A Model for Opinion Agreement and Confidence in Multi-Expert Multi-Criteria Decision Making

G. Canfora and L. Troiano
RCOST - University of Sannio
Viale Traiano — 82100 Benevento, Italy
e-mail:{ canfora, troiano} @unisannio.it

Abstract

In multi-expert multi-criteria decision making problems, we often have to deal with different opinions, different importance of criteria and experts, missing data, unexpressed opinions and experts who are not fully confident with their judgment. All these factors make the problem more difficult to solve, and run the risk of making the model logic less transparent. In this paper, we present a model based on simple assumptions described by logical rules, in order to maintain the model transparency and verifiability. In particular the model explicitly considers the level of agreement of experts, such as their importance and confidence.

1 Introduction

In some decision problems, the decision maker considers a set of alternatives, that are assessed by a pool of experts regarding a set of criteria. The comparison of alternatives, and consequently the selection of the best alternative, must be done considering all this information. This problem is known as Multiple Expert - Multiple Criteria Decision Making (ME-MCDM).

This problem is generally approached as a two-stage process. In the first stage, we get an alternative ranking from each expert; experts' rankings are then aggregated to get an overall score. The issue regarding how to aggregate information has been widely investigated [7, 15]. Generally, investigation has focused on means, due to their compensation property. Smolíková and Wachowiak [9] present a study aimed at comparing the behavior of different aggregation operators when applied to ME-MCDM problems. Currently, Valls and Torra [11, 12] are investigating clustering as a means for solving the disagreement between experts/criteria. Among the aggregation operators, the Ordered Weighted Averaging (OWA) operators, introduced by Yager [13], have shown an interesting property because they are a weighted mean of statistic orders, and their weights are associated with linear nondecreasing quantifiers [14].

Whatever the aggregation operator is, any ME-MCDM model can be regarded as a formal method for inferring a summary score for each alternative: we can regard an ME-MCDM model as a rational agent, capable of aggregating different sources of information according to some logical and mathematical assumptions. Thus, the model should be a white box, so that its assumptions can be verified and validated [1].

In ME-MCDM, difficulty arises when:

- data is not fully available

- experts deal with the same problem with different opinions

- experts are not fully confident in their own opinions

- experts and criteria are considered with different relevance

These items contribute to making the decision more unpredictable, as they increase the uncertainty of inferred information.

In this paper, we present a model which addresses those difficulties by means of inferential rules and numbers with indeterminateness. In particular the model explicitly considers the level of agreement among experts and their confidence. This paper illustrates the problem of selecting Ph.D. students as an example of application. This example represents a simple case-study which has already been investigated [7, 6, 5]. The remainder of this paper is organized as follows: Section 2 demonstrates the aggregation model; Section 3 shows the applicability of model to a case study; Section 4 includes conclusions and future work.

2 Aggregation model

2.1 Numbers with inderminateness

There are several causes that contribute to increasing the uncertainty of inferred information, such as vagueness, incompleteness and relevance of sources. It has been argued [2] that although these causes contribute to making decision effects more unpredictable, it is useful to keep their effects separate. Information is vague when we are not sure about the "value" assumed. Information is incomplete when it is not fully available, as is the case when some data is missing. Moreover, when information is inferred by different sources, relevance of sources plays a role in uncertainty: if sources are not relevant, we are not able to deduce any information. All these reasons contribute to making aggregated information not inferable. *Numbers with indeterminateness* [4] have been proposed to model such a situation.

A number with indeterminateness is defined as

$$F = \xi \cdot \mathbf{I} + \zeta \cdot K \tag{1}$$

where K is the *numeric component*, $\mathbf{I}$ is the *indeterminate element*, the coefficient ξ is called *indeterminateness*, and $\zeta = 1 - \xi$ is called *determinateness*. The numeric component K maintains all the available quantitative information, in a numeric

realm as real numbers $\mathbb{R}$ and ordinary fuzzy numbers $\mathcal{F}(\mathbb{R})$. In contrast, **I** represents "total ignorance": we assume that it is not possible to describe it by any membership function. Therefore we consider that **I** does not belong to any realm of ordinary numbers. We can refer to **I** as a primitive entity and consider it only symbolically. Its coefficient ξ provides a relevance index related to how factors leading to total ignorance affect information. It is a measure which concerns our ability to infer information according to available data. When its value is maximum ($\xi = 1$) we are not able to infer any information from available data. Such a situation is modeled as total ignorance. In contrast, complementary coefficient ζ is a measure of confidence which we can assign to the numeric value K when inferred by formal methods.

