The Wasserstein Distance for Multigranular Fuzzy Linguistic Term Sets

Amal Esmail[1,2] , **Eddy Soria**[3] , **Jordi Pascual**[1,2] , **Antonio Moreno**[1,2] and **Aida Valls**[1,2]

[1]Dept. Enginyeria Informàtica i Matemàtiques, Universitat Rovira i Virgili, Tarragona, Spain
[2]TecnATox: Centre of Environmental, Food and Toxicological Technology, Tarragona, Spain
[3]Dept. Business Management, Universitat Rovira i Virgili, Reus, Spain
{amalesmailqasem.al-zubairi,eddy.soria, jordi.pascual, antonio.moreno, aida.valls}@urv.cat

Abstract

Linguistic vocabularies used to assess the quality of alternatives in decision support systems may be different depending on the person who makes the evaluation. Fixing a single set of terms can cause problems due to differing semantic interpretations among participants. In this paper, we address the case of having multiple and multigranular linguistic term sets, based on trapezoidal fuzzy sets. A new dissimilarity metric between linguistic term sets based on the Wasserstein distance is proposed. It is then used to measure the consensus in group decision making in an illustrative case study.

1 Introduction

Linguistic vocabularies may be used for qualitative assessment of the available alternatives in decision support systems [Villani, 2009; Abuasaker *et al.*, 2023; Ren and Hao, 2025; Cabrerizo *et al.*, 2010]. In group decision making, giving the flexibility to use different linguistic term sets facilitates the evaluation task of the participants. In fact, if the same terms are fixed in advance, the problem of different people considering different semantic interpretations appears [Abuasaker *et al.*, 2023; Cabrerizo *et al.*, 2010]. The use of fuzzy sets to represent the meaning of the linguistic terms in qualitative reasoning is widespread [Cabrerizo *et al.*, 2010].

When different and multigranular sets of fuzzy linguistic terms are allowed in a system, it may be required to know their similarity to compare them to a consensus linguistic term set, or to evaluate the degree of consensus among the experts. In both cases, a distance measure is needed.

In this paper, we propose the use of the Wasserstein distance between two linguistic term sets using trapezoidal fuzzy membership functions. The paper gives first a definition of distance between trapezoidal fuzzy sets, which is later used as cost in the optimal transport construction of the Wassertstein distance. The Wasserstein distance is known for comparing different probability distributions over the same scale [Villani, 2009]. In our proposal, we interpret the linguistic terms as a uniform probability distribution on a fixed numerical range $[a, b]$. The paper presents a short case study to illustrate the results of this distance.

2 Preliminaries

Let $U = [a, b] \subset R$ with $a < b$. A **Trapezoidal Fuzzy Number** (TrFN) on U is a 4-tuple $x = (x^{(1)}, x^{(2)}, x^{(3)}, x^{(4)})$ satisfying $a \leq x^{(1)} \leq x^{(2)} \leq x^{(3)} \leq x^{(4)} \leq b$.

The interval $[x^{(2)}, x^{(3)}]$ is the core (membership equal to 1) and $[x^{(1)}, x^{(4)}]$ is the support (positive membership). When $x^{(2)} = x^{(3)}$, we have a triangular fuzzy number, which is considered a special case of a trapezoidal function.

We denote by $\tau_{[a,b]}$ the set of all such TrFNs.

A **Linguistic Term Set** over U is a finite collection $A = \{A_1, \ldots, A_m\}$ with $A_i \in \tau_{[a,b]}$, where $i \neq j \rightarrow A_i \neq A_j$.

Having two linguistic terms sets $A = \{A_1, \ldots, A_m\}$ and $B = \{B_1, \ldots, B_n\}$ of possibly different cardinalities, the remainder of this paper develops a distance between them. Note that this corresponds to a distance between multigranular unbalanced term sets defined in the same numerical interval U. An important characteristic of this measure is that it does not require any correspondence between their labels.

3 The Trapezoidal Fuzzy Distance

We first define a distance between individual trapezoidal fuzzy numbers, with values in the range $[0, M]$, being $M > 0$ a fixed scale parameter

Definition 1. Trapezoidal Fuzzy Distance.

