

A Statistical Comparative Study of Different Similarity Measures of Consensus in Group Decision Making

F. Chiclana^{1a}, J. M. Tapia García^b, M. J. del Moral^c, E. Herrera-Viedma^d

^a*DIGITS, Department of Informatics, Faculty of Technology, De Montfort University, Leicester LE1 9BH, UK.*

^b*Department of Quantitative Methods in Economic and Business, University of Granada, 18071 Granada Spain.*

^c*Department of Statistics and Operational Research, University of Granada, 18071 Granada Spain.*

^d*Department of Computer Science and A.I., University of Granada, 18071 Granada Spain.*

chiclana@dmu.ac.uk; jmtaga@ugr.es; delmoral@ugr.es; viedma@decsai.ugr.es

Abstract

An essential aim in Group Decision Making (GDM) problems is to achieve a high level of consensus among experts. Consensus is defined as general or widespread agreement, and it is usually modelled mathematically via a similarity function measuring how close experts' opinions or preferences are. Similarity functions are defined based on the use of a metric describing the distance between experts' opinions or preferences. In the literature, different metrics or distance functions have been proposed to implement in consensus models, but no study has been carried out to analyse the influence the use of different distance functions can have in the GDM process.

This paper presents a comparative study of the effect of the application of some different distance functions for measuring consensus in GDM. By using the nonparametric Wilcoxon matched-pairs signed-ranks test, it is concluded that different distance functions can produce significantly different results. Moreover, it is also shown that their application also has a significant effect on the speed of achieving consensus. Finally, these results are analysed and used to derive decision support rules, based on a convergent criterion, that can be used to control the convergence speed of the consensus process using the compared distance functions.

Keywords: Group decision making, fuzzy preferences, similarity, consensus, decision support rules, Wilcoxon test

1. Introduction

In order to reach a decision, experts have to express their opinions or preferences by means of a set of evaluations over a set of alternatives. Consensus is defined as the full and unanimous agreement of all the experts regarding all the feasible alternatives. In practice, this definition is inconvenient because it only allows differentiating between two states, namely, the existence and absence of consensus. Also, the chances for reaching such a full agreement are rather low. Indeed, unanimity is not necessary in most real life situations. A second meaning of the concept of consensus refers to the judgement arrived at by 'most of' those concerned (<http://www.merriam-webster.com>), which has led to the definition and use of a new concept of consensus degree referred to as 'soft' consensus degree [16, 18, 26, 30, 31, 33, 40, 50].

Based on the use of such soft consensus measure, the consensus process can be modelled as a dynamic and iterative group discussion process, coordinated by a moderator, who helps the experts to make their opinions closer. In each step of this process, the moderator knows

¹Corresponding author email: chiclana@dmu.ac.uk

Cite as: F. Chiclana, J.M. Tapia Garcia, M.J. del Moral, E. Herrera-Viedma, A Statistical Comparative Study of Different Similarity Measures of Consensus in Group Decision Making, *Information Sciences* (2012), doi: <http://dx.doi.org/10.1016/j.ins.2012.09.014>

the actual level of consensus between the experts, by means of the soft consensus measure, which establishes the distance to an ‘ideal’ state of consensus. If the consensus level is not acceptable, that is, if it is lower than a specified threshold, which means that there exists a great discrepancy between the experts’ opinions, then the moderator would urge the experts to discuss their opinions further in an effort to make them closer. On the contrary, when the consensus level is acceptable, the moderator would apply a selection process in order to obtain the final consensus solution to the group decision making (GDM) problem [1, 27].

Soft consensus measures represent the level of agreement among experts, and therefore their definition is based on the concept of similarity between their opinions (preferences). The evaluation of consensus necessarily implies the computation and aggregation of the ‘distance’ representing disagreement between the opinions (preferences) of each pair of experts on each pair of alternatives [3]. An issue here is that the convergence of the consensus process towards a solution acceptable by most of the experts could be affected by the particular metric, i.e. distance function, used to measure disagreement and subsequently to compute the soft consensus measure.

The aim of this paper is to present a comparative study between five of the most commonly used distance functions in modelling soft consensus in GDM problems: Manhattan, Euclidean, Cosine, Dice, and Jaccard distance functions. Two-sample statistical tests are used to establish whether the application of two distance functions is different, or whether one distance function is ‘better’ than another. Indeed, the statistical comparative study carried out, and reported in this paper, covers both aspects mentioned above: it is tested whether the application of different distance functions produces significant differences in the measuring of the consensus, and also it is analysed which distance functions are better than others with respect to the speed of convergence of the decision making process towards the achievement of an acceptable level of consensus by the group of experts. Using nonparametric Wilcoxon tests [45, 49], significant differences were found in many cases between the behaviour of the compared distance functions. This behaviour was further analysed using a convergent criterion and a set of rules were identified for their application to control the speed of convergence towards consensus.

The remainder of the paper is structured as follows. Section 2 introduces concepts essential to the understanding of the rest of the paper: the GDM problem (Subsection 2.1), the selection process (Subsection 2.2) and the consensus process (Subsection 2.3). Following that, Section 3 describes the design of the experiment used to evaluate the different distance functions for measuring consensus in GDM problems. Section 4 presents and discusses the results of the experiment. Section 5 includes a practical application of the use of different compared distance functions for the same GDM problem to illustrate the application of the identified rules affecting the consensus process convergence. Lastly, Section 6 concludes the paper.

2. Preliminaries

To make the paper self-contained, the main concepts that will be used are introduced here.

2.1. The GDM Problem

GDM problems consist in finding the best alternative(s) from a set of feasible alternatives $X = \{x_1, \dots, x_n\}$ according to the preferences provided by a group of experts $E = \{e_1, \dots, e_m\}$. Different preference elicitation methods were compared in [38], where it was concluded that pairwise comparison methods are more accurate than non-pairwise methods. Among the different representation formats that experts may use to express their opinions, fuzzy preference relations [8, 9, 29, 34, 39, 46] are one of the most used because of their utility and usability as a tool for modelling experts’ preferences and in their aggregation into group ones in decision processes [10, 28, 29, 47].

Definition 1 (Fuzzy Preference Relation). A fuzzy preference relation P on a finite set of alternatives X is characterized by a membership function $\mu_P: X \times X \rightarrow [0, 1]$, $\mu(x_i, x_j) = p_{ij}$, verifying

$$p_{ij} + p_{ji} = 1 \quad \forall i, j \in \{1, \dots, n\}. \quad (1)$$

When cardinality of X is small, the fuzzy preference relation may be conveniently denoted by the matrix $P = (p_{ij})$. The following interpretation is also usually assumed:

- $p_{ij} = 1$ indicates the maximum degree of preference for x_i over x_j .
- $p_{ij} \in]0.5, 1[$ indicates a definite preference for x_i over x_j .
- $p_{ij} = 0.5$ indicates indifference between x_i and x_j .

Two different processes are applied in GDM problems before a final solution can be obtained [5, 22, 27, 32, 41]: (1) *the consensus process* and (2) *the selection process*. The consensus process refers to how to obtain the maximum degree of consensus or agreement between the set of experts. The selection process obtains the final solution according to the preferences [8] given by the experts.

2.2. Selection Process

The selection process involves two different steps [24, 43]: (i) *aggregation* of individual preferences and (ii) *exploitation* of the collective preference.

Aggregation phase. This phase defines a collective preference relation, $P^c = (p_{ij}^c)$, obtained by means of the aggregation of all individual fuzzy preference relations $\{P^1, P^2, \dots, P^m\}$, and indicates the global preference between every pair of alternatives according to the majority of experts' opinions. Currently, at least 90 different families of aggregation operators have been studied [11, 12, 17, 19, 21, 35, 51, 52, 55, 56]. Among them the Ordered Weighted Averaging (OWA) operator proposed by Yager [52] is the most widely used.