Although, an arithmetic of such numbers has not been developed yet, we can define a convex combination of number with indeterminateness axiomatically as

$$\sum_{i=1}^{n} \lambda_i F_i = \sum_{i=1}^{n} \lambda_i \left(\xi_i \cdot \mathbf{I} + \zeta_i \cdot K_i \right) = \left(\sum_{i=1}^{n} \lambda_i \xi_i \right) \cdot \mathbf{I} + \left(\sum_{i=1}^{n} \lambda_i \zeta_i \right) \cdot K \qquad (2)$$

where

$$K = \sum_{i=1}^{n} \lambda_i K_i \qquad (3)$$

Eq.(2) expresses a combination of numbers with indeterminateness, so that the higher the coefficient λ_i is, the more similar $\sum_{i=1}^{n} \lambda_i F_i$ is to F_i.

For example,

$$0.4 \cdot (0.3 \cdot \mathbf{I} + 0.7 \cdot [0.8]) + 0.6 \cdot (0.6 \cdot \mathbf{I} + 0.4 \cdot [0.3]) =$$
$$= (0.12 + 0.36) \cdot \mathbf{I} + (0.28 + 0.24) \cdot [0.32 + 0.18] =$$
$$= 0.48 \cdot \mathbf{I} + 0.52 \cdot [0.5]$$

2.2 Agreement

Another source of uncertainty which can lead to the impossibility of inferring aggregated information is related to different opinions from experts. In particular, we expect that the result of aggregation is coherent with the experts' opinions whatever the adopted aggregation model is. The level of coherence of aggregated information with the experts' opinions depends on the distance between the aggregation result and the opinions: the further the aggregation is from opinions, the lesser it is coherent with them.

We can measure the level of coherence by means of the *agreement function* δ^+. The agreement of the aggregated score with the opinions is maximum ($\delta^+ = 1$) when the aggregation result is fully coherent (coincides) with all the opinions. This is the case when all the experts declare the same opinion. The aggregation result is fully coherent with the experts' opinions when it coincides with them. When the aggregated result is far from the experts' opinions, the level of coherence decreases, because of a lower agreement between the two. The agreement of the aggregated

result with opinions can be computed on the pairwise basis. Disagreement is complementary to agreement and defined as

$$\delta^- = 1 - \delta^+ \tag{4}$$

The disagreement between the two values x (aggregated score) and x_i (expert i's opinion) can be directly linked to the distance between them, as stated by the following rules

if the distance between values x and x_i is high then disagreement is high; otherwise disagreement is low.

In particular we assume disagreement to be maximum $(\delta^- = 1)$ when the distance is maximum; disagreement is minimum $(\delta^- = 0)$ when the distance is minimum.

There are several definitions of distance. However, if values are expressed as real (crisp) numbers, it is natural to assume the Euclidean distance between them. Predicate *distance is high* can be evaluated by a proper membership function

$$h : \mathbb{R} \to \mathbb{R} \tag{5}$$
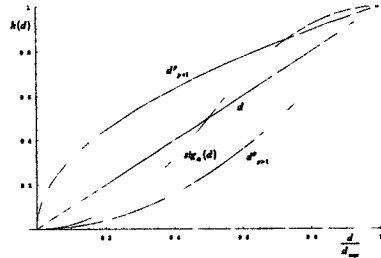
so that

$$\delta^-(x, x_i) = h(d(x, x_i)) \tag{6}$$

Figure 1: Membership function $h(d)$

The easiest way to map agreement on distance is

$$h(d) = \frac{d}{d_{\max}} \tag{7}$$

which in the case of $d_{\max} = 1$, becomes $high(d) = d$. Other choices are possible (Fig.1). For example we could judge the distance d according to the function

$$h(d) = \left(\frac{d}{d_{\max}} \right)^p \tag{8}$$

for any given $p \in \mathbb{R}^+$. When $p > 1$ we emphasize distant values as a source of disagreement, minimizing differences between close values. Whereas with regard

to $p < 1$, we entail a different semantic, in accordance with the concept of linguistic modifiers [8]. Another possible function is

$$h(d) = sig_\alpha(d) = \begin{cases} \frac{1}{2}(2\frac{d}{d_{max}})^\alpha & d \leq 0.5 \\ 1 - \frac{1}{2}\left(2(1 - \frac{d}{d_{max}})\right)^\alpha & d > 0.5 \end{cases} \tag{9}$$

with $\alpha \geq 1$, which has the property of emphasizing the variation of disagreement moving from near opinions to far opinions. This property becomes more evident by increasing parameter α.

