Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 55 (2015) 1000 – 1008

Procedia Computer Science

Information Technology and Quantitative Management (ITQM 2015)

Analyzing Consensus Measures in Group Decision Making

F. Chiclana[a], J. M. Tapia García[b], M. J. del Moral[c], E. Herrera-Viedma[d,*]

[a]*CCI, 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

In Group Decision Making (GDM) problems before to obtain a solution a high level of consensus among experts is required. Consensus measures are usually built using similarity functions 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 distance functions have been proposed to implement consensus measures. This paper presents analyzes 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.

Peer-review under responsibility of the Organizing Committee of ITQM 2015

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

1. Introduction

For reaching 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. A second meaning of the concept of consensus refers to the judgement arrived at by 'most of' those concerned, which has led to the definition and use of a new concept of consensus degree referred to as 'soft' consensus degree [1].

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.

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

*Corresponding author.
E-mail address: viedma@decsai.ugr.es.

Peer-review under responsibility of the Organizing Committee of ITQM 2015
doi:10.1016/j.procs.2015.07.103

the computation and aggregation of the 'distance' representing disagreement between the opinions (preferences) of each pair of experts on each pair of alternatives [2]. 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 analyze five of the most commonly used distance functions in modelling soft consensus measures: Manhattan, Euclidean, Cosine, Dice, and Jaccard distance functions. Using nonparametric Wilcoxon tests [3] and [4], 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. Lastly, Section 5 concludes the paper.

2. Preliminaries

In this section, we will introduce the basic notions of GDM problem and Consensus Model.

2.1. The GDM Problem

GDM problems consist in finding the best alternative(s) from a set of feasible alternatives $X = \{x_1, \ldots, x_n\}$ according to the preferences provided by a group of experts $E = \{e_1, \ldots, e_m\}$. Different preference elicitation methods were compared in [5], where it was concluded that pairwise comparison methods are more accurate than non-pairwise methods.

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

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

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 [6]: (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 [7] given by the experts.

2.2. Selection Process

The selection process involves two different steps [8]: (i) *aggregation* of individual preferences and (ii) *exploitation* of the collective preference.

Aggregation phase. This phase defines a collective preference relation, $P^c = \left(p_{ij}^c\right)$, obtained by means of the aggregation of all individual fuzzy preference relations $\left\{P^1, P^2, \ldots, P^m\right\}$, and indicates the global preference between every pair of alternatives according to the majority of experts' opinions.

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

$$p_{ij}^c = \phi_Q\left(p_{ij}^1, \ldots, p_{ij}^m\right) = \sum_{k=1}^{m} w_k \cdot p_{ij}^{\sigma(k)}, \tag{2}$$

where σ is a permutation function such that $p_{ij}^{\sigma(k)} \geq p_{ij}^{\sigma(k+1)}$, $\forall\ k = 1, \cdots, m-1$; Q is a fuzzy linguistic quantifier [9] of fuzzy majority and it is used to calculate the weighting vector of ϕ_Q, $W = (w_1, \ldots, w_n)$ such that, $w_k \in [0, 1]$ and $\sum_{k=1}^{n} w_k = 1$, according to the following expression [10]:

$$w_k = Q\left(k/n\right) - Q\left((k-1)/n\right), \forall k \in \{1, \ldots, n\}. \tag{3}$$

Some examples of linguistic quantifiers are *"at least half"*, *"most of"* and *"as many as possible"*, which can be represented by the following function

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

using the values $(0, 0.5)$, $(0.3, 0.8)$ and $(0.5, 1)$ for (a, b), respectively [11].

Alternative representations for the concept of fuzzy majority can be found in the literature [12].

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.

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

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 [13]:

Definition 2 (Distance). *Let A be a set. A function* $d\colon A \times A \longrightarrow \mathbb{R}$ *is called a distance (or disimilarity) 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* $s\colon A \times A \longrightarrow \mathbb{R}$ *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 an a similarity s bounded by 1 are [13]:

$$d = 1 - s;\ d = \frac{1-s}{s};\ d = \sqrt{1-s}; d = \sqrt{2(1-s^2)};\ d = \arccos s;\ d = -\ln s \tag{5}$$

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 (6) 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. This consensus model has been widely investigated in [6] and [14].

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*, $SM^r = \left(sm^r_{ij}\right)$, with
$$sm^r_{ij} = s(\mathbf{p^r_{ij}}, \mathbf{p_{ij}}) \tag{6}$$
where $\mathbf{p^r_{ij}} = (p^r_{ij}, \ldots, p^r_{ij})$, $\mathbf{p_{ij}} = (p^1_{ij}, \ldots, p^{r-1}_{ij}, p^{r+1}_{ij}, \ldots, p^m_{ij})$ and $s\colon [0,1]^{m-1} \times [0,1]^{m-1} \to [0,1]$ a similarity function. The closer sm^r_{ij} to 1 the more similar $\mathbf{p^r_{ij}}$ and $\mathbf{p_{ij}}$ are, while the closer sm^r_{ij} to 0 the more distant $\mathbf{p^r_{ij}}$ 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 (6):
$$\forall i, j \in \{1, \ldots, n\} : \; cm_{ij} = \phi(sm^1_{ij}, \ldots, sm^m_{ij}) \tag{7}$$

