a way that a non-uniform FS is represented by a function of two uniform T2 FSs.
- A new efficient design is proposed for designing all non-uniform T2 FLSs based on only two uniform T2 FLSs.
- A 2uFunction-based trapezoid T2 FLS is applied to forecasting Mackey–Glass chaotic time series.

This paper is organized as follows. Section 2 briefly reviews all representations that have been introduced to date. Section 3 demonstrates why the concepts of probability theory lead to our proposed representation. Section 4 presents the proposed representation and its application for designing non-uniform T2 FLSs. Section 5 applies it to the problem of forecasting Mackey–Glass chaotic time series. Section 6 discusses the proposed representation. Finally, Section 7 gives conclusions and suggestions for future works.

2. A brief review on the representations of non-uniform T2 FSs

The methods for representing non-uniform T2 FSs have a great effect on developing non-uniform T2 FLSs and on reducing their complexities. Hence, the development of these methods is one of important topics in the researches related to T2 FSs. To date, several representation methods have been proposed. This section provides a brief review on these methods and their advantages and applications.

2.1. Vertical slice representation

Let $\tilde{A}$ denotes a non-uniform T2 FS; $\mu_{\tilde{A}}(x, u)$ denotes the membership function (MF) of $\tilde{A}$; and $\mu_{\tilde{A}}(x)$ denotes the secondary MF of $\tilde{A}$. Then, $\tilde{A}$ can be represented by its secondary MFs [27,40]:

$$\tilde{A} = \{(x, \mu_{\tilde{A}}(x)) | \forall x \in X\}. \quad (1)$$

or,

$$\tilde{A} = \int_{x\in X} \mu_{\tilde{A}}(x)/x = \int_{x\in X} \left[\int_{u\in J_x} f_x(u)/u\right]/x \quad J_x \subseteq [0, 1], \tag{2}$$

in which u indicates the primary membership grade of $\tilde{A}$; $f_x(u)$ indicates the secondary one; and J_x is the domain of $\mu_{\tilde{A}}(x)$.

This representation is very useful for T2 computations. It can also be useful for theoretical studies. However, by using this representation, computing T2 operations (such as meet, join and type-reduction) is still complex and time-consuming.

2.2. α-Level set representation

This representation includes two stages. First is to extend T1 definitions to uniform T2 definitions. Second is to generalize uniform T2 definitions to non-uniform T2 definitions by the use of the level-set form of the extension principle [29].

Assume that $\tilde{A}$ and $\tilde{B}$ are two non-uniform T2 FSs with secondary MFs $\mu_{\tilde{A}}(x)$ and $\mu_{\tilde{B}}(x)$; the α-level sets of these functions are denoted by $\mu^{\alpha}_{\tilde{A}}(x)$ and $\mu^{\alpha}_{\tilde{B}}(x)$. Now, using the representation, the intersection of these two sets is computed as follows. Firstly, $\mu^{\alpha}_{\tilde{A}\cap\tilde{B}}(x)$, the intersection of $\mu^{\alpha}_{\tilde{A}}(x)$ and $\mu^{\alpha}_{\tilde{B}}(x)$, is computed. If $\mu^{\alpha}_{\tilde{A}}(x) = [a_1, a_2]$ and $\mu^{\alpha}_{\tilde{B}}(x) = [b_1, b_2]$, then $\mu^{\alpha}_{\tilde{A}\cap\tilde{B}}(x)$ is computed as:

$$\mu^{\alpha}_{\tilde{A}\cap\tilde{B}}(x) = [a_1 \wedge b_1, a_2 \wedge b_2]. \tag{3}$$

Finally, the intersection of $\tilde{A}$ and $\tilde{B}$, i.e. $\mu_{\tilde{A}\cap\tilde{B}}(x)$, can be expressed as the union of $\mu^{\alpha}_{\tilde{A}\cap\tilde{B}}(x)$ [29]:

$$\mu_{\tilde{A}\cap\tilde{B}}(x) = \int_0^1 \alpha\mu^{\alpha}_{\tilde{A}\cap\tilde{B}}(x). \tag{4}$$

Comparing (3) and (4), it can be seen that all computations related to the intersection of two non-uniform T2 FSs can be indirectly done by T1 computations. Hence, this representation has the potential to reduce the complexities involved in T2 operations. It is very useful for T2 computations and theoretical studies. However, an encyclopedic theory was not presented for it. Also, no application has been reported for it.

2.3. Wavy slice representation

Let $\tilde{A}^j_e$ denote the jth T2 embedded set for a T2 FS $\tilde{A}$, i.e.,

$$\tilde{A}^j_e \equiv \left\{\left(u^j_i, f_{x_i}\left(u^j_i\right)\right),\ i = 1, \ldots, N\right\}, \tag{5}$$

in which $u^j_i \in \{u_{ik}, k = 1, \ldots, M_i\}$; N is the number of points on x-axis; and M_i is the number of points on the vertical slice x_i.

Then, $\tilde{A}$ can be represented as the union of its T2 embedded sets, i.e.,

$$\tilde{A} = \sum_{j=1}^{n} \tilde{A}^j_e, \tag{6}$$

in which $n \equiv \prod_{i=1}^{N} M_i$ [4].

This representation is not recommended for computational purposes, because the number of embedded sets is astronomical. However, it is very useful in the theoretical studies. By utilizing it, Mendel et al. [2] showed that uniform T2 operations can be obtained through two T1 MFs, namely upper and lower MFs. For example, the intersection of two uniform T2 FSs $\tilde{A}$ and $\tilde{B}$ is computed by:

$$\tilde{A} \cap \tilde{B} = 1/[\underline{\mu}_{\tilde{A}}(x) \wedge \underline{\mu}_{\tilde{B}}(x), \overline{\mu}_{\tilde{A}}(x) \wedge \overline{\mu}_{\tilde{B}}(x)], \quad \forall x \in X. \tag{7}$$

Note that (3) is obtained by using Zadeh's extension principle [42], but (7) is obtained by using wavy slice representation.

2.4. Geometric representation

This representation represents a non-uniform T2 FS by a collection of planar triangles. It utilizes the geometry to make the inference operations efficient. Consequently, it leads to the operations to be performed over a continuous domain [32,43]. It is applied for control of a mobile robot [44]. However, the geometric inference is somewhat slower than discrete inference. In addition, it is limited to the minimum t-norm and the maximum t-conorm. Moreover, by using this representation, the type-reduction is departed from T2 FLSs, and only the result of the defuzzification is available. Therefore, it may not be useful when more information is required about the output.

2.5. Triangular representation

This representation approximates the extended t-norms on triangular FTNs by FTNs of the triangular shape. It also approximates a triangular type-reduced set by the three-node linear interpolation of the general result for centroid type-reduction [30,31]. Using this representation, a triangular T2 FLS can be efficiently designed by a T1 FLS and a uniform T2 FLS. This system is applied to the classification of Iris and Wine datasets [45]. However, this representation is just presented for triangular uncertainties.

2.6. GFOU representation

The idea behind this representation is captured from the way of plotting non-uniform T2 FSs. In the plot of a non-uniform T2 FS, the intensity of the shading is proportional to secondary membership grades. Indeed, this representation is an extension of Footprint of Uncertainty that is introduced in [4]. Also, it is a simple way to display non-uniform T2 FSs in a two-dimensional plane. As a result, it makes the understanding of T2 operations easier [33,46]. It applies on the problem of land cover classification [46]. However, this representation does not provide an efficient design for non-uniform T2 FLSs.

2.7. α-Plane representation

The key idea of this representation is derived from the α-cuts of a T1 FS. Since a T2 FS has a three-dimensional MF, a term α-plane is used instead of the α-cut. It is clear that the α-plane of a non-uniform T2 FS is associated with a uniform T2 FS of level α. Therefore, a non-uniform T2 FS can be decomposed to several uniform T2 FSs [34–36]. This makes the design of non-uniform T2 FLSs easier. It also reduces the computational complexities of T2 operations. For example, computing the intersection of two non-uniform T2 FSs $\tilde{A}$ and $\tilde{B}$ is performed by the computation of their α-planes intersection:

$$\tilde{A} \cap \tilde{B} = \bigcup_{\alpha \in [0,1]} \alpha / \tilde{A}_\alpha \cap \tilde{B}_\alpha. \tag{8}$$

Note that although (8) is similar to (4), this representation is independently developed with different terminologies.

Using this representation, it can be obtained non-uniform T2 operations by uniform T2 operations. Hence, it is very useful for computational purposes. Also, this representation provides important properties for the centroid of a convex T2 FS. These properties confirmed by the means of examples are as follows [35,36]:

- The shape of the centroid is predictable and depends on the shape of the chosen secondary MF. For example, "when all secondary MFs are normal triangles, then, the centroid is triangle-looking". Also, "when some or all of the secondary MFs are (normal) trapezoids, then, the centroid is trapezoid-looking".
- The centroid is almost symmetric.
- The centroid of a triangular (or trapezoid) T2 FS can be approximated by linear interpolation of the centroids related to two planes of $\alpha = 0$ and $\alpha = 1$.

By using these properties, Mendel et al. [35] proposed two efficient designs for triangular T2 FLSs and trapezoid T2 FLSs. These designs are based on two uniform T2 FLSs.The applicability of the triangular T2 FLS was shown for forecasting Mackey–Glass chaotic time series. However, other uncertainties await an application.

2.8. zSlice representation

This representation utilizes the discretization of z-axis (the third dimension of a non-uniform T2 FS) [37,38]. From the continuous nature of the variable u_i in (9) in [38], this representation is an alternative form of the α-cut representation. Indeed, if a secondary MF is denoted by A, the continuous interval $[l_i, r_i]$ and $\tilde{Z}_i$ in (9) of [38] are equal to A_α and $\alpha.A_\alpha$, respectively. Therefore, this is similar to α-plane representation [35], but here, a triangular T2 FLS is approximated by four zSlices instead of two α-planes. The proposed system was used for control of a two-wheeled mobile robot. Nevertheless, this representation does not provide an efficient design for other non-uniform uncertainties.

