A New Similarity Function for Generalized Trapezoidal Fuzzy Numbers

E. Vicente, A. Mateos, and A. Jiménez

Artificial Intelligence Department, Technical University of Madrid, Spain
e.vicentecestero@upm.es, {amateos, ajimenez}@fi.upm.es

Abstract. Numerous authors have proposed functions to quantify the degree of similarity between two fuzzy numbers using various descriptive parameters, such as the geometric distance, the distance between the centers of gravity or the perimeter. However, these similarity functions have drawbacks for specific situations. We propose a new similarity measure for generalized trapezoidal fuzzy numbers aimed at overcoming such drawbacks. This new measure accounts for the distance between the centers of gravity and the geometric distance but also incorporates a new term based on the shared area between the fuzzy numbers. The proposed measure is compared against other measures in the literature.

1 Introduction

The *theory of fuzzy sets* was first introduced by Zadeh [14]. It is a multivalued logic developed to deal with imprecise or vague data based on degrees of truth rather than the usual Boolean true or false logic. It is useful for modeling concepts in a environment concerning inaccurate or vague measurements.

Fuzzy logic is useful for building a linguistic terms scale that experts will use to measure imprecise parameters. For instance, a nine-member linguistic terms set is introduced in [12]. These linguistic terms are usually associated with a triangular or trapezoidal fuzzy number [15]. Fuzzy number arithmetic, defined in conformity with the model in question, is then used to make computations (addition, multiplication, subtraction, ranking...) using the fuzzy information provided by experts (see, e.g. the arithmetic proposed in [13] for linguistic values trapezoidal fuzzy numbers or the one in [4,6] for generalized trapezoidal fuzzy numbers). For advance in research in fuzzy number arithmetic and logical operators, see [11].

Partial or final results of computations with fuzzy numbers lead to new (triangular or trapezoidal) fuzzy numbers that often need to be expressed again by a linguistic term. Consequently, we have to identify the linguistic term on the previously defined scale whose associated fuzzy number is most similar to the one derived from computations.

Different metrics can be used to establish the similarity between fuzzy numbers, based on their distance, form or size. These parameters can be aggregated in mathematical expressions that define the degree of similarity between two fuzzy numbers.

However, all similarity measures proposed by different authors have drawbacks, because the parameters used are not always best suited to the circumstances of the problem and the type of fuzzy number that the model uses. In this paper, we propose a similarity measure for generalized trapezoidal fuzzy numbers with good properties that overcomes the drawbacks of other similarity measures proposed in the literature.

In Section 2, we review the similarity measures proposed in the literature, analyzing their advantages and drawbacks. In Section 3, we propose a new similarity measure. We demonstrate that the new similarity measure has the same good properties as earlier measures and other additional properties that overcome their drawbacks. In Section 4, we compare the proposed measure with the measures outlined in this section, taking as a reference the set of 30 pairs of generalized fuzzy numbers provided in [13]. Finally, some conclusions are provided in Section 5.

2 Overview of Similarity Measures

First we introduce preliminary concepts to formalize similarity measures. We then review the major similarity measures proposed in the literature and more recent measures derived from them in chronological order, identifying their most interesting properties, as well as their drawbacks.

Generalized trapezoidal fuzzy numbers were first proposed by Chen [4,5]. A generalized trapezoidal fuzzy number with support in the interval $[0, 1]$ is a tuple $(a_1, a_2, a_3, a_4; w_{\tilde{A}})$ with $0 \leq a_1 \leq a_2 \leq a_3 \leq a_4 \leq 1$, and $w_{\tilde{A}} \in [0, 1]$ together with a membership function $(\mu_{\tilde{A}} : \mathbb{R} \rightarrow [0, w_{\tilde{A}}])$,

$$\mu_{\tilde{A}} = \begin{cases} 0 & \text{if } x < a_1 \\ \frac{w_{\tilde{A}}(x-a_1)}{a_2-a_1} & \text{if } a_1 < x < a_2 \\ w_{\tilde{A}} & \text{if } a_2 < x < a_3 \\ \frac{w_{\tilde{A}}(x-a_4)}{a_3-a_4} & \text{if } a_3 < x < a_4 \\ 0 & \text{if } a_4 < x \end{cases}$$

indicating the degree of membership of any value $x \in \mathbb{R}$ to the fuzzy number $\tilde{A}$. We denote by $TF[0, 1]$ the set of all these fuzzy numbers.

In particular, if $w_{\tilde{A}} = 1$, then we say that $\tilde{A}$ is a *normalized fuzzy number*, and denote by $TF[0, 1; 1]$ the set of these fuzzy numbers.

A *similarity measure* is a function $S : TF[0, 1] \times TF[0, 1] \rightarrow [0, 1]$ indicating the degree of similarity between two fuzzy numbers. This value must match the intuitive perception that we have of the fuzzy numbers that we are comparing. The closer this value is to 1, the more similar the fuzzy numbers will be.