The aggregation operation by means of a quantifier guided OWA operator, ϕ_Q , is carried out as follows:

$$p_{ij}^c = \phi_Q(p_{ij}^1, \dots, p_{ij}^m) = \sum_{k=1}^m w_k \cdot p_{ij}^{\sigma(k)}, \quad (2)$$

where σ is a permutation function such that $p_{ij}^{\sigma(k)} \geq p_{ij}^{\sigma(k+1)}$, $\forall k = 1, \dots, m-1$; Q is a fuzzy linguistic quantifier [54] that represents the concept of fuzzy majority and it is used to calculate the weighting vector of ϕ_Q , $W = (w_1, \dots, w_m)$ such that, $w_k \in [0, 1]$ and $\sum_{k=1}^m w_k = 1$, according to the following expression [52]:

$$w_k = Q(k/n) - Q((k-1)/n), \quad \forall k \in \{1, \dots, m\}. \quad (3)$$

Some examples of linguistic quantifiers, depicted in Figure 1, are “*at least half*”, “*most of*” and “*as many as possible*”, which can be represented by the following function

$$Q(r) = \begin{cases} 0 & \text{if } 0 \leq r < a \\ \frac{r-a}{b-a} & \text{if } a \leq r \leq b \\ 1 & \text{if } b < r \leq 1 \end{cases} \quad (4)$$

using the values $(0, 0.5)$, $(0.3, 0.8)$ and $(0.5, 1)$ for (a, b) , respectively [29].

Alternative representations for the concept of fuzzy majority can be found in the literature. For example, Yager in [53] considered the parameterized family of RIM quantifiers $Q(r) = r^a$ ($a \geq 0$) for such representation. This family of functions guarantees that [12]: (i) all the experts contribute to the final aggregated value (strict monotonicity property), and (ii) associates, when $a \in [0, 1]$, higher weight values to the aggregated values with associated higher importance values (concavity property).

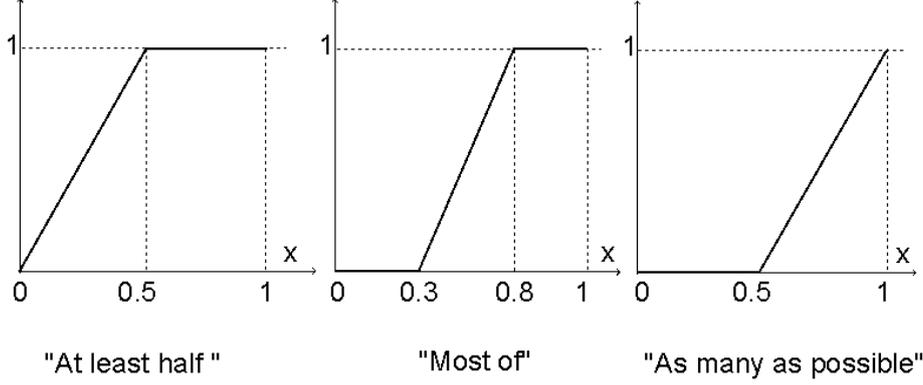


Figure 1: Linguistic quantifiers “at least half”, “most of” and “as many as possible”

Exploitation phase. This phase transforms the global information about the alternatives into a global ranking of them, from which the set of solution alternatives is obtained. The global ranking is obtained by applying two choice degrees of alternatives to the collective fuzzy preference relation [20]: the *quantifier guided dominance degree* and the *quantifier guided non dominance degree*.

1. *Quantifier Guided Dominance Degree:* For the alternative x_i we calculate the quantifier guided dominance degree, $QGDD_i$, used to quantify the dominance that alternative x_i has over all the others in a fuzzy majority sense as follows:

$$GDD_i = \phi_Q(p_{ij}^c, j = 1, \dots, n). \quad (5)$$

2. *Quantifier Guided Non Dominance Degree:* We also calculate the quantifier guided non dominance degree, $QGNDD_i$, according to the following expression:

$$QGNDD_i = \phi_Q(1 - p_{ji}^s, j = 1, \dots, n), \quad (6)$$

where $p_{ji}^s = \max\{p_{ji}^c - p_{ij}^c, 0\}$ represents the degree to which x_i is strictly dominated by x_j . In our context, $QGNDD_i$, gives the degree in which each alternative is not dominated by a fuzzy majority of the remaining alternatives.

Finally, the solution X_{sol} is obtained by applying these two choice degrees and selecting those alternatives with maximum choice degrees.

Clearly, it is preferable that the experts achieve a high level of consensus concerning their preferences before applying the selection process.

2.3. Consensus Model

Initially, in any GDM problem the experts could disagree in their opinions so that consensus can be viewed as an iterative process, which means that agreement is obtained only after a number of rounds of consultation. In each round a consensus support system calculates two consensus parameters [4, 22, 23]: (i) the *consensus measure* to guide the consensus process and (ii) the *proximity measure* to support the group discussion phase of the consensus process. These measures are computed and used to find out the consensus state of the process at the three different levels of a fuzzy preference relation: the pairs of alternatives, the alternatives and the relation levels. This will allow us, for example, to identify which experts are close to the consensus solution, or in which alternatives the experts are having more trouble to reach consensus.

The computation of the level of agreement among experts involves necessarily the measurement of the distance or, equivalently, the similarity between their preference values. In the following, we provide the formal definition of distance and similarity functions as given in [15]:

Definition 2 (Distance). Let A be a set. A function $d: A \times A \rightarrow \mathbb{R}$ is called a distance (or dissimilarity) on A if, for all $x, y \in A$, there holds

1. $d(x, y) \geq 0$ (non-negativity)
2. $d(x, y) = d(y, x)$ (symmetry)
3. $d(x, x) = 0$ (reflexivity)

Definition 3 (Similarity). Let A be a set. A function $s: A \times A \rightarrow \mathbb{R}$ is called a similarity on A if s is non-negative, symmetric, and if $s(x, y) \leq s(x, x)$ holds for all $x, y \in A$, with equality if and only if $x = y$.

The main transforms between a distance d and a similarity s bounded by 1 are [15]:

$$d = 1 - s; \quad d = \frac{1 - s}{s}; \quad d = \sqrt{1 - s}; \quad d = \sqrt{2(1 - s^2)}; \quad d = \arccos s; \quad d = -\ln s$$

In this paper, we use the first transform to go from a distance function to a similarity function.

The similarity function is used for measuring both consensus degrees and proximity measures. The first ones are calculated by fusing the similarity of the preference values of all the experts on each pair of alternatives as per the expression (7) below. The second ones are calculated by measuring the similarity between the preferences of each expert in the group and the collective preferences, previously obtained by fusing all the individual experts' preferences.

The main problem is how to find a way of making individual positions converge and, therefore, how to support the experts in obtaining and agreeing with a particular solution. To do this, a consensus level required for that solution is fixed in advance. When the consensus measure reaches this level then the decision making session is finished and the solution is obtained. If that is not the case, the experts' opinions must be modified. This is done in a group discussion session in which the consensus support system uses the proximity measure to propose a feedback mechanism based on a set of recommendation rules to support the experts in changing their opinions. This consensus model has been widely investigated in [2, 13, 22, 25, 27, 28, 37] and it is represented in Figure 2.

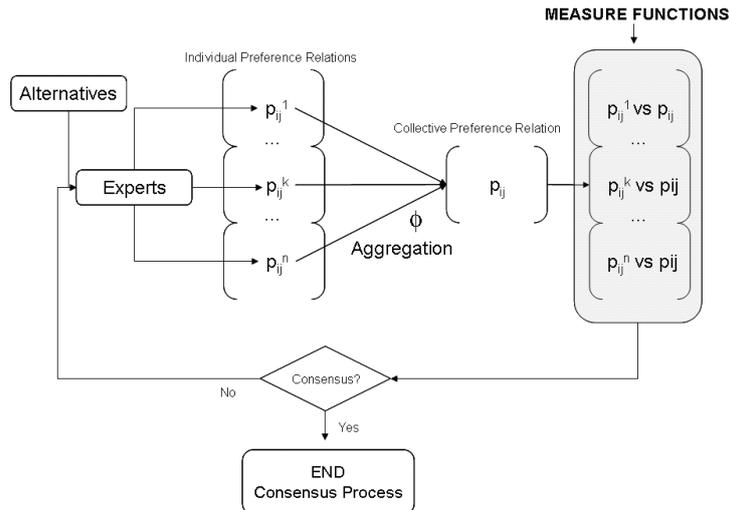


Figure 2: Consensus Model

The computation of consensus degrees is carried out as follows:

1. The proximity between the preference values provided by each expert, r , and the corresponding preference values of the rest of the experts in the group is measured and recorded in a *similarity matrix*, $SM^r = (sm_{ij}^r)$, with

$$sm_{ij}^r = s(\mathbf{p}_{ij}^r, \mathbf{p}_{ij}) \quad (7)$$

where $\mathbf{p}_{ij}^r = (p_{ij}^r, \dots, p_{ij}^r)$, $\mathbf{p}_{ij} = (p_{ij}^1, \dots, p_{ij}^{r-1}, p_{ij}^{r+1}, \dots, p_{ij}^m)$ and $s: [0, 1]^{m-1} \times [0, 1]^{m-1} \rightarrow [0, 1]$ a similarity function. The closer sm_{ij}^r to 1 the more similar \mathbf{p}_{ij}^r and \mathbf{p}_{ij} are, while the closer sm_{ij}^r to 0 the more distant \mathbf{p}_{ij}^r and \mathbf{p}_{ij} are.