3. From the probability theory to a new representation

As explained in Section 2, there are some representations for non-uniform T2 FSs. These representations capture the ideas from the way of plotting, computational geometry, triangular FTNs, slicing, or α-cut. This paper also proposes a new representation by inspiring probability theory. Hence, this section provides a probability background. To do this, it firstly provides a comparison between some definitions used in theories of probability and fuzzy sets. Then, it introduces the topics of probability that could help to reduce the complexities of non-uniform T2 FLSs. Finally, it demonstrates how these topics lead to introduce a new representation for non-uniform T2 FSs.

Table 1
A comparison between three definitions in probability logic and Fuzzy logic.

Probability logic	Fuzzy logic
Probability density function	Fuzzy membership function
Locus of constant probability [47]	Level-sets of a FS [29]
Modeling random uncertainty by the mean and variance [45]	Modeling linguistic uncertainty by T2 FSs [45]

By comparing the definitions of the probability and fuzzy set theories, it can be seen that some definitions are somewhat close together. Some of these similarities are as follows:

- A *probability density function* is a measure to define a probability distribution of a random variable (RV). Similarly**,** a *fuzzy membership function* is a measure to define a possibility distribution of a fuzzy variable.
- *Locus of constant probability* shows the locus of values of the RV x where the probability density function is greater than or equal to a given pre-specified value K [47]. Similarly, *Level-sets of a fuzzy set* show a non-fuzzy set which comprises all values whose membership grade is greater than or equal to a given pre-specified value α [29].
- Studies show that a random uncertainty can be modeled by using *at least the first of two moments of its probability density function,* i.e. the mean and variance [45]. Similarly, linguistic uncertainty can be modeled by using *T2 FSs* that provide a measure of dispersion about the mean [45].

Table 1 shows these similarities, in brief. The existence of such resemblances motivates us to search the topics of probability that could help us to optimally design a non-uniform T2 FLS and to establish probability-like theorems in fuzzy logic. In the probability theory, there are two topics close to our aims: non-uniform RV generation and one function of two RVs. In the following, these two topics are briefly introduced:

- **Non-uniform RV generation** [39]**.** The generation of RVs with arbitrary distributions cannot be done directly. The available algorithms only generate uniform RVs. Hence, the generation of non-uniform RVs is indirectly obtained by a variety of methods involving the uniform RVs [39,40, Chapter 8–5]. For example, a non-uniform RV is generated by simple transformations of k uniform RVs. "It is remarkable that one can obtain the normal and indeed all stable distributions using simple transformations of two uniform RVs" [39]. For example, any Gaussian RV is obtained by applying Box-Muller method [41] to two uniform RVs.
- **One function of two RVs** [40, p. 135]**.** Given two RVs x and y and a function $g(x, y)$, the RV z is defined as $z = g(x, y)$. The statistics of z is defined by:

$$F_z(z) = \int\int_{g(x,y)<z} f_{xy}(x, y)\, dx\, dy, \tag{9}$$

which $f_{xy}(x, y)$ denotes the joint statistics of x and y. If two RVs are independent, then

$$F_z(z) = \int\int_{g(x,y)<z} f_x(x) f_y(y)\, dx\, dy. \tag{10}$$

For $z = x + y$, this formula is changed to the convolution of the probability density functions of x and y:

$$f_z(z) = \int_{-\infty}^{\infty} f_x(z - x) f_y(y) dy. \tag{11}$$

For example, if the RVs x and y are uniform in intervals (a, b) and (c, d), respectively, then the density of their sum $z = x + y$ is a trapezoid. If, in particular, $b - a = d - c$, then $f_z(z)$ is a triangle.

In summary, there are many algorithms for the generation of uniform RVs. However, the generation of non-uniform RVs is difficult. To tackle this difficulty, there are some solutions which utilize uniform RVs such as a function of two RVs.

Similarly, in fuzzy set theory, there are many methods for the implementation and design of uniform T2 FLSs. These systems have very manageable computations. Hence, they have been utilized in many applications. However, the design of non-uniform T2 FLSs is faced with problems. Consequently, these systems are rarely used in engineering applications. The analysis of non-uniform T2 FLSs is still an open question. This paper proposes the solution similar to what probability theory proposes for generating non-uniform RVs. The proposed solution is to represent a non-uniform T2 FS by a function of two uniform T2 FSs. As well, the analysis of non-uniform T2 FLSs can be indirectly obtained by two uniform T2 FLSs.

4. A new representation for non-uniform T2 FSs

This section introduces a new representation for non-uniform T2 FSs and gives an application for it. Using this representation, any non-uniform T2 FS is represented by a function of two uniform T2 FSs. This representation is abbreviated to *2uFunction*. Consequently, the represented T2 FSs are called *2uFunction-based T2 FSs*. Rule-based FLSs using these sets are called *2uFunction-based T2 FLSs*, as well. Subsection 4.1 gives all needed definitions for 2uFunction-based T2 FSs. These definitions are inspired from the formulas presented in the second topic of Section 3. Subsection 4.2 applies this representation for designing non-uniform T2 FLSs.

4.1. 2uFunction-based T2 FSs

Definition 1 (*2uFunction Representation*)**.** By the application of the extension principle for interactive fuzzy numbers [48], every secondary MF in the vertical slice x, $\mu_{\tilde{A}}(u|x)$,[1] could be described in terms of a function of two uniform MFs $\mu_{\tilde{A}_1}(u_1|x)$ and $\mu_{\tilde{A}_2}(u_2|x)$ as follows:

$$\mu_{\tilde{A}}(u|x) = \bigvee_{\substack{u_1,u_2 \\ u=g(u_1,u_2)}} \mu_{\tilde{A}_1\tilde{A}_2}(u_1, u_2|x), \tag{12}$$

or,

$$\mu_{\tilde{A}}(u|x) = \underset{\substack{u_1,u_2 \\ u=g(u_1,u_2)}}{\upsilon} \mu_{\tilde{A}_1\tilde{A}_2}(u_1, u_2|x), \tag{13}$$

in which $\vee$ and υ indicate union and averaging operators, respectively. $\mu_{\tilde{A}_1\tilde{A}_2}(u_1, u_2|x)$ denotes the joint possibility distribution of $\tilde{A}_1$ and $\tilde{A}_2$.

Definition 2. A 2uFunction-based T2 FS $\tilde{A}$ can be represented by the union of $\mu_{\tilde{A}}(u|x)$ computed by (12) or (13):

$$\tilde{A} = \int_{x\in X} \mu_{\tilde{A}}(u|x)/x. \tag{14}$$

Remark 1. If u_1 and u_2 are non-interactive, (12) and (13) turn into the following relations, respectively:

$$\mu_{\tilde{A}}(u|x) = \bigvee_{\substack{u_1,u_2 \\ u=g(u_1,u_2)}} \mu_{\tilde{A}_1}(u_1|x) \wedge \mu_{\tilde{A}_2}(u_2|x), \tag{15}$$

and,

$$\mu_{\tilde{A}}(u|x) = \underset{\substack{u_1,u_2 \\ u=g(u_1,u_2)}}{\upsilon} \mu_{\tilde{A}_1}(u_1|x) \wedge \mu_{\tilde{A}_2}(u_2|x), \tag{16}$$

which $\wedge$ is an intersection operator. Indeed, these relations are the applications for the extension principle of Zadeh [42].

Note that a non-uniform uncertainty could be obtained by using two uniform uncertainties provided that function g; the operators of υ, $\vee$ and $\wedge$; and the interaction between two uniform fuzzy variables are determined. For example, if two uniform fuzzy variables are non-interactive, the following settings are proposed for triangular and trapezoid uncertainties:

$$\begin{aligned} &g(u_1, u_2) = w_1u_1 + w_2u_2, \quad 0 \le w_1, w_2 \le 1, \quad w_1 + w_2 = 1, \\ &u_1 \,\upsilon\, u_2 = u_1 \oplus_\gamma u_2 = (1-\gamma)(u_1 \wedge u_2) + \gamma(u_1 \vee u_2), \quad 0 \le \gamma \le 1, \\ &u_1 \vee u_2 = u_1 + u_2 - u_1u_2, \\ &u_1 \wedge u_2 = u_1u_2. \end{aligned} \tag{17}$$

Definition 3 (*weighted fuzzy convolution*)**.** If $u = u_1 + \eta u_2$, the weighted fuzzy convolution of $\mu_{\tilde{A}_1}(u_1|x)$ and $\mu_{\tilde{A}_2}(u_2|x)$ is defined as:

$$(\mu_{\tilde{A}_1} \overline{\nabla}_\eta \mu_{\tilde{A}_2})(u|x) = \underset{u_2}{\upsilon}\, \mu_{\tilde{A}_1}(u - \eta u_2|x) \wedge \mu_{\tilde{A}_2}(u_2|x), \tag{18}$$

in which $\overline{\nabla}_\eta$ denotes the weighted convolution operator.

Remark 2. By using (17), the weighted convolution (the convolution) of two uniform fuzzy variables is a trapezoid. Hence, if the fuzzy variables u_1 and u_2 are uniform in the intervals (a, b), and (c, d), respectively; and the coefficient γ is defined by $1/\max(b-a, d-c)$, then the MF of their weighted sum (the sum) is a trapezoid (see Fig. 1). If , in particular, $b-a = d-c$, then the result is a 2uFunction-based triangular MF.

[1] Since every vertical slice of $\tilde{A}$ is a function of a secondary variable u, we use the notation $\mu_{\tilde{A}}(u|x)$ instead of a widely used notation $\mu_{\tilde{A}}(x)$. The used notation was, for the first time, introduced by Mendel et al. [35], but it has not been used to date.