The first ideas about the similarity of normalized fuzzy numbers with support in $[0, 1]$ stemmed, of course, from the distance notion. If we define a distance $d(\tilde{A}, \tilde{B}) \in [0, 1]$ between fuzzy numbers $\tilde{A}, \tilde{B} \in TF[0, 1; 1]$, then generally $S = 1 - d$ is a similarity function. Chen [5] defined the degree of similarity between

two normalized fuzzy numbers $\tilde{A} = (a_1, a_2, a_3, a_4; 1)$ and $\tilde{B} = (b_1, b_2, b_3, b_4; 1)$ using the geometric distance as

$$S(\tilde{A}, \tilde{B}) = 1 - \frac{\sum_{i=1}^4 |a_i - b_i|}{4}. \quad (1)$$

This measure has a number of interesting properties: *Property 1* ($S(\tilde{A}, \tilde{B}) = S(\tilde{B}, \tilde{A})$), *Property 2* ($S(\tilde{A}, \tilde{B}) = 1 \Leftrightarrow \tilde{A} = \tilde{B}$) and *Property 3* (if $\tilde{A} = (a, a, a, a; 1)$, $\tilde{B} = (b, b, b, b; 1)$ then $S(\tilde{A}, \tilde{B}) = 1 - |a - b|$).

Tran and Duckstein [11] defined a distance, which was computed as a weighted sum of distances between two intervals across all the α -cuts from 0 to 1. This distance was also used in [8] to measure the intensity of dominance between trapezoidal fuzzy weights representing the preferences of DMs within MAUT. However, neither Chen's nor Tran and Duckstein's measures can be used to determine the similarity between generalized fuzzy numbers.

Chen and Chen [2] extended the similarity measure to the set $TF[0, 1]$ adding to Eq. (1) the distance between the centers of gravity of the compared numbers [1]. Thus, the similarity measure between the numbers $\tilde{A} = (a_1, a_2, a_3, a_4; w_{\tilde{A}})$ and $\tilde{B} = (b_1, b_2, b_3, b_4; w_{\tilde{B}})$ is determined by the expression

$$S(\tilde{A}, \tilde{B}) = \left[1 - \frac{\sum_{i=1}^4 |a_i - b_i|}{4} \right] \times [1 - |X_{\tilde{A}} - X_{\tilde{B}}|]^{B(S_{\tilde{A}}, S_{\tilde{B}})} \times \left[\frac{\min\{Y_{\tilde{A}}, Y_{\tilde{B}}\}}{\max\{Y_{\tilde{A}}, Y_{\tilde{B}}\}} \right], \quad (2)$$

where $(X_{\tilde{A}}, Y_{\tilde{A}})$ and $(X_{\tilde{B}}, Y_{\tilde{B}})$ are the centroids of $\tilde{A}$ and $\tilde{B}$, respectively, i.e.

$$X_{\tilde{A}} = \begin{cases} \frac{Y_{\tilde{A}}(a_3 + a_2) + (w_{\tilde{A}} - Y_{\tilde{A}})(a_4 + a_1)}{2w_{\tilde{A}}}, & \text{if } w_{\tilde{A}} \neq 0 \\ \frac{a_4 + a_1}{2}, & \text{if } w_{\tilde{A}} = 0 \end{cases}, \quad Y_{\tilde{A}} = \begin{cases} \frac{w_{\tilde{A}} \left(\frac{a_3 - a_2}{a_4 - a_1} + 2 \right)}{6}, & \text{if } a_4 \neq a_1 \\ \frac{w_{\tilde{A}}}{2}, & \text{if } a_4 = a_1 \end{cases}, \quad (3)$$

$$B(S_{\tilde{A}}, S_{\tilde{B}}) = \begin{cases} 1, & \text{if } S_{\tilde{A}} + S_{\tilde{B}} > 0 \\ 0, & \text{otherwise} \end{cases}, \quad S_{\tilde{A}} = a_4 - a_1 \text{ and } S_{\tilde{B}} = b_4 - b_1.$$

The factor $[1 - |X_{\tilde{A}} - X_{\tilde{B}}|]^{B(S_{\tilde{A}}, S_{\tilde{B}})}$ is used to distinguish pairs of the form $\tilde{A} = (a, a, a, a; w_{\tilde{A}})$ and $\tilde{B} = (b, b, b, b; w_{\tilde{B}})$ from the remaining pairs of generalized fuzzy numbers. This extends the previous measures published in 1996 to generalized fuzzy numbers, and provides a fourth property: *Property 4* (if $\tilde{A} = (a, a, a, a; 0)$, $\tilde{B} = (a, a, a, a; 1)$ then $S(\tilde{A}, \tilde{B}) = 0$).

Indeed, the first fuzzy number is clearly not the real number a , whereas the second fuzzy number clearly is the real number a . So, the similarity between them is evidently zero.

However, this measure has a small drawback since it assigns a degree of similarity $S(\tilde{A}, \tilde{B}) = 0$ to fuzzy numbers $\tilde{A} = (a, a, a, a; 0)$ and $\tilde{B} = (a, a, a, a; 10^{-10^{10}})$.

Can we be sure that these numbers are completely different?, i.e. as different as the numbers of property 4? Obviously not. Then, we need a measure that distinguishes these numbers in a fairer way.

