A Linguistic Group Best–Worst Method for Measuring Good Governance in the Third Sector: A Spanish Case Study

Ana Licerán-Gutiérrez[1] · Cristina Ortega-Rodríguez[1] · Antonio Moreno-Albarracín[1] ·
Álvaro Labella[2] · Rosa M. Rodríguez[2] · Luis Martínez[2]

Abstract The need of Non-profit Organizations (NPOs) of generating trust and credibility, to their stakeholders by an efficient management of their resources, lead them to openly show that they develop adequate good governance practices. But this is not a simple task and few research has been done on measuring methods of good governance in this field; without achieving an agreement about the best procedure. This paper aims at facilitating the measurement of good governance practices in NPOs by a fuzzy linguistic consensus-based group multi-criteria decision-making (MCGDM) model that will provide agreed and easy-understanding weights for a list of indicators proposed by the stakeholders and entities in such good governance practices. To do that, a linguistic 2-tuple BWM method with a consensus reaching process (CRP) will be developed and then applied to a real-world case in Spain, in which a group of experts from significant Spanish NPOs will assess the list of indicators proposed by the most representative entities (the alliance between the non-governmental organizations (NGO) Platform for Social Action, and the NGO Coordinator for Development (CONGDE) to obtain a prioritization of such indicators for measuring the good governance practices in Spanish NPOs.

Keywords BWM · Linguistic 2-tuple · Consensus · Good governance · Non-profit organizations

1 Introduction

Recently, Third Sector[1] has grown rapidly in significance and size in many countries [3] and it is usual to consider non-profit organizations (NPOs) as one of the most important agents in their society [4]. Classically, NPOs aimed at enhancing in charge of improving the quality of disadvantaged people life [5]. Thus, with its growing global importance entails responsibility towards the community. Such a responsibility was established in the Code of Ethics and Conduct of the World Association of Non-Governmental Organizations where it is stated that NPOs have a responsibility to be transparent, honest, responsible, and ethical and disclose accurate information [6]. Hence, NPOs worldwide are increase demanded for accountability and improved transparency [7, 8].

✉ Luis Martínez
martin@ujaen.es

Ana Licerán-Gutiérrez
aliceran@ujaen.es

Cristina Ortega-Rodríguez
corodrig@ujaen.es

Antonio Moreno-Albarracín
almoreno@ujaen.es

Álvaro Labella
alabella@ujaen.es

Rosa M. Rodríguez
rmrodrig@ujaen.es

1   Department of Financial Economics and Accounting, University of Jaén, Jaén, Spain

2   Computer Science Department, University of Jaén, Jaén, Spain

---

[1] The conceptualization of the Third Sector is certainly plural [1, 2]. The next concepts are often used as synonyms, generating an immense disparity in terms of terminology: "third sector", "non-profit sector", "voluntary sector", even "charitable sector". The same with the organizations that compose it: "Non-profit organizations (NPOs)", "Third Sector Organizations (TSOs)" or "Nongovernmental Organization (NGOs)".

Within the 2030 Agenda for Sustainable Development issued in 2015, the United Nations (UN) promulgated 17 Sustainable Development Goals (SDGs) that play a key role in its evaluation and implementation. Our study is focused on the goals "16.6 Develop effective, accountable and transparent institutions at all levels" and "17.19 Build on existing initiatives to develop measurements of progress on sustainable development that complement gross domestic product, and support statistical capacity building in developing countries".

Accountability is useful for stakeholders forming a clear idea about the funds administration by NPOs. It is precisely at this point, where the terms "non-profit responsibility" together with "good governance practices" arise. However, an analysis on good governance in the Third Sector reveals a series of concerns:

- It shows how literature on accountability issues in the Third Sector is limited [9]. Furthermore, much remains to be understood in terms of accountability mechanisms [10, 11], given the scarce research on the analysis of governance mechanisms [12–14].
- The scant number of researches that empirically measure the levels of non-profit good governance. Different institutions have prepared their own proposals based on lists of good governance indicators [15] and most of these lists do not weight the indicators and, if they are weighted, there is not any theoretical or empirical validation for such weights.

Previous issues drive us to conclude that there is no a clear consensus on the use of a certain procedure for measuring the degree of good governance in the field of NPOs.

To fill previous gaps up, we propose a novel linguistic consensus-based multi-criteria group decision-making (MCGDM) approach that models the problem of weighting a list of indicators for good governance in NPOs and aims at achieving agreed and easy-understanding results. Among the different proposals that can be found in the literature to deal with MCGDM weighting criteria [16–18], our proposal takes as basis the Best–Worst method (BWM) presented by Rezaei in [19] and introduces a new consensus-based linguistic group BWM in order to:

- Smooth out the conflicts among experts involved in the weighting of indicators for Good Governance in NPOs by applying a CRP [20–22].
- Weighting indicators for Good Governance in NPOs by the BWM method that gathers information from multiple decision-makers. To facilitate the elicitation of such information, it is used, modified and extended, the linguistic 2-tuple BWM [23] that will manage the inherent uncertainty, related to the vagueness of meaning of knowledge expressions

provided by decision-makers. Additionally, it will also provide easy-understanding results for all stakeholders in spite of their different knowledge and background.

Notice the BWM derives criteria weights from pairwise comparisons provided by an expert. Concretely, the expert should choose the best and worst criterion and then compare these ones with the remainder. However, in MCGDM problems, as the one studied in this proposal, experts may disagree on the selection of the best and/or worst criterion or on the pairwise comparisons, resulting in different weights for the same criterion for each expert. To face this issue, we extend the BWM by including a consensus approach based on a comprehensive minimum cost consensus model (CMCC) [24]. A CMCC model is a non-linear programming model that guarantees to achieve a collective consensual solution by preserving as much as possible the initial preferences of each expert. In this way, we are able to obtain global consensual weights for all the criteria. Additionally, we make use of an extension of the BWM, so-called 2-tuple BWM [23], which allows modeling the experts' preferences by 2-tuple linguistic values [25]. However, this method does not represent properly, from a linguistic point of view, the criteria weights, making difficult the interpretation of the results from the experts. Keeping in mind the previous limitations, the main novelties of our proposal are listed below:

- A linguistic consensus-based approach that aims at removing disagreements in the final weighting of the indicators. It is remarkable that no previous BWM proposals within MCGDM problems have considered consensus approaches yet.
- The recent extension of BWM, so-called 2-tuple BWM [23] that models input and output information by linguistic 2-tuples [25] and facilitates the interpretability and keeps accuracy of the results, will be modified to model the linguistic output by building an adequate syntax and semantics for the linguistic term set used to express the weights obtained using a fuzzy unbalanced linguistic approach [26].

Eventually, once the model has been developed to show the performance of the consensus-based 2-tuple BWM it will be undertaken an empirical analysis about good governance in NPOs, empirically validating indicators with the decision-makers' opinions in accountability of these entities. Specifically, such an empirical analysis will focus on the Spanish social economy context, in which Spain plays a key role within the European Union, and in which the list of indicators to be weighted is the proposed by the alliance between the Social Action NGO Platform and the NGO

Coordinator for Development (CONGDE, hereinafter) in 2019 [27]. This case study focuses on the Spanish context since, in the field of the social economy, Spain has become a pioneer country, being the first member of the European Union to develop a law on social economy, the Law 5/2011, of March 29 [28]. This law gives a relevant value to this model of provision of services to citizens, granting it a unique recognition to this sector. Besides, Spain has been the first member state of the European Union to implement a Social Economy Strategy 2017–2020 based on 63 measures that are supported by 11 strategic axes. For the development of the 2030 Agenda, Spain places the Spanish Social Economy Strategy 2017–2020 as an essential element to achieve the Sustainable Developments Goals (SDGs) promulgated by the United Nations. Here is the relevance of studying the Spanish case.

The paper is set up as follows: In the Sect. 2, different concepts of the 2-tuple linguistic model, 2-tuple BWM and the CRPs that will be used in our proposal are revised and then a review of related works about good governance in the Third Sector is done to understand the importance of studying the challenges facing the sector. Sect. 3 introduces the novel consensus-based 2-tuple BWM that will be used in Sect. 4 to provide a proper and consensual weighting for the good governance indicators in the Spanish case. Afterwards, the results obtained are analyzed in Sect. 5. Then, the proposal is compared with other similar studies in Sect. 6. Finally, Sect. 7 concludes this paper.

2 Background

This section revises basic concepts about 2-tuple linguistic model, 2-tuple BWM and CRPs that are necessary to understand our proposal. Eventually, a related works section about Good governance in Third Sector is presented to understand the importance and application of our proposal.

2.1 2-Tuple Linguistic Model

Fuzzy linguistic approach [29] has been broad and successfully used for dealing with uncertainty in decision-making. Among the fuzzy linguistic models used in decision-making, outstands the 2-tuple linguistic model [25] because it facilitates the accomplishment of linguistic computations with high readability and precision. High readability is related to the use of a computing with words (CW) approach [30], in which both decision-makers' preferences and results are represented linguistically. High precision is related to the use of the symbolic translation concept. The 2-tuple linguistic model represents the information by a 2-tuple (s_i, α) where s_i is a linguistic term belonging to a predefined linguistic term set $S =$

$\{s_1, \ldots, s_g\}$ and $\alpha \in [-0.5, 0.5)$ a numerical value that represents the symbolic translation of the membership function of the linguistic term s_i:

$$\alpha = \begin{cases} [-0.5, 0.5) & \text{if } s_i \in \{s_1, s_2, \ldots, s_{g-1}\} \\ [0, 0.5) & \text{if } s_i = s_0 \\ [-0.5, 0] & \text{if } s_i = s_g \end{cases}$$

To accomplish CW processes with linguistic 2-tuple values, different functions were established.