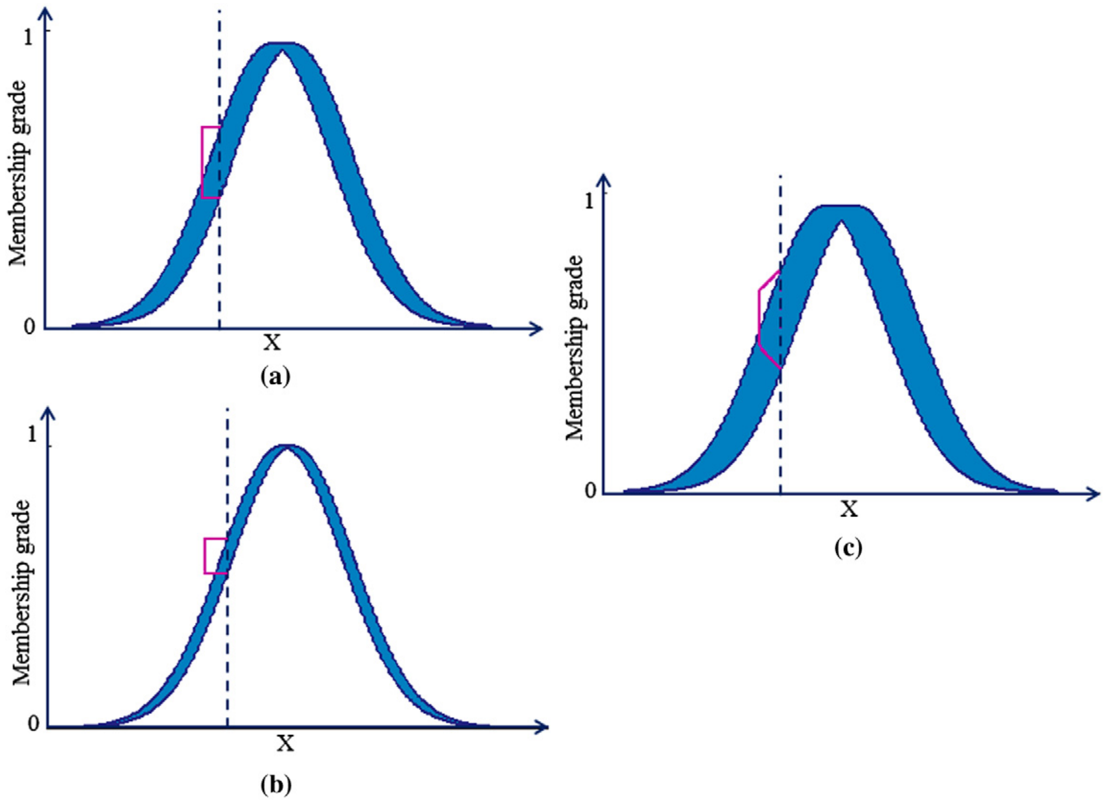

Fig. 1. FOUs of two uniform T2 FSs are shown in (a) and (b); FOU of a 2uFunction-based trapezoid T2 FS are also shown in (c). In addition, a vertical slice of each T2 FS accompanying with its secondary MF is displayed.

Example 1. if $w_1 = w_2 = 1$, $\mu_{\tilde{A}_1}(u_1|x) = \frac{1}{1} + \frac{1}{2} + \frac{1}{3}$ and $\mu_{\tilde{A}_2}(u_2|x) = \frac{1}{0} + \frac{1}{1} + \frac{1}{2} + \frac{1}{3}$, then according to Remark 2, $\mu_{\tilde{A}}(u|x)$ will be obtained as follows:

$$\begin{aligned}\mu_{\tilde{A}}(1|x) &= (\mu_{\tilde{A}_1} \bigtriangledown \mu_{\tilde{A}_2})(1|x) = \underset{u_2}{\upsilon}(\mu_{\tilde{A}_1}(1-u_2|x) \wedge \mu_{\tilde{A}_2}(u_2|x))\\ &= \upsilon(\mu_{\tilde{A}_1}(1|x) \wedge \mu_{\tilde{A}_2}(0|x), \mu_{\tilde{A}_1}(0|x) \wedge \mu_{\tilde{A}_2}(1|x), \mu_{\tilde{A}_1}(-1|x) \wedge \mu_{\tilde{A}_2}(2|x), \mu_{\tilde{A}_1}(-2|x) \wedge \mu_{\tilde{A}_2}(3|x))\\ &= \upsilon(1, 0, 0, 0) = \left(1 - \frac{1}{3}\right)(1 \wedge 0 \wedge 0 \wedge 0) + \frac{1}{3}(1 \vee 0 \vee 0 \vee 0) = 0.3333.\end{aligned}$$

$$\begin{aligned}\mu_{\tilde{A}}(2|x) &= (\mu_{\tilde{A}_1} \bigtriangledown \mu_{\tilde{A}_2})(2|x) = \underset{u_2}{\upsilon}(\mu_{\tilde{A}_1}(2-u_2|x) \wedge \mu_{\tilde{A}_2}(u_2|x))\\ &= \upsilon(\mu_{\tilde{A}_1}(2|x) \wedge \mu_{\tilde{A}_2}(0|x), \mu_{\tilde{A}_1}(1|x) \wedge \mu_{\tilde{A}_2}(1|x), \mu_{\tilde{A}_1}(0|x) \wedge \mu_{\tilde{A}_2}(2|x), \mu_{\tilde{A}_1}(-1|x) \wedge \mu_{\tilde{A}_2}(3|x))\\ &= \upsilon(1, 1, 0, 0) = \left(1 - \frac{1}{3}\right)(1 \wedge 1 \wedge 0 \wedge 0) + \frac{1}{3}(1 \vee 1 \vee 0 \vee 0) = 0.6667.\end{aligned}$$

$$\begin{aligned}\mu_{\tilde{A}}(3|x) &= (\mu_{\tilde{A}_1} \bigtriangledown \mu_{\tilde{A}_2})(3|x) = \underset{u_2}{\upsilon}(\mu_{\tilde{A}_1}(3-u_2|x) \wedge \mu_{\tilde{A}_2}(u_2|x))\\ &= \upsilon(\mu_{\tilde{A}_1}(3|x) \wedge \mu_{\tilde{A}_2}(0|x), \mu_{\tilde{A}_1}(2|x) \wedge \mu_{\tilde{A}_2}(1|x), \mu_{\tilde{A}_1}(1|x) \wedge \mu_{\tilde{A}_2}(2|x), \mu_{\tilde{A}_1}(0|x) \wedge \mu_{\tilde{A}_2}(3|x))\\ &= \upsilon(1, 1, 1, 0) = \left(1 - \frac{1}{3}\right)(1 \wedge 1 \wedge 1 \wedge 0) + \frac{1}{3}(1 \vee 1 \vee 1 \vee 0) = 1.\end{aligned}$$

$$\begin{aligned}\mu_{\tilde{A}}(4|x) &= (\mu_{\tilde{A}_1} \bigtriangledown \mu_{\tilde{A}_2})(4|x) = \underset{u_2}{\upsilon}(\mu_{\tilde{A}_1}(4-u_2|x) \wedge \mu_{\tilde{A}_2}(u_2|x))\\ &= \upsilon(\mu_{\tilde{A}_1}(4|x) \wedge \mu_{\tilde{A}_2}(0|x), \mu_{\tilde{A}_1}(3|x) \wedge \mu_{\tilde{A}_2}(1|x), \mu_{\tilde{A}_1}(2|x) \wedge \mu_{\tilde{A}_2}(2|x), \mu_{\tilde{A}_1}(1|x) \wedge \mu_{\tilde{A}_2}(3|x))\\ &= \upsilon(0, 1, 1, 1) = \left(1 - \frac{1}{3}\right)(0 \wedge 1 \wedge 1 \wedge 1) + \frac{1}{3}(0 \vee 1 \vee 1 \vee 1) = 1.\end{aligned}$$

$$\begin{aligned}\mu_{\tilde{A}}(5|x) &= (\mu_{\tilde{A}_1} \triangledown \mu_{\tilde{A}_2})(5|x) = \underset{u_2}{\upsilon}(\mu_{\tilde{A}_1}(5-u_2|x) \wedge \mu_{\tilde{A}_2}(u_2|x)) \\ &= \upsilon(\mu_{\tilde{A}_1}(5|x) \wedge \mu_{\tilde{A}_2}(0|x), \mu_{\tilde{A}_1}(4|x) \wedge \mu_{\tilde{A}_2}(1|x), \mu_{\tilde{A}_1}(3|x) \wedge \mu_{\tilde{A}_2}(2|x), \mu_{\tilde{A}_1}(2|x) \wedge \mu_{\tilde{A}_2}(3|x)) \\ &= \upsilon(0,0,1,1) = \left(1-\frac{1}{3}\right)(0 \wedge 0 \wedge 1 \wedge 1) + \frac{1}{3}(0 \vee 0 \vee 1 \vee 1) = 0.6667.\end{aligned}$$

$$\begin{aligned}\mu_{\tilde{A}}(6|x) &= (\mu_{\tilde{A}_1} \triangledown \mu_{\tilde{A}_2})(6|x) = \underset{u_2}{\upsilon}(\mu_{\tilde{A}_1}(6-u_2|x) \wedge \mu_{\tilde{A}_2}(u_2|x)) \\ &= \upsilon(\mu_{\tilde{A}_1}(6|x) \wedge \mu_{\tilde{A}_2}(0|x), \mu_{\tilde{A}_1}(5|x) \wedge \mu_{\tilde{A}_2}(1|x), \mu_{\tilde{A}_1}(4|x) \wedge \mu_{\tilde{A}_2}(2|x), \mu_{\tilde{A}_1}(3|x) \wedge \mu_{\tilde{A}_2}(3|x)) \\ &= \upsilon(0,0,0,1) = \left(1-\frac{1}{3}\right)(0 \wedge 0 \wedge 0 \wedge 1) + \frac{1}{3}(0 \vee 0 \vee 0 \vee 1) = 0.3333.\end{aligned}$$

Finally, $\mu_{\tilde{A}}(u|x)$ will be $\mu_{\tilde{A}}(u|x) = \frac{0.3333}{1} + \frac{0.6667}{2} + \frac{1}{3} + \frac{1}{4} + \frac{0.6667}{5} + \frac{0.3333}{6}$.