Wei and Chen [12] proposed a new measure using the perimeter concept of generalized trapezoidal fuzzy numbers:

$$S(\tilde{A}, \tilde{B}) = \left[1 - \frac{\sum_{i=1}^4 |a_i - b_i|}{4} \right] \times \left[\frac{\min\{P(\tilde{A}), P(\tilde{B})\} + \min\{w_{\tilde{A}}, w_{\tilde{B}}\}}{\max\{P(\tilde{A}), P(\tilde{B})\} + \max\{w_{\tilde{A}}, w_{\tilde{B}}\}} \right], \quad (4)$$

where $P(\tilde{A}) = \sqrt{(a_1 - a_2)^2 + w_{\tilde{A}}^2} + \sqrt{(a_3 - a_4)^2 + w_{\tilde{A}}^2} + (a_3 - a_2) + (a_4 - a_1)$, and analogously for $P(\tilde{B})$.

Like the measure proposed by Chen and Chen [2], this new measure also verifies the four properties but again has the above drawback. This was the ground proffered by Xu et al. [13] for proposing a new measure based, like the measure published by Chen and Chen [2], on the concept of center of gravity. Xu et al. consider two weights $w, 1 - w \in (0, 1)$ to attach more or less importance to the concepts used:

$$S_w(\tilde{A}, \tilde{B}) = 1 - w \frac{\sum |a_i - b_i|}{4} - (1 - w) \frac{\sqrt{(X_{\tilde{A}} - X_{\tilde{B}})^2 + (Y_{\tilde{A}} - Y_{\tilde{B}})^2}}{\sqrt{1.25}}.$$

This action effectively mitigates the drawback of the measures by Chen and Chen[2] and Wei and Chen [12], since it assigns a high degree of similarity to fuzzy numbers of the form $(a, a, a, a, 0)$ and (a, a, a, a, ϵ) with ϵ near zero, and also provides a new property: *Property 5* ($S(\tilde{A}, \tilde{B}) = 0$ (and $\tilde{A} < \tilde{B}$) $\Leftrightarrow \tilde{A} = (0, 0, 0, 0; 0)$ and $\tilde{B} = (1, 1, 1, 1; 1)$).

However, Xu et al. sacrificed properties 3 and 4 to achieve this fifth property. For example, if we consider $w = 0.5$, then the degree of similarity of $(0.1, 0.1, 0.1, 0.1; 1)$ and $(0.1, 0.1, 0.1, 0.1; 0)$ is 0.7763 with respect to the fourth property. For $(a, a, a, a; 1)$ and $(b, b, b, b; 1)$, we have $S(\tilde{A}, \tilde{B}) = 1 - 0.5 |a - b| - 0.5 \frac{|a - b|}{\sqrt{1.25}} \neq 1 - |a - b|$, with respect to the third property.

Another drawback of Xu et al.'s measure is as follows. If we consider the trapezoidal fuzzy numbers shown in Fig. 1, $\tilde{A} = (0, 0.1, 0.3, 0.4; 1)$, $\tilde{B} = (0.25, 0.4, 0.6, 0.75; 1)$ and $\tilde{C} = (0.75, 0.775, 0.825, 0.85; 1)$, then the degree of similarity of $\tilde{C}$ and $\tilde{A}$ with respect to $\tilde{B}$, with $w = 0.5$, is 0.7156 in both cases. Therefore, the numbers $\tilde{C}$ and $\tilde{A}$ are just similar to the central number $\tilde{B}$. However, $\tilde{B}$ should clearly be more similar to $\tilde{A}$ than $\tilde{C}$ on the basis of its shape, size, and more importantly, the shared area.

Apart from Chen and Xu et al.'s measures, numerous authors have defined the degree of similarity between two trapezoidal fuzzy numbers, without giving up any of the five described properties. Note, for example, the Sridevi and Nadarajan's extension [10], a fuzzy distance that replaces the geometric distance

Fig. 1. A drawback of Xu et al.'s measure

in the measure proposed by Chen and Chen[2],

$$S(\tilde{A}, \tilde{B}) = \left[\frac{\sum_{i=1}^4 \mu_d(x)}{4} \right] \times [1 - |X_{\tilde{A}} - X_{\tilde{B}}|]^{B(S_{\tilde{A}}, S_{\tilde{B}})} \times \left[\frac{\min\{Y_{\tilde{A}}, Y_{\tilde{B}}\}}{\max\{Y_{\tilde{A}}, Y_{\tilde{B}}\}} \right],$$

with $\mu_d(x) = \begin{cases} 1 - \frac{x}{d}, & \text{if } 0 \leq x \leq d \\ 0, & \text{otherwise} \end{cases}$, $d \in (0, 1]$, $x = |a_i - b_i|$ and $(X_{\tilde{A}}, Y_{\tilde{A}})$ and $(X_{\tilde{B}}, Y_{\tilde{B}})$ the centroids of the compared fuzzy numbers. Parameter d represents the level of precision required to measure the similarity between the two fuzzy numbers. The measure by Sridevi and Nadarajan sacrifices the third and the fifth property.