The *Trapezoidal Fuzzy Distance* $d_{\mathrm{TrF}} : \tau_{[a,b]} \times \tau_{[a,b]} \rightarrow [0, M]$ is defined as:

$$d_{\mathrm{TrF}}(x, y) = \frac{M}{6(b-a)}\Big(|x^{(1)} - y^{(1)}| + 2|x^{(2)} - y^{(2)}| + 2|x^{(3)} - y^{(3)}| + |x^{(4)} - y^{(4)}|\Big). \tag{1}$$

The weights $(1, 2, 2, 1)$ assign greater importance to the core vertices than to the support boundaries, and the factor $1/6$ normalizes the weights so that d_{TrF} is bounded by M (Proposition 2). The parameter M defines the maximum value of the output scale.This definition extends the one of the Triangular Fuzzy Rescaling Distance defined in [Soria *et al.*, 2026].

The coefficients (1, 2, 2, 1) are those proposed in [Kaufmann and Gil-Aluja, 1988], prioritizing core vertices over support boundaries.

3.1 Properties

Proposition 1. Metric. The function d_{TrF} is a metric on $\tau_{[a,b]}$.

Proof. Non-negativity, symmetry, and implication $(x = y) \to (d_{\mathrm{TrF}}(x,y) = 0)$ are immediately derived from the corresponding properties of the absolute value. For the converse implication, $d_{\mathrm{TrF}}(x,y) = 0$ forces the equality $|x^{(k)} - y^{(k)}| = 0$ for every $k \in \{1,2,3,4\}$, and hence $x = y$. Finally, the triangle inequality is proven by applying the scalar triangle inequality vertex-wise $k \in \{1,2,3,4\}$:

$$|x^{(k)} - y^{(k)}| \leq |x^{(k)} - z^{(k)}| + |z^{(k)} - y^{(k)}|$$

and averaging them with the positive weights $(1,2,2,1)$. $\square$

Proposition 2. Boundedness. For every $x, y \in \tau_{[a,b]}$, $0 \leq d_{\mathrm{TrF}}(x,y) \leq M$.

Proof. Non-negativity has already been established. For the upper bound, since

$$x^{(k)}, y^{(k)} \in [a,b] \quad \text{for every } k \in \{1,2,3,4\},$$

we have

$$|x^{(k)} - y^{(k)}| \leq b - a,$$

and using the weights $1+2+2+1 = 6$, it follows that

$$d_{\mathrm{TrF}}(x,y) \leq \frac{M}{6(b-a)} \cdot 6(b-a) = M.$$

$\square$

Proposition 3. Topological structure of the ground space. The metric space $\left(\tau_{[a,b]}, d_{\mathrm{TrF}}\right)$ is complete and separable.

Proof. The set $\tau_{[a,b]}$ is the subset of $\mathbb{R}^4$ defined by the closed conditions $a \leq x^{(1)} \leq x^{(2)} \leq x^{(3)} \leq x^{(4)} \leq b$, and is therefore closed and bounded in $\mathbb{R}^4$. The metric d_{TrF} is a positively weighted ℓ^1 norm on the differences of coordinates and is therefore equivalent to the Euclidean norm on $\mathbb{R}^4$ when restricted to $\tau_{[a,b]}$. As a closed and bounded subset of a finite-dimensional Euclidean space, it is complete and separable. $\square$

4 The Wasserstein Distance for Multigranular Fuzzy Linguistic Term Sets

We now extend the distance from individual TrFNs to entire linguistic term sets. The idea is to represent each label set as an empirical probability measure on $(\tau_{[a,b]}, d_{\mathrm{TrF}})$ and to define the distance between linguistic term sets as the optimal transport cost between these measures, with d_{TrF} as the ground cost.

To each finite set $A = \{A_1, \ldots, A_m\}$, we associate a uniform probability distribution that assigns to each element the mass as:

$$\mu_A(A_i) = \frac{1}{m},\ i = 1, ..., m. \tag{2}$$

Definition 2. Multigranular Fuzzy Wasserstein Distance. The *Multigranular Fuzzy Wasserstein Distance* between μ_A and μ_B is defined as

$$d_{\mathrm{MFW}}(\mu_A, \mu_B) = \min_{\pi \in \Pi(\mu_A, \mu_B)} \sum_{i=1}^{m} \sum_{j=1}^{n} \pi_{ij}\, d_{\mathrm{TrF}}(A_i, B_j) \tag{3}$$

where $\Pi(\mu_A, \mu_B)$ denotes the set of transport matrices

$$\Pi(\mu_A, \mu_B) = \left\{ \begin{array}{c} \pi \in \mathbb{R}^{m \times n}_{\geq 0} \ : \ \sum_j \pi_{ij} = \dfrac{1}{m}\ \forall i \in [1,m], \\ and \sum_i \pi_{ij} = \dfrac{1}{n}\ \forall j \in [1,n] \end{array} \right\} \tag{4}$$

Each π_{ij} represents the fractional mass transferred from label A_i to label B_j. The marginal constraints encode mass conservation: each label in A contributes mass $1/m$, each label in B receives mass $1/n$. Notice that the fuzzy membership functions of the labels are not considered as probability distributions; instead, each label is considered an element with equal probability with regards to the mass transference, while the membership function is used when calculating the trapezoidal fuzzy distance.