4.2. 2uFunction-based T2 FLSs

This subsection introduces a design for a non-uniform T2 FLS. To do this, the properties of its centroid are utilized. As mentioned in Subsection 2.7, the shape of the centroid of a non-uniform T2 FLS is similar to the shape of its secondary MFs. In addition, using the 2uFunction representation, the centroid can be represented by two uniform FSs. This subsection proposes that these two sets can be the centroids of two arbitrary uniform T2 FLSs.

Assume that B_1 and B_2 are the centroids of two uniform T2 FLSs, and B is the centroid of a non-uniform T2 FLS. Also, y_1, y_2 and y are output fuzzy variables with MFs $\mu_{B_1}(y_1)$, $\mu_{B_2}(y_2)$, and $\mu_B(y)$, respectively. By utilizing 2uFunction representation, $\mu_B(y)$ is computed by:

$$\mu_B(y) = \bigvee_{\substack{y_1, y_2 \\ y=g(y_1,y_2)}} \mu_{B_1B_2}(y_1, y_2) \tag{19}$$

or,

$$\mu_B(y) = \underset{\substack{y_1, y_2 \\ y=g(y_1,y_2)}}{\upsilon} \mu_{B_1B_2}(y_1, y_2). \tag{20}$$

From these formulas, a non-uniform T2 FLS can be designed by two uniform T2 FLSs. Fig. 2a shows this design, schematically. The design can be used for any non-uniform uncertainty provided that the following conditions are established. Firstly, theoretic operators such as t-norm, s-norm and averaging are defined. Secondly, the function g is defined in such a way that it satisfies design considerations for a T2 FLS. In other words, "when all sources of uncertainty disappear, a T2 FLS must reduce to a T1 FLS" [35]. Finally, how the uniform variables interact with each other is determined. For example, a design for trapezoid uncertainties can be done by the settings used in Remark 2. In this case, its centroid could be computed by:

$$\mu_B(y) = (\mu_{B_1} \triangledown_\eta \mu_{B_2})\left(\frac{1}{w_1}y\right). \tag{21}$$

Fig. 2b shows the design for trapezoid uncertainties. In both designs, the final output is obtained by defuzzifying $\mu_B(y)$. This paper uses centroid defuzzification. Hence, the output in Fig. 2a is obtained by:

$$y_c = g(y_{c1}, y_{c2}), \tag{22}$$

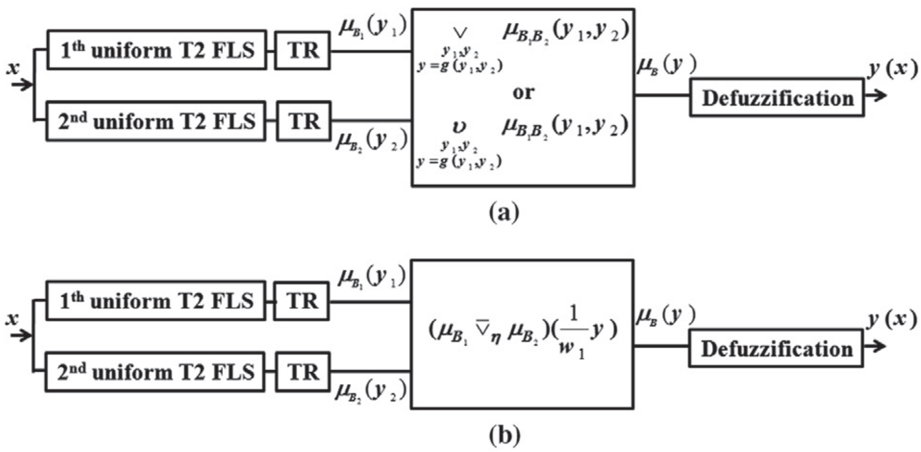

Fig. 2. New designs for: (a) a non-uniform T2 FLS and (b) a trapezoid T2 FLS.

in which y_{c1} and y_{c2} denote the defuzzified output and are computed by $(y_{l1}+y_{r1})/2$ and $(y_{l2}+y_{r2})/2$, respectively. Besides, y_l and y_r are obtained by the Karnik–Mendel algorithm [6] as follows:

$$y_l = \frac{\sum_{i=1}^{L} \bar{f}^i y_l^i + \sum_{j=L+1}^{M} \underline{f}^j y_l^j}{\sum_{i=1}^{L} \bar{f}^i + \sum_{j=L+1}^{M} \underline{f}^j}, \tag{23}$$

and

$$y_r = \frac{\sum_{i=1}^{R} \underline{f}^i y_r^i + \sum_{j=R+1}^{M} \bar{f}^j y_r^j}{\sum_{i=1}^{R} \underline{f}^i + \sum_{j=R+1}^{M} \bar{f}^j}, \tag{24}$$

in which R is a switch point from lower MF to upper MF and L is vice versa; $\underline{f}^i$ and $\bar{f}^i$ are the bounds for the strength of ith rule in uniform T2 FLSs; and, y_l^i and y_r^i are the bounds for the output of ith rule.

According to Remark 2, the output in Fig. 2b is obtained by:

$$y_c = w_1 y_{c1} + w_2 y_{c2}, \tag{25}$$

in which w_1 and w_2 are the weights for uniform T2 FLSs. In both designs, the parameters of two uniform T2 FLSs are simultaneously optimized, and because of (22) or (25), a coupled optimization happens.

As it can be seen in Fig. 2, the proposed designs use only two uniform T2 FLSs for a non-uniform T2 FLS. As a result, they can utilize the results of all researches on uniform T2 FLSs. For example, Mendel et al. [2] showed that a uniform T2 FLS can be obtained by T1 computations. Consequently, the computations of a 2ufunction-based T2 FLS are greatly decreased. As well, the proposed designs are much simpler than full-blown T2 FLSs. Therefore, the 2uFunction representation can provide a good next step in the hierarchy of FLSs from T1 to uniform T2 to non-uniform T2.

5. Forecasting Mackey–Glass chaotic time series

T1 FLSs have been extensively used in time-series forecasting [49–52]. Besides, studies showed that uniform T2 FLSs obtain better predictions than T1 ones [3,22,23,25,35,53–55]. Since a non-uniform T2 FLS has more design degrees of freedom than uniform T2 FLSs, it is conjectured that it outperforms existing fuzzy systems. Hence, this section evaluates the ability of a 2uFunction-based trapezoid T2 FLS (2uF-T2 FLS) to forecast Mackey–Glass chaotic time series. Besides, it compares the performance of the proposed design with T1 FLS, uniform T2 FLS and the studies done in [3,35,54]. To do this, three different experiments are used. Subsections 5.1– 5.3 briefly describe the required settings for these experiments. Subsection 5.4 also introduces the designs for the FLSs.

5.1. The Mackey–Glass chaotic time series

The Mackey–Glass time series has been widely considered as a benchmark for forecasters [49,56]. This series models white blood cells production in leukemia patients, and is generated by the following differential equation:

$$\frac{\mathrm{d}s(t)}{\mathrm{d}t} = \frac{0.2s(t-\tau)}{1+s^{10}(t-\tau)} - 0.1s(t), \tag{26}$$

which, for values of τ greater than 17, exhibits chaos [57]. This paper focuses on $\tau = 30$, which is used in [3,35,54].

To compare our method with the methods studied in [3,35,54], this paper obtains Mackey–Glass time series values in the following conditions:

- For experiments 1 & 2, Euler's method [58] is applied to (26) in such a way that the time step is set to 1, and initial values of $s(k)$ for $k \le \tau$ are set randomly. Finally, $s(k)$ is derived for $0 \le k \le 3004$.
- For experiment 3, the fourth-order Runge–Kutta method [59] is applied to (26). The time step is set to 1. The following initial conditions are used: $s(0) = 1.2$ and $s(k) = 0$ for $k < 0$. Finally, $s(k)$ is derived for $0 \le k \le 1200$.

5.2. Noise

Our experiments assume that the noise has a uniform distribution with zero-mean and a given signal-to-noise ratio (SNR). In experiments 1 and 3, SNR is a certain value such as 0, 4, 6, or 10 dB, while, in experiment 2, SNR varies from 0 to 10 dB. In other words, experiments 1 and 3 use a stationary noise, whereas experiment 2 uses a non-stationary noise.

Assume that the range of SNR is uniformly distributed into 100 levels. Now, for a given SNR, the variance of noise is firstly computed by:

$$10\log_{10}\sigma_n^2 = 10\log_{10}\left(\frac{1}{N}\sum_{k=1}^{N} s^2(k)\right) - SNR, \tag{27}$$

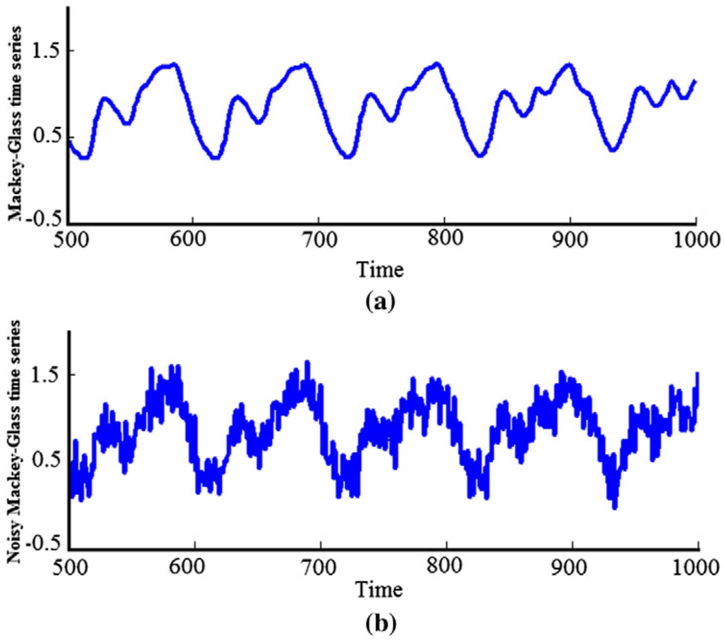

Fig. 3. Mackey–Glass chaotic time series: (a) the noise-free data and (b) One realization of the noise-corrupted data.

in which σ_n^2 denotes the variance of noise; $s(k)$ denotes the time series values; and N is the total number of samples in the time series. Then, the noise values $n(k)$ can be simulated by a uniform distribution with zero mean and variance σ_n^2. After that, the noise is added to the time series:

$$x(k) = s(k) + n(k), \tag{28}$$

which $x(k)$ denotes noise-corrupted data. Fig. 3 shows a realization of the noise-corrupted data in which SNR ranges between 0 and 10 dB.