The function Δ_S, from a value of $\beta \in [0, g]$, returns an equivalent 2-tuple linguistic value, (s_i, α):

Definition 1 ([25]) Let $S = \{s_0, \ldots, s_g\}$ be a set of linguistic terms and $\overline{S}$ *the 2-tuple set associated with S defined as* $\overline{S} = S \times [-0.5, 0.5)$:

$$\Delta_S(\beta) = (s_i, \alpha), \text{ with } \begin{cases} i = \text{round}(\beta) \\ \alpha = \beta - i \end{cases}$$

where $\text{round}(\cdot)$ assigns the closest integer number $i \in \{0, \ldots, g\}$ to β.

The function Δ_S^{-1}, from a 2-tuple linguistic value (s_i, α), returns its equivalent numerical value β in the interval of granularity of S, [0, g].

Preposition 1 ([25]) *Let* $S = \{s_0, \ldots s_g\}$ *be a linguistic term set and* $(s_i, \alpha) \in \overline{S}$ *be a 2-tuple linguistic value*:

$$\Delta^{-1} : \overline{S} \to [0, g]$$
$$\Delta_S^{-1}(s_i, \alpha) = \alpha + i = \beta$$

Remark 1 A linguistic term $s_i \in S$ can be transformed into a 2-tuple linguistic value in $\overline{S}$ by including a symbolic translation equal to zero:

$$s_i \in S \to (s_i, 0) \in \overline{S}$$

The 2-tuple linguistic model has been extended to facilitate computations. Specifically, Tai et al. [31] proposed an approach in which any value $\beta \in [0.1]$ can be mapped to the linguistic 2-tuple terms set $\overline{S'}$ via the mapping $\Delta_S' : [0.1] \to \overline{S'} = S \times [-\frac{1}{2g}, \frac{1}{2g})$.

Definition 2 [31] Let $S = \{s_0, \ldots, s_g\}$ be a set of linguistic terms and $\overline{S'}$ *the 2-tuple set associated with S defined as* $\overline{S'} = S \times [-\frac{1}{2g}, \frac{1}{2g})$. The function Δ_S' is given by:

$$\Delta_S'(\beta) = (s_i, \alpha), \text{ with } \begin{cases} i = \text{round}(\beta \cdot g) \\ \alpha = \beta - \frac{i}{g} \end{cases} \text{with} \alpha \in [-\frac{1}{2g}, \frac{1}{2g})$$

Again, the 2-tuple linguistic value (s_i, α) could be translated into an equivalent numerical value $\beta \in [0.1]$:

Definition 3 [31] Let $S = \{s_0, \ldots s_g\}$ be a linguistic term set and $(s_i, \alpha) \in \overline{S'}$ be a 2-tuple linguistic value. There is a function $\Delta_S'^{-1}$:

$$\Delta_S'^{-1}(s_i, \alpha) = \beta = \frac{i}{g} + \alpha.$$

with $\beta \in [0.1] \subset \Re$.

2.2 2-Tuple Best–Worst Method

BWM was introduced by Rezaei in [19] as a MCDM method for computing the criteria weights by reducing the number of pairwise comparisons among criteria and the inconsistency of decision-makers' preferences. In the BWM, such pairwise comparisons are so-called *reference comparisons*. In the original proposal, decision-makers use a numerical scale to provide their preferences. However uncertain contexts definition are very common in real-world MCGDM problems and cannot be easily handled using numerical precise assessments. For this reason, Labella et al. [23] proposed the 2-tuple BWM able to model both decision-makers' preferences and results by linguistic information. This model consists of the below steps based on the classical BWM:

- **Step 1**: To determine a set of decision criteria, $C = \{C_1, \ldots, C_n\}$.
- **Step 2**: To select the best criterion C_B and the worst criterion C_W. In case that there are several best and worst criteria, these can be selected arbitrary.
- **Step 3**: To make linguistic pairwise comparisons among C_B and the rest of the criteria using the linguistic scale S^{BWM} represented in Table 1 and obtaining the Best to Others (BO) vector, $BO = \{a_{B1}, a_{B2}, \ldots a_{Bn}\}$, where a_{Bj} denotes the preference degree of C_B over the criterion C_j and $a_{Bj} \geq 1, j = 1, 2, \ldots n, \ j \neq B$.
- **Step 4**: To make analogous linguistic pairwise comparisons among C_W and the rest of the criteria and obtaining the Others to Worst (OW) vector,

$OW = \{a_{1W}, a_{2W}, \ldots a_{nW}\}$, where a_{jW} denotes the preference degree of the criterion C_j over C_W and $a_{jW} \geq 1, j = 1, 2, \ldots n, \ j \neq B$ or W.

- **Step 5**: The linguistic pairwise comparisons are collected into a Linguistic Best–Worst Comparison Matrix (LBWCM) by transforming the linguistic terms into 2-tuple linguistic values according to Remark 1.

$$LBWCM = \begin{matrix} & c_1 & c_2 & \cdots & c_n \\ c_B & \begin{pmatrix} (d_{B1}, \alpha_{B1}) & (d_{B2}, \alpha_{B2}) & \cdots & (d_{Bn}, \alpha_{Bn}) \\ c_W & (d_{1W}, \alpha_{1W}) & (d_{2W}, \alpha_{2W}) & \cdots & (d_{nW}, \alpha_{nW}) \end{pmatrix} \end{matrix}$$

where (d_{Bj}, α_{Bj}) is the decision-maker's linguistic preference over the comparison of C_B with c_j and (d_{jW}, α_{jW}) denotes the linguistic preference for the comparison of C_W with c_j.

- **Step 6**: To compute the criteria weights by an optimization model from the LBWCM in which the maximum absolute differences $|w_B/w_j - \Delta_{S^{BWM}}^{-1}((d_{Bj}, \alpha_{Bj}))|$ and $|w_j/w_W - \Delta_{S^{BWM}}^{-1}((d_{jW}, \alpha_{jW}))|$ should be minimized:

$(\mathbf{M-1})$

$$\min \xi$$

$$\text{s.t.} \begin{cases} \sum_{j=1}^n w_j = 1, \\ |\dfrac{w_B}{w_j} - \Delta_{S^{BWM}}^{-1}((d_{Bj}, \alpha_{Bj}))| \leq \xi, \\ |\dfrac{w_j}{w_W} - \Delta_{S^{BWM}}^{-1}((d_{jW}, \alpha_{jW}))| \leq \xi, \\ w_j \geq 0, \ \text{forall } j = 1, 2, \ldots, n. \end{cases} \quad (1)$$

At this point, the priority weights that are numerical, are transformed into linguistic weights. These linguistic weights are represented through a linguistic term set (see Table 2) since any value from the interval $\beta \in [0, 1]$ can be linguistically interpreted by Definition 2.

This approach presents an important limitation from the syntax point of view in the fuzzy linguistic scale for representing criteria weights. Because the current fuzzy linguistic scale, S_w^{BWM}, is symmetric and equally distributed and due to the fact that the numerical weights should sum 1, the linguistic representation of most of weights obtained

Table 1 Fuzzy linguistic scale, S^{BWM}, for preferences in 2-tuple BWM

Linguistic terms	Membership function
Equally Importance (EI)	(1, 1, 1)
Weakly Important (WI)	(2/3, 1, 3/2)
Fairly Important (FI)	(3/2, 2, 5/2)
Very Important (VI)	(5/2, 3, 7/2)
Absolutely Important (AI)	(7/2, 4, 9/2)

Table 2 Fuzzy linguistic scale, S_w^{BWM}, for weights in 2-tuple BWM

Linguistic terms	Membership function
Very Unimportant (VU)	(0, 0, 0.25)
Unimportant (U)	(0, 0.25, 0.5)
Fair (F)	(0.25, 0.5, 0.75)
Important (I)	(0.5, 0.75, 1)
Very Important (VI)	(0.75, 1, 1)

by the method are usually modeled by either *VU* or *U* which could be a bit confuse. Such a problem is further detailed in the below example:

Example 1 Let us suppose a MCDM problem composed by six criteria whose resulting numerical weights obtained from (M-1) are $w_1 = 0.26$, $w_2 = 0.282$, $w_3 = 0.141$, $w_4 = 0.141$, $w_5 = 0.035$ and $w_6 = 0.141$. Their linguistic representation according to the fuzzy linguistic scale represented in Table 2 is $w_1 = (U, 0.01)$, $w_2 = (U, 0.032)$, $w_3 = (U, -0.109)$, $w_4 = (U, -0.109)$, $w_5 = (VU, 0.035)$, $w_6 = (U, -0.109)$. Keeping in mind the number of criteria, in this case six, the equal importance for all of them would be $1/6 = 0.167$ and it does not seem logical to provide a "unimportant" importance for weights above the average mean, for instance w_2.

The previous limitation will be fixed in the proposal introduced in Sect. 3 using a fuzzy unbalanced linguistic scale for representing the linguistic weights [26].