2. A *consensus matrix*, $CM = (cm_{ij})$, is obtained by aggregating, using an OWA operator (ϕ), all the similarity matrices obtained via Equation (7):

$$\forall i, j \in \{1, \dots, n\} : cm_{ij} = \phi(sm_{ij}^1, \dots, sm_{ij}^m) \quad (8)$$

3. *Consensus degrees* are defined in each one of the three different levels of a fuzzy preference relation:

Level 1. *Consensus on the pairs of alternatives*, cp_{ij} . It measures the agreement among all experts on the pair of alternatives (x_i, x_j) :

$$\forall i, j = 1, \dots, n \wedge i \neq j : cp_{ij} = cm_{ij} \quad (9)$$

Level 2. *Consensus on alternatives*, ca_i . It measures the agreement among all experts on the alternative x_i , and it is obtained by aggregating the consensus degrees of all the pairs of alternatives involving it:

$$ca_i = \phi(cp_{ij}, cp_{ji}; j = 1, \dots, n \wedge j \neq i) \quad (10)$$

Level 3. *Consensus on the relation*, cr . It measures the global agreement among all experts, and it is obtained by aggregating all the consensus degrees at the level of pairs of alternatives:

$$cr = \phi(ca_i; i = 1, \dots, n) \quad (11)$$

It was mentioned before that an issue here is that the convergence of the consensus process towards a solution acceptable by most of the experts could be affected by the particular distance function implemented to measure the level of consensus. In the next sections we present a statistical study carried out to ascertain whether the application of different distance functions produces significant differences in the measuring of the consensus, and also to analyse which distance functions are better in speeding up the convergence of the consensus process. This statistical study is based on the use of the Wilcoxon matched-pairs signed-ranks nonparametric test [36, 45].

3. Statistical Comparative Study: Experimental Design

Given a GDM problem, the similarity function used to measure consensus plays a fundamental role in the convergence of the consensus process. It is therefore worth conducting research to ascertain whether or not the use of different similarity functions could affect the consensus process. Furthermore, if this was the case, the production of recommendations to the group of experts in the GDM problem on the different strategies available regarding the convergence of the consensus process based on the implementation of different similarity functions could prove to be an important decision support tool.

Given two vectors of real numbers $\mathbf{a} = (a_1, \dots, a_n)$ and $\mathbf{b} = (b_1, \dots, b_n)$, the following five distance functions have been considered in our study [6, 7, 15, 48]:

$$\begin{aligned}
\text{Manhattan: } d_1(\mathbf{a}, \mathbf{b}) &= \sum_{i=1}^n |a_i - b_i| \\
\text{Euclidean: } d_2(\mathbf{a}, \mathbf{b}) &= \sqrt{\sum_{i=1}^n |a_i - b_i|^2} \\
\text{Cosine: } d_3(\mathbf{a}, \mathbf{b}) &= \frac{\sum_{i=1}^n a_i \cdot b_i}{\sqrt{\sum_{i=1}^n a_i^2} \cdot \sqrt{\sum_{i=1}^n b_i^2}} \\
\text{Dice: } d_4(\mathbf{a}, \mathbf{b}) &= \frac{2 \cdot \sum_{i=1}^n a_i \cdot b_i}{\sum_{i=1}^n a_i^2 + \sum_{i=1}^n b_i^2} \\
\text{Jaccard: } d_5(\mathbf{a}, \mathbf{b}) &= \frac{\sum_{i=1}^n a_i \cdot b_i}{\sum_{i=1}^n a_i^2 + \sum_{i=1}^n b_i^2 - \sum_{i=1}^n a_i \cdot b_i}
\end{aligned}$$

The hypothesis that we are testing in this paper can be stated as follows:

The application of the Manhattan, Euclidean, Cosine, Dice and Jaccard distance functions in GDM problems do not produce significant differences in the measurement of consensus

Note that d_3 , d_4 , and d_5 have been satisfactorily applied in vectorial models of information retrieval by Salton and McGill [44] to measure the similarity of two documents, with the value of 1 measuring the highest similarity value. This ‘similarity’ interpretation is also taken in this paper, and therefore d_3 , d_4 , and d_5 are not subject to the transform mentioned above but d_1 and d_2 are.

To test the above hypothesis, twelve (12) sets of fuzzy preference relations were randomly generated for each possible combination of experts ($m = 4, 6, 8, 10, 12$) and alternatives ($n = 4, 6, 8$), and the different distance functions were applied in turn to measure consensus at the three possible levels (pairs of alternatives, alternatives and relation), using the three different quantifier guided OWA operators presented in Subsection 2.2. All distance functions were tested in pairs, d_i vs d_j ($i = 1, \dots, 4$, $j = i + 1, \dots, 5$), and therefore we ended having repeated measurements on a single sample [45].

For each pair of distance functions to compare we have to analyse two related samples. The usual parametric test to use in these cases is the t test applied to the difference scores. However, this test requires for its application the assumption of normality and independence distribution of the difference scores in the population from which the random sample of fuzzy preference relation is drawn. On the one hand, we consider these assumption unrealistic in our context as no evidence can be provided to support them, i.e. we do not possess any information that can lead to the identification of the nature of the population from which the random sample of fuzzy preference relations is drawn nor we have any knowledge about any of its parameters. On the other hand, by not requiring these stringent assumptions we can achieve greater generality with our conclusions. Therefore, we conclude that nonparametric test are most appropriate in our experimental study [36, 45, 49].

For continuous data and two related samples, the main nonparametric tests available are the sign test and the Wilcoxon signed-rank test [14, 36, 42, 45, 49]. The sign test calculates

the differences between two variables and classifies the differences as positive, negative, or zero (tied). If two variables have the same distribution (null hypothesis), the median of the differences between the two variable scores is zero, i.e. if the null hypothesis is true we would expect about half the differences to be negative and half to be positive. The null hypothesis is rejected if ‘too few’ differences of one sign occur. An alternative test to apply in this context is the Wilcoxon signed-rank test, which takes into account information from the sign of the differences and also their magnitude so that they are appropriately ranked in order of absolute magnitude. Since this test incorporates more information about the data it is more powerful than the sign test, and therefore preferable to use in our study. In what follows we describe the Wilcoxon matched-pairs signed-ranks statistical test in detail to understand its application in our experimental study to ascertain whether the application of the different distance functions given above produces significant differences in measuring consensus.

Wilcoxon Matched-pairs Signed-ranks Statistical Test. Let X_1, X_2, \dots, X_n be a random sample of size n from some unknown continuous distribution function F . Let p be a positive real number, $0 < p < 1$, and let $\xi_p(F)$ denote the quantile of order p for the distribution function F , that is, $\xi_p(F)$ is a solution of $F(x) = p$. For $p = 0.5$, $\xi_{0.5}(F)$ is known as the median of F .

A problem of location is set up by testing the null hypothesis $H_0 : \xi_p(F) = \xi_0$ against one of the alternatives $\xi_p(F) > \xi_0$, $\xi_p(F) < \xi_0$ or $\xi_p(F) \neq \xi_0$. A problem of location and symmetry consists of testing the null hypothesis $H_0 : \xi_{0.5}(F) = \xi_0$ and F is symmetric against $\xi_{0.5}(F) \neq \xi_0$ and F is not symmetric. The Wilcoxon signed-ranks test provides a statistical hypothesis test which takes into account the magnitude of the difference between the observations and the hypothesized quantile in order to carry out a problem of location and symmetry.