Once pairwise agreement is defined, we can get an aggregated measure of agreement $\overline{\delta^+}$ (disagreement $\overline{\delta^-}$) according to the rule

if the opinion of expert i is relevant then $\overline{\delta^+} = \delta_i^+$ ($\overline{\delta^-} = \delta_i^-$);

where $\delta_i^- = \delta^-(x, x_i)$. We can compute inference as a weighted average [10]

$$\overline{\delta^+} = \frac{\sum_{i=1}^{m} v_i \delta_i^+}{\sum_{i=1}^{m} v_i} \quad \left(\overline{\delta^-} = \frac{\sum_{i=1}^{m} v_i \delta_i^-}{\sum_{i=1}^{m} v_i}\right) \tag{10}$$

When all opinions are considered with the same relevance

$$\overline{\delta^+} = \frac{1}{m}\sum_{i=1}^{m} \delta_i^+ \quad \left(\overline{\delta^-} = \frac{1}{m}\sum_{i=1}^{m} \delta_i^-\right) \tag{11}$$

In conformity to Eq.4, the property of complementarity is still verified.

Because $h(\cdot)$ is monotonic, the value

$$d^* = h^{-1}\left(\frac{\sum_{i=1}^{m} v_i h(d_i)}{\sum_{i=1}^{m} v_i}\right) \tag{12}$$

is obtained by a generalized weighted mean, and it represents the distance at which we can image all opinions to be placed according to a given level of disagreement $\overline{\delta^-}$. Thus, we can express $\overline{\delta^-}$ as a function of d^*, as

$$\overline{\delta^-} = \overline{\delta^-}(x, x_1, \ldots, x_m) = h(d^*) \tag{13}$$

Then, we must ask ourselves what is the aggregated value x_m^* which minimizes disagreement $\overline{\delta^-}$ given the opinions $(x_1, \ldots, x_m)$ and the function $h(\cdot)$. For the sake of simplicity we will consider all the experts' opinions as having relevance $v_i = 1$.

Let $\chi_1, \ldots, \chi_m$ be the increasing ordered permutation of the opinions $x_1, \ldots, x_m$. If $h(\cdot)$ is defined as in Eq.7, then

$$\overline{\delta^-} = \frac{1}{m \cdot d_{max}} \sum_{i=1}^{m} |x - \chi_i| \tag{14}$$

It is easy to verify that, if m is odd then the minimum x_m^* is reached at the median opinion $x_m^* = \chi_{(m+1)/2}$; on the other hand, if m is even then $x_m^* \in [\chi_{m/2}, \chi_{m/2+1}]$.

This has a semantic interpretation. Let us consider the opinion $\overline{x} \in [\chi_i, \chi_{i+1}]$, then move it to $\overline{x} + \Delta \in [\chi_i, \chi_{i+1}]$, for some $\Delta > 0$. Due to the direct link between disagreement and distance, we get a variation

$$\Delta h = \frac{\Delta}{d_{\max}} \tag{15}$$

Since the opinion $\overline{x}+\Delta$ is distancing itself from the opinions $\chi_1, \ldots, \chi_i$, and drawing nearer the opinions $\chi_{i+1}, \ldots, \chi_m$, then variation Δh is positive with regard to the opinions which come before, and negative with regard to the opinions which come after. Thus, the overall agreement at $\overline{x} + \Delta$ will decrease only if the opinions which come before are less than the opinions which come after; otherwise, it will increase. The minimum is reached when the number of opinions which come before equals the number of opinions that come after, regardless of their value. This is in accordance with a scheme of decision making based on the majority: the compromise score which minimizes the disagreement is reached when the number of experts who would give a lower score equals the number of experts who would give a higher score.

In some cases this scheme is not appropriate. If we choose

$$h(d) = \left(\frac{d}{d_{\max}}\right)^2 \tag{16}$$

then

$$\overline{\delta^-} = \frac{1}{m \cdot d_{\max}^2} \sum_{i=1}^{m} (x - x_i)^2 \tag{17}$$

and minimum disagreement is reached at

$$x_m^* = \frac{x_1 + \cdots + x_m}{m} \tag{18}$$

This entails a compensation decision making scheme: the arithmetic mean of opinions is the compromise score which minimizes the overall disagreement δ^-.

2.3   Aggregation of criteria and opinions

In ME-MCDM problems, experts and criteria can have different importance in determining the result of aggregation. It is intuitive that the more important a criterion or an expert opinion is, the more it should affect the aggregated score. As the issue of considering importance criteria and experts can be formulated in general terms, we will refer to experts and criteria with the common term of *factors*. In particular we will briefly outline the general importance model which is fully discussed in reference [4].

We can describe the importance of factor C_i by means of $v_i \in [0,1]$. If C_i is important, then $v_i = 1$; viceversa, if C_i is unimportant, then $v_i = 0$.