5.3. Training and test datasets

Assume the time series split into a training set of D points and a test set of $N - D$ points. Recall that, in the forecasting of time series, the objective is to predict a future point of time series, $s(k + l)$, from a window of p past measurements of $s(k)$ with a time interval Δ, i.e. $x(k - (p - 1)\Delta)$, ..., $x(t - \Delta)$, $x(k)$. Consequently, training and test datasets are the matrices whose dimensions are $(D - p\Delta) \times (p + 1)$ and $(N - D - p\Delta) \times (p + 1)$, respectively [56].

To obtain training and test dataset, the values of $N = 1000$, $D = 500$, and $p = 4$ are used. The training dataset consists of points t_0 to $t_0 + 499$ (i.e. 500 training data). The following 500 data points are used as test dataset. In addition, the values of $l = \Delta = 1$, and $t_0 = 2000$ are used for experiments 1 and 2; and the values of $l = \Delta = 6$, and $t_0 = 124$ are used for experiment 3.

5.4. FLS forecasters

One of the common methods in the problem of forecasting is FLSs. This paper evaluates three FLSs with three types of uncertainty: singleton, uniform, and trapezoid. These forecasters are called T1 FLS, uniform T2 FLS, and 2uF-T2 FLSs, respectively. They are based on the following settings:

- four antecedents and one consequent per rule;
- the Wang–Mendel algorithm [60] for fuzzy rule extraction;
- two MFs for each antecedent (as in [3,35,49]), so the number of rules is 16;
- product implication (as in [3,31,35,55]);
- center-of-sets defuzzification for the T1 FLSs, and the center-of-sets type-reduction (as in [6–8,36,61]) for uniform T2 FLSs and 2uF-T2 FLSs;
- Singleton secondary MFs for the T1 FLSs, uniform secondary MFs for the uniform T2 FLSs, and trapezoid secondary MFs defined by (21) for the 2uF-T2 FLSs;

Table 2
The design details of T1 FLS, uniform T2 FLS and 2uF-T2 FLS.

Design items	T1	Uniform T2	2uF-T2
Fuzzification	Singleton	Singleton	Singleton
The number of primary MFs	2	2	2
The shape of primary MFs	Triangular	Triangular	Triangular
The shape of secondary MFs	Singleton	Uniform	Trapezoid
FOU	–	Uncertain centre	Uncertain centre
Fuzzy rule extraction	Wang–Mendel algorithm	Wang–Mendel algorithm	Wang–Mendel algorithm
Implication	Product	Product	Product
Defuzzification	Center-of-sets	Center-of-sets	Center-of-sets
Type-reduction	–	Center-of-sets	Center-of-sets
Optimization method	Descent gradient	Descent gradient	Descent gradient

- Triangular primary MFs for the T1 FLSs, and triangular FOUs with uncertain centre but certain spread for the uniform T2 FLSs and 2uF-T2 FLSs. Triangular MFs are defined as follows:

$$\mu(x, b, \Delta) = \begin{cases} 1 + (x - b)/\Delta & b - \Delta \le x \le b, \\ 1 - (x - b)/\Delta & b \le x \le b + \Delta, \\ 0 & otherwise. \end{cases} \quad (29)$$

in which b and Δ denote centre and spread, respectively.

Table 2 shows the above settings, briefly. T1 FLSs and uniform T2 FLSs are designed by the methods presented in [62] and [3], respectively; A 2uF-T2 FLS is designed by the design presented in Fig. 2b. The descent gradient method is used for parameter optimization. It is based on the minimization of the following objective function:

$$J = \frac{1}{2}(d - y)^2, \quad (30)$$

which d denotes the desired output in FLS forecasters; In other words, it is the noise-free data; and y denotes the defuzzified output in FLS forecasters. For a uniform T2 FLS, y is defined by:

$$y = \frac{y_l + y_r}{2}, \quad (31)$$

in which y_l and y_r are lower and upper bounds of the output, respectively. These bounds are computed by the Karnik–Mendel algorithm [6]. For a 2uF-T2 FLS forecaster, y is defined as follows:

$$y = w_1 \frac{y_{l1} + y_{r1}}{2} + w_2 \frac{y_{l2} + y_{r2}}{2}, \quad (32)$$

in which y_{l1} and y_{r1} are lower and upper bounds of the output in the first uniform T2 FLS forecaster, respectively; y_{l2} and y_{r2} are the output bounds in the second uniform T2 FLS forecaster.

Using the available input–output data, the parameters of MFs can be optimized as follows:

$$\theta(q + 1) = \theta(q) - \alpha \frac{\partial e}{\partial \theta}\bigg|_q, \quad (33)$$

which $q = 0, 1, 2, 3, \ldots$; α is a learning coefficient; and θ is a parameter that will be optimized. The performance of each FLS forecaster is evaluated by the Root Mean Square Error (RMSE).

5.5. Simulation results

This subsection compares the performance of the following three FLS forecasters: T1 FLS, Uniform T2 FLS, and 2uF-T2 FLS. It also compares the proposed method with ANFIS [54], IT2FNN [54], IT2SFLS [3] and triangular Q-T2 FLS [35]. To do this, three different experiments are done in noisy conditions. In addition, the settings reported in [3,35,54] are used for simulating Mackey–Glass time series and designing FLSs.

The initial locations of primary MFs in antecedent are determined by the mean m_t and standard deviation σ_t of training set. In other words, the initial values of the centre are $m_t \pm 2\sigma_t$; and Initial values of the spread is $4s\sigma_t$, which s varies in $[0, 1]$. Also, the values of consequent is randomly used in $[0, 1]$. These settings are close to the settings in [3]. These settings could be originated from the interval estimation in probability theory. Recall that a confidence interval for a parameter such as m_t is a random interval as $[m_t - z\sigma_t, m_t + z\sigma_t]$ constructed from data and depends on z. For $z = 1.96$, the probability that the interval contains the true value of the parameter is 96% .[40]. In this paper, the value of $z = 2$ is applied. Also, the left-sided and right-sided values in this interval are used as initial values for antecedent MFs.

Table 3
Initial conditions for FLS forecasters.

Type of FLS	Antecedents	Consequent
T1	$b_{F_k^i} = m_t - 2\sigma_t$ and $b_{F_k^i} = m_t + 2\sigma_t$, $\Delta_{F_k^i} = 4s\sigma_t$	$\bar{y}^i \in [0, 1]$
Uniform T2	$b_{\tilde{F}_k^i} = [m_t - 2\sigma_t - 2c_1\sqrt{3}\sigma_n, m_t - 2\sigma_t]$ and $b_{\tilde{F}_k^i} = [m_t + 2\sigma_t, m_t + 2\sigma_t + 2c_1\sqrt{3}\sigma_n]$, $\Delta_{F_k^i} = 4s\sigma_t$	$y_l^i = \bar{y}^i, y_r^i = \bar{y}^i +2c_2\sqrt{3}\sigma_n$
1st Uniform T2 in 2uF-T2	$b_{\tilde{F}_{k1}^i} = [m_t - 2\sigma_t - 2c_3\sqrt{3}\sigma_n, m_t - 2\sigma_t]$ and $b_{\tilde{F}_{k1}^i} = [m_t + 2\sigma_t, m_t + 2\sigma_t + 2c_3\sqrt{3}\sigma_n]$, $\Delta_{F_{k1}^i} = 4s\sigma_t$	$y_{l1}^i = \bar{y}^i,\ y_{r1}^i = \bar{y}^i +2c_4\sqrt{3}\sigma_n$
2nd Uniform T2 in 2uF-T2	$b_{\tilde{F}_{k2}^i} = [m_t - 2\sigma_t - 2c_5\sqrt{3}\sigma_n, m_t - 2\sigma_t]$ and $b_{\tilde{F}_{k2}^i} = [m_t + 2\sigma_t, m_t + 2\sigma_t + 2c_5\sqrt{3}\sigma_n]$, $\Delta_{F_{k2}^i} = 4s\sigma_t$	$y_{l2}^i = \bar{y}^i,\ y_{r2}^i = \bar{y}^i +2c_6\sqrt{3}\sigma_n$