Let $H_0 : \xi_{0.5}(F) = \xi_0$ be the null hypothesis. Consider the differences $D_i = X_i - \xi_0$, $i = 1, 2, \dots, n$. Under H_0 : (i) the expected number of negative differences will be $n/2$ and, (ii) negative and positive differences of equal absolute magnitude should occur with equal probability. Consider the absolute values $|D_1|, |D_2|, \dots, |D_n|$ and rank them from 1 to n . Let T_+ be the sum of ranks assigned to those D_i 's that are positive and T_- be the sum of ranks assigned to those D_i 's that are negative. Then it is

$$T_+ + T_- = \sum_{k=1}^n k = \frac{n(n+1)}{2}$$

so that T_+ and T_- are linearly related and offer equivalent criteria. A large value of T_+ indicates that most of the larger ranks are assigned to positive D_i 's. It follows that large values of T_+ support $H_1 : \xi_{0.5}(F) > \xi_0$. A similar analysis applies to the other two alternatives. So, the test rejects $H_0 : \xi_{0.5}(F) = \xi_0$ to accept $H_1 : \xi_{0.5}(F) > \xi_0$ if $T_+ > c_1$; it rejects H_0 to accept $H_1 : \xi_{0.5}(F) < \xi_0$ if $T_- > c_2$; and it rejects H_0 to accept $H_1 : \xi_{0.5}(F) \neq \xi_0$ if $T_+ > c_3$ or $T_- > c_4$, being c_1, c_2, c_3 and c_4 the corresponding critical region values.

Under H_0 , the common distribution of T_+ and T_- is symmetric with mean $E[T_+] = n(n+1)/4$ and variance $\text{var}[T_+] = n(n+1)(2n+1)/24$. For large n , the standardised T_+ has approximately a standard normal distribution.

In the case of matched-paired data, $\{(X_1, Y_1), (X_2, Y_2), \dots, (X_n, Y_n)\}$, obtained from the application of two treatments (in our case – pair of different distance functions) to the same set of subjects (in our case – the set of random fuzzy preference relations constructed), in order to test $H_0 : \xi_{0.5}(F_{X_i - Y_i}) = \xi_0$ against one-sided or two-sided alternatives, the Wilcoxon matched-pairs signed-ranks statistical test is performed exactly as above by taking $D_i = X_i - Y_i - \xi_0$. In our study, we want to test whether the application of the different distance function does not affect significantly the measurement of consensus in GDM, i.e. we are testing a null hypothesis with a value $\xi_0 = 0$.

We assume that two measures with test p -value under the null hypothesis lower than or equal to 0.05 (α) will be considered as significantly different; we refer to it as the test being significant and therefore we conclude that the hypothesis tested is to be rejected.

4. Statistical Comparative Study: Experimental Results

A total of twelve (12) random GDM problems were generated for each one of the possible combinations of experts ($m = 4, 6, 8, 10, 12$) and alternatives ($n = 4, 6, 8$). Each one of these random GDM problems was executed three (3) times, each time using one of the three different OWA operators given in Subsection 2.2 to compute the consensus degrees at the three different levels of a fuzzy preference relation: the pairs of alternatives, the alternatives and the relation levels. In the following, we summarise the percentage of cases that were found to be significantly different according to the Wilcoxon matched-pairs signed-ranks statistical test when all the five different distance functions were compared in pairs, as per the description given in Section 3.

4.1. Pairs of Alternatives Level

Table 1, depicted in Figure 3, shows the percentage of tests with p -value lower than or equal to 0.05 (α) for each one of the linguistic quantifier guided OWA operators used in our experimental study. The application of different distance functions to measure consensus at the level of the pairs of alternatives produces significantly different results in at least 70% of all possible combinations of all the parameters used in the experiment (number of alternatives, number of experts and OWA operators). In particular, we observe that when the number of experts is fixed, the percentage of significantly different results never decreases when the number of alternatives increases. A similar behaviour is observed with respect to the number of experts, when the number of alternatives is fixed. There is an exception to this rule for the case of 8 alternatives, where the percentage decreases when we go from 10 to 12 experts.

We also note that there was a combination of distance functions for which the coincidence of consensus values were highest in their application: d_1 vs d_2 (Manhattan and Euclidean distance functions). In the case of the linguistic quantifier “*as many as possible*”, the Manhattan and Euclidean distance functions produce no significant differences.

In summary, we conclude that at the level of the pairs of alternatives the measurement of consensus is not affected significantly if the Manhattan or the Euclidean distance functions are used, but not for a different pair of distance functions. Obviously, the application of different distance functions, for which significant variation has been established, could affect the convergence of the consensus process at this level, something that will be discussed in more detail in Subsection 4.4.

Table 1: Percentage of significantly different results when different distance functions are applied to measure consensus at the level of the pairs of alternatives

“At least half” OWA						“Most of” OWA					
A\E	4	6	8	10	12	A\E	4	6	8	10	12
4	80%	80%	90%	90%	90%	4	70%	80%	90%	90%	90%
6	80%	90%	90%	90%	90%	6	70%	70%	90%	90%	90%
8	90%	90%	90%	100%	90%	8	70%	90%	90%	100%	90%

“As many as possible” OWA					
A\E	4	6	8	10	12
4	80%	90%	90%	90%	90%
6	90%	90%	90%	90%	90%
8	90%	90%	90%	90%	90%

4.2. Alternatives Level

Table 2, depicted in Figure 4, shows the percentage of tests with p -value lower than or equal to 0.05 (α) for each one of the linguistic quantifier guided OWA operators used in our

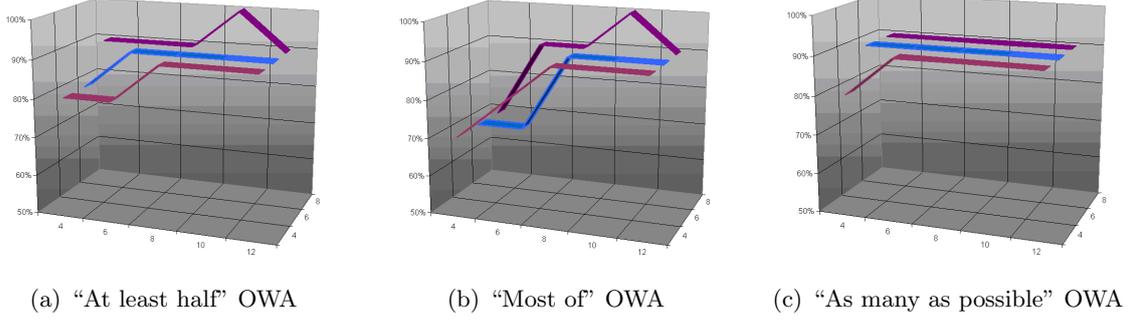


Figure 3: Graphical representation of Table 1

experimental study at the alternatives level. The application of different distance functions to measure consensus at the level of the alternatives produces significantly different results in at least 50% of all possible combinations of all the parameters used in the experiment (number of alternatives, number of experts and OWA operators). We observe that in most cases the greater the number of experts, the greater tends to be the percentage of results classed as significantly different. This analysis is not that apparent when the number of experts is fixed; in fact, in some cases the effect is the opposite as it can be seen for: (a) the linguistic quantifier “*at least half*” and $m = 4, 8, 12$; (b) the linguistic quantifier “*most of*” and $m = 12$; (c) the linguistic quantifier “*as many as possible*” and $m = 6, 10$.

We also note that the pairs of distance functions for which the coincidence of consensus values were highest corresponded to d_1 vs d_2 (Manhattan and Euclidean distance functions) and d_3 vs d_4 (Cosine and Dice distance functions), although there were combinations where all distance functions were found to be significantly different in the totality of cases: (a) the linguistic quantifier “*at least half*” and $(n, m) \in \{(4, 8), (4, 12)\}$; (b) the linguistic quantifier “*most of*” and $(n, m) \in \{(4, 8), (4, 12)\}$; (c) the linguistic quantifier “*as many as possible*” and $(n, m) \in \{(6, 12), (8, 12)\}$.

In summary, we conclude that the measurement of consensus at the level of the alternatives does not seem to be significantly affected if the Manhattan or the Euclidean distance functions are used, nor it is when the Cosine or the Dice distance functions are used; otherwise the contrary can be asserted. Therefore, at this level, the application of different distance functions, for which significant variation has been established, could affect the convergence of the consensus process.