Regardless of whatever aggregation operator is chosen, an unimportant factor should not be considered at all; otherwise that factor must be taken into account. We can describe this by the following logical rules

if C_i is relevant then the aggregation operator should consider C_i;
otherwise, the aggregation operator can ignore C_i.

Ignoring a factor means that the aggregation result is independent on it: the result does not change no matter what value is given to that factor. Aggregation is then restricted to the resulting subset of criteria. We can formalize this by recursion. Let $N = \{1, \ldots, n\}$ be the index set of factors $C_1, \ldots, C_n$, and $M_{A|B}$ the generic aggregation element such that A is the index subset of criteria which is surely considered by aggregation M, while B is the index subset of criteria which is certainly not considered. Consequently, $A \cap B = \varnothing$ and $A \cup B \subseteq N$. Rules related to factor importance, can be rewritten as

$$r_{i,1}: \text{imp}(C_i) \text{ is high} \Rightarrow M_{A'|B'} = M_{A \cup \{i\}|B}$$

$$r_{i,2}: \text{imp}(C_i) \text{ is low} \Rightarrow M_{A'|B'} = M_{A|B \cup \{i\}} \tag{19}$$

$$\forall i, A, B | A \cup B \cup \{i\} \subseteq N, A \cap B = \varnothing, i \notin A \cup B$$

We can evaluate these rules by applying Sugeno-Takagi's method [10]. If v_i is the measure of factor importance, then $\tau_{i,1} = v_i$ and $\tau_{i,2} = 1 - v_i$ (so that $\tau_{i,1} + \tau_{i,2} = 1$) are respectively the firing levels for $r_{i,1}$ and $r_{i,2}$. Therefore

$$M_{A'|B'} = \frac{\tau_{i,1} M_{A \cup \{i\}|B} + \tau_{i,2} M_{A|B \cup \{i\}}}{\tau_{i,1} + \tau_{i,2}} =$$

$$= \tau_{i,1} M_{A \cup \{i\}|B} + \tau_{i,2} M_{A|B \cup \{i\}} \tag{20}$$

$$i \notin A \cup B$$

Iterating from $i = 1$ to n, we get

$$F = \sum_{\substack{A \subseteq N \\ A \neq \varnothing}} \tau_A M_A + \tau_\varnothing M_\varnothing \tag{21}$$

where

$$M_A = M_{A|N-A}$$

$$\tau_A = \prod_{i=1}^{n} \tau_i \tag{22}$$

$$\tau_i = \begin{cases} \tau_{i,1} = v_i & i \in A \\ \tau_{i,2} = 1 - v_i & i \in N - A \end{cases}$$

$M_\varnothing$ represents the result of aggregation when all factors are unimportant. In this case it is impossible to infer an aggregated score. Thus, we can put

$$M_\varnothing = \mathbf{I} \tag{23}$$

Therefore, aggregation in Eq.(21) results in a number with indeterminateness (see Eq.(1)). If all factors are unimportant ($v_i = 0, \forall i$), the aggregation result cannot be determined, otherwise the result coincides with the aggregation of relevant factors. This can be described by the following rules

if C_i is not relevant $\forall i \in N$ then $F = \mathbf{I}$;
otherwise $F = K$.

so that

$$F = \tau_{1,2}\tau_{2,2}\cdots\tau_{n,2} \cdot \mathbf{I} + (1 - \tau_{1,2}\tau_{2,2}\cdots\tau_{n,2}) \cdot K =$$
$$= \xi_J \cdot \mathbf{I} + (1 - \xi_J) \cdot K \tag{24}$$

By equaling Eq.(24) and Eq.(21), we obtain

$$K = \frac{1}{1-\xi} \sum_{\substack{A \subseteq N \\ A \neq \emptyset}} \tau_A M_A \tag{25}$$

and

$$\xi = \prod_{i=1}^{n} (1 - v_i)$$
$$\zeta = 1 - \xi \tag{26}$$

2.4 Missing opinions and respondent confidence

In decision making or analysis problems, we often deal with missing or unexpressed opinions. In this case, we can assume an indeterminate value $x_i = \mathbf{I}$. In other cases, experts can be only partially confident of their opinions. We can still use numbers with indeterminateness, and assume $x_i' = \xi \cdot I + \zeta \cdot x_i$, where ζ is a measure of respondent confidence in the score x_i. Moreover, multi-stage decision problems such as ME-MCDM require aggregation of scores from lower layers to higher aggregation layers. Based on the outlined model, aggregation at higher levels could entail numbers with indeterminateness as input values.

Again, when a factor is fully indeterminate ($\xi_i = 1$), we should not consider it. This assumption, combined to criteria relevance, leads to the following set of rules:

if C_i is relevant and determined then aggregation should consider it;
otherwise, aggregation can ignore it.