More recently, Gomathi and Sivaraman [7] proposed a new measure. This measure sacrifices only the fifth property but again has the same the drawback as the measures by Chen and Wei and Chen. It modifies the measure proposed by Wei and Chen by using the geometric instead of the arithmetic mean of the difference of the vertices of the fuzzy numbers under comparison. Moreover, it considers a straightforward function including the vertices and heights of the fuzzy numbers rather than their perimeters in order to reduce computational time with respect to the measure by Wei and Chen but achieve similar results:

$$S(\tilde{A}, \tilde{B}) = \left[\sqrt[4]{\prod_{i=1}^4 (1 - |a_i - b_i|)} \right] \times \left[\frac{\min\{Q(\tilde{A}), Q(\tilde{B})\} + \min\{w_{\tilde{A}}, w_{\tilde{B}}\}}{\max\{Q(\tilde{A}), Q(\tilde{B})\} + \max\{w_{\tilde{A}}, w_{\tilde{B}}\}} \right], \quad (5)$$

where $Q(\tilde{A}) = \sqrt{(a_2 - a_1)^2 + (a_3 - a_2)^2 + (a_4 - a_3)^2 + w_{\tilde{A}}^2}$.

3 A New Similarity Function

The most common parameters in the similarity measures are the geometric distance, the distance between the centers of gravity and the perimeter. In the measure that we propose, we incorporate the shared area between the generalized fuzzy numbers with respect to the total area of these fuzzy numbers. The closer this value is to 1, the more similar are the compared fuzzy numbers.

We also directly use the difference between the height of the generalized fuzzy numbers, since although the distance between the centroids to some extent

already considers this parameter, failures have been observed when measuring the similarity between some pairs of fuzzy numbers whose height is close to zero, as discussed in Section 2.

We define the degree of similarity of the generalized trapezoidal fuzzy numbers $\tilde{A} = (a_1, a_2, a_3, a_4; w_{\tilde{A}})$ and $\tilde{B} = (b_1, b_2, b_3, b_4; w_{\tilde{B}})$ as follows:

- if $\max\{(a_4 - a_1), (b_4 - b_1)\} \neq 0$, then

$$S(\tilde{A}, \tilde{B}) = (1 - |w_{\tilde{A}} - w_{\tilde{B}}|) \times \left(1 - (1 - \alpha - \beta) \times \left(1 - \frac{\int_0^1 \mu_{\tilde{A} \cap \tilde{B}}(x) dx}{\int_0^1 \mu_{\tilde{A} \cup \tilde{B}}(x) dx} \right) - \alpha \frac{\sum |a_i - b_i|}{4} - \beta \frac{d[(X_{\tilde{A}}, Y_{\tilde{A}}), (X_{\tilde{B}}, Y_{\tilde{B}})]}{M} \right),$$

- otherwise,

$$S(\tilde{A}, \tilde{B}) = (1 - |w_{\tilde{A}} - w_{\tilde{B}}|) \times \left(1 - \left(\frac{1 - \alpha - \beta}{2} + \alpha \right) \times \frac{\sum |a_i - b_i|}{4} - \left(\frac{1 - \alpha - \beta}{2} + \beta \right) \times \frac{d[(X_{\tilde{A}}, Y_{\tilde{A}}), (X_{\tilde{B}}, Y_{\tilde{B}})]}{M} \right),$$

where $\alpha + \beta < 1$, $\mu_{\tilde{X}}$ is the membership function of $\tilde{X}$, $M = \max_{[0,1] \times [0, \frac{1}{2}]} \{d(x, y), (x', y')\}$,

$\mu_{\tilde{A} \cap \tilde{B}}(x) = \min_{0 \leq x \leq 1} \{\mu_{\tilde{A}}(x), \mu_{\tilde{B}}(x)\}$, $\mu_{\tilde{A} \cup \tilde{B}}(x) = \max_{0 \leq x \leq 1} \{\mu_{\tilde{A}}(x), \mu_{\tilde{B}}(x)\}$, $(X_{\tilde{A}}, Y_{\tilde{A}})$, $(X_{\tilde{B}}, Y_{\tilde{B}})$ are computed as in Eqs. (3), and d is a distance in $\mathbb{R}^2$.

From now on, we analyze the properties of the proposed similarity measure

Proposition 1. $S(\tilde{A}, \tilde{B}) \in [0, 1]$.

Proof. Since the weights sum 1, it suffices to see that $(1 - |w_{\tilde{A}} - w_{\tilde{B}}|) \leq 1$, $\frac{\int_0^1 \mu_{\tilde{A} \cap \tilde{B}}(x) dx}{\int_0^1 \mu_{\tilde{A} \cup \tilde{B}}(x) dx} \leq 1$, $\frac{\sum |a_i - b_i|}{4} \leq 1$, $\frac{d[(X_{\tilde{A}}, Y_{\tilde{A}}), (X_{\tilde{B}}, Y_{\tilde{B}})]}{M} \leq 1$, which is trivial. $\square$

Proposition 2. $S(\tilde{A}, \tilde{B}) = S(\tilde{B}, \tilde{A})$.

Proof. Trivial. $\square$

Proposition 3. $S(\tilde{A}, \tilde{B}) = 1 \Leftrightarrow \tilde{A} = \tilde{B}$.