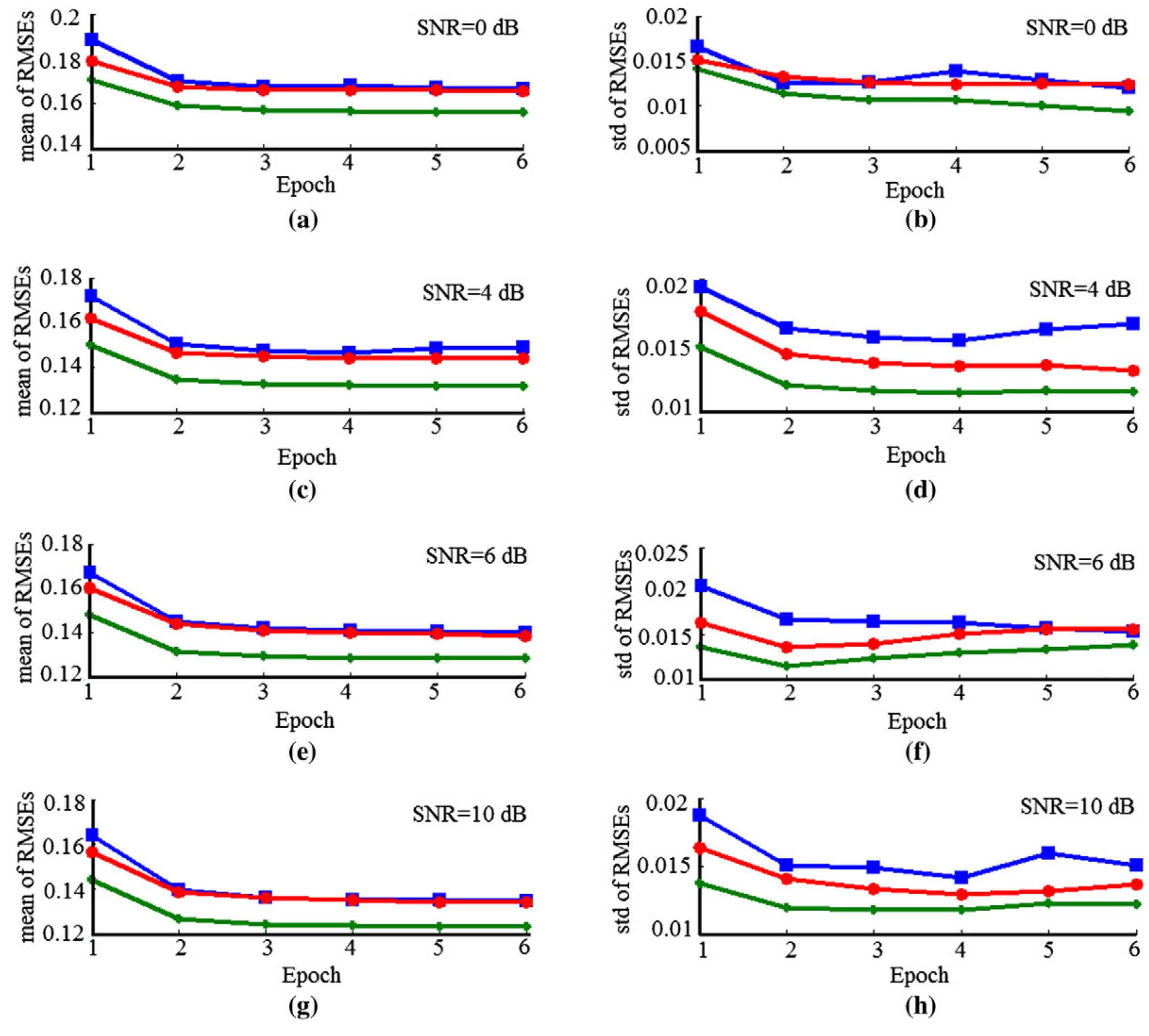

Fig. 4. Comparing the results of each design for 6 epochs when different levels of noise are added during 20 realizations: the mean values of RMSEs are shown in (a), (c), (e), and (g); the std of RMSEs are shown in (b), (d), (f), and (h). The lines marked by square, circle, and asterisk indicate T1 FLS, uniform T2 FLS, 2uF-T2 FLS, respectively.

The range of uncertainty in the centre of primary MFs is based on the standard deviation σ_n of noise. It is set to $2c\sqrt{3}\sigma_n$, which comes from the characteristics of the noise. Recall that a uniform RV with the mean equal to zero and standard deviation σ_n can take values in the interval $[-\sqrt{3}\sigma_n, +\sqrt{3}\sigma_n]$ [22]. Therefore, the length of the interval in which the centre can vary is $2c\sqrt{3}\sigma_n$. Table 3 briefly shows the initial conditions of FLS forecasters. For 2uF-T2 FLS, the values of $w_1 = 0.5$, and $w_2 = 0.5$ are used. The value of learning coefficient α is set to 0.5 and remains unchanged during the learning process.

Note that this paper does not try to determine the best initial parameters, because, the goal is to evaluate the effect of considering more uncertainties in a FLS and to compare the performance of three types of FLSs.

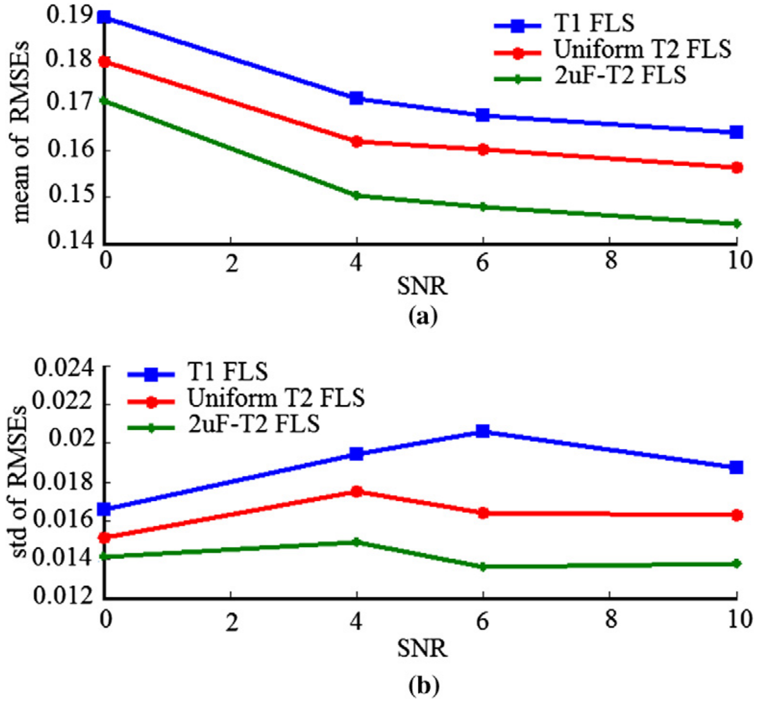

Fig. 5. Mackey–Glass time series prediction in 20 noisy realizations for first epoch: (a) the mean values of RMSEs, and (b) the std of RMSEs.

Table 4
The mean and std of RMSEs for T1, uniform T2, and 2uF-T2 FLS in the first epoch.

SNR (dB)	T1	Uniform T2	2uF-T2
0	0.1885 ± 0.0166	0.1789 ± 0.0151	0.1705 ± 0.0141
4	0.1711 ± 0.0194	0.1619 ± 0.0175	0.1502 ± 0.0149
6	0.1673 ± 0.0206	0.1602 ± 0.0164	0.1478 ± 0.0136
10	0.1639 ± 0.0187	0.1564 ± 0.0163	0.1442 ± 0.0138

Table 5
Comparison of percentage improvement of mean of RMSEs between the designs.

SNR	Uniform T2 versus T1	2uF-T2 versus Uniform T2	2uF-T2 versus T1
0	5.09%	4.69%	9.81%
4	5.76%	7.23%	12.5%
6	4.24%	7.74%	11.65%
10	4.58%	7.80%	12.02%

Experiment 1: Mackey–Glass time series prediction for $\tau = 30$, $l = \Delta = 1$ *corrupted by a noise whose SNR is 0, 4, 6, or 10 dB.*

Each FLS forecaster is designed over 20 realizations of the time series corrupted by a noise whose SNR is 0, 4, 6, or 10 dB. For all realizations, c_1 and c_2 are set to 0.1443 and 0.5774 (as in [3]), respectively; c_3, c_4, c_5, and c_6 are set to 0.0722, 0.6495, 0.0722, and 0.8776, respectively; and the value of $s = 1$ is chosen. For each realization, each FLS is trained by using the gradient descent algorithm with a fixed learning rate 0.5 for six epochs. After each epoch, the values of RMSEs are computed by applying the test data to each FLS. Fig. 4 depicts the mean and standard deviation (std) of RMSEs for each design for 6 epochs when different levels of noise are added. For the first epoch, Fig. 5 depicts the mean and std of the RMSEs for each design under different noisy conditions. Table 4 numerically shows these results. Table 5 provides the comparative results. From the results, the following points can be observed:

- All FLS forecasters provide better predictions when SNR increases.
- Increasing uncertainty from singleton to uniform to trapezoid improves the performance of FLS forecaster.
- 2uF-T2 FLSs outperform other designs and show higher improvement than those for all SNRs.
- Since 2uF-T2 FLSs have a smaller std than other designs, they are much more robust to the noise than those.
- Because 2uF-T2 FLSs achieve good predictions at the first epoch, they are a good candidate for real-time applications.

The existing literature on T2 FSs confirms the above points. For example, studies showed that uniform T2 FLSs obtain better predictions than T1 ones [3,22,23,25,35,53–55]. Besides, Mendel et al. [3] showed that a uniform T2 FLS

Table 6
Comparison of the mean and std of RMSEs in Q-T2 FLS [35] and 2uF-T2 FLS.

SNR	Q-T2	2uF-T2
0	0.2695 ± 0.017840	0.1705 ± 0.0141
4	0.2276 ± 0.008670	0.1502 ± 0.0149
6	0.2026 ± 0.006768	0.1478 ± 0.0136
10	0.1550 ± 0.005076	0.1442 ± 0.0138

Table 7
Comparison of the percent of improvement in results of Q-T2 FLS [35] and 2uF-T2 FLS to T1 FLS.