The modified set of rules requires the computation of firing levels as

$$\tau_{i,1}' = \zeta_i v_i$$
$$\tau_{i,2}' = 1 - \zeta_i v_i \tag{27}$$

Eq.(24) becomes

$$F = \xi_J' \cdot \mathbf{I} + \zeta_J' \cdot K \tag{28}$$

where

$$\xi' = \prod_{i=1}^{n} (1 - \zeta_i v_i)$$

$$\zeta' = 1 - \xi'$$

(29)

Therefore

$$K = \frac{1}{1 - \xi'} \sum_{\substack{A \subseteq N \\ A \neq \emptyset}} \tau_A' M_A$$

(30)

where

$$\tau_A' = \prod_{i=1}^{n} \tau_i'$$

$$\tau_i' = \begin{cases} \tau_{i,1}' & i \in A \\ \tau_{i,2}' & i \in N - A \end{cases}$$

(31)

It can be easily proven that $\xi' \geq \xi$ and $\zeta' \leq \zeta$.

If all aggregation factors are undetermined ($\xi_i = 1$) or unimportant ($v_i = 0$), we are not able to infer the aggregated score, and aggregation results into indetermination. The more important factors are determined, the more we can infer the aggregated result. We take this effect into account by means of the *level of determinateness* [2], defined as

$$\lambda^+ = \frac{\sum_{i=1}^{n} v_i \zeta_i}{\sum_{i=1}^{n} v_i}$$

(32)

and the complementary *level of indeterminateness* as

$$\lambda^- = 1 - \lambda^+$$

(33)

The level of determinateness is minimum ($\lambda^+ = 0$) if all relevant criteria are undetermined ($\zeta_i = 0$). It is maximum ($\lambda^+ = 1$) when all relevant criteria are determined. Let us consider the following set of rules

> *if C_i is irrelevant or C_i is indeterminated $\forall i \in N$ then $F = \mathbf{I}$*
> *if C_i is relevant and C_i is determined $\exists i \in N$ and level λ^- is high then $F = \mathbf{I}$*
> *if C_i is relevant and C_i is determined $\exists i \in N)$ and level λ^+ is high then $F = K$*

In this case, the result of aggregation is

$$F = (\xi' + \lambda^- \zeta') \cdot \mathbf{I} + \lambda^+ \zeta' \cdot K$$

(34)

where K is still as Eq.(30)

Finally we should consider the level of disagreement δ^- among experts when we aggregate opinions. Aggregation elements M_A are aggregated scores on opinion

subsets $A \subseteq O$, where O is the opinion index set. Let us consider the following rules

> *if disagreement is high then $M_A = \mathbf{I}$*
> *otherwise $M_A = K_{MA}$*

where K_{MA} is the aggregated numeric score of opinions indexed by A. This means that each opinion aggregation element M_A will result in a number with indeterminateness

$$M_A = \delta_A^- \cdot \mathbf{I} + \delta_A^+ \cdot K_{MA} \tag{35}$$

where

$$\delta_A^- = \frac{1}{m} \sum_{i=1}^{m} \delta^-(K_{MA}, x_i) \tag{36}$$

since opinion relevance has already been taken into account by the firing level τ_A defined as in Eq.(31). Thus Eq.(21) results in a convex combination of numbers with indeterminateness as defined in Eq.(2). The result of aggregation is a number with indeterminateness

$$\xi_T = \xi' + \lambda^- \zeta' + \lambda^+ \sum_{\substack{A \subseteq N \\ A \neq \varnothing}} \tau_A' \delta_A^- \tag{37}$$

where

- ξ' takes into account the relevance and determinateness of expert opinions or aggregation criteria;

- $\lambda^- \zeta'$ takes into account the level of indeterminateness in expert opinions or criteria;

- $\lambda^- \sum_{\substack{A \subseteq N \\ A \neq \varnothing}} \tau_A' \delta_A^-$ takes into account the disagreement among expert opinions.

The numeric component K is computed as in Eq.(30)

3 Case Study

We illustrate the applicability of the proposed aggregation model using the case study known as the doctoral student selection problem of the Graduate School of the Turku Centre for Computer Science [5].

Opinions of a pool of 11 experts coming from different groups are solicited for selecting young promising doctoral researchers according to the following 6 criteria

1. Fit in research groups

2. On the frontier of research

3. Contributions

4. University

5. Grade average

6. Time for acquiring degree

The evaluation consists in assigning a score chosen from the scale 1, 2, 3, where 3 stands for excellent, 2 stands for average and 1 means weak performance. Thus, experts provide a 6-tuple $(a_{i,1}, \ldots, a_{i,6})$ for each applicant.