Another key issue, when a MCDM method deals with pairwise comparisons, is the evaluation of their consistency, since inconsistent preferences implies unreliable results. In [23] was introduced a consistency ratio based on a *random average consistency index* (RACI). The RACI represents the average value of consistency ξ generated by selecting N random configurations of LBWCMs for a given n and a_{BW}, thus it depends on both the best to worst ratio a_{BW} and the number of the objects under consideration. Table 3 shows the RACI values depending on the number of objects to compare (rows) and the best to worst ratios (columns).

$$\text{RACI}(n, a_{BW}) = \frac{1}{N} \sum_{k=1}^{N} \xi^{(k)}(n, a_{BW}) \tag{2}$$

Based on the notion of $\text{RACI}(n, a_{BW})$, a consistency ratio for a given LBWCM is defined as follows:

$$CR(n, a_{BW}) = \frac{\xi(n, a_{BW})}{RACI(n, a_{BW})} \tag{3}$$

According to [23], $CR(n, a_{BW}) \leq 0.35$ is an acceptable consistency threshold because of it is not very

restrictive and guarantees sufficient consistency to generate reasonable results.

2.3 Consensus Reaching Process

Decision-making often involves groups to obtain better decisions. Group decisions are usually better accepted when the solution is agreed by all decision-makers involved in the group decision problem [20]. The achievement of this agreement usually implies a CRP before selecting the best alternative for the group decision-making problem.

In the specialized literature different CRPs models have been proposed [21, 32–36] and their classification depends on several features [35]. In this paper, we focus on CRPs without feedback mechanism that aim at achieving the agreement in an automatic way [37–39].

Specifically, we propose the use of a *minimum cost consensus* (MCC) model to optimize the CRP and obtain quick agreed solutions. MCC was introduced by Ben-Arieh and Easton [40] aiming at minimize the overall cost of moving all decision-makers' opinions to achieve the agreement. Taking into account the MCC concept, Zhang et al. studied how the level of agreement in the group can be different according to the selected aggregation operator for computing the collective opinion and proposed a new MCC model [38]. The properties of the latter model were investigated under the situations that the weighted average operator or the ordered weighted average operator are used to compute the collective opinion [38].

Recently, some researchers have paid much attention on the model proposed by Zhang et al. [38] and have introduced some new MCC approaches [24, 41–43]. Particularly, Labella et al. [24] introduced a MCC approach in which the consensus computation within the decision-makers group is considered. According to their research, the classical MCC cannot guarantee a minimum level of agreement for the group of decision-makers, but just a maximum distance between each decision-maker's opinion and collective opinion. Therefore, a CMCC that modifies the model introduced by Zhang et al. in [38] was developed by including the computation of consensus level:

Table 3 RACI

$RACI(n, a_{BW})$	2	3	4	5	6	7	8	9
3	0.18361	0.37665	0.55037	0.75272	0.95964	1.18496	1.40357	1.64334
4	0.29530	0.54800	0.81162	1.09939	1.40172	1.71159	2.05341	2.40471
5	0.36089	0.64679	0.96388	1.30009	1.64452	2.04431	2.42486	2.848
6	0.40466	0.71746	1.06277	1.45598	1.84730	2.27373	2.71134	3.14877
7	0.43129	0.76699	1.14422	1.55603	1.98077	2.43942	2.90589	3.41465
8	0.44958	0.80825	1.20708	1.64420	2.09304	2.58759	3.07247	3.56621
9	0.77678	1.37058	2.00741	2.62732	3.27161	3.93522	4.67623	5.33078

Definition 4 [24] Let $(o_1, \ldots, o_m)$ be the original assessments provided by a set of decision-makers $E = \{e_1, e_2, \ldots, e_m\}$ over an alternative. Suppose that after CRP, the decision-makers' assessments are modified into $(\overline{o}_1, \ldots, \overline{o}_m)$, and a collective opinion $\overline{o}$ is obtained based on the modified assessments, and $(c_1, \ldots, c_m)$ are the cost of moving each decision-maker's opinion 1 unit, respectively. The parameter ε is the maximum acceptable distance of each decision-maker to the collective opinion. The MCC model based on a linear cost function is given as follows:

$$(\mathbf{M-2})$$

$$\min \sum_{k=1}^{m} c_k |\overline{o}_k - o_k|$$

$$\text{s.t.} \begin{cases} \overline{o} = F(\overline{o}_1, \ldots, \overline{o}_m) \\ |\overline{o}_k - \overline{o}| \leq \varepsilon, \quad k = 1, 2, \ldots, m \\ \text{consensus}(\overline{o}_1, \ldots, \overline{o}_m) \leq \gamma, \end{cases}$$

where F is an aggregation function, $\text{consensus}(\cdot)$ represents the consensus level achieved and γ is equal to $1 - \mu$, being $\mu \in [0.1]$ a consensus threshold defined a priori.

Labella et al. proposed several CMCC models based on different measures to compute the group agreement level. Here we introduce the one used in this proposal:

$$consensus(\overline{o}_1, \ldots, \overline{o}_m) = \sum_{i=1}^{m} \phi_i |\overline{o}_i - \overline{o}| \tag{4}$$

where $\phi_k \in [0.1]$ is the decision-makers e_k's weight and $\sum_{k=1}^{m} \phi_k = 1$.

2.4 Related Works on Good Governance in the Third Sector

A brief but insightful review of main researches done in good governance in the Third Sector pointing out its importance and challenges is provided here.

Accountability in the Third Sector is crucial [44] to achieve NPO's social mission [45], and as a way to strengthen trust, improving relationships and demonstrating transparency in its activities within the community [11]. In this sense, accountability practices are considered for users as NPO's overall commitment to transparency, based on legitimacy [46] which reaffirms their relevant contribution to the society.

To carry out activities for the community, NPOs are financially dependent on several internal actors, such as their own members and beneficiaries, as well as external founders [47]. Due to its social nature, NPOs obtain funding [48]; however, that money must not be used for their personal benefit [6]. Besides, taken into account the competitive environment, where donors have multiple options, maintaining positive perceptions of trustworthiness have been proved to be decisive for the existence of the sector as a whole [49]. Consequently, trust in the Third Sector is essential. Good governance has become the most valuable element for NPOs to achieve the social credibility that allows ensuring their future.

Unfortunately, fraud cases have triggered a crisis of confidence in the sector [50, 51]. Founders are concerned with allegations of corruption among NPOs [52]. Furthermore, the controversial behavior of some NPOs has resulted in increased efforts to analyze and improve the reputation of the sector [53]. In summary, it is observed the increasing need of accountability due to the inappropriate behavior of some NPOs, which have damaged the credibility of the organizations that conform the Third Sector [54].

To solve this, it is necessary to carry out ethical practices that are visible for their stakeholders. Under these circumstances, the concept "good governance" arises. Taking this view, non-profit governance should relate to all stakeholders involved in a NPO [44, 55, 56]. Non-profit governance and accountability are social and dynamic processes [57], becoming a central concern for NPOs [3]. This is reinforced with the view in prior literature that NPOs are perceived as more effective when they manage to align the diversity of expectations of stakeholders with good governance [56].

3 A Consensus Model Based on Linguistic BWM for MCGDM

With the aim of obtaining agreed weights for a MCGDM such as we intend in the good governance for the Third Sector problem, it is necessary to apply a consensus process to smooth disagreements in the final weighting of the criteria. Therefore, our proposal consists of a CMCC model based on the 2-tuple BWM that will obtain the importance of the criteria in a consensual way. The steps of this new approach are further described below and graphically depicted in Fig. 1:

- **Step 1**: decision-makers select the best and worst criterion according to their view and provide linguistic pairwise comparisons between them and the rest of the criteria using the linguistic scale represented in Table 1. Then, the pairwise comparisons are transformed into linguistic 2-tuple values.
- **Step 2**: for each decision-maker e_k the criteria weights are obtained $(w_1^k, w_2^k, \ldots, w_n^k)$ by applying the 2-tuple BWM (see Sect. 2.2). Due to the 2-tuple BWM allows to represent the weights linguistically, the decision-makers may choose whether they prefer
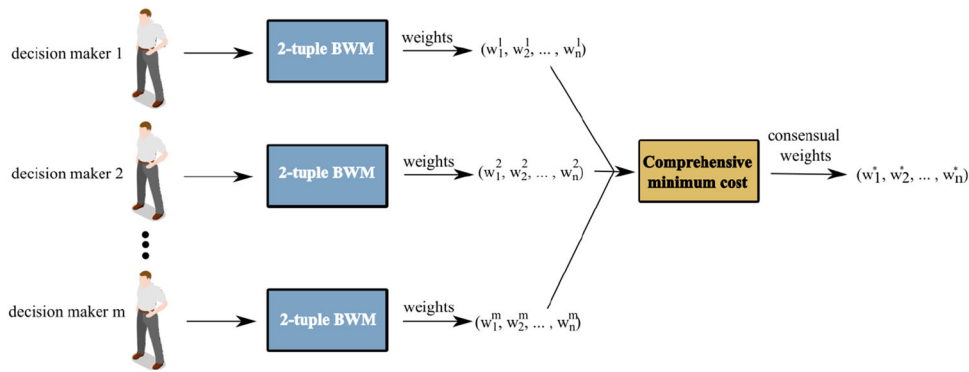
Fig. 1 Proposal scheme