Proof. The reverse implication is obvious. Consider the direct implication. If $\max\{(a_4 - a_1), (b_4 - b_1)\} \neq 0$, then $S(\tilde{A}, \tilde{B}) = 1$, and, since both factors of $S(\tilde{A}, \tilde{B})$ are less than or equal to 1, theoretically

$$1 - (1 - \alpha - \beta) \left(1 - \frac{\int_0^1 \mu_{\tilde{A} \cap \tilde{B}}(x) dx}{\int_0^1 \mu_{\tilde{A} \cup \tilde{B}}(x) dx} \right) - \alpha \frac{\sum |a_i - b_i|}{4} - \beta \frac{d[(X_{\tilde{A}}, Y_{\tilde{A}}), (X_{\tilde{B}}, Y_{\tilde{B}})]}{M} = 1$$

$$\Rightarrow (1 - \alpha - \beta) \left(1 - \frac{\int_0^1 \mu_{\tilde{A} \cap \tilde{B}}(x) dx}{\int_0^1 \mu_{\tilde{A} \cup \tilde{B}}(x) dx} \right) + \alpha \frac{\sum |a_i - b_i|}{4} + \beta \frac{d[(X_{\tilde{A}}, Y_{\tilde{A}}), (X_{\tilde{B}}, Y_{\tilde{B}})]}{M} = 0,$$

and, as the three summands are positive or zero, necessarily: $\frac{\int_0^1 \mu_{\tilde{A} \cap \tilde{B}}(x) dx}{\int_0^1 \mu_{\tilde{A} \cup \tilde{B}}(x) dx} = 1$, $\frac{\sum |a_i - b_i|}{4} = 0$ and $\frac{d[(X_{\tilde{A}}, Y_{\tilde{A}}), (X_{\tilde{B}}, Y_{\tilde{B}})]}{M} = 0$. Thus, $\tilde{A} = \tilde{B}$.

If $\max\{(a_4 - a_1), (b_4 - b_1)\} = 0$, then we have analogously that $\frac{\sum |a_i - b_i|}{4} = 0$ and $\frac{d[(X_{\tilde{A}}, Y_{\tilde{A}}), (X_{\tilde{B}}, Y_{\tilde{B}})]}{M} = 0$. Thus, $\tilde{A} = \tilde{B}$. $\square$

Proposition 4. If $M = 1$ and $\tilde{A} = (a, a, a, a; 1)$ and $\tilde{B} = (b, b, b, b; 1)$, then $S(\tilde{A}, \tilde{B}) = 1 - \lfloor a - b \rfloor$.

Proof. Trivial. $\square$

We will see afterwards that the use of distances with $M=1$ has additional advantages.

Proposition 5. If $\tilde{A} = (a, a, a, a; 0)$ and $\tilde{B} = (a, a, a, a; 1)$, then $S(\tilde{A}, \tilde{B}) = 0$.

Proof. Trivial. $\square$

Proposition 6. If $S(\tilde{A}, \tilde{B}) = 0$ and $\tilde{A} \leq \tilde{B}$, then $|w_{\tilde{A}} - w_{\tilde{B}}| = 1$ or $\tilde{A} = (0, 0, 0, 0; w_{\tilde{A}})$ and $\tilde{B} = (1, 1, 1, 1; w_{\tilde{B}})$, with $w_{\tilde{A}}, w_{\tilde{B}} \in [0, 1]$.

Proof. Let us assume that $|w_{\tilde{A}} - w_{\tilde{B}}| \neq 1$, then if $\max\{(a_4 - a_1), (b_4 - b_1)\} \neq 0$

$$S(\tilde{A}, \tilde{B}) = 0 \Rightarrow (1 - \alpha - \beta) \left(1 - \frac{\int_0^1 \mu_{\tilde{A} \cap \tilde{B}}(x) dx}{\int_0^1 \mu_{\tilde{A} \cup \tilde{B}}(x) dx} - 1 \right) + \alpha \left(\frac{\sum |a_i - b_i|}{4} - 1 \right) + \beta \left(\frac{d[(X_{\tilde{A}}, Y_{\tilde{A}}), (X_{\tilde{B}}, Y_{\tilde{B}})]}{M} - 1 \right) = 0$$

and, since they are summands of $[0, 1]$, they must each be zero, i.e.

$$\frac{\int_0^1 \mu_{\tilde{A} \cap \tilde{B}}(x) dx}{\int_0^1 \mu_{\tilde{A} \cup \tilde{B}}(x) dx} = 0, \quad \frac{\sum |a_i - b_i|}{4} = 1 \quad \text{and} \quad \frac{d[(X_{\tilde{A}}, Y_{\tilde{A}}), (X_{\tilde{B}}, Y_{\tilde{B}})]}{M} = 1.$$

It follows from the second expression that $|a_i - b_i| = 1 \forall i$, and, as $\tilde{A} \leq \tilde{B}$, we have $b_i = a_i + 1 \forall i$, and, since $a_i, b_i \in [0, 1]$, necessarily $a_i = 0, b_i = 1 \forall i$.