Table 2: Percentage of significantly different results when different distance functions are applied to measure consensus at the level of the alternatives

“At least half” OWA						“Most of” OWA					
A \ E	4	6	8	10	12	A \ E	4	6	8	10	12
4	90%	70%	100%	90%	100%	4	80%	50%	100%	90%	100%
6	80%	60%	90%	50%	90%	6	60%	50%	80%	90%	90%
8	80%	80%	90%	80%	90%	8	80%	80%	90%	90%	80%

“As many as possible” OWA					
A \ E	4	6	8	10	12
4	90%	90%	80%	90%	90%
6	50%	80%	90%	90%	100%
8	80%	80%	80%	80%	100%

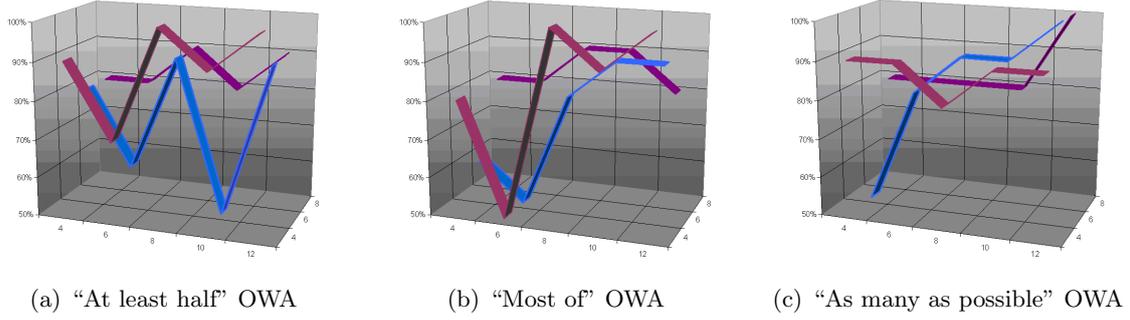


Figure 4: Graphical representation of Table 2

Table 3: Consensus degree in percentages for all GDM problems at the level of the relation

“At least half” OWA						“Most of” OWA					
$d_i \setminus E$	4	6	8	10	12	$d_i \setminus E$	4	6	8	10	12
d_1	63,08	82,40	91,04	96,45	99,86	d_1	69,49	85,09	92,90	97,22	100,00
d_2	64,60	82,82	91,60	96,71	100,00	d_2	67,94	85,25	92,51	96,89	99,71
d_3	93,05	93,61	94,53	94,83	94,93	d_3	84,66	84,06	84,69	84,57	85,05
d_4	95,96	94,62	94,82	94,91	94,48	d_4	86,59	85,41	85,37	84,76	84,70
d_5	81,44	80,71	78,86	79,81	79,89	d_5	69,02	68,12	65,96	67,60	66,94

“As many as possible” OWA					
$d_i \setminus E$	4	6	8	10	12
d_1	73,17	87,12	94,13	97,65	100,00
d_2	70,40	86,47	93,27	97,21	99,76
d_3	78,66	77,20	76,65	76,74	77,67
d_4	80,00	78,06	78,19	77,36	77,64
d_5	60,75	59,79	58,50	59,49	58,45

4.3. Relation Level

Table 3, depicted in Figure 5, shows the level of consensus in percentage achieved by the different distance functions for each GDM problem, showing only the number of experts as the variable parameter, and for each one the linguistic quantifier guided OWA operators used in our experimental study. The greater a value in this table the greater the global level of consensus achieved by the experts in the corresponding GDM problem. The comparison of column entries could be used to find out which distance function returns the largest values and therefore could lead to a faster convergence of the consensus process.

From Table 3 we can conclude the following:

1. The Manhattan (d_1) and the Euclidean (d_2) distance functions increase the global consensus level as the number of experts increases. Also, the values of consensus returned by both distance functions are quite similar, which was already evidenced by the results obtained in the pair of alternatives and the alternative levels.
2. The Cosine (d_3) and the Dice (d_4) distance functions result in fairly similar and stable global consensus levels regardless of the number of experts. For low number of experts both tend to produce higher values of consensus than the Manhattan and the Euclidean distance functions, which reverse when the number of experts is 8 or higher.
3. The Jaccard distance function (d_5) produces the lowest global consensus levels, being fairly stable in value regardless of the number of experts.

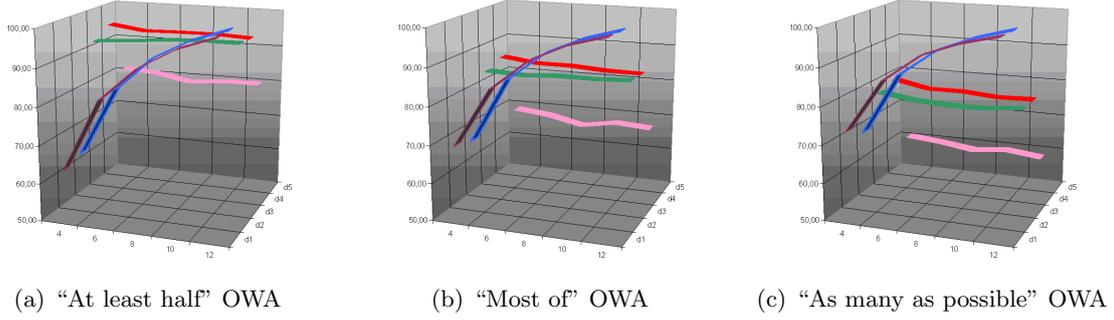


Figure 5: Graphical representation of Table 3

The following conclusions are drawn from the application of the Wilcoxon matched-pairs signed-ranks statistical test to the differences between the global consensus values:

“At least half” guided OWA operator The Cosine (d_3) and the Dice (d_4) distance functions never produce significant different global consensus values. Also, the highest consensus values are obtained with these two distance functions when the number of experts is below 10, otherwise it is the Manhattan (d_1) and the Euclidean (d_2) distance functions. The Jaccard distance function d_5 results in the smallest consensus values but for the case of four (4) experts.

“Most of” guided OWA operator Again in this case we have that the Cosine (d_3) and the Dice (d_4) distance functions never produce significant different global consensus values. In most cases, the highest global consensus values are obtained when the Manhattan (d_1) or the Euclidean (d_2) distance functions are applied. The Jaccard distance function (d_5) always yields the lowest global consensus values.

“As many as possible” guided OWA operator All distances show to produce significant different global consensus values. Again, in most cases, the highest global consensus values are obtained when the Manhattan distance function (d_1) or the Euclidean distance function (d_2) are applied. The Jaccard distance function (d_5) always yields the lowest global consensus values.

4.4. Consensus Process Convergence Rules

Based on the above analysis we can draw rules to speed up or slow down the convergence of the consensus that could prove an important decision support tool in GDM problems.

- The Manhattan (d_1) and the Euclidean (d_2) distance functions help the consensus process to converge faster than the rest as they consistently produce the highest consensus values for almost all possible combinations of number of experts and linguistic quantifier guided OWA operators.
- The Jaccard distance function (d_5) contributes the least to the speed of convergence of the consensus process.
- The Cosine (d_3) and the Dice (d_4) distance functions are placed in a mid term position in terms of helping speed up the convergence of the consensus process.
- The Manhattan (d_1) and the Euclidean (d_2) distance functions are quite sensitive to the number of experts, i.e. they produce significant different consensus values when the number of expert changes.