SNR	Q-T2 versus T1	2uF-T2 versus T1
0	4.40%	9.81%
4	2.76%	12.5%
6	1.65%	11.65%
10	2.39%	12.02%

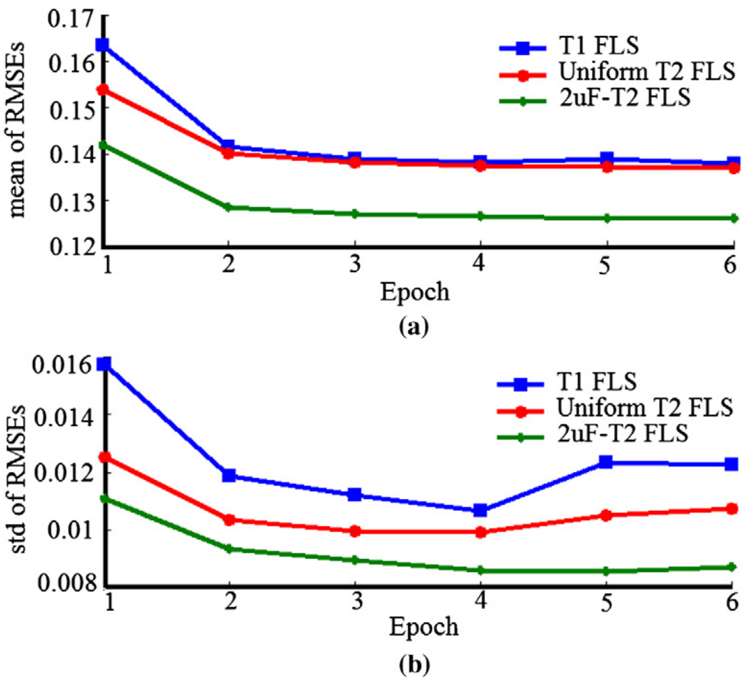

Fig. 6. A comparison of RMSEs between T1 FLS, uniform T2 FLS, and 2uF-T2 FLS in 50 noisy realizations for 6 epochs: (a) the mean values of RMSEs, and (b) the std of RMSEs.

in comparison with a T1 FLS is an appropriate candidate for applications of real-time signal processing and adaptive filtering.

In the following, the performance of 2uF-T2 FLSs are compared with triangular Q-T2 FLSs [35]. Table 6 compares the values of the mean and std of RMSEs in Q-T2 FLSs [35] and 2uF-T2 FLSs. It shows that the performance of our proposed method is better than Q-T2 FLS. Table 7 provides the comparative results. To compare these two systems, the relative results in Table 7 are more appropriate than the absolute results in Table 6, because:

- Using the random initial conditions for simulating the chaotic Mackey–Glass time series affects the results.
- Type of uncertainty in these systems is different. In Q-T2 FLSs, triangular is used, while in 2uF-T2 FLSs, trapezoid is used.
- Different initial conditions are used. This paper designs all three forecasters under the same initial conditions. This allows us to have a fair comparison between the designs. However, Mendel et al. [35] used the trained T1 and uniform T2 FLSs to initialize uniform T2 and Q-T2 FLSs, respectively.
- This paper uses the fixed learning coefficient, while Mendel et al. [35] used the adaptive learning coefficient: it decreases during the training if the RMSE increases.

Experiment 2: Mackey–Glass time series prediction for $\tau = 30$, $l = \Delta = 1$ *corrupted by a non-stationary noise whose SNR ranges between 0 and 10 dB.*

Table 8
Comparison of the prediction errors in the studies in [54] and the proposed method for SNR 0dB.

Prediction method	Secondary MF	Training set	Test set	RMSE
ANFIS [54]	Singleton	500	500	0.4000
IT2FNN-1 [54]	Uniform	500	500	0.3053
The proposed method	Trapezoid	500	500	0.2180

The design of each FLS forecaster is performed over 50 realizations of the noise-corrupted Mackey–Glass time series when SNR is uncertain and changes in the interval [0, 10]. Initial values for parameters $c_1 - c_6$ and s are similar to Experiment 1. Since standard deviation of noise changes in $[\sigma_{n_{0\,\text{dB}}}, \sigma_{n_{10\,\text{dB}}}]$, the value of σ_n is set to $(\sigma_{n_{0\,\text{dB}}} + \sigma_{n_{10\,\text{dB}}})/2$.

For each design, the mean and std of RMSEs are computed for 6 epochs under the non-stationary noise. Fig. 6 displays these results. It shows that 2uF-T2 FLSs outperform other designs and are more robust to a non-stationary noise. These systems also have the smallest std of RMSEs. By comparing Fig. 6 with Fig. 4 in [3], our percentage improvement of the mean of RMSEs to T1 FLS is higher than T2SFLSs in [3].

Experiment 3: Mackey–Glass time series prediction for $\tau = 30$, $l = \Delta = 6$ *corrupted by a noise with SNR value of 0 dB.*

To design 2uF-T2 FLS forecasters, c_3, c_4, c_5, and c_6 are set to 0.1443, 0.2887, 0.3, and 1, respectively; s is set to 1.5. After 50 epochs, the results are compared with ANFIS [54] and IT2FNN-1 [54]. Since this paper utilizes singleton fuzzifiers, the results are not compared with IT2FNN-2&3 [54] utilizing T2 non-singleton fuzzifiers with uncertain standard deviation. Table 8 shows the comparative results. As expected, the proposed method has the best performance under very noisy conditions.

6. Discussion

As explained in Section 2, some of the representations utilize uniform T2 FSs to represent a non-uniform T2 FS, such as α-level set representation [29], α-plane representation [34–36], triangular representation [31], and zSlice representation [37,38]. Our proposed representation is also based on uniform T2 FSs. Hence, this section provides a comparison between our representation and four recent representations.

In comparison with the α-level set representation, our representation provides both an efficient design for non-uniform T2 FLSs and an application for trapezoid T2 FLSs, while α-level set representation is not used in practice. It is only used for defining T2 operations.

The comparison between our representation and the α-plane representation [34–36] leads to the following results:

- Our method is inspired by probability topics (see Section 3), while Liu's [35] is the extended form of α-cut in T1 fuzzy mathematics;
- Our design is efficient for any non-uniform T2 FLSs, while Liu's [35] provides two efficient designs for only triangular and trapezoid T2 FLSs;
- In ours, two arbitrary uniform T2 FLSs are used to handle any non-uniform uncertainty, while, in Liu's [35], two planes of $\alpha = 0$ and $\alpha = 1$ are used to handle triangular and trapezoid uncertainties;
- In ours, according to probability topics, the number of uniform T2 FLSs is set to 2, while, in Liu's [35], it is experimentally set to two planes.
- In ours, a trapezoid type-reduced set is obtained from the weighted convolution (the convolution) of two uniform type-reduced sets with settings (17), while, in Liu's [35], it is obtained by the linear interpolation of two uniform type-reduced sets.
- This paper presents the application for a trapezoid T2 FLS, while Liu [35] do not. To our knowledge, to date, no application exists for trapezoid T2 FLS.

From the comparison of our method with triangular representation [31], it can be concluded that (1) these two methods have different origins: probability and triangular FTNs; (2) ours can be utilized to design any non-uniform uncertainties, whiles Starczewsk's [31] is limited to triangular T2 FLSs; (3) In ours, a triangular type-reduced set is obtained from the weighted convolution (the convolution) of two uniform type-reduced sets with settings (17), while, in Starczewsk's [31], it is obtained by the three-node linear interpolation of the general results for centroid type-reduction.

By using the zSlice representation, Wagner et al. [38] presented a triangular T2 FLS approximated by four zSlices (four uniform T2 FLSs), while ours utilizes two uniform T2 FLSs for a triangular T2 FLS. Also, the triangular type-reduced set is obtained by the linear interpolation of uniform T2 FSs, while, in ours, it is obtained from the weighted convolution (the convolution) of two uniform type-reduced sets with settings (17). Moreover, the idea of this representation is from the discretization of z-axis, which is completely different from ours.

Overall, our representation is an entirely new method for representing non-uniform T2 FSs. By utilizing it, any non-uniform T2 FLS could be indirectly analyzed in terms of a function of two uniform T2 FLSs provided that the function g, theoretic operators (such as t-norm, s-norm, averaging), and how the uniform fuzzy variables interact with each other are determined.

7. Conclusions and future works

This paper introduces a new representation for non-uniform T2 FSs called 2uFunction representation. This representation is originated from two topics in probability theory: (1) generating non-uniform RVs based on uniform ones and (2) a function of two RVs. The main advantage of this representation is to reduce the complexities of non-uniform T2 FSs. This representation is applied to the design of FLS forecasters. Results show that increasing uncertainty in FLSs leads to improve the prediction accuracy and shows superiority of 2uF-T2 FLSs in comparison with T1 FLSs, uniform T2 FLSs, ANFIS [54], IT2FNN-1 [54], T2SFLSs [3] and Q-T2 FLSs [35]. Future works will consist in developing a 2uFunction-based Gaussian T2 FLS. To do this, we will use two following approaches. First is to introduce appropriate definitions for the function g and the theoretic operations. Second is to use a trapezoidal approximation for representing any non-uniform T2 FSs.