If $\max\{(a_4 - a_1), (b_4 - b_1)\} = 0$, then:

$$1 - \left(\frac{1 - \alpha - \beta}{2} + \alpha \right) \frac{\sum |a_i - b_i|}{4} - \left(\frac{1 - \alpha - \beta}{2} + \beta \right) \frac{d[(X_{\tilde{A}}, Y_{\tilde{A}}), (X_{\tilde{B}}, Y_{\tilde{B}})]}{M} = 0$$

and an analogous analysis would be applied. $\square$

Other noteworthy observations about the proposed measure are:

1. For $\alpha + \beta = 1$, and d the Euclidean distance on $\mathbb{R}^2$, we have Xu et al.'s measure for trapezoidal fuzzy numbers such that $|w_{\tilde{A}} - w_{\tilde{B}}| = 0$.
2. The measure penalizes the fact that two sets are disjoint using a weight $(1 - \alpha - \beta)$.

3. As demonstrated in propositions 2-5, the proposed measure verifies the first four properties if we use a distance d with $M = \max_{[0,1] \times [0, \frac{1}{2}]} \{d((x, y), (x', y'))\} = 1$, like, for example, the distance $l_\infty((x_1, y_1), (x_2, y_2)) = \max\{|x_1 - x_2|, |y_1 - y_2|\}$. The fifth property holds only partially, since there are other pairs of fuzzy numbers whose similarity is zero in addition to the numbers $\tilde{A} = (0, 0, 0, 0; 0)$ and $\tilde{B} = (1, 1, 1, 1; 1)$
4. The proposal has the following property: $\lim_{\varepsilon \rightarrow 0} S((a, a, a, a; 0), (a, a, a, a; \varepsilon)) =$

1 and $\lim_{\varepsilon \rightarrow 1} S((a, a, a, a; 0), (a, a, a, a; \varepsilon)) = 0$, which overcomes that drawback of the measures proposed by Chen and Chen[2] and Wei and Chen [12], outlined in Section 2.

For example, the similarity between $\tilde{A} = (a, a, a, a; 0)$ and $\tilde{B} = (a, a, a, a; 10^{-10^{10}})$ is $S(\tilde{A}, \tilde{B}) \approx 1$, which appears to be more reasonable than the null value assigned by the measures proposed by Chen and Chen[2] and Wei and Chen [12].

5. The set $TF[0, 1; 1] = \{(a, b, c, d; 1) \in TF[0, 1]\}$ is a subset of $TF[0, 1]$, especially interesting in many domains of decision theory, since experts will often identify a linguistic term scale represented by numbers in $TF[0, 1; 1]$. However, certain considerations are required regarding the distance d used in the similarity measure. For convenience's sake we write $(a, b, c, d) \equiv (a, b, c, d; 1)$ to denote the elements in $TF[0, 1; 1]$. Suppose that the spheres identified by the distance d are not rectangular¹. If we restrict to $TF[0, 1; 1]$, a good measure of similarity should identify $(0, 0, 0, 0)$ and $(1, 1, 1, 1)$ as the most different elements. However, there exists $\tilde{A} \in TF[0, 1; 1]$ such that $S((0, 0, 0, 0), \tilde{A}) < S((0, 0, 0, 0), (1, 1, 1, 1))$. For example, if we take Xu et al.'s measure (2010), S_w , with $w = 0.5$, $S_{0.5}((0, 0, 0, 0), (1, 1, 1, 1)) = 0.052$.

As the spheres in the Euclidean distance in $\mathbb{R}^2$ are circles, if we represent the circle centered at the centroid of $(0, 0, 0, 0)$, i.e. at $(0, 0.5)$, whose radius is the distance to the centroid of $(1, 1, 1, 1)$, i.e. at $(1, 0.5)$, we obtain a region beyond this circumference containing the centroid of another number in $TF[0, 1; 1]$. This number will be farther from $(0, 0, 0, 0)$ than $(1, 1, 1, 1)$ itself.

More specifically, we know that the centroid of any number in $TF[0, 1; 1]$ is located in the band $[0, 1] \times [1/3, 1/2]$, [4]. The circle intersects the line $y = \frac{1}{3}$ at $x = \sqrt{\frac{35}{36}}$. Then, any number $(a, 1, 1, 1)$ with $a > \sqrt{\frac{35}{36}}$ is less similar to $(0, 0, 0, 0)$ than $(1, 1, 1, 1)$ itself. For instance, $S_{0.5}((0.99, 1, 1, 1), (0, 0, 0, 0)) = 0.049 < 0.52 = S_{0.5}((0, 0, 0, 0), (1, 1, 1, 1))$. However, this is not a problem if the spheres defined by a distance are rectangular. For example, with the distance l_∞ , whose spheres are square, we can ensure that the elements that differ most from $TF[0, 1; 1]$ are $(0, 0, 0, 0)$ and $(1, 1, 1, 1)$.

Another appropriate pseudo-distance is $d((x_1, y_1), (x_2, y_2)) = |x_1 - x_2|$, whose spheres are vertical bands. We denote this distance by l_* . As the

¹ The sphere with center a and radius r with distance d is $A(a, r) = \{x \in \mathbb{R}^n : d(x, a) = r\}$.

centroids are located in the band $[0, 1] \times [1/3, 1/2]$ [1], the range of variation on the abscissa is much greater than the ordinate. These distances mostly attach more importance to the position on the horizontal axis of the trapezoidal fuzzy numbers, which is, together with the shared area, one of the main parameters to be taken into account when identifying a linguistic term from the given fuzzy scale.