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) :
$$\forall i, j = 1, \ldots, n \; \wedge \; i \neq j : \; cp_{ij} = cm_{ij} \tag{8}$$

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:
$$ca_i = \phi(cp_{ij}, cp_{ji}; j = 1, \ldots, n \; \wedge \; j \neq i) \tag{9}$$

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:
$$cr = \phi(ca_i; i = 1, \ldots, n) \tag{10}$$

3. Statistical Comparative Study: Experimental Design

Given two vectors of real numbers $\mathbf{a} = (a_1, \ldots, a_n)$ and $\mathbf{b} = (b_1, \ldots, b_n)$, the following five distance functions have been considered in our study [13]:

Manhattan:
$$d_1(\mathbf{a}, \mathbf{b}) = \sum_{i=1}^{n} |a_i - b_i| \tag{11}$$

Euclidean:
$$d_2(\mathbf{a}, \mathbf{b}) = \sqrt{\sum_{i=1}^{n} |a_i - b_i|^2} \tag{12}$$

Cosine:
$$d_3(\mathbf{a}, \mathbf{b}) = \frac{\sum_{i=1}^{n} a_i \cdot b_i}{\sqrt{\sum_{i=1}^{n} a_i^2} \cdot \sqrt{\sum_{i=1}^{n} b_i^2}} \tag{13}$$

Dice:
$$d_4(\mathbf{a}, \mathbf{b}) = \frac{2 \cdot \sum_{i=1}^{n} a_i \cdot b_i}{\sum_{i=1}^{n} a_i^2 + \sum_{i=1}^{n} b_i^2} \tag{14}$$

Jaccard:
$$d_5(\mathbf{a}, \mathbf{b}) = \frac{\sum_{i=1}^{n} a_i \cdot b_i}{\sum_{i=1}^{n} a_i^2 + \sum_{i=1}^{n} b_i^2 - \sum_{i=1}^{n} a_i \cdot b_i} \tag{15}$$

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 [15] 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 ($i = 1, \ldots, 4,\ j = i + 1, \ldots, 5$), and therefore we ended having repeated measurements on a single sample [16]. For each pair of distance functions to compare we have to analyse two related samples.

Nonparametric test are more appropriate in our experimental study [16] and [4]. For continuous data and two related samples, the main nonparametric tests available are the sign test and the Wilcoxon signed-rank test [17], [16], [3] and [4]. Since the last test incorporates more information about the data it is more powerful than the sign test, and therefore preferable to use in our study.

Wilcoxon Matched-pairs Signed-ranks Statistical Test. Let $X_1, X_2, \ldots, X_n$ 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 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.

Let $H_0 : \xi_{0.5}(F) = \xi_0$ be the null hypothesis. Consider the differences $D_i = X_i - \xi_0$, $i = 1, 2, \ldots, n$. 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 $|D_1|, |D_2|, \ldots, |D_n|$ 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

$$T_+ + T_- = \sum_{k=1}^{n} k = \frac{n(n+1)}{2} \tag{16}$$

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 $H_1 : \xi_{0.5}(F) > \xi_0$, large values of T_- support $H_1 : \xi_{0.5}(F) < \xi_0$ and extreme values of T_+ and T_- support $H_1 : \xi_{0.5}(F) \neq \xi_0$.

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

In the case of matched-paired data, $\{(X_1, Y_1), (X_2, Y_2), \ldots, (X_n, Y_n)\}$, 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 $H_0 : \xi_{0.5}(F_{X_i - Y_i}) = \xi_0$, 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 shows the percentage of tests with p-value lower than or equal to 0.05 (α) for each one 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 summary, 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.

4.2. Alternatives Level

Table 2 shows the percentage of tests with p-value lower than or equal to 0.05 (α) for each one the linguistic quantifier guided OWA operators used in our experimental study at the alternatives level. The application of different distance functions to measure consensus at the level of the alternatives produces significantly different

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						"As many as possible" OWA					
A\E	4	6	8	10	12	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	4	80	90	90	90	90
6	80	90	90	90	90	6	70	70	90	90	90	6	90	90	90	90	90
8	90	90	90	100	90	8	70	90	90	100	90	8	90	90	90	90	90

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).

In summary, 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						"As many as possible" OWA					
A\E	4	6	8	10	12	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	4	90	90	80	90	90
6	80	60	90	50	90	6	60	50	80	90	90	6	50	80	90	90	100
8	80	80	90	80	90	8	80	80	90	90	80	8	80	80	80	80	100

4.3. Relation Level

Table 3 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.

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.