to obtain weights represented numerically or linguistically. However, as it was aforementioned, the fuzzy linguistic scale used for representing the weights in the 2-tuple BWM and represented by fuzzy triangular numbers it is not suitable either syntactically or semantically to represent the real importance of the weights. In order to overcome such a weakness, it is reasonable to develop an unbalanced linguistic scale [26] that takes into account that the average weight for all elements cannot be considered as unimportant but rather than as "*average*". In spite of multiple proposals for building unbalanced linguistic term sets [26, 58], such proposals proposed methods that fix the semantics prior to the number of elements to be assessed. However, in the BWM the importance of the elements are related to the number of elements, i.e., the average importance for 5 elements is 0.2 but for 9 elements is 0.11. For this reason, in order to express properly the importance of the elements in a linguistic way, we propose the use of a fuzzy unbalanced linguistic scale whose semantics is built according to the number of the elements that will be assessed in the 2-tuple BWM. Such a fuzzy unbalanced linguistic scale, $U_w^{\text{BWM}}(n)$, aims at building a terms set that will be a fuzzy partition in the sense of Ruspini [59] and then develop the terms set taking into account that the "*Average Important* (AI)" linguistic term represents the average importance value thus, its core is the value of the universe of discourse ([0, 1]) that is equal to $1/n$, being n is the number of elements to compare. Hence, if all the elements in the problem solved by the BWM were of equal importance, their linguistic weights would be $(AI, 0)$. Then, the "*Unimportant*" and "*Average Important*" linguistic terms are distributed from the average importance value $1/n$. The remaining terms are adjusted to build the fuzzy partition such that their central values are symmetrically distributed in

the interval $[1/n, 1]$. In other words, the distance between the core of the 3 remaining triangular fuzzy terms is derived as $\frac{1-(1/n)}{3} = \frac{n-1}{3n}$. Note that, in spite of the fuzzy unbalanced scale semantics changes depending on n, its syntax remains unchanged. This syntax is composed by triangular fuzzy linguistic terms whose meaning facilitates the readability of the linguistic weights. (see Table 4).

Figures 2 and 3 show fuzzy unbalanced linguistic scales with four ($U_w^{\text{BWM}}(4)$) and six ($U_w^{\text{BWM}}(6)$) elements to compare respectively

The numerical weights obtained by the 2-tuple BWM, can be translated to 2-tuple linguistic weights in $U_w^{\text{BWM}}(n)$, by means of the below process (see Fig. 4):

Preposition 2 *Let w_i the numerical weight associated to the criterion c_i and (s_j, α) its corresponding 2-tuple linguistic representation in $U_w^{\text{BWM}}(n)$.*

$$j = \arg\min_h |w_i - \bar{x}_h|, \ h \in \{0, \dots, g\}$$
$$\alpha = \frac{w_i - \bar{x}_j}{\bar{x}_j - \bar{x}_k} \in [0.5, 0.5), \ k = j+1 \ if \ w_i > \bar{x}_j \ or \ k = j-1 \ if \ w_i < \bar{x}_j$$

where $\bar{x}_h$ represents the value of the coordinate x of the centroid [60] of the linguistic term s_h, $g+1$ the cardinality

Table 4 Fuzzy unbalanced linguistic scale, $U_w^{\text{BWM}}(n)$, for weights in proposal

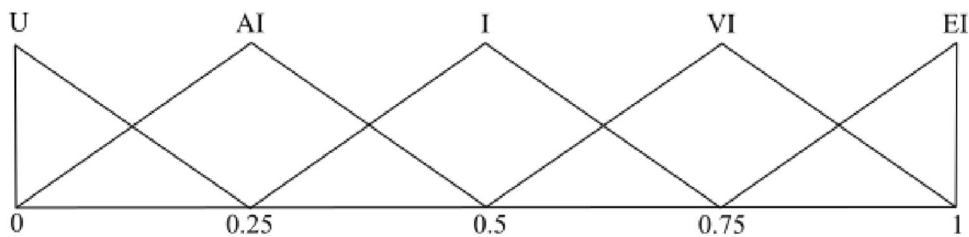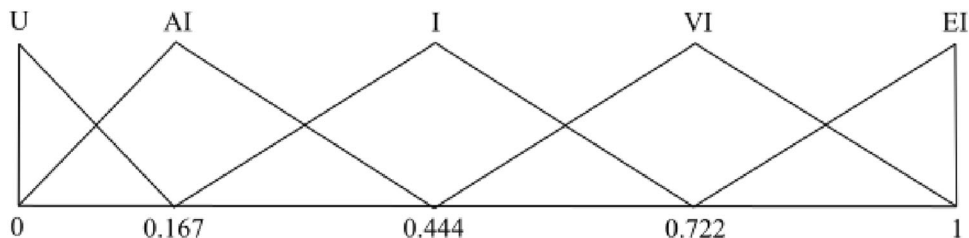| Linguistic terms	Membership function
Unimportant (U)	$(0, 0, \frac{1}{n})$
Average Important (AI)	$(0, \frac{1}{n}, \frac{n+2}{3n})$
Important (I)	$(\frac{1}{n}, \frac{n+2}{3n}, \frac{2n+1}{3n})$
Very Important (VI)	$(\frac{n+2}{3n}, \frac{2n+1}{3n}, 1)$
Extremely Important (EI)	$(\frac{2n+1}{3n}, 1, 1)$

Fig. 2 Fuzzy unbalanced linguistis scale $U_w^{\mathrm{BWM}}(4)$

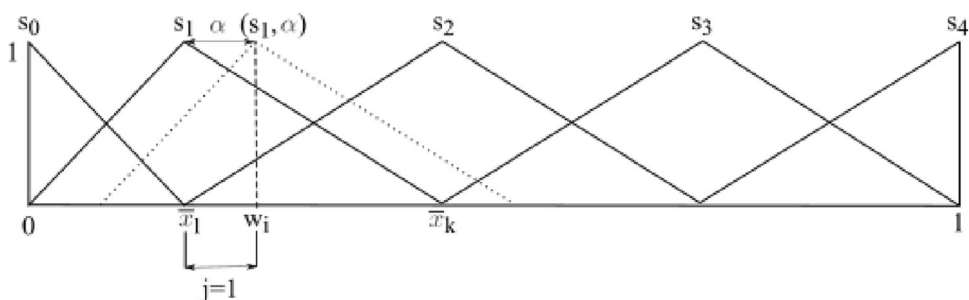**Fig. 3** Fuzzy unbalanced linguistis scale $U_w^{\mathrm{BWM}}(6)$

Fig. 4 2-tuple linguistic weight computation

of $U_w^{\mathrm{BWM}}(n)$ and j is the index of the linguistic term whose centroid is the closest one to w_i.

At this point, from each decision-maker e_k's preferences we obtain their corresponding criteria weights $(w_1^k, w_2^k, \ldots, w_n^k)$ represented linguistically to improve their readability. However, the decision-makers may present different points of view and disagreements among the importance of the criteria. To smooth such disagreements and achieve an agreed solution, a CRP is included [20–22] in the resolution process. Particularly, we apply a CMCC model that guarantees to achieve a desired level of agreement between the decision-makers by changing as less as possible their preferences regarding the criteria weights. In this way, we will obtain a collective consensual weight for each criterion in which decision-makers agree, noted as $(\overline{w}_1^*, \overline{w}_2^*, \ldots, \overline{w}_n^*)$.

- **Step 3**: to obtain the consensual weights, the computed linguistic criteria weights for each

decision-maker e_k, $(w_1^k, w_2^k, \ldots, w_n^k)$ are transformed into their numerical representation in [0, 1] using Def. 3 and then used as inputs in the consensus model to achieve the agreed solution. Among the different automatic consensus models pointed out in Sect. 2.3, we have chosen the CMCC model, because it models not only the distance but also the level of agreement. Therefore, the CMCC model (M-2) is used to obtain the consensual collective weights for the criteria $(\overline{w}_1^*, \overline{w}_2^*, \ldots, \overline{w}_n^*)$. According to the decision-maker choice, these weights can be represented numerically or transformed again into 2-tuple linguistic values using the fuzzy unbalanced scale shown in Table 4 and Preposition 2 improving the interpretability of the results. The CMCC model used in the proposal is described in further detail below:

$(\mathbf{M} - 3)$

$$\min \sum_{k=1}^{m} \sum_{i=1}^{n} c_k |\overline{w}_i^k - w_i^k|$$

$$\text{s.t.} \begin{cases} \overline{w}_i^* = \sum_{i=n}^{n} \sum_{k=1}^{m} \phi^k \overline{w}_i^k \\ |\overline{w}_i^k - \overline{w}_i^*| \leq \varepsilon, \ k = 1, 2, \ldots, m, \ i = 1, 2, \ldots n \\ \sum_{k=1}^{m} \sum_{i=1}^{n} \phi_i |\overline{w}_i^k - \overline{w}_i^*| \leq \gamma, \\ \sum_{i=1}^{n} \overline{w}_i^* = 1 \end{cases}$$

Note that this proposal uses the CMCC model for deriving

model preserves as much as possible the initial experts' preferences.

Therefore, the model (M-2) modifies the initial weights of each decision-maker obtained from the 2-tuple BWM and guarantees to compute global weights for the criteria by preserving as much as possible the decision-makers' views regarding their importance in a quick and easy way.

All the previous steps have been summarized in the following algorithm:

Algorithm 1: Proposal algorithm.