4 Comparative Analysis

We have compared the proposed measure using the distance l_∞ and the pseudo-distance l_* , with $\alpha = \beta = \frac{1}{3}$, with the measures by Chen and Chen [2], Wei and Chen [12], Xu et al. [13] and Gomathi and Sivaraman [7], outlined in Section 1. We have applied the measures to compute the similarity of 30 pairs of fuzzy numbers previously proposed by Xu et al. in [13], see Fig. 2. The results are shown in Table 1.

Fig. 2. Sets of fuzzy pairs for comparison

First, there are not great differences in the proposed measure when using l_* and l_∞ , see Table 1, since

$$l_\infty((X_{\tilde{A}}, Y_{\tilde{A}}), (X_{\tilde{B}}, Y_{\tilde{B}})) \neq l_*((X_{\tilde{A}}, Y_{\tilde{A}}), (X_{\tilde{B}}, Y_{\tilde{B}})) \Leftrightarrow |Y_{\tilde{A}} - Y_{\tilde{B}}| > |X_{\tilde{A}} - X_{\tilde{B}}|.$$

However, $|Y_{\tilde{A}} - Y_{\tilde{B}}| < \frac{1}{2} - \frac{1}{3} = \frac{1}{6}$, and therefore $|X_{\tilde{A}} - X_{\tilde{B}}| > \frac{1}{6}$, is a sufficient condition for both measures to coincide [1]. Also, whenever $|X_{\tilde{A}} - X_{\tilde{B}}| < \frac{1}{6}$, both measures will be very similar but not necessarily equal.

We also realize that there are missing values in Table 1. The similarity of the fuzzy numbers involved in cases 24, 25, 26 and 28 cannot be computed using the measure proposed by Chen and Chen [2] since the height of both fuzzy numbers is 0. Then, the term $\max\{Y_{\tilde{A}}, Y_{\tilde{B}}\}$ would be 0, leading to a division by 0, see Eq. (2). On the other hand, cases 26 and 28 cannot be addressed by the measures proposed by Wei and Chen [12] and Gomathi and Sivaraman [7], since the perimeter of both fuzzy numbers is also 0 and, again, we would have division by 0, see Eqs. (4) and (5), respectively.

In Fig. 5 we graphically compare the similarity measures. The proposed measures (with l_* and l_∞ , respectively) are always located at the ordinate axe, while the compared measure is located at the abscissa. Each point represent the degree of similarity output by the two compared measures for one out of the 30 pairs of fuzzy numbers. The farther the points are from the bisector of the first quadrant, the greater the difference between the measures compared.

Fig. 3. Charts comparing measures

Most of points in the graphs are located between the lines $y = x$ and $y = x - \frac{1}{3}$. It can be easily explained since the proposed measures penalize with weight $1 - \alpha - \beta = \frac{1}{3}$ the similarity of pairs of disjoint fuzzy numbers, as pointed out in the observations about the proposed method in Section 3.

We see that the biggest controversy is output when comparing the proposed measure with the measure proposed by Xu et al. [13], in which a large number of points considerably away from the bisector of the first quadrant. This matches up with the data in Table 1, in which values in the column corresponding to this measure are quite higher than the corresponding to the other. The higher differences appear when the two fuzzy numbers considered have very different heights, like cases 14, 19 and 24, where the values output by all measures are quite similar but for the measure by Xu et al.

Table 1. Comparison with other similarity measures

Set	[2]	[12]	[13]	[7]	l_*	l_∞	Set	[2]	[12]	[13]	[7]	l_*	l_∞
1	0.84	0.95	0.96	0.97	0.91	0.89	16	1	1	1	1	1	1
2	0.31	0.58	0.62	0.60	0.40	0.40	17	0.75	0.78	0.97	0.76	0.73	0.72
3	0.55	0.78	0.81	0.80	0.54	0.54	18	0.48	0.63	0.81	0.61	0.56	0.55
4	0.17	0.31	0.84	0.24	0.14	0.12	19	0	0.17	0.75	0.10	0.10	0.10
5	0	0.16	0.77	0.09	0	0	20	0	0.79	1	0.95	0.66	0.66
6	0	0	0.67	0	0	0	21	0.76	0.59	0.92	0.66	0.49	0.49
7	1	1	1	1	1	1	22	0	0	0.83	0	0.12	0.11
8	0.36	0.6	0.62	0.6	0.40	0.40	23	0	0.19	0.82	0.12	0.12	0.18
9	0.64	0.8	0.81	0.8	0.55	0.55	24	*	0.9	0.90	0.73	0.60	0.5998
10	0.8	0.82	0.96	0.80	0.75	0.72	25	*	0	0.93	0	0.61	0.62
11	0	0.16	0.77	0.08	0	0	26	*	*	0.81	*	0.53	0.53
12	0.44	0.32	0.86	0.36	0.25	0.23	27	0.45	0.45	0.89	0.45	0.48	0.48
13	0.40	0.52	0.89	0.46	0.56	0.56	28	*	*	1	*	1	1
14	0	0.19	0.80	0.10	0.12	0.11	29	0	0	1	0	1	0.99
15	0	0.98	1	0.94	0.66	0.66	30	0	0	0	0	0	0

We also note the large discrepancy with the measure by Chen and Chen [2] in cases 15, 20 and 29. This discrepancy is due to the drawback associated to this measure, outlined in Section 2, is overcome in the proposed one (Section 3), i.e., a high degree of similarity should be output to the fuzzy numbers of the form $(a, a, a, a, 0)$ and (a, a, a, a, ϵ) with ϵ near zero.