References

[1] J.M. Mendel, IEEE Transactions on Fuzzy Systems Special Issue on Advances in Type-2 Fuzzy Sets and Systems, in: IEEE Computational Intelligence Magazine, 2012, pp. 75.
[2] J.M. Mendel, R.I. John, F. Liu, Interval type-2 fuzzy logic systems made simple, IEEE Transactions on Fuzzy Systems 14 (2006) 808–821.
[3] Q. Liang, J.M. Mendel, Interval type-2 fuzzy logic systems: theory and design, IEEE Transactions on Fuzzy Systems 8 (2000) 535–550.
[4] J.M. Mendel, R.I.B. John, Type-2 fuzzy sets made simple, IEEE Transactions on Fuzzy Systems 10 (2002) 117–127.
[5] S. Coupland, R. John, Geometric interval type-2 fuzzy systems, in: Proceedings of the Joint EUSFLAT-LFA, Barcelona, Spain, 2005, pp. 449–454.
[6] N.N. Karnik, J.M. Mendel, Centroid of a type-2 fuzzy set, Information Sciences 132 (2001) 195–220.
[7] S. Greenfield, F. Chiclana, S. Coupland, R. John, The collapsing method of defuzzification for discretised interval type-2 fuzzy sets, Information Sciences 179 (2009) 2055–2069.
[8] H. Wu, J.M. Mendel, Uncertainty bounds and their use in the design of interval type-2 fuzzy logic systems, IEEE Transactions on Fuzzy Systems 10 (2002) 622–639.
[9] C. Hwang, F.C.-H. Rhee, Uncertain fuzzy clustering: interval type-2 fuzzy approach to c-means, IEEE Transaction on Fuzzy Systems 15 (2007) 107–120.
[10] K. Santiago-Sánchez, C.A. Reyes-García, P. Gómez-Gil, Type-2 fuzzy sets applied to pattern matching for the classification of cries of infants under neurological risk, in: Emerging Intelligent Computing Technology and Applications Lecture Notes in Computer Science, 2009, pp. 201–210.
[11] S.F. Molaeezadeh, M. Salarian, M.H. moradi, Type-2 fuzzy pattern matching for classifying hunger and pain cries of healthy full-term infants in: 16th CSI International Symposium on Artificial Intelligence and Signal Processing, Shiraz, 2012.
[12] Q. Liang, MPEG VBR video traffic classification using Bayesian and nearest neighbor classifiers in: IEEE International Symposium on Circuits and Systems, 2002, pp. II-77–II-80.
[13] J. Zeng, Z.-Q. Liu, Type-2 fuzzy sets for pattern classification: a review, in: Proceedings of the 2007 IEEE Symposium on Foundations of Computational Intelligence (FOCI 2007), 2007, pp. 193–200.
[14] U. Mahmood, A. Al-Jumaily, M.m. Al-Jaafreh, Type-2 fuzzy classification of blood pressure parameters in: IEEE ISSNIP, 2007, pp. 595–600.
[15] A. Celikyilmaz, I.B. Turksen, Genetic type-2 fuzzy classifier functions, in: Annual Meeting of the North American Fuzzy Information Processing Society (NAFIPS 2008), 2008, pp. 1–6.
[16] J. Sanz, A. Fernández, H. Bustince, F. Herrera, A genetic tuning to improve the performance of Fuzzy Rule-Based Classification Systems with Interval-Valued Fuzzy Sets: Degree of ignorance and lateral position, International Journal of Approximate Reasoning 52 (2011) 751–766.
[17] O. Castillo, P. Melin, Type-2 Fuzzy Logic: Theory and Applications, springer, 2008.
[18] P. Melin, J. Urias, D. Solano, M. Soto, M. Lopez, O. Castillo, Voice recognition with neural networks, type-2 fuzzy logic and genetic algorithms, Journal of Engineering Letters 13 (2006) 108–116.
[19] P. Melin, O. Mendoza, O. Castillo, Face recognition with an improved interval type-2 fuzzy logic sugeno integral and modular neural networks, IEEE Transactions on Systems, Man, and Cybernetics – Part A: Systems and Humans 41 (2011) 1001–1012.
[20] J. Zeng, L. Xie, Z.h.-Q. Liu, Type-2 fuzzy Gaussian mixture models, Pattern Recognition 4 (2008) 3636–3643.
[21] J. Zeng, Z.-Q. Liu, Type-2 fuzzy hidden Markov models and their application to speech recognition, IEEE Transactions on Fuzzy Systems 4 (2006) 454–467.
[22] N.N. Karnik, J.M. Mendel, Applications of type-2 fuzzy logic systems to forecasting of time-series, Information Sciences 120 (1999) 89–111.
[23] D. Hidalgo, P. Melin, O. Castillo, An optimization method for designing type-2 fuzzy inference systems based on the footprint of uncertainty using genetic algorithms, Expert Systems with Applications 39 (2012) 4590–4598.
[24] A. Mencattini, M. Salmeri, S. Bertazzoni, R. Lojacono, E. Pasero, W. Moniaci, Local Meteorological forecasting by type-2 fuzzy systems time series prediction, in: IEEE International Conference on Computational Intelligence for Measurement Systems and Applications, Giardini Naxos, Italy, 2005.
[25] A. Mencattini, M. Salmeri, Meteorological forecasting using type-2 fuzzy logic systems, in: J.J. Zhu, G.P.C. Fung (Eds.), Forecasting Models – Methods and Applications, Createspace, 2010, pp. 13–28.
[26] M.m.O. Al-Jaafreh, A.A. Al-Jumaily, Type-2 fuzzy system based blood pressure parameters estimation in: Second Asia International Conference on Modelling & Simulation, 2008, pp. 953–958.
[27] O. Castillo, P. Melin, A review on the design and optimization of interval type-2 fuzzy controllers, Applied Soft Computing 12 (2012) 1267–1278.
[28] C. Leal-Ramírez, O. Castillo, P. Melin, A. Rodríguez-Díaz, Simulation of the bird age-structured population growth based on an interval type-2 fuzzy cellular structure, Information Sciences 181 (2011) 519–535.
[29] L.A. Zadeh, The concept of a linguistic variable and its application to approximate reasoning-I, Information Sciences 8 (1975) 199–249.
[30] J.T. Starczewski, A triangular type-2 fuzzy logic system, in: IEEE International Conference on Fuzzy Systems, Vancouver, BC 2006, pp. 1460–1467.
[31] J.T. Starczewski, Efficient triangular type-2 fuzzy logic systems, International Journal of Approximate Reasoning 50 (2009) 799–811.
[32] S. Coupland, R. John, New geometric inference techniques for type-2 fuzzy sets, International Journal of Approximate Reasoning 49 (2008) 198–211.
[33] L.A. Lucas, T.M. Centeno, M.R. Delgado, General type-2 fuzzy inference systems: analysis, design and computational aspects, in: IEEE International Conference on Fuzzy Systems, London 2007, pp. 1–6.
[34] J.M. Mendel, F. Liu, On new quasi-type-2 fuzzy logic systems, in: IEEE International Conference on Fuzzy Systems, Hong Kong, 2008.
[35] J.M. Mendel, F. Liu, D. Zhai, α-Plane representation for type-2 fuzzy sets: theory and applications, IEEE Transactions on Fuzzy Systems 17 (2009) 1189–1207.
[36] F. Liu, An efficient centroid type-reduction strategy for general type-2 fuzzy logic system, Information Sciences 178 (2008) 2224–2236.
[37] C. Wagner, H. Hagras, zSlices — towards bridging the gap between interval and general type-2 fuzzy logic, in: IEEE International Conference on Fuzzy Systems, Hong Kong, 2008, pp. 489–497.
[38] C. Wagner, H. Hagras, Toward general type-2 fuzzy logic systems based on zSlices, IEEE Transactions on Fuzzy Systems 18 (2010) 637–660.
[39] L. Devroye, Non-Uniform Random Variate Generation, Springer-Verlag, New York, 1986.
[40] A. Papoulis, Probability, Random Variables, and Stochastic Processes, third ed., McGraw-Hill, New York, 1991.
[41] G.E.P. Box, M.E. Muller, A note on the generation of random normal deviates, Annals of Mathematical Statistics 29 (1958) 610–611.
[42] L.A. Zadeh, Fuzzy sets, Information and Control 8 (1965) 338–353.
[43] S. Coupland, R. John, Geometric type-1 and type-2 fuzzy logic systems, IEEE Transaction on Fuzzy Systems 15 (2007) 3–15.
[44] S. Coupland, R. John, A fast geometric method for defuzzification of type-2 fuzzy sets, IEEE Transactions on Fuzzy Systems 16 (2008) 929–941.

[45] J.M. Mendel, Type-2 fuzzy sets and systems: an overview, in: IEEE Computational Intelligence Magazine, 2007, pp. 20–29.
[46] L.A. Lucas, T.M. Centeno, M.R. Delgado, General type-2 fuzzy classifiers to land cover classification, in: Proceedings of the 2008 ACM symposium on Applied Computing (SAC'08), 2008, pp. 1743–1747.
[47] M.I. Ribeiro, Gaussian Probability Density Functions: Properties and Error Characterization, Institute for Systems and Robotics, Lisboa, Portugal, Technical Report, 2004.
[48] C. Carlsson, R. Fuller, P. Majlender, An extension principle for interactive fuzzy numbers, in: Proceedings of the Fourth International Symposium of Hungarian Researchers on Computational Intelligence, Budapest, November 13–14, 2003, pp. 113–118.
[49] J.-S.R. Jang, ANFIS: adaptive-network-based fuzzy inference system, IEEE Transactions on Systems, Man and Cybernetics 23 (1993) 665–685.
[50] A. Gholipour, B.N. Araab, C. Lucas, Predicting chaotic time series using neural and neurofuzzy models: a comparative study, Neural Processing Letters 24 (2006) 217–239.
[51] L. Herrera, H. Pomares, I. Rojas, O. Valenzuela, A. Prieto, Tase, a taylor series based fuzzy system model that combines interpretability and accuracy, Fuzzy Sets and Systems 153 (2005) 403–427.
[52] Q. Liang, Ad hoc wireless network traffic—self-similarity and forecasting, IEEE Comunications Letters 6 (2002)
[53] H. Saima, J. Jaafar, S. Belhaouari, T.A. Jillani, ARIMA based interval type-2 fuzzy model for forecasting, International Journal of Computer Applications 28 (2011) 17–21.
[54] J.R. Castro, O. Castillo, P. Melin, A. Rodríguez-Díaz, A hybrid learning algorithm for a class of interval type-2 fuzzy neural networks, Information Sciences 179 (2009) 2175–2193.
[55] M. Almaraashi, R. John, Tuning of type-2 fuzzy systems by simulated annealing to predict time series, in: Proceedings of the World Congress on Engineering 2011.
[56] J.M. Mendel, Uncertain Rule-Based Fuzzy Logic Systems: Introduction and New Directions, Prentice-Hall, 2001.
[57] M.C. Mackey, L. Glass, Oscillation and chaos in physiological control systems, Science 197 (1977) 287–289.
[58] D. Quinney, An Introduction to the Numerical Solution of Differential Equations, Research Studies Press, 1987.
[59] R. But, Introduction to Numerical Analysis Using Matlab, first ed., Jones and Bartlett Publishers, 2009.
[60] L.X. Wang, J.M. Mendel, Generating fuzzy rules by learning from examples, IEEE Transactions on Systems, Man, and Cybernetics 22 (1992) 1414–1427.
[61] Z. Daoyuan, J.M. Mendel, Computing the centroid of a general type-2 fuzzy set by means of the centroid-flow algorithm, IEEE Transactions on Fuzzy Systems 19 (2011) 401–422.
[62] L.-X. Wang, A Course in Fuzzy Systems and Control, Prentice-Hall International, 1997.