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

"At least half" OWA					
d_i\E	4	6	8	10	12
d_1	63,08	82,40	91,04	96,45	99,86
d_2	64,60	82,82	91,60	96,71	100,00
d_3	93,05	93,61	94,53	94,83	94,93
d_4	95,96	94,62	94,82	94,91	94,48
d_5	81,44	80,71	78,86	79,81	79,89

"Most of" OWA					
d_i\E	4	6	8	10	12
d_1	69,49	85,09	92,90	97,22	100,00
d_2	67,94	85,25	92,51	96,89	99,71
d_3	84,66	84,06	84,69	84,57	85,05
d_4	86,59	85,41	85,37	84,76	84,70
d_5	69,02	68,12	65,96	67,60	66,94

"As many as possible" OWA					
d_i\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

- 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.
- 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 1.

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.

5. Conclusion

In this paper we have analyzed different distance functions used to compute consensus measures. 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.

Acknowledgements

The authors would like to acknowledge FEDER financial support from the project TIN2013-40658-P and the Andalusian Excellence Project TIC-05991 and also the research project MTM2012-35591.

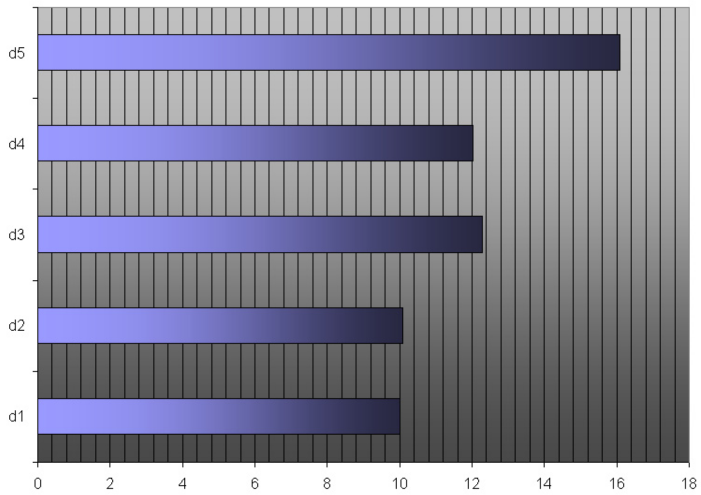

Fig. 1. Number of consensus rounds necessary for each distance function to reach consensus threshold in a GDM problem: 8 experts and *most of* guided OWA operator

References

[1] J. Kacprzyk, M. Fedrizzi. 'Soft' consensus measures for monitoring real consensus reaching processes under fuzzy preferences. *Control and Cybernetics* 1986;**15**:309–323.

[2] F.J. Cabrerizo, J.M. Moreno, I.J. Pérez, E. Herrera-Viedma. Analyzing consensus approaches in fuzzy group decision making: advantages and drawbacks. *Soft Computing* 2010; **14**(5):451–463.

[3] V.K. Rohatgi, A.K.M.E. Saleh. *An introduction to probability and statistics*. 2nd Edition. New York: Wiley-Interscience; 2011.

[4] F. Wilcoxon. Individual comparisons by ranking methods. *Biometrics Bulletin* 1945;**1**(6):80–83.

[5] I. Millet. The effectiveness of alternative preference elicitation methods in the analytic hierarchy process. *Journal of Multi-Criteria Decision Analysis* 1997;**6**(1):41–51.

[6] E. Herrera-Viedma, F. Herrera, 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* 2002;**32**(3):394–402.

[7] F. Chiclana, F. Herrera, E. Herrera-Viedma. Integrating three representation models in fuzzy multipurpose decision making based on fuzzy preference relations. *Fuzzy Sets and Systems* 1998;**97**(1):33–48.

[8] F. Herrera, E. Herrera-Viedma, J.L. Verdegay. Choice processes for non-homogeneous group decision making in linguistic setting. *Fuzzy Sets and Systems* 1998;**94**(3):287–308.

[9] L.A. Zadeh. A computational approach to fuzzy quantifiers in natural languages. *Computers & Mathematics with Applications* 1983;**9**(1):149–184.

[10] R.R. Yager. On ordered weighted averaging aggregation operators in multicriteria decisionmaking. *IEEE Transactions on Systems, Man and Cybernetics* 1988;**18**(1):183–190.

[11] J. Kacprzyk. Group decision making with a fuzzy linguistic majority. *Fuzzy Sets and Systems* 1986,**18**(2):105–118.

[12] F. Chiclana, E. Herrera-Viedma, F. Herrera, 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* 2007;**182**(1):383–399.

[13] M.M. Deza, E. Deza. *Encyclopedia of Distances*. Berlin: Springer Berlin Heidelberg; 2009.

[14] F. Chiclana, F. Mata, L. Martinez, E. Herrera-Viedma, S. Alonso. Integration of a consistency control module within a consensus model. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems* 2008;**16** (Supplement 1):35–53.

[15] G. Salton, M.J. McGill. *Introduction to modern information retrieval*. New York: Mc-Graw Hill; 1983.

[16] E. Lehmann. *Nonparametrics: Statistical Methods Based on Ranks*. San Francisco: Holden-Day; 1975.

[17] M.J. Del Moral Avila, J.M. Tapia García. *Técnicas estadísticas aplicadas*. Granana: Grupo Editorial Universitario; 2006.