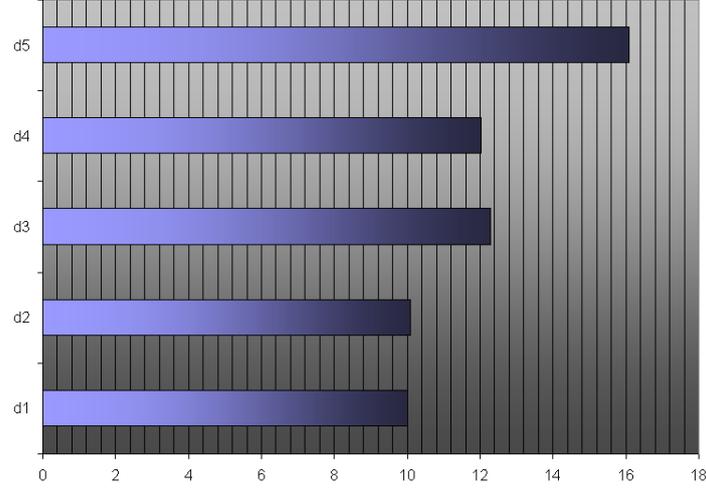


Figure 6: Number of consensus rounds necessary for each distance function to reach consensus threshold in a GDM problem: 8 experts and “most of” guided OWA operator

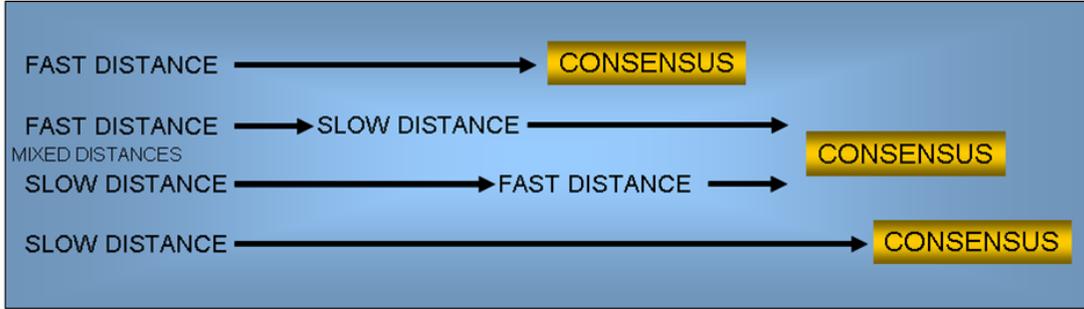


Figure 7: Distance functions to use for consensus reaching

- The Cosine (d_3), the Dice (d_4) and the Jaccard (d_5) distance functions are quite stable in the global consensus values they produce with respect to changes in the number of experts.

To corroborate the above rules, we run a GDM problem with 8 experts using the “most of” guided OWA operator with the five different distance functions and record the number of rounds necessary for the consensus process to reach the threshold consensus level acceptable for the GDM to reach a solution of consensus. This is graphically represented in Figure 6.

Figure 7 summarises the use of the above classification of the distance functions compared in relation to the speed of convergence of the consensus process towards the acceptable consensus level by the group of experts in a GDM problem. It seems reasonable that in the early stages of the consensus reaching process fairly stable distance functions should be used, with the application of less tolerant distance functions in later stages of the consensus process to speed up its convergence towards the threshold consensus level.

In the following we illustrate the use of the above rules in one of the examples randomly generated for the experimental study.

5. Example

Before providing any preferences the group of experts agree on a consensus threshold $\gamma \in [0, 1]$ such that when $cr \geq \gamma$ the consensus process will stop and the selection process will

be applied to obtain the solution of consensus. Otherwise, the consensus process continues and a (new) discussion round would be necessary for the experts to change preferences in an attempt to increase their global consensus level. The value γ depends on the particular problem dealt with. When the consequences of the decision making are of a significant importance, the minimum level of agreement required should be set as very high. On the contrary, or when it is urgent to obtain a solution of consensus, this value might not be set very high.

We assume a GDM with four alternatives $X = \{x_1, x_2, x_3, x_4\}$ and four experts $E = \{e_1, e_2, e_3, e_4\}$. We will be using the OWA operator guided by the linguistic quantifier “as many as possible”, with a fixed minimum threshold consensus value of $\gamma = 0.9$. We are assuming that the initial set of individual fuzzy preference relations are:

$$\begin{pmatrix} 0.50 & 0.11 & 0.23 & 0.32 \\ 0.89 & 0.50 & 0.44 & 0.12 \\ 0.77 & 0.66 & 0.50 & 0.21 \\ 0.68 & 0.88 & 0.79 & 0.50 \end{pmatrix} \begin{pmatrix} 0.50 & 0.65 & 0.91 & 0.70 \\ 0.35 & 0.50 & 0.82 & 0.63 \\ 0.09 & 0.18 & 0.50 & 0.34 \\ 0.30 & 0.37 & 0.66 & 0.50 \end{pmatrix} \begin{pmatrix} 0.50 & 0.43 & 0.81 & 0.64 \\ 0.57 & 0.50 & 0.72 & 0.91 \\ 0.19 & 0.28 & 0.50 & 0.43 \\ 0.36 & 0.09 & 0.57 & 0.50 \end{pmatrix} \begin{pmatrix} 0.50 & 0.32 & 0.41 & 0.63 \\ 0.68 & 0.50 & 0.62 & 0.93 \\ 0.59 & 0.38 & 0.50 & 0.81 \\ 0.37 & 0.07 & 0.19 & 0.50 \end{pmatrix}$$

As mentioned above, when $cr \geq \gamma$ the consensus process will stop and the selection process will be applied to obtain the solution of consensus. Otherwise, the alternatives and their preference values with consensus degrees below the global consensus level are identified. A feedback mechanism suggests the experts the right direction of the changes if the global consensus is to be increased. This is done via simple “advice rules” based on a broad comparison between the individual and collective preferences:

DR.1. If $p_{ij}^t - p_{ij}^c < 0$, expert e_t will be recommended to increase p_{ij}^t and decrease p_{ji}^t in the same quantity.

DR.2. If $p_{ij}^t - p_{ij}^c > 0$, expert e_t will be recommended to decrease p_{ij}^t and increase p_{ji}^t in the same quantity.

DR.3. If $p_{ij}^t - p_{ij}^c = 0$, expert e_t will not receive a recommendation of change for p_{ij}^t and p_{ji}^t .

More details can be consulted in [13, 25, 27].

First round. Using the stable Jaccard distance function the consensus degree at the relation level is 0.43.

Second round. Because the global consensus degree is lower than the threshold consensus level, experts receive feedback to modify their preference relations. The new fuzzy preference relations are:

$$\begin{pmatrix} 0.50 & 0.51 & 0.63 & 0.32 \\ 0.49 & 0.50 & 0.44 & 0.12 \\ 0.37 & 0.36 & 0.50 & 0.21 \\ 0.68 & 0.88 & 0.19 & 0.50 \end{pmatrix} \begin{pmatrix} 0.50 & 0.65 & 0.91 & 0.80 \\ 0.35 & 0.50 & 0.82 & 0.63 \\ 0.09 & 0.18 & 0.50 & 0.34 \\ 0.20 & 0.37 & 0.66 & 0.50 \end{pmatrix} \begin{pmatrix} 0.50 & 0.43 & 0.81 & 0.90 \\ 0.57 & 0.50 & 0.72 & 0.91 \\ 0.19 & 0.28 & 0.50 & 0.43 \\ 0.10 & 0.01 & 0.57 & 0.50 \end{pmatrix} \begin{pmatrix} 0.50 & 0.32 & 0.41 & 0.70 \\ 0.68 & 0.50 & 0.62 & 0.93 \\ 0.59 & 0.38 & 0.50 & 0.81 \\ 0.30 & 0.07 & 0.19 & 0.50 \end{pmatrix}$$

The Jaccard distance function results in a consensus degree level of 0.53.

Third round. New fuzzy preference relations:

$$\begin{pmatrix} 0.50 & 0.51 & 0.63 & 0.32 \\ 0.49 & 0.50 & 0.60 & 0.12 \\ 0.37 & 0.40 & 0.50 & 0.21 \\ 0.68 & 0.88 & 0.19 & 0.50 \end{pmatrix} \begin{pmatrix} 0.50 & 0.65 & 0.91 & 0.80 \\ 0.35 & 0.50 & 0.60 & 0.63 \\ 0.09 & 0.40 & 0.50 & 0.34 \\ 0.20 & 0.37 & 0.66 & 0.50 \end{pmatrix} \begin{pmatrix} 0.50 & 0.43 & 0.81 & 0.90 \\ 0.57 & 0.50 & 0.60 & 0.70 \\ 0.19 & 0.40 & 0.50 & 0.70 \\ 0.10 & 0.30 & 0.30 & 0.50 \end{pmatrix} \begin{pmatrix} 0.50 & 0.60 & 0.80 & 0.70 \\ 0.40 & 0.50 & 0.62 & 0.93 \\ 0.20 & 0.38 & 0.50 & 0.81 \\ 0.30 & 0.07 & 0.19 & 0.50 \end{pmatrix}$$

The Jaccard distance function results in a consensus degree level of 0.597.