Tuples are aggregated into an overall score e_i for every expert for each applicant. At this stage, for each applicant there is a 11-tuple $(e_1, \ldots, e_{11})$. Scores from experts are aggregated in order to rank each applicant.

In [5], the authors use OWA operators to solve the selection problem. Weights are determined using a regular non-decreasing quantifier $Q_\alpha = r^\alpha, \alpha \geq 0$. The exponent value is chosen in order to satisfy a set of additional requirements:

1. If there are more than 2 weak performances, then the overall performance should be < 2; otherwise it should be ≥ 2.

2. If all but one of the performances are excellent, then the overall performance should be about 2.75;

3. If there are 3 weak performances and one of them is on the frontier of research, [1] then the overall performance should not be above 1.5

4. If an applicant has all but one excellent score, then the final score should be about 2.75.

The first three requirements are used to weigh the OWA aggregation at the first stage, whilst the last requirement is adopted for weighing the OWA aggregation at the second stage. The OWA operator weights used to aggregate scores by a given expert are

$$(0.116, 0.151, 0.168, 0.180, 0.189, 0.196)$$

Then, unit scores are further aggregated in order to derive an overall assessment for each applicant. This task is made by OWA aggregation with the following weights

$$(0.526, 0.199, 0.150, 0.125, 0, 0, 0, 0, 0, 0, 0)$$

so that only the four best unit scores are considered for each applicant. Let us consider an applicant with the scores reported in Tab.1.

In the following example we will measure disagreement as defined by Eq.(16).

Example 1

Expert aggregated scores are reported in Tab.2. The applicant in our example obtains a final score of 2.476 and (s)he has a good chance of obtaining the scholarship.

We would have obtained the same result if all the experts had expressed the same aggregated opinion equal to 2.476. Instead there are opinions varying between minimum 1.615 and maximum 2.615. The disagreement between the resulting score

Criteria	C_1	C_2	C_3	C_4	C_5	C_6
Expert 1	3	2	3	2	3	1
Expert 2	2	3	3	2	3	2
Expert 3	2	2	3	2	2	1
Expert 4	3	2	3	3	3	2
Expert 5	2	2	3	2	3	1
Expert 6	3	2	3	2	3	1
Expert 7	1	2	3	2	3	2
Expert 8	1	2	3	2	3	1
Expert 9	1	2	2	2	3	2
Expert 10	1	2	2	3	3	1
Expert 11	1	2	2	2	2	1

Table 1: Applicant's scores

	OWA
Expert 1	2.239
Expert 2	2.435
Expert 3	1.920
Expert 4	2.615
Expert 5	2.071
Expert 6	2.239
Expert 7	2.071
Expert 8	1.882
Expert 9	1.920
Expert 10	1.882
Expert 11	1.615
Final Score	2.476

Table 2: Experts' aggregated scores

and expert opinions is $\delta^{(-)} = 0.210$. Thus, because there is no other source of indeterminateness, the result of aggregation is

$$F = 0.210 \cdot \mathbf{I} + 0.790 \cdot [2.476]$$

The low level of disagreement is due to the fact that the disagreement defined in Eq.(16) emphasizes only relevant differences.

In the selection problem [5], all experts have the same importance for determining the applicant's final score. However, we should consider situations entailing experts with different importance.

Example 2

Let us suppose that experts are considered with importance according to Tab.3 The final score is computed as

$$F = 0.3M_N + 0.3M_{N-\{2\}} + 0.2M_{N-\{5\}} + 0.2M_{N-\{2,5\}}$$

Experts	1	2	3	4	5	6	7	8	9	10	11
Relevance	1.0	**0.5**	1.0	1.0	**0.6**	1.0	1.0	1.0	1.0	1.0	1.0

Table 3: Importance of experts

where
$$N = \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11\}$$

and

$$M_N = 0.210 \cdot \mathbf{I} + 0.790 \cdot [2.476]$$
$$M_{N-\{2\}} = 0.205 \cdot \mathbf{I} + 0.795 \cdot [2.416]$$
$$M_{N-\{5\}} = 0.211 \cdot \mathbf{I} + 0.789 \cdot [2.476]$$
$$M_{N-\{2,5\}} = 0.209 \cdot \mathbf{I} + 0.791 \cdot [2.416]$$

so that

$$F = 0.209 \cdot \mathbf{I} + 0.791 \cdot [2.446]$$

The other complication we should add is that some experts might not be fully confident with their judgment, and some opinions could be missing or unexpressed.