Data: criteria $C = \{C_1, \ldots, C_n\}$, experts $E = \{e_1, e_2, \ldots e_m\}$, experts' costs $(c_1, \ldots, c_m)$, experts' weights $(\phi_1, \ldots, \phi_m)$, CMCC distance threshold ε and CMCC consensus level γ

Result: the global consensus weights for the criteria $(\overline{w}_1^*, \overline{w}_2^*, \ldots, \overline{w}_n^*)$

for $k \leftarrow 1 : m$ **do**
  $C_B^k \leftarrow e_k$ *selects the best criterion*;
  $C_W^k \leftarrow e_k$ *selects the worst criterion*;
  $BO^k \leftarrow \{a_{B1}, \ldots, a_{Bn}\};$ /* best to others comparisons */
  $OW^k \leftarrow \{a_{1W}, \ldots, a_{nW}\};$ /* others to worst comparisons */
  $W^K \leftarrow Apply(M-1);$ /* $W^k \leftarrow (w_1^k, w_2^k, \ldots, w_n^k)$ */
end
$\overline{W}^* \leftarrow Apply(M3);$ /* $\overline{W}^* \leftarrow (\overline{w}_1^*, \overline{w}_2^*, \ldots, \overline{w}_n^*)$ */

---

the consensual weights because of the following reasons:

1. To achieve a consensual solution implies to modify the initial experts' preferences. The CMCC aims at minimizing the cost of modifying experts' preferences thus, it guarantees to reach a consensual solution by changing as less as possible the initial experts' opinions, which is, in fact, the best solution that can be obtained for the problem.
2. Classical MCC models guaranteed only a maximal distance between each expert and the collective opinion, the smaller the distance the higher the consensus within the group. However, it does not imply to reach a desired level of consensus. The CMCC model [24] adds an additional constraint related to the consensus computation that guarantees to achieve a solution with a required level of consensus represented by the parameter γ.
3. The CMCC is an automatic optimization model that does not requires the participation of the experts in the consensus process. In this way, the consensus process is faster and possible deadlocks are avoided. In addition, a real consensus process would not get a better solution since the CMCC

4 Weighting Good Governance Indicators: A Spanish Case Study

The CONGDE (2019) tool defines a set of indicators necessary to regulate and guarantee an adequate level of good governance of the NPOs, in a responsible exercise of self-regulation [27]. For each indicator, CONGDE proposes the following information: (1) the weight that indicator has within the block of which it is part and (2) the degree of importance is a field in which we classify the indicators into two categories: Some are of "inexcusable compliance" and other indicators are "relevant". This is where certain inconsistencies are observed because there is no numerical connection (weightings) with linguistics (relevant or inexcusable compliance). This, together with the fact that the CONGDE does not explain why some indicators weight more than others, nor does it empirically contrast it with the decision-makers, leads us to apply this BWM method in a double sense: (1) To check the weights numerically, and (2) To check if linguistically, what the decision-makers think coincides with that of the CONGDE in what is more or less important and, accordingly, what a stronger degree of compliance is required in qualitative terms.
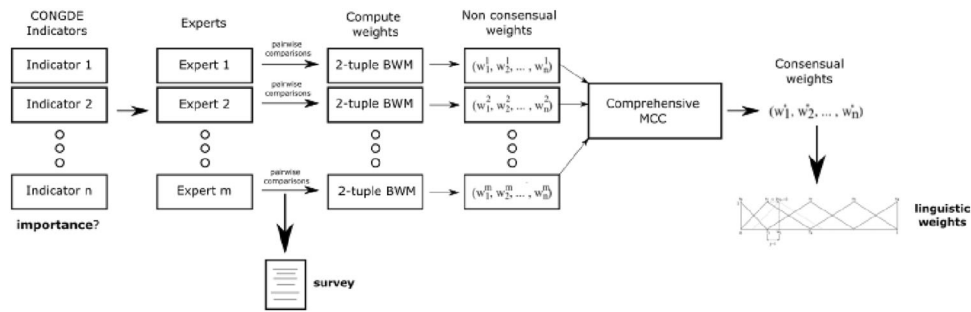
Fig. 5 Resolution process for the case study

Table 5 Parameters

Parameter	Description	Value
γ	Desired consensus level	0.85
ε	Maximum acceptable distance	0.15
$c_k,\ k = 1\ldots m$	Experts' cost	1
$\phi_k k = 1\ldots m$	Experts' weights	$\frac{1}{m}$

Therefore, the aim of this case study is to derive the priority of the list of indicators proposed by CONGDE using the linguistic BWM approach introduced in Sect. 3. Figure 5 represents a general system diagram with the different steps carried out to solve the case study. Table 5 shows the values of the parameters that are necessary such an approach in the resolution of this case study and the different steps that compose the process are described below.

Remark 2 Note that in the resolution of this case study, the opinions of all experts have been considered equally important as the cost of modifying them.

4.1 Step 1

First, the decision-makers should select the best and worst indicators according to their expertise and then compare each one with the rest of indicators. In this case study, we ask to 5 decision-makers, who representing some of the most important Spanish social entities, about these issues by means of a questionnaire. This questionnaire is available online[2]. The decision-makers' preferences are shown in Tables 6, 7, 8, 9, 10, 11 and 12. These tables show, for each decision-maker, the best indicator (C_B), the worst indicator (C_W), the pairwise comparisons between the best indicator and the rest of them (BO), the pairwise comparisons of the rest of indicators with the worst one (OW) and the consistency of the opinions obtained from Eq. 3.

Remark 3 Note that all the decision-makers' preferences are consistent according to the consistency threshold defined in [23].

4.2 Step 2

Lately, the importance of the indicators are derived from the decision-makers' opinions using the non-linear programming model (M-1). The indicators weights for each decision-maker are shown in Tables 13, 14, 15, 16, 17, 18 and 19, represented both numerically (N) and linguistically (L).

4.3 Step 3

From the resulting individual weights, it is clear to see that decision-makers disagree in the importance of some indicators. In order to achieve an agreed solution that satisfies all the decision-makers, we apply a CMCC model (M-2) to smooth automatically such disagreements and obtain consensual collective weights for each indicator. The consensual weights are shown in the column 'Consensus' both in a numerical and linguistic way, together with the weights assigned to each indicator by the CONGDE, in Tables 13, 14, 15, 16, 17, 18 and 19.

5 Discussion of the Results

This section discusses the resulting weights of each block of indicators obtained in the previous section. Table 13 shows general weights for the six blocks (BG1, BG2, BG3, BG4, BG5 and BG6) of general aspects to assess good governance in the Third Sector and Tables 14, 15, 16, 17, 18 and 19 show the individual decision-makers' weights for each indicator.

---

[2] https://sinbad2.ujaen.es/sites/default/files/2021-06/questionnaire.pdf.

Table 6 Decision-makers' preferences for "general aspects"

General aspects	C_B	C_W	BO	OW	Consistency
e_1	BG3	BG1	8 3 1 5 7 6	1 6 8 4 2 3	0.31
e_2	BG2	BG5	1 1 3 3 7 3	7 7 5 5 1 5	0.27
e_3	BG2	BG1	5 1 2 3 2 4	1 5 3 2 3 2	0.315
e_4	BG6	BG3	3 3 7 5 3 1	5 5 1 3 5 7	0.27
e_5	BG2	BG1	5 1 1 3 4 5	1 5 5 3 2 1	0.27

Table 7 Decision-makers' preferences for "management bodies operation"

BG1	C_B	C_W	BO	OW	Consistency
e_1	BG1.9	BG1.1	8 7 6 3 6 4 5 2 1	1 2 1 4 2 3 1 2 8	0.164
e_2	BG1.5	BG1.6	5 5 3 7 1 7 7 5 1	3 3 5 1 7 1 1 3 7	0.254
e_3	BG1.8	BG1.6	2 2 2 2 3 5 3 1 3	4 4 4 4 2 1 2 5 2	0.175
e_4	BG1.8	BG1.4	7 7 5 9 5 3 3 1 7	3 3 5 1 5 7 7 9 3	0.324
e_5	BG1.5	BG1.9	6 6 6 3 1 4 3 6 7	2 2 2 5 7 4 5 2 1	0.291

Table 8 Decision-makers' preferences for "adequate management of the mission, vision and values"

BG2	C_B	C_W	BO	OW	Consistency
e_1	BG2.7	BG2.3	5 7 9 3 6 4 1 5	2 1 1 6 3 5 9 3	0.24
e_2	BG2.1	BG2.7	1 5 1 5 1 5 7 7	7 3 7 3 7 3 1 1	0.27
e_3	BG2.1	BG2.8	1 2 3 2 3 3 3 7	7 4 3 4 3 3 3 1	0.115
e_4	BG2.8	BG2.3	3 3 9 7 3 5 5 1	7 7 1 3 7 4 4 9	0.24
e_5	BG2.6	BG2.1	7 1 4 4 4 1 3 2	1 7 4 4 4 7 5 6	0.3

Table 9 Decision-makers' preferences for "strategic and operational planning operation"

BG3	C_B	C_W	BO	OW	Consistency
e_1	BG3.8	BG3.9	6 3 5 7 8 4 6 1 9	2 4 2 2 1 3 2 9 1	0.099
e_2	BG3.8	BG3.7	3 3 1 1 3 1 7 1 1	5 5 7 7 5 7 1 7 7	0.257
e_3	BG3.3	BG3.8	3 2 1 2 3 2 3 4 3	2 3 4 3 2 3 2 1 2	0.176
e_4	BG3.8	BG3.4	3 5 3 9 3 7 7 1 5	7 4 7 1 7 3 3 9 4	0.247
e_5	BG3.2	BG3.9	2 1 3 4 4 5 7 4 8	7 8 6 5 5 4 2 3 1	0.298

Table 10 Decision-makers' preferences for "economic management"