Finally, the measure by Gomathi and Sivaraman [7] significantly differs from the proposed in the sets 25 and 29. This measure also differs with the measures by Chen (even more than with the proposed) in sets 20 and 15.

5 Conclusions

Quantifying the degree of similarity between two fuzzy numbers is necessary in a great variety of applications of fuzzy logic, specially when a linguistic terms scale has been defined and a fuzzy number resulting from different computations has to be compared with the fuzzy numbers associated to the linguistic terms to identify the most similar one.

We have proposed a new similarity measure for generalized trapezoidal fuzzy numbers based on the difference of heights and the shared area between the numbers involved in relation to the total area of both, in addition to the distance between the centers of gravity and the geometric distance, which have already been considered by other authors. The result is a measure of similarity with many good properties, which outperforms the other measures in the sense that it can properly compare pairs of fuzzy numbers that the other methods can not address or do not fit well. Specifically, the proposed measure outperforms the measure by Chen since fuzzy numbers with null height can now be compared. It also outperforms measures by Wei and Chen and by Gomathi and Sivaraman

since it can compare fuzzy numbers with null perimeter. Moreover, the proposed measure keeps good properties of the other measures and establishes a more realistic similarity when comparing fuzzy numbers of the form $(a, a, a, a, 0)$ and (a, a, a, a, ϵ) , with ϵ near zero. Regarding the measure by Xu et al., it is outperformed by the proposed measure since properties 3 and 4 are accomplished and the additional drawback of this method illustrated in Section 2 is overcome.

However, the proposed measure does not necessarily identify the fuzzy numbers $\tilde{A} = (0, 0, 0, 0, 0)$ and $\tilde{B} = (1, 1, 1, 1, 1)$ as the most different, i.e., there are other numbers in $TF[0, 1]$ whose similarity is zero as well, i.e. property 5 is only satisfied in one direction. In any case, both the number and quality of the properties we won and the difficulties that the proposed measure exceeds represent benefits greater than the losses from the partial satisfaction of property 5.

Acknowledgment. The paper was supported by Madrid Government project S-2009/ESP-1685 and the Ministry of Science project MTM2011-28983-CO3-03.

References

1. Chen, S.J., Chen, S.M.: A new simple center-of-gravity method for handling the fuzzy ranking and the defuzzification problems. In: 8th National Conf. Fuzzy Theory Applications, Taipei, Republic of China (2000)
2. Chen, S.J., Chen, S.M.: Fuzzy risk analysis based on similarity measures of generalized fuzzy numbers. *IEEE Fuzzy Syst.* 11, 45–56 (2003)
3. Chen, S.J., Chen, S.M.: Fuzzy risk analysis based on the ranking of generalized trapezoidal fuzzy numbers. *Appl. Intell.* 26, 1–11 (2007)
4. Chen, S.M.: Operations on fuzzy numbers with function principle. *Tamkang J. Manage. Sc.* 6, 13–25 (1985)
5. Chen, S.M.: New methods for subjective mental workload assessment and fuzzy risk analysis. *Cybernet. Syst.* 27, 449–472 (1996)
6. Chen, S.M.: Ranking generalized fuzzy number with graded mean integration. In: *Proceedings of IFSA 1999*, vol. 2, pp. 899–902 (1999)
7. Gomathi Nayagam, V.L., Sivaraman, G.: A novel similarity measure between generalized fuzzy numbers. *Int. J. Comput. Theor. Eng.* 4, 448–450 (2012)
8. Jiménez, A., Mateos, A., Sabio, P.: Dominance intensity measure within fuzzy weight oriented MAUT: an application. *OMEGA* 41, 397–405 (2013)
9. Ross, T.J.: *Fuzzy logic with engineering applications*. John Wiley & Sons, Chichester (2010)
10. Sridevi, B., Nadarajan, R.: Fuzzy similarity measure for generalized fuzzy numbers. *Int. J. Open Probl. Comput. Sc. Math.* 2, 111–116 (2009)
11. Tran, L., Dukstein, L.: Comparison of fuzzy numbers using a fuzzy distance measure. *Fuzzy Set. Syst.* 130, 331–341 (2002)
12. Wei, S.H., Chen, S.M.: A new approach for fuzzy risk analysis based on similarity measures of generalized fuzzy numbers. *Expert Syst. Appl.* 36, 589–598 (2009)
13. Xu, Z., Shang, S., Qian, W., Shu, W.: A method for fuzzy risk analysis based on the new similarity of trapezoidal fuzzy numbers. *Expert Syst. Appl.* 37, 1920–1927 (2010)
14. Zadeh, L.A.: Fuzzy sets. *Inform. Control* 8, 338–353 (1965)
15. Zhang, W.R.: *Knowledge representation using linguistic fuzzy relations*. PhD. Dissertation, University of South Carolina, USA (1986)