Example 3

Let us suppose that the determinateness of Expert 3's opinion is $\zeta_{e3} = 0.7$, whilst Expert 8's opinion is unknown ($\xi_{e8} = 1$). The other data is as in Example 2. Any aggregation element $M_{\{8\} \in A}$ considering Expert 8's opinion is undeterminate, due to the fact that the corresponding firing level is zero ($v_8 \zeta_8 = 0$). In this case, the aggregation formula is reduced to

$$F = \lambda^{(-)} \cdot \mathbf{I} + \lambda^{(+)} \cdot \left(0.21 M_{N-\{8\}} + 0.21 M_{N-\{2,8\}} + 0.09 M_{N-\{3,8\}} + 0.14 M_{N-\{5,8\}} + \right.$$
$$\left. + 0.09 M_{N-\{2,3,8\}} + 0.14 M_{N-\{2,5,8\}} + 0.06 M_{N-\{3,5,8\}} + 0.06 M_{N-\{2,3,5,8\}} \right)$$

where
$$\lambda^{(-)} = 0.13$$
$$\lambda^{(+)} = 0.87$$

and

$$M_{N-8\} = 0.201 \cdot \mathbf{I} + 0.799 \cdot [2.476]$$
$$M_{N-\{2,8\}} = 0.198 \cdot \mathbf{I} + 0.802 \cdot [2.416]$$
$$M_{N-\{3,8\}} = 0.193 \cdot \mathbf{I} + 0.807 \cdot [2.476]$$
$$M_{N-\{5,8\}} = 0.201 \cdot \mathbf{I} + 0.799 \cdot [2.476]$$
$$M_{N-\{2,3,8\}} = 0.192 \cdot \mathbf{I} + 0.808 \cdot [2.416]$$
$$M_{N-\{2,5,8\}} = 0.201 \cdot \mathbf{I} + 0.799 \cdot [2.416]$$
$$M_{N-\{3,5,8\}} = 0.192 \cdot \mathbf{I} + 0.808 \cdot [2.476]$$
$$M_{N-\{2,3,5,8\}} = 0.195 \cdot \mathbf{I} + 0.805 \cdot [2.416]$$

Therefore, the result of aggregation is

$$F = 0.301 \cdot \mathbf{I} + 0.699 \cdot [2.446]$$

Although the final score is numerically equal to the previous one, we have a lower level of confidence with it, because of the increased sources of indeterminateness.

We can adopt another aggregation scheme: first we aggregate the experts' opinions with regard to each criterion, then we aggregate criteria scores into the applicant's final score. Although both models look similar, the result differs in the two cases. In the first scheme, experts opinions are aggregated into an expert overall score, then the model looks for consensus among experts about whether or not an applicant should gain the scholarship. In the second scheme we are looking for consensus among experts regarding each evaluation criterion of evaluation; then the model applies a multi-criteria decision scheme to derive the applicant's final score.

Example 4

Let us consider the scores as in Tab.1 to aggregate experts' opinions concerning each criterion. We obtain a result as described in Tab.4.

Criteria	C_1	C_2	C_3	C_4	C_5	C_6
Expert 1	3	2	3	2	3	1
Expert 2	2	3	3	2	3	2
Expert 3	2	2	3	2	2	1
Expert 4	3	2	3	3	3	2
Expert 5	2	2	3	2	3	1
Expert 6	3	2	3	2	3	1
Expert 7	1	2	3	2	3	2
Expert 8	1	2	3	2	3	1
Expert 9	1	2	2	2	3	2
Expert 10	1	2	2	3	3	1
Expert 11	1	2	2	2	2	1
Aggregated Opinion	2.875	2.526	3.000	2.725	3.000	2.000
Opinion Agreement	0.563	0.261	0.136	0.322	0.091	0.318

Table 4: Criteria aggregation

Weights used to aggregate criteria have an attitudinal character of $\sigma = 0.447$. Aggregation of criteria can be done by means of recursive weights [3] respecting the same character. Thus, the final score is

$$F = 0.282 \cdot I + 0.718 \cdot [2.667]$$

4   Conclusions

In this paper we have presented an ME-MCDM aggregation model based on logical assumptions aimed at considering real situations in which experts provide opinions with different degrees of confidence and relevance with regard to some evaluation criteria. In particular some opinions might not be solicited or might be missing. Different opinions, relevance of sources and missing information make decision effects more unpredictable. The model proposed in this paper attempts to deals with these problems making some logical assumptions in order to make the aggregation model more transparent. The result of the aggregation gathers sources

of uncertainty in an index named *indeterminateness*, which might make the result non-inferrable. This coefficient depends on the disagreement of opinions. Disagreement is function of distance. Future work will investigate in more detail such a dependency. Moreover, the MCDM model has been defined on fuzzy scores in order to model the vagueness of human judgments. We aim at extending the proposed model to the case when fuzzy experts opinion are described by fuzzy scores.