BG4	C_B	C_W	BO	OW	Consistency
e_1	BG4.6	BG4.9	2 7 4 3 5 1 6 7 9	8 1 5 8 3 9 2 1 1	0.266
e_2	BG4.5	BG4.8	1 1 1 5 1 3 1 7 7	7 7 7 3 7 5 7 1 1	0.254
e_3	BG4.4	BG4.8	2 3 2 1 5 4 4 7 5	4 3 4 7 3 2 2 1 2	0.254
e_4	BG4.3	BG4.8	7 7 1 5 5 3 3 9 7	3 3 9 4 4 7 7 1 3	0.236
e_5	BG4.3	BG4.5	1 5 1 5 7 4 5 5 3	7 2 7 3 1 2 2 2 5	0.254

5.1 Results from the General Aspects of Good Governance

Appendix A shows the variables definition from the CONGDE (2019) where six blocks of general aspects can be analyzed to assess good governance in NPOs: (BG1) Governing bodies; (BG2) Mission, vision and values; (BG3) Planning and evaluation; (BG4) Economic management; (BG5) Human resources; (BG6) Stakeholders. Regarding the general aspects blocks (BG1, BG2, ...), the CONGDE does not offer any weighting on the importance of each block as a whole, but rather a weighting of each of the indicators individually. Therefore, one of our contributions is presented in Table 13 as it shows the consensual weights (represented numerically (N) and linguistically (L)) that decision-makers consider most

Table 11 Decision-makers' preferences for "human resources management"

BG5	C_B	C_W	BO	OW	Consistency
e_1	BG5.5	BG5.2	3 9 7 2 1 6 5 7 4	5 1 2 7 9 1 1 1 3	0.136
e_2	BG5.4	BG5.8	5 3 1 1 1 3 5 7 3	3 5 7 7 7 5 3 1 5	0.254
e_3	BG5.6	BG5.9	2 2 2 2 2 1 2 3 4	3 3 3 3 3 4 3 2 1	0.176
e_4	BG5.1	BG5.4	1 3 5 9 5 3 7 7 5	9 7 4 1 4 7 3 3 4	0.236
e_5	BG5.5	BG5.8	4 3 2 1 1 3 2 8 4	5 6 7 8 8 6 7 1 5	0.298

Table 12 Decision-makers' preferences for "relationship and communication with stakeholders"

BG6	C_B	C_W	BO	OW	Consistency
e_1	BG6.6	BG6.8	2 7 4 3 5 1 6 9	8 2 3 3 2 9 1 1	0.19
e_2	BG6.2	BG6.1	7 1 3 3 1 5 3 5	1 7 5 5 7 3 5 3	0.27
e_3	BG6.1	BG6.2	1 5 2 4 2 4 3 3	5 1 4 2 4 2 3 3	0.27
e_4	BG6.1	BG6.2	1 9 3 5 5 3 7 3	9 1 7 4 4 7 3 7	0.24
e_5	BG6.7	BG6.4	4 5 5 7 5 3 1 4	4 3 3 1 3 5 7 4	0.3

representative of good governance in NPOs for each of the 6 blocks. Besides, each of the institutions represented by each decision-maker is indicated by "e" (for example, expert 1 = e_1).

According to Table 13, the six general aspects blocks (BG1, BG2, ...) receive the same linguistic term regarding their importance (L) by the decision-makers (AI; average important). One of the aspects is considered to be more representative of good governance as compared with the other five aspects, the block BG2 (numerical weight (N) = 0.296). This block contains eight indicators that measure the adequate management of the mission, vision and values of the organization. Moreover, with the exception of an individual respondent who considers the aspect to have a lower weight (e_4, weight = 0.174), the remaining institutions give this block of general aspect a highest weight, thus confirming the superiority of this block as compared to the others in numeric terms. Second, the blocks BG4 "Economic management" and BG5 "Human resources". These blocks show, both linguistically (L) and numerically (N), the same weighting (AI, -0.162) as shown in the last column of "Consensus". Finally, we must note the lower extent of importance given to the BG6 "Stakeholders" aspects in terms of good governance, where only one individual respondent (e_4; 0.348) shows a very high weighting in this block.

5.2 Results from BG1: Governing Bodies

Table 14 presents the results from each indicator from the block BG1. Following the CONGDE order, we have started with the issues related to the governing body. As the CONGDE offers a possible weighting for each indicator, we will compare these indicative weights (CONGDE column) with the weights that have been validated with the

BWM considering the decision-makers' opinions (Consensus column).

A comparison between the BWM and the CONGDE values shows an overvaluation of several indicators (BG1.1, BG1.2, BG1.6 y BG1.9) by the CONGDE. This fact is most notably observed in two indicators (BG1.2 and BG1.6) by the CONGDE. BG1.2 referring to the proportion of women in the governing body and BG1.6 regarding the government body members that receive remuneration for other positions. CONGDE considers that both are the most important aspect regarding this block (weight: 0.15), while the decision-makers offer a considerably lower consensus value for the valuation of these indicators (consensus weight value: 0.071 and 0.104, respectively). Moreover, in the CONGDE set of indicators, BG1.2 indicator is reinforced by "relevant" as a degree of importance, while BG1.6 is categorized as "inexcusable compliance". These differences proposed by the CONGDE between the degree of importance and not between weighting values leads us to think that a revision of these indicators would be advisable. Regarding linguistics values, both decision-makers opinions (consensus weight) and the CONGDE present the same linguistic term (AI) for all the indicators that compose the block BG1. Finally, we draw our attention to the indicator BG1.5 that represents the fact that 80% of the members, minimum, attends at least 50% of the meetings. We observe here a great difference in the weight proposed by the CONGDE (0.1) and the one given by the decision-makers (consensus value: 0.176). Moreover, the extent of compliance degree suggested by the CONGDE is also lower, they only considered it to be "relevant", rather than a criteria of "inexcusable compliance".

Table 13 Decision-makers' weights for "general aspects' indicator using the fuzzy unbalanced linguistic scale $U_w^{\text{BWM}}(6)$

General aspects	Type	e_1	e_2	e_3	e_4	e_5	Consensus
BG1	N	0.047	0.26	0.064	0.174	0.06	0.174
	L	(U, 0.281)	(AI, 0.336)	(U, 0.383)	(AI, 0.025)	(U, 0.359)	(AI, 0.025)
BG2	N	0.246	0.282	0.347	0.174	0.338	0.296
	L	(AI, 0.285)	(AI, 0.415)	(I, − 0.348)	(AI, 0.025)	(I, − 0.38)	(AI, 0.466)
BG3	N	0.438	0.141	0.186	0.043	0.3	0.188
	L	(I, − 0.022)	(AI, − 0.094)	(AI, 0.069)	(U, 0.257)	(AI, 0.48)	(AI, 0.076)
BG4	N	0.122	0.141	0.119	0.087	0.142	0.122
	L	(AI, − 0.162)	(AI, − 0.094)	(AI, − 0.173)	(AI, − 0.289)	(AI, − 0.09)	(AI, − 0.162)
BG5	N	0.065	0.035	0.186	0.174	0.092	0.122
	L	(U, 0.389)	(U, 0.21)	(AI, 0.069)	(AI, 0.025)	(AI, − 0.271)	(AI, − 0.162)
BG6	N	0.083	0.141	0.098	0.348	0.068	0.098
	L	(U, 0.497)	(AI, − 0.094)	(AI, − 0.249)	(I, − 0.344)	(U, 0.407)	(AI, − 0.249)

Table 14 Decision-makers' weights for "governing bodies" indicator using the fuzzy unbalanced linguistic scale $U_w^{\text{BWM}}(9)$

BG1	Type	e_1	e_2	e_3	e_4	e_5	Consensus	CONGDE
BG1.1	N	0.045	0.071	0.133	0.048	0.055	0.071	0.1
	L	(U, 0.405)	(AI, − 0.135)	(AI, 0.074)	(U, 0.432)	(U, 0.495)	(AI, − 0.135)	(AI, − 0.037)
BG1.2	N	0.055	0.071	0.133	0.048	0.055	0.071	0.15
	L	(U, 0.495)	(AI, − 0.135)	(AI, 0.074)	(U, 0.432)	(U, 0.495)	(AI, − 0.135)	(AI, 0.132)
BG1.3	N	0.057	0.142	0.133	0.091	0.055	0.133	0.1
	L	(AI, − 0.182)	(AI, 0.105)	(AI, 0.074)	(AI, − 0.068)	(U, 0.495)	(AI, 0.074)	(AI, − 0.037)
BG1.4	N	0.149	0.04	0.133	0.028	0.159	0.133	0.1
	L	(AI, 0.128)	(U, 0.36)	(AI, 0.074)	(U, 0.252)	(AI, 0.162)	(AI, 0.074)	(AI, − 0.037)
BG1.5	N	0.058	0.283	0.07	0.091	0.315	0.176	0.1
	L	(AI, − 0.179)	(I, − 0.419)	(AI, − 0.139)	(AI, − 0.068)	(I, − 0.311)	(AI, 0.22)	(AI, − 0.037)
BG1.6	N	0.103	0.035	0.038	0.174	0.11	0.104	0.15
	L	(AI, − 0.027)	(U, 0.315)	(U, 0.342)	(AI, 0.213)	(AI, − 0.003)	(AI, − 0.024)	(AI, 0.132)
BG1.7	N	0.066	0.04	0.079	0.174	0.159	0.113	0.1
	L	(AI, − 0.152)	(U, 0.36)	(AI, − 0.108)	(AI, 0.213)	(AI, 0.162)	(AI, 0.007)	(AI, − 0.037)
BG1.8	N	0.124	0.071	0.206	0.298	0.055	0.124	0.1
	L	(AI, 0.044)	(AI, − 0.135)	(AI, 0.321)	(I, − 0.368)	(U, 0.495)	(AI, 0.044)	(AI, − 0.037)
BG1.9	N	0.343	0.248	0.075	0.048	0.039	0.075	0.1
	L	(I, − 0.216)	(AI, 0.463)	(AI, − 0.122)	(U, 0.432)	(U, 0.351)	(AI, − 0.122)	(AI, − 0.037)

5.3 Results from BG2: Mission, Vision and Values

The second aspect that the CONGDE proposes to represent good governance is the adequate management of the mission, vision and values. Moreover, as we indicated above, this is the most highly weighted block of indicators according to the decision-makers. Table 15 shows the results from the BWM for the weighting of the indicators defined by this block, as well as those proposed by the CONGDE.