This definition is the standard discrete Wasserstein-1 distance between empirical probability measures with uniform mass in the metric space $(\tau_{[a,b]}, d_{\mathrm{TrF}})$ [Villani, 2009]. Because linguistic labels are on a common ordered 1D scale and their assigned mass is uniform, the optimal Wasserstein-1 coupling is the comonotone (quantile) coupling [Abdellatif *et al.*, 2025]. Therefore, we can compute π^* by monotonous cumulative mass matching (order-preserving, no crossings). Otherwise, a generic linear program could be used.

4.1 Properties

Proposition 4. Boundedness. For every pair of label sets, $0 \leq d_{\mathrm{MFW}}(\mu_A, \mu_B) \leq M$.

Proof. Non-negativity follows from $d_{\mathrm{TrF}} \geq 0$ and $\pi_{i,j} \geq 0$. For the upper bound, by Proposition 2, the maximum value of the distance between to labels is $d_{\mathrm{TrF}}(\mu_A, \mu_B) = M$ for all i, j, and by marginal constraints we have $\sum_{i,j} \pi_{i,j} = 1$, hence in this case we have that the upper bound is

$$\sum_{i,j} \pi_{ij}\, d_{\mathrm{TrF}}(A_i, B_j) = M.$$

$\square$

Proposition 5. Metric. The function d_{MFW} is a metric on the set of empirical probability measures induced by finite label sets in $\tau_{[a,b]}$. That is, for all label sets A, B, C:

1. $d_{\mathrm{MFW}}(\mu_A, \mu_B) \geq 0$, with equality if and only if $\mu_A = \mu_B$;
2. $d_{\mathrm{MFW}}(\mu_A, \mu_B) = d_{\mathrm{MFW}}(\mu_B, \mu_A)$;
3. $d_{\mathrm{MFW}}(\mu_A, \mu_C) \leq d_{\mathrm{MFW}}(\mu_A, \mu_B) + d_{\mathrm{MFW}}(\mu_B, \mu_C)$.

Proof. By Proposition 1, d_{TrF} is a metric on $\tau_{[a,b]}$. By Proposition 3, $(\tau_{[a,b]}, d_{\mathrm{TrF}})$ is a complete and separable metric space. The function d_{MFW} is the Wasserstein distance W with d_{TrF} as ground cost. By [Villani, 2009], W is a metric on the space of probability measures with finite first moment over a complete and separable metric space. It is satisfied by the empirical measures μ_A, μ_B, μ_C, with finite support contained in the bounded set $\tau_{[a,b]}$, hence having finite first moment. □

Proposition 6. Reduction to the classical Wasserstein-1 distance. Suppose that every label in A and B is a degenerate TrF, that is,

$$A_i = (a_i, a_i, a_i, a_i), \qquad B_j = (b_j, b_j, b_j, b_j),$$

with $a_i, b_j \in [a, b]$. Then,

$$d_{\mathrm{MFW}}(\mu_A, \mu_B) = \frac{M}{b-a} W_1^{\mathbb{R}}(\tilde{\mu}_A, \tilde{\mu}_B),$$

where

$$\tilde{\mu}_A = a_i, \qquad \tilde{\mu}_B = b_j$$

are empirical measures on $\mathbb{R}$, and $W_1^{\mathbb{R}}$ denotes the classical Wasserstein-1 distance on $\mathbb{R}$ with ground cost $|x - y|$.

Proof. For degenerate TrFNs, the four absolute differences in the definition of d_{TrF} all collapse to $|a_i - b_j|$, giving

$$d_{\mathrm{TrF}}(A_i, B_j) = \frac{M}{6(b-a)} \cdot 6\,|a_i - b_j| = \frac{M}{b-a}\,|a_i - b_j|.$$

Substituting this expression into Definition 2 and factoring the constant $\frac{M}{b-a}$ out of the minimization yields the stated identity. □

5 Case Study

In this section we will briefly outline one of the possible applications of the distance defined in this paper. Let us assume we have a group of 10 experts $(A, B, \ldots, J)$. Each one uses a different fuzzy linguistic term set (Figure 1) for the qualitative evaluation of some objects. As the number and the interpretation of the terms are different, we want to calculate the degree of consensus in the terms. Following [Cabrerizo *et al.*, 2010], we calculate the collective dissimilarity matrix as the distance between each of the term sets (Table 1). In this example, G and H are the most different vocabularies (distance of 32.5), while I and J are the most similar (2.5), with a very similar distribution as can be seen in Figure 1).