A large part of the proposed model was applied to different assessment contexts in four case studies. The first and the second case study dealt with tool assessment in which a software configuration management tool and an information retrieval system were selected respectively from a set of alternatives. The third case study regarded the assessment of software maintainability for a device driver. The study was made in comparison with another traditional maintainability assessment technique. The fourth case study faced the evaluation of security policies. This experimentation aimed at understanding whether a more detailed management of uncertainty could improve the process of decision making. From our experience we learned some lessons. The most valuable in this context is that model assumptions should be made clear. This is in accordance with The ACM Code of Ethics and Professional Conduct [1], which highlights how model transparency is important because it can build trust on decision support.

The proposed model makes clear assumptions based on logical rules and computes the result by inference. However, although the structure of dependencies in the proposed model is simple, the overall structure was not always completely understood, which limited the aim of the proposed model to be white box. In fact the model was mainly perceived as black box. This suggests that although fuzzy models are white-box, their transparency can be obfuscated by the complexity of model logic complexity. This requires attention to be paid to the question of how to make the model more easily understandable.

References

[1] R. E. Anderson, D. G. Johnson, D. Gotterbarn, and J. Perrolle. Using the ACM code of ethics in decision making. *Comm. of the ACM*, 36(2):98, 1993.

[2] G. Canfora and L. Troiano. Dealing with the "don't know" answer in risk assessment. In *Proc. Int. Conf. on Enterprise Information Systems - ICEIS'03*, volume Volume II - Artificial Intelligence and Decision Support Systems, pages 229–237. ACM, 2003.

[3] G. Canfora and L. Troiano. Recursive owa. In O. Kaynak, editor, *Proc. IFSA World Congress*, volume 776, pages 167–170, Istanbul, June 2003. IFSA, Figuür Grafik and Maatbacilik Sanayi Ticaret Ltd. Sti. ISBN: 975-518-208-X.

[4] G. Canfora and L. Troiano. A rule-based model to aggregate criteria with different relevance. In T. Bigiç, B. D. Baets, and O. Kaynak, editors, *Fuzzy Sets and Systems - IFSA 2003*, volume LNAI 2715 of *Lecture Notes in Artificial Intelligence*, pages 311–318, Istanbul, 2003. IFSA, Springer Verlag.

[5] C. Carlsson, R. Fullér, and S. Fullén. OWA operators for doctoral student selection problem. In R.R.Yager and J.Kacprzyk, editors, *The ordered weighted averaging operators: Theory, Methodology, and Applications*, pages 167–178. Kluwer Academic Publishers, Norwell, MA, 1997.

[6] A. Davey, D. Olson, and J. Wallenius. The process of multiattribute decision making: a case study of selecting applicants for a Ph.D. program. *European J. Oper. Res.*, 72:469–484, 1994.

[7] G. P. G. Bordogna, M. Fedrizzi. A linguistic modeling of consensus in group decision making based on OWA operators. *IEEE Trans. Systems, Man, Cybernet.-Pt. A: Systems Humans*, 27(1):126–132, 1997.

[8] E. Kerre and M. D. Cock. Linguistic modifiers: an overview. Fuzzy Logic and Soft Computing, pages 69–86. Kluwer Academic Publishers, Boston, 1999.

[9] R. Smolíková and M. P. Wachowiak. Aggregation operators for selection problems. *Fuzzy Sets Syst.*, 131(1):23–34, 2002.

[10] M. Sugeno and T. Takagi. A new approach to design of fuzzy controllers. pages 325–334. Plenum Press, New York, Wang, P.P. edition, 1983.

[11] A. Valls and V. Torra. Using classification as an aggregation tool for MCDM. *Fuzzy Sets and Systems*, 115(1):159–168, 2000.

[12] A. Valls and V. Torra. Fusion of qualitative preferences with different vocabularies. In M. E. Monferrer and F. T. Lobo, editors, *CCIA 2002*, volume LNAI 2504, pages 137–144. Springer-Verlag, 2002.

[13] R. R. Yager. On ordered weighted averaging aggregation operators in multi-criteria decision making. *IEEE Trans. on Systems, Man, and Cybernetics*, 18(1):183–190, 1988.

[14] R. R. Yager. Families of owa operators. *Fuzzy Sets and Systems*, (59):125–148, 1993.

[15] R. R. Yager. Non-Numeric Multi-Criteria Multi-Person Decision Making. *International Journal of Group Decision Making and Negotiation*, 2:81–93, 1993.