According to BG2, we find similarities in the weighting proposed by the CONGDE and the weighting drawn from BWM expert's analysis. This fact is reflected in a similar weight for the review of the mission every time the strategic plan is updated (BG2.4), the values review indicators every 10 years (BG2.6), and the definition and review of the mission, vision and values refer to the Coordinator's Code of Conduct and the Third Strategic Plan Social Action Sector (BG2.7). Regarding the degree of importance, the three indicators are proposed as

Table 15 Decision-makers' weights for "mission, vision and values" indicator using the fuzzy unbalanced linguistic scale $U_w^{\mathrm{BWM}}(8)$

BG2	Type	e_1	e_2	e_3	e_4	e_5	Consensus	CONGDE
BG2.1	N	0.073	0.266	0.271	0.16	0.029	0.266	0.15
	L	(AI, − 0.178)	(AI, 0.483)	(I, − 0.502)	(AI, 0.12)	(U, 0.232)	(AI, 0.483)	(AI, 0.086)
BG2.2	N	0.051	0.066	0.145	0.16	0.197	0.145	0.1
	L	(U, 0.408)	(AI, − 0.202)	(AI, 0.068)	(AI, 0.12)	(AI, 0.247)	(AI, 0.068)	(AI, − 0.086)
BG2.3	N	0.036	0.232	0.1	0.028	0.084	0.1	0.15
	L	(U, 0.288)	(AI, 0.366)	(AI, − 0.086)	(U, 0.224)	(AI, − 0.14)	(AI, − 0.086)	(AI, 0.086)
BG2.4	N	0.195	0.066	0.145	0.049	0.084	0.11	0.1
	L	(AI, 0.24)	(AI, − 0.202)	(AI, 0.068)	(U, 0.392)	(AI, − 0.14)	(AI, − 0.051)	(AI, − 0.086)
BG2.5	N	0.066	0.232	0.1	0.162	0.084	0.1	0.15
	L	(AI, − 0.202)	(AI, 0.366)	(AI, − 0.086)	(AI, 0.127)	(AI, − 0.14)	(AI, − 0.086)	(AI, 0.086)
BG2.6	N	0.134	0.066	0.1	0.077	0.241	0.1	0.1
	L	(AI, 0.031)	(AI, − 0.202)	(AI, − 0.086)	(AI, − 0.164)	(AI, 0.397)	(AI, − 0.086)	(AI, − 0.086)
BG2.7	N	0.366	0.033	0.1	0.077	0.117	0.1	0.1
	L	(I, − 0.175)	(U, 0.264)	(AI, − 0.086)	(AI, − 0.164)	(AI, − 0.027)	(AI, − 0.086)	(AI, − 0.086)
BG2.8	N	0.079	0.038	0.039	0.286	0.164	0.079	0.15
	L	(AI, − 0.158)	(U, 0.304)	(U, 0.312)	(I, − 0.45)	(AI, 0.134)	(AI, − 0.158)	(AI, 0.086)

Table 16 Decision-makers' weights for "planning and evaluation" indicator using the fuzzy unbalanced linguistic scale $U_w^{\mathrm{BWM}}(9)$

BG3	Type	e_1	e_2	e_3	e_4	e_5	Consensus	CONGDE
BG3.1	N	0.066	0.078	0.083	0.152	0.187	0.128	0.1
	L	(AI, − 0.152)	(AI, − 0.111)	(AI, − 0.095)	(AI, 0.139)	(AI, 0.257)	(AI, 0.057)	(AI, − 0.037)
BG3.2	N	0.148	0.078	0.133	0.072	0.26	0.148	0.1
	L	(AI, 0.125)	(AI, − 0.111)	(AI, 0.074)	(AI, − 0.132)	(I, − 0.497)	(AI, 0.125)	(AI, − 0.037)
BG3.3	N	0.078	0.147	0.219	0.152	0.142	0.149	0.05
	L	(AI, − 0.111)	(AI, 0.122)	(AI, 0.365)	(AI, 0.139)	(AI, 0.105)	(AI, 0.128)	(U, 0.45)
BG3.4	N	0.06	0.147	0.133	0.027	0.1	0.1	0.1
	L	(AI, − 0.172)	(AI, 0.122)	(AI, 0.074)	(U, 0.243)	(AI, − 0.037)	(AI, − 0.037)	(AI, − 0.037)
BG3.5	N	0.048	0.078	0.083	0.159	0.1	0.093	0.15
	L	(U, 0.432)	(AI, − 0.111)	(AI, − 0.095)	(AI, 0.162)	(AI, − 0.037)	(AI, − 0.061)	(AI, 0.132)
BG3.6	N	0.106	0.147	0.133	0.047	0.072	0.106	0.15
	L	(AI, − 0.017)	(AI, 0.122)	(AI, 0.074)	(U, 0.423)	(AI, − 0.132)	(AI, − 0.017)	(AI, 0.132)
BG3.7	N	0.066	0.02	0.083	0.047	0.039	0.047	0.1
	L	(AI, − 0.152)	(U, 0.18)	(AI, − 0.095)	(U, 0.423)	(U, 0.351)	(U, 0.423)	(AI, − 0.037)
BG3.8	N	0.387	0.156	0.05	0.273	0.072	0.156	0.1
	L	(I, − 0.068)	(AI, 0.152)	(U, 0.45)	(I, − 0.453)	(AI, − 0.132)	(AI, 0.152)	(AI, − 0.037)
BG3.9	N	0.041	0.148	0.083	0.073	0.028	0.073	0.15
	L	(U, 0.369)	(AI, 0.125)	(AI, − 0.095)	(AI, − 0.128)	(U, 0.252)	(AI, − 0.128)	(AI, 0.132)

"relevant" by the CONGDE. One more similarity is that the consensus value of the decision-makers and the CONGDE present the same linguistic value (AI). Regarding notable differences, it should be noted how the CONGDE values BG2.1, BG2.3, BG2.5, BG2.8 as the most important indicators (0.15) within the block. However, the opinion of decision-makers manifests that only BG2.1, about the mission formulation approved by the governing body, obtains the maximum weighting (0.266) of this block, in an outstanding way compared to the rest of the indicators. Additionally, the BG2.8 indicator regarding a Code of Conduct on abuse, harassment, fraud and illegal

Table 17 Decision-makers' weights for "economic management" indicator using the fuzzy unbalanced linguistic scale $U_w^{\mathrm{BWM}}(9)$

BG4	Type	e_1	e_2	e_3	e_4	e_5	Consensus	CONGDE
BG4.1	N	0.213	0.159	0.145	0.052	0.24	0.168	0.15
	L	(AI, 0.345)	(AI, 0.162)	(AI, 0.115)	(U, 0.468)	(AI, 0.436)	(AI, 0.193)	(AI, 0.132)
BG4.2	N	0.035	0.159	0.103	0.052	0.057	0.103	0.1
	L	(U, 0.315)	(AI, 0.162)	(AI, − 0.027)	(U, 0.468)	(AI, − 0.182)	(AI, − 0.027)	(AI, − 0.037)
BG4.3	N	0.105	0.159	0.145	0.305	0.276	0.264	0.15
	L	(AI, − 0.02)	(AI, 0.162)	(AI, 0.115)	(I, − 0.345)	(I, − 0.443)	(I, − 0.483)	(AI, 0.132)
BG4.4	N	0.182	0.045	0.291	0.081	0.069	0.113	0.15
	L	(AI, 0.24)	(U, 0.405)	(I, − 0.392)	(AI, − 0.101)	(AI, − 0.142)	(AI, 0.007)	(AI, 0.132)
BG4.5	N	0.062	0.181	0.073	0.081	0.035	0.081	0.1
	L	(AI, − 0.166)	(AI, 0.236)	(AI, − 0.128)	(AI, − 0.101)	(U, 0.315)	(AI, − 0.101)	(AI, − 0.037)
BG4.6	N	0.288	0.091	0.074	0.173	0.07	0.091	0.1
	L	(I, − 0.402)	(AI, − 0.068)	(AI, − 0.125)	(AI, 0.209)	(AI, − 0.139)	(AI, − 0.068)	(AI, − 0.037)
BG4.7	N	0.048	0.159	0.074	0.173	0.057	0.091	0.15
	L	(U, 0.432)	(AI, 0.162)	(AI, − 0.125)	(AI, 0.209)	(AI, − 0.182)	(AI, − 0.068)	(AI, 0.132)
BG4.8	N	0.04	0.023	0.036	0.03	0.057	0.036	0.05
	L	(U, 0.36)	(U, 0.207)	(U, 0.324)	(U, 0.27)	(AI, − 0.182)	(U, 0.324)	(U, 0.45)
BG4.9	N	0.028	0.026	0.059	0.052	0.138	0.052	0.05
	L	(U, 0.252)	(U, 0.234)	(AI, − 0.176)	(U, 0.468)	(AI, 0.091)	(U, 0.468)	(U, 0.45)