The average distance of an expert from the rest $\bar{X}$, (last row in Table 1) can be used to determine which of the term sets could be used as representative, that is, the one that requires fewer changes in the vocabularies of the other experts. In this example, it corresponds to expert J, with an average of 11.8.

The overall degree of consensus is the average of all distances in the matrix, which corresponds to 15.02. If we consider some subgroups of experts; for example, the degree of consensus of the group $\{G, H, I, J\}$ is 18.55, while for the group $\{C, D, E, F\}$ is 11.01. It may help to build subgroups of experts that have more similar qualitative scales, like the second group, in this case.

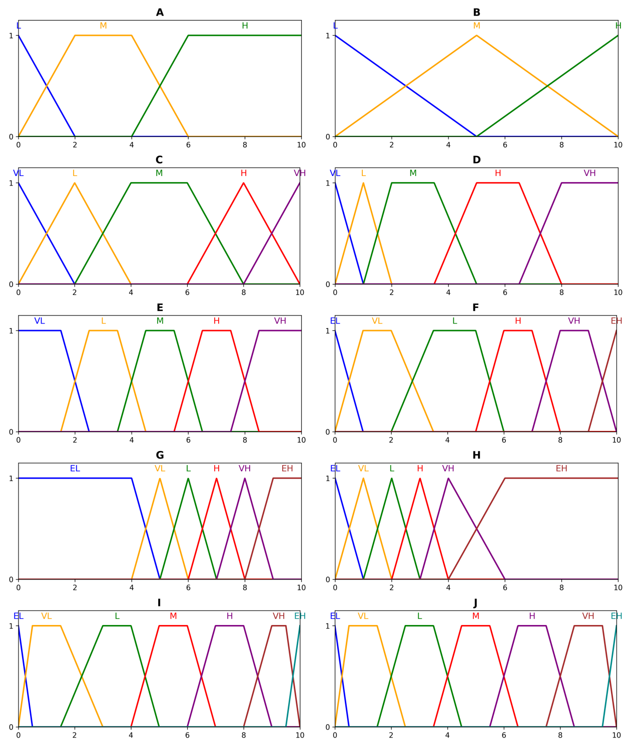

Figure 1: Fuzzy linguistic term sets of 10 experts in the case study

6 Conclusions and Future Work

In this paper, we proposed a new distance between multi-granular fuzzy linguistic term sets represented by trapezoidal fuzzy numbers. It combines a distance defined between individual trapezoidal fuzzy sets with the Wasserstein optimal transport method to compare complete linguistic vocabularies.

One of the main advantages of the method is that it does not require any predefined correspondence between labels, which makes it suitable for comparing linguistic term sets with different numbers of terms and different semantic interpretations. It only requires the labels to be defined in a common numerical interval $[a, b]$.

	A	B	C	D	E	F	G	H	I	J
A	0.0	13.3	16.4	10.1	16.8	16.3	28.6	11.7	18.1	16.3
B	13.3	0.0	12.7	17.9	16.2	12.6	21.4	23.9	14.0	14.4
C	16.4	12.7	0.0	13.0	8.0	7.4	16.8	20.7	7.7	7.8
D	10.1	17.9	13.0	0.0	13.0	14.5	26.1	9.9	16.2	13.8
E	16.8	16.2	8.0	13.0	0.0	10.3	13.6	20.2	9.1	7.0
F	16.3	12.6	7.4	14.5	10.3	0.0	14.6	21.3	6.6	6.4
G	28.6	21.4	16.8	26.1	13.6	14.6	0.0	32.5	15.5	16.9
H	11.7	23.9	20.7	9.9	20.2	21.3	32.5	0.0	23.2	20.7
I	18.1	14.0	7.7	16.2	9.1	6.6	15.5	23.2	0.0	2.5
J	16.3	14.4	7.8	13.8	7.0	6.4	16.9	20.7	2.5	0.0
$\bar{X}$	16.4	16.3	12.3	14.9	12.7	12.2	20.7	20.4	12.6	**11.8**

Table 1: Collective distance matrix and Expert's average distance.