Fourth round. New fuzzy preference relations:

$$\begin{pmatrix} 0.50 & 0.60 & 0.63 & 0.32 \\ 0.40 & 0.50 & 0.60 & 0.12 \\ 0.37 & 0.40 & 0.50 & 0.70 \\ 0.68 & 0.88 & 0.30 & 0.50 \end{pmatrix} \begin{pmatrix} 0.50 & 0.65 & 0.91 & 0.80 \\ 0.35 & 0.50 & 0.60 & 0.63 \\ 0.09 & 0.40 & 0.50 & 0.70 \\ 0.20 & 0.37 & 0.30 & 0.50 \end{pmatrix} \begin{pmatrix} 0.50 & 0.60 & 0.81 & 0.90 \\ 0.40 & 0.50 & 0.60 & 0.70 \\ 0.19 & 0.40 & 0.50 & 0.70 \\ 0.10 & 0.30 & 0.30 & 0.50 \end{pmatrix} \begin{pmatrix} 0.50 & 0.60 & 0.80 & 0.90 \\ 0.40 & 0.50 & 0.62 & 0.70 \\ 0.20 & 0.38 & 0.50 & 0.81 \\ 0.10 & 0.30 & 0.19 & 0.50 \end{pmatrix}$$

The Jaccard distance function results in a consensus degree level of 0.691.

Fifth round. New fuzzy preference relations are:

$$\begin{pmatrix} 0.50 & 0.60 & 0.80 & 0.50 \\ 0.40 & 0.50 & 0.60 & 0.12 \\ 0.20 & 0.40 & 0.50 & 0.70 \\ 0.50 & 0.88 & 0.30 & 0.50 \end{pmatrix} \begin{pmatrix} 0.50 & 0.65 & 0.91 & 0.80 \\ 0.35 & 0.50 & 0.60 & 0.63 \\ 0.09 & 0.40 & 0.50 & 0.70 \\ 0.20 & 0.37 & 0.30 & 0.50 \end{pmatrix} \begin{pmatrix} 0.50 & 0.60 & 0.81 & 0.90 \\ 0.40 & 0.50 & 0.60 & 0.70 \\ 0.19 & 0.40 & 0.50 & 0.70 \\ 0.10 & 0.30 & 0.30 & 0.50 \end{pmatrix} \begin{pmatrix} 0.50 & 0.60 & 0.80 & 0.90 \\ 0.40 & 0.50 & 0.62 & 0.70 \\ 0.20 & 0.38 & 0.50 & 0.70 \\ 0.10 & 0.30 & 0.30 & 0.50 \end{pmatrix}$$

Using the Cosine distance function we would have had a consensus degree level of 0.876, which could be considered quite close to the threshold consensus level as for the consensus process to stop. However, using again the Jaccard distance function we need to continue with the consensus reaching process as the consensus degree level would be 0.78.

Sixth round. New fuzzy preference relations:

$$\begin{pmatrix} 0.50 & 0.60 & 0.80 & 0.90 \\ 0.40 & 0.50 & 0.60 & 0.12 \\ 0.20 & 0.40 & 0.50 & 0.70 \\ 0.10 & 0.88 & 0.30 & 0.50 \end{pmatrix} \begin{pmatrix} 0.50 & 0.60 & 0.80 & 0.80 \\ 0.40 & 0.50 & 0.60 & 0.63 \\ 0.20 & 0.40 & 0.50 & 0.70 \\ 0.20 & 0.37 & 0.30 & 0.50 \end{pmatrix} \begin{pmatrix} 0.50 & 0.60 & 0.81 & 0.90 \\ 0.40 & 0.50 & 0.60 & 0.70 \\ 0.19 & 0.40 & 0.50 & 0.70 \\ 0.10 & 0.30 & 0.30 & 0.50 \end{pmatrix} \begin{pmatrix} 0.50 & 0.60 & 0.80 & 0.90 \\ 0.40 & 0.50 & 0.62 & 0.70 \\ 0.20 & 0.38 & 0.50 & 0.70 \\ 0.10 & 0.30 & 0.30 & 0.50 \end{pmatrix}$$

The Cosine distance function results in a consensus degree level of 0.917, the consensus reaching process stops and the selection process is activated to derive the solution of consensus.

6. Conclusion

In this paper we have considered some of the most widely used distance functions used in consensus process for GDM problems. We have presented detailed comparative experimental study based on the use of the nonparametric Wilcoxon statistical test. The results are interesting in that our experimental study has shown that the compared distance functions produce significant different results in most of the GDM problems carried out. The analysis of the results allows for the draw of a set of rules for the application of the compared distance functions that can be used to control the convergence speed of the consensus process using the compared distance functions.

In future, we will address this problem from a theoretical point of view by conducting an in-depth investigation to find out the intrinsic features of the distance functions that can be responsible for the significant differences in their application.

Acknowledgements

The authors would like to acknowledge FEDER financial support from the project FUZZYLING-II Project TIN2010-17876; the financial support from the Andalusian Excellence Projects TIC-05299 and TIC-05991, and also from the research project MTM2009-08886. Prof. Francisco Chiclana would like to acknowledge the financial support from the University of Granada 2012 GENIL Strengthening through Short-Visits research program (Ref. GENIL-SSV).

- [1] S. Alonso, E. Herrera-Viedma, F. Chiclana, and F. Herrera. A web based consensus support system for group decision making problems and incomplete preferences. *Information Sciences*, 180(23):4477–4495, 2010.

- [2] F.J. Cabrerizo, S. Alonso, and E. Herrera-Viedma. A consensus model for group decision making problems with unbalanced fuzzy linguistic information. *International Journal of Information Technology & Decision Making*, 8(1):109–131, 2009.
- [3] F.J. Cabrerizo, J.M. Moreno, I.J. Pérez, and E. Herrera-Viedma. Analyzing consensus approaches in fuzzy group decision making: advantages and drawbacks. *Soft Computing*, 14(5):451–463, 2010.
- [4] F.J. Cabrerizo, I.J. Pérez, and E. Herrera-Viedma. Managing the consensus in group decision making in an unbalanced fuzzy linguistic context with incomplete information. *Knowledge-Based Systems*, 23(2):169–181, 2010.
- [5] C. Carlsson, D. Ehrenberg, P. Eklund, M. Fedrizzi, P. Gustafsson, P. Lindholm, G. Merkuryeva, T. Riissanen, and A.G.S. Ventre. Consensus in distributed soft environments. *European Journal of Operational Research*, 61(1-2):165–185, 1992.
- [6] Sung-Hyuk Cha. Comprehensive survey on distance/similarity measures between probability density functions. *International Journal of Mathematical Models and Methods in Applied Sciences*, 4(1):300–3007, 2007.
- [7] S.M. Chen, M.S. Yeh, and P.Y. Hsiao. A comparison of similarity measures of fuzzy values. *Fuzzy Sets and Systems*, 72(1):79–89, 1995.
- [8] F. Chiclana, F. Herrera, and E. Herrera-Viedma. Integrating three representation models in fuzzy multipurpose decision making based on fuzzy preference relations. *Fuzzy Sets and Systems*, 97(1):33–48, 1998.
- [9] F. Chiclana, F. Herrera, and E. Herrera-Viedma. A note on the internal consistency of various preference representations. *Fuzzy Sets and Systems*, 131(1):75–78, 2002.
- [10] F. Chiclana, F. Herrera, E. Herrera-Viedma, and L. Martínez. A note on the reciprocity in the aggregation of fuzzy preference relations using owa operators. *Fuzzy Sets and Systems*, 137(1):71–83, 2003.
- [11] F. Chiclana, E. Herrera-Viedma, F. Herrera, and S. Alonso. Induced ordered weighted geometric operators and their use in the aggregation of multiplicative preference relations. *International Journal of Intelligent Systems*, 19(3):233–255, 2004.
- [12] F. Chiclana, E. Herrera-Viedma, F. Herrera, and S. Alonso. Some induced ordered weighted averaging operators and their use for solving group decision-making problems based on fuzzy preference relations. *European Journal of Operational Research*, 182(1):383–399, 2007.
- [13] F. Chiclana, F. Mata, L. Martinez, E. Herrera-Viedma, and S. Alonso. Integration of a consistency control module within a consensus model. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, 16 (Supplement 1):35–53, 2008.
- [14] M.J. Del Moral Avila and J.M. Tapia García. *Técnicas estadísticas aplicadas*. Grupo Editorial Universitario, Granada, 2006.
- [15] Michel Marie Deza and Elena Deza. *Encyclopedia of Distances*. Springer Berlin Heidelberg, 2009.
- [16] Y. Dong, Y. Xu, H. Li, and B. Feng. The owa-based consensus operator under linguistic representation models using position indexes. *European Journal of Operational Research*, 203(2):455–463, 2010.