This distance can be useful in problems where different linguistic term sets must be compared. The case of finding similarities between the scales used by different experts has been taken as an illustrative example. It may also be useful to compare linguistic scales with respect to a reference one (e.g., one generated as a centroid). This problem appears for example in data fusion processes as presented in [Herrera *et al.*, 2000; Wan *et al.*, 2025], and in group decision aiding applications when the information about the alternatives is given by different experts using different linguistic vocabularies, such as for job selection [Nguyen *et al.*, 2018], student's competencies assessment [Porro *et al.*, 2021] or energy storage evaluation [Liang *et al.*, 2022], among many others.

Acknowledgments

This research was funded by Universitat Rovira i Virgili (2023PFR-URV-00114), Instituto de Salud Carlos III and FEDER (PI21/00064). Mrs. Esmail is funded by a predoctoral grant from URV (2023PMF-PIPF22). Mr. Soria is funded by AGAUR predoctoral grant (2023 FI-1 00622), co-funded by European Union.

References

[Abdellatif *et al.*, 2025] Mariem Abdellatif, Peter Kuchling, Barbara Rüdiger, and Irene Ventura. Wasserstein distance in terms of the comonotonicity copula. *Stochastics*, 97(8):1097–1108, 2025.

[Abuasaker *et al.*, 2023] Walaa Abuasaker, Jennifer Nguyen, Núria Agell, Francisco Javier Ruiz, and Mónica Sánchez. An application to measure consensus on customers' ratings using hesitant fuzzy linguistic term sets. In *Artificial Intelligence Research and Development. Proc. 25th Int. Conf. of the Catalan Association for Artificial Intelligence*, volume 375, pages 291–300. IOS Press, 2023.

[Cabrerizo *et al.*, 2010] Francisco Javier Cabrerizo, Juan Manuel Moreno, Ignacio J Pérez, and Enrique Herrera-Viedma. Analyzing consensus approaches in fuzzy group decision making: advantages and drawbacks. *Soft Computing*, 14(5):451–463, 2010.

[Herrera *et al.*, 2000] F. Herrera, E. Herrera-Viedma, and Luis Martınez. A fusion approach for managing multi-granularity linguistic term sets in decision making. *Fuzzy Sets and Systems*, 114(1):43–58, 2000.

[Kaufmann and Gil-Aluja, 1988] Arnold Kaufmann and Jaime Gil-Aluja. *Modelos para la investigación de efectos olvidados*. Milladoiro, Vigo, Spain, 1988.

[Liang *et al.*, 2022] Yuanyuan Liang, Yanbing Ju, Peiwu Dong, Luis Martínez, Xiao-Jun Zeng, Ernesto D.R. Santibanez Gonzalez, Mihalis Giannakis, Jinhua Dong, and Aihua Wang. Sustainable evaluation of energy storage technologies for wind power generation: A multistage decision support framework under multi-granular unbalanced hesitant fuzzy linguistic environment. *Applied Soft Computing*, 131:109768, 2022.

[Nguyen *et al.*, 2018] J. Nguyen, G. Sánchez-Hernández, A. Armisen, N. Agell, X. Rovira, and C. Angulo. A linguistic multi-criteria decision-aiding system to support university career services. *Applied Soft Computing*, 67:933–940, 2018.

[Porro *et al.*, 2021] Olga Porro, Núria Agell, Mónica Sánchez, and Francisco Javier Ruiz. A multi-attribute group decision model based on unbalanced and multi-granular linguistic information: An application to assess entrepreneurial competencies in secondary schools. *Applied Soft Computing*, 111:107662, 2021.

[Ren and Hao, 2025] Fangling Ren and Fei Hao. A new fusion method of fuzzy numbers and linguistic terms based on individual semantics in mixed decision making problems. *International Journal of Fuzzy Systems*, 27(6):1887–1903, 2025.

[Soria *et al.*, 2026] Eddy Soria, Aida Valls, and Ana Beatriz Hernández-Lara. Triangular fuzzy rescaling distance. In Vicenç Torra, Yasuo Narukawa, and Josep Domingo-Ferrer, editors, *Modeling Decisions for Artificial Intelligence*, pages 117–129, Cham, 2026. Springer Nature Switzerland.

[Villani, 2009] Villani. *Optimal transport: old and new*, volume 338. Springer, 2009.

[Wan *et al.*, 2025] Shu-Ping Wan, Wen-Chang Zou, Jiu-Ying Dong, and Yu Gao. Dual strategies consensus reaching process for ranking consensus based probabilistic linguistic multi-criteria group decision-making method. *Expert Systems with Applications*, 262:125342, 2025.