- [17] D. Dubois and H. Prade. A review of fuzzy set aggregation connectives. *Information Sciences*, 36(1–2):85–121, 1985.
- [18] C. Elzinga, H. Wang, Z. Lin, and Y. Kumar. Concordance and consensus. *Information Sciences*, 2011.
- [19] János Fodor and Marc Roubens. *Fuzzy Preference Modelling and Multicriteria Decision Support*. Kluwer Academic Publishers, 1994.
- [20] F. Herrera and E. Herrera-Viedma. Choice functions and mechanisms for linguistic preference relations. *European Journal of Operational Research*, 120(1):144–161, 2000.
- [21] F. Herrera, E. Herrera-Viedma, and F Chiclana. A study of the origin and uses of the ordered weighted geometric operator in multicriteria decision making. *International Journal of Intelligent Systems*, 18(6):689–707, 2003.
- [22] F. Herrera, E. Herrera-Viedma, et al. A model of consensus in group decision making under linguistic assessments. *Fuzzy Sets and Systems*, 78(1):73–87, 1996.
- [23] F. Herrera, E. Herrera-Viedma, and JL Verdegay. Linguistic measures based on fuzzy coincidence for reaching consensus in group decision making. *International Journal of Approximate Reasoning*, 16(3-4):309–334, 1997.
- [24] F. Herrera, E. Herrera-Viedma, and JL Verdegay. Choice processes for non-homogeneous group decision making in linguistic setting. *Fuzzy Sets and Systems*, 94(3):287–308, 1998.
- [25] E. Herrera-Viedma, S. Alonso, F. Chiclana, and F. Herrera. A consensus model for group decision making with incomplete fuzzy preference relations. *IEEE Transactions on Fuzzy Systems*, 15(5):863–877, 2007.
- [26] E. Herrera-Viedma, J.L. García-Lapresta, J. Kacprzyk, M. Fedrizzi, H. Nurmi, and S. Zadrozny. *Consensual Processes*, volume 267 of *Studies in Fuzziness and Soft Computing*. Springer, 2011.
- [27] E. Herrera-Viedma, F. Herrera, and F. Chiclana. A consensus model for multiperson decision making with different preference structures. *IEEE Transactions on Systems, Man and Cybernetics, Part A: Systems and Humans*, 32(3):394–402, 2002.
- [28] E. Herrera-Viedma, L. Martinez, F. Mata, and F. Chiclana. A consensus support system model for group decision-making problems with multigranular linguistic preference relations. *IEEE Transactions on Fuzzy Systems*, 13(5):644–658, 2005.
- [29] J. Kacprzyk. Group decision making with a fuzzy linguistic majority. *Fuzzy Sets and Systems*, 18(2):105–118, 1986.
- [30] J. Kacprzyk and M. Fedrizzi. ‘Soft’ consensus measures for monitoring real consensus reaching processes under fuzzy preferences. *Control and Cybernetics*, 15:309–323, 1986.
- [31] J. Kacprzyk and M. Fedrizzi. A ‘soft’ measure of consensus in the setting of partial (fuzzy) preferences. *European Journal of Operational Research*, 34(3):316–325, 1988.
- [32] J. Kacprzyk, M. Fedrizzi, and H. Nurmi. Group decision making and consensus under fuzzy preferences and fuzzy majority. *Fuzzy Sets and Systems*, 49(1):21–31, 1992.
- [33] J. Kacprzyk, H. Nurmi, and M. Fedrizzi. *Consensus under fuzziness*, volume 10. Kluwer Academic Publishers, 1997.

- [34] C. Kahraman, D. Ruan, and I. Dogan. Fuzzy group decision-making for facility location selection. *Information Sciences*, 157:135–153, 2003.
- [35] G. J. Klir and T. A. Folger. *Fuzzy Sets, Uncertainty and Information*. Prentice Hall, New Jersey, 1988.
- [36] E. Lehmann. *Nonparametrics: Statistical Methods Based on Ranks*. Holden-Day, San Francisco, 1975.
- [37] F. Mata, L. Martínez, and E. Herrera-Viedma. An adaptive consensus support model for group decision-making problems in a multigranular fuzzy linguistic context. *IEEE Transactions on Fuzzy Systems*, 17(2):279–290, 2009.
- [38] I. Millet. The effectiveness of alternative preference elicitation methods in the analytic hierarchy process. *Journal of Multi-Criteria Decision Analysis*, 6(1):41–51, 1997.
- [39] S.A. Orlovsky. Decision-making with a fuzzy preference relation. *Fuzzy Sets and systems*, 1(3):155–167, 1978.
- [40] R.O. Parreiras, P.Y. Ekel, J.S.C. Martini, and R.M. Palhares. A flexible consensus scheme for multicriteria group decision making under linguistic assessments. *Information Sciences*, 180(7):1075–1089, 2010.
- [41] I.J. Pérez, F.J. Cabrerizo, and E. Herrera-Viedma. A mobile decision support system for dynamic group decision-making problems. *IEEE Transactions on Systems, Man and Cybernetics, Part A: Systems and Humans*, 40(6):1244–1256, 2010.
- [42] V.K. Rohatgi and A.K.M.E. Saleh. *An introduction to probability and statistics, 2nd Edition*. Wiley-Interscience, 2011.
- [43] M. Roubens. Fuzzy sets and decision analysis. *Fuzzy sets and systems*, 90(2):199–206, 1997.
- [44] G. Salton and M.J. McGill. *Introduction to modern information retrieval*. Mc-Graw Hill, 1983.
- [45] S. Siegel. *Nonparametric statistics for the behavioral sciences*. New York: McGraw-Hill, 1956.
- [46] Z. Switalski. General transitivity conditions for fuzzy reciprocal preference matrices. *Fuzzy Sets and Systems*, 137(1):85–100, 2003.
- [47] T. Tanino. Fuzzy preference orderings in group decision making. *Fuzzy Sets and Systems*, 12(2):117–131, 1984.
- [48] X. Wang, B. De Baets, and E. Kerre. A comparative study of similarity measures. *Fuzzy Sets and systems*, 73(2):259–268, 1995.
- [49] F. Wilcoxon. Individual comparisons by ranking methods. *Biometrics Bulletin*, 1(6):80–83, 1945.
- [50] Z. Xu and X. Cai. Group consensus algorithms based on preference relations. *Information Sciences*, 181(1):150–162, 2011.
- [51] Z. S. Xu and Q. L. Da. An overview of operators for aggregating information. *International Journal of Intelligent Systems*, 18(9):953–969, 2003.

- [52] R.R. Yager. On ordered weighted averaging aggregation operators in multicriteria decisionmaking. *IEEE Transactions on Systems, Man and Cybernetics*, 18(1):183–190, 1988.
- [53] R.R. Yager. Quantifier guided aggregation using owa operators. *International Journal of Intelligent Systems*, 11(1):49–73, 1996.
- [54] L.A. Zadeh. A computational approach to fuzzy quantifiers in natural languages. *Computers & Mathematics with Applications*, 9(1):149–184, 1983.
- [55] S-M. Zhou, F. Chiclana, R.I. John, and J.M. Garibaldi. Type-1 owa operators for aggregating uncertain information with uncertain weights induced by type-2 linguistic quantifiers. *Fuzzy Sets and Systems*, 159(24):3281–3269, 2008.
- [56] S-M. Zhou, F. Chiclana, R.I. John, and J.M. Garibaldi. Alpha-level aggregation: a practical approach to type-1 owa operation for aggregating uncertain information with applications to breast cancer treatments. *IEEE Transactions on Knowledge and Data Engineering*, 23(10):1455–1468, 2011.