Table 18 Decision-makers' weights for "human resources" indicator using the fuzzy unbalanced linguistic scale $U_w^{\mathrm{BWM}}(9)$

BG5	Type	e_1	e_2	e_3	e_4	e_5	Consensus	CONGDE
BG5.1	N	0.14	0.051	0.116	0.295	0.073	0.132	0.15
	L	(AI, 0.098)	(U, 0.459)	(AI, 0.017)	(I, − 0.378)	(AI, − 0.128)	(AI, 0.071)	(AI, 0.132)
BG5.2	N	0.033	0.103	0.116	0.167	0.104	0.104	0.1
	L	(U, 0.297)	(AI, − 0.027)	(AI, 0.017)	(AI, 0.189)	(AI, − 0.024)	(AI, − 0.024)	(AI, − 0.037)
BG5.3	N	0.047	0.179	0.116	0.079	0.138	0.116	0.1
	L	(U, 0.423)	(AI, 0.23)	(AI, 0.017)	(AI, − 0.108)	(AI, 0.091)	(AI, 0.017)	(AI, − 0.037)
BG5.4	N	0.222	0.205	0.116	0.029	0.161	0.161	0.1
	L	(AI, 0.375)	(AI, 0.318)	(AI, 0.017)	(U, 0.261)	(AI, 0.169)	(AI, 0.169)	(AI, − 0.037)
BG5.5	N	0.319	0.179	0.116	0.079	0.19	0.179	0.1
	L	(I, − 0.297)	(AI, 0.23)	(AI, 0.017)	(AI, − 0.108)	(AI, 0.267)	(AI, 0.23)	(AI, − 0.037)
BG5.6	N	0.051	0.103	0.19	0.169	0.104	0.104	0.05
	L	(U, 0.459)	(AI, − 0.027)	(AI, 0.267)	(AI, 0.196)	(AI, − 0.024)	(AI, − 0.024)	(U, 0.45)
BG5.7	N	0.056	0.051	0.116	0.051	0.138	0.08	0.15
	L	(AI, − 0.186)	(U, 0.459)	(AI, 0.017)	(U, 0.459)	(AI, 0.091)	(AI, − 0.105)	(AI, 0.132)
BG5.8	N	0.045	0.026	0.072	0.051	0.02	0.045	0.15
	L	(U, 0.405)	(U, 0.234)	(AI, − 0.132)	(U, 0.459)	(U, 0.18)	(U, 0.405)	(AI, 0.132)
BG5.9	N	0.087	0.103	0.044	0.079	0.073	0.079	0.1
	L	(AI, − 0.081)	(AI, − 0.027)	(U, 0.396)	(AI, − 0.108)	(AI, − 0.128)	(AI, − 0.108)	(AI, − 0.037)

conduct, which is considered one of the most important indicators by the CONGDE, is underestimated by decision-makers (consensus weight: 0.079), being the indicator that presents the lowest consensus value within the BG2 block.

5.4 Results from BG3: Planning and Evaluation

This block of indicators is considered, as we pointed out above, the second most important block according to the decision-makers. Once again, it is a block that presents

Table 19 Decision-makers' weights for "stakeholders" indicator using the fuzzy unbalanced linguistic scale $U_w^{\text{BWM}}(8)$

BG6	Type	e_1	e_2	e_3	e_4	e_5	Consensus	CONGDE
BG6.1	N	0.255	0.032	0.259	0.284	0.115	0.255	0.1
	L	(AI, 0.445)	(U, 0.256)	(AI, 0.459)	(I, − 0.457)	(AI, − 0.034)	(AI, 0.445)	(AI, − 0.086)
BG6.2	N	0.052	0.257	0.046	0.028	0.079	0.052	0.1
	L	(U, 0.416)	(AI, 0.452)	(U, 0.368)	(U, 0.224)	(AI, − 0.158)	(U, 0.416)	(AI, − 0.086)
BG6.3	N	0.091	0.128	0.169	0.163	0.079	0.128	0.15
	L	(AI, − 0.116)	(AI, 0.01)	(AI, 0.151)	(AI, 0.13)	(AI, − 0.158)	(AI, 0.01)	(AI, 0.086)
BG6.4	N	0.108	0.128	0.07	0.076	0.04	0.083	0.15
	L	(AI, − 0.058)	(AI, 0.01)	(AI, − 0.188)	(AI, − 0.168)	(U, 0.32)	(AI, − 0.144)	(AI, 0.086)
BG6.5	N	0.068	0.198	0.169	0.076	0.08	0.08	0.1
	L	(AI, − 0.195)	(AI, 0.25)	(AI, 0.151)	(AI, − 0.168)	(AI, − 0.154)	(AI, − 0.154)	(AI, − 0.086)
BG6.6	N	0.337	0.064	0.07	0.163	0.165	0.165	0.15
	L	(I, − 0.275)	(AI, − 0.209)	(AI, − 0.188)	(AI, 0.13)	(AI, 0.137)	(AI, 0.137)	(AI, 0.086)
BG6.7	N	0.055	0.128	0.109	0.049	0.327	0.128	0.1
	L	(U, 0.44)	(AI, 0.01)	(AI, − 0.055)	(U, 0.392)	(I, − 0.309)	(AI, 0.01)	(AI, − 0.086)
BG6.8	N	0.035	0.064	0.109	0.163	0.115	0.109	0.15
	L	(U, 0.28)	(AI, − 0.209)	(AI, − 0.055)	(AI, 0.13)	(AI, − 0.034)	(AI, − 0.055)	(AI, 0.086)

Table 20 Main features of the analyzed proposal

Method	Groups	Consensus	Linguistic
Rezaei [19]	✗	✗	✗
Labella et al. [23]	✔	✗	✔
Safarzadeh et al. [61]	✔	✗	✗
Proposal	✔	✔	✔

many differences between the weightings of the decision-makers and those proposed by the CONGDE. A detailed comparison is shown in Table 16.

According to the decision-makers, differences in practically all the indicators are observed in this block BG3. Moreover, this block has the greatest number of indicators for which the values proposed by the CONGDE differs more in magnitude with the decision-makers' opinions. Firstly, one of the indicator for which the CONGDE gives a higher weight (0.15) and degree of importance (inexcusable compliance) is the one referred to the fact that the strategic planning is specified in periodic operational schedules approved by the governing body (BG3.5). However, decision-makers do not consider this fact so important (consensus value: 0.093). Furthermore, decision-makers consider this indicator with one of the lowest weights. Another indicator for which the CONGDE gives a higher weight (0.1) and a higher degree of importance (inexcusable compliance) is the indicator BG3.7: Governing body monitors and evaluates operational schedules.

Nevertheless, decision-makers do not consider this fact so important (consensus value: 0.047). Furthermore, decision-makers consider this indicator as the lowest weight for the whole block BG3. In this sense, the indicator that the decision-makers consider to be most important (consensus value 0.156), is the public availability of a document that reflects the policy, system or procedure for monitoring and evaluating the NPO's own activity projects and programs, directly linked to the mission's fulfillment (BG3.8). This indicator is given a lower value by the CONGDE (weight: 0.1) and degree of importance (relevant). Lastly, the most valued indicator by the decision-makers in contrast to the CONGDE are the following: BG3.2 (consensus value: 0.148; CONGDE: 0.1) and BG3.3. (Consensus value: 0.149; CONGDE: 0.05). It is very remarkable how this last indicator as the strategic planning explicitly includes references to the mission, vision and values and ethical recommendations of the Third Sector, obtains the lowest weighting of the block by the CONGDE. Against this, it gets the second highest weighting of the BG3 block by the

decision-makers. Regarding linguistic values, we highlight two differences. First, BG3.3 (Consensus value: AI (average important); CONGDE: U (Unimportant)) and second, BG3.7 (Consensus value: U (Unimportant)); CONGDE: AI (average important)). The consensus value of the decision-makers and the CONGDE present the same linguistic value (AI) for the remaining indicators.

5.5 Results from BG4: Economic Management

The forth aspect that the CONGDE proposes to represent good governance is about the economic management as the block BG4. Table 17 shows the results from the BWM for the weighting of the indicators defined by this block, as well as those proposed by the CONGDE.

Most of the indicators reflecting a higher extent of good governance in terms of the economic management by the NPO show similar weights between decision-makers' opinions and those proposed by CONGDE. Both decision-makers and the CONGDE agree that the most important aspects are BG4.1, about the public availability of the annual income and expense budget approved by the governing body (consensus value: 0.168; CONGDE: 0.15); and BG4.3, referred to the public availability of the annual budget settlement executed, reviewed and approved by the governing body (consensus value: 0.264; CONGDE: 0.15). This is also consistent with the higher degree of importance (inexcusable compliance) of these indicators. However, despite the fact that the CONGDE considers BG4.3 to have the highest weight (CONGDE: 0.15), it would be more logical to revise this magnitude and increase the weight for this indicator, approaching it to the decision-makers opinions (consensus value: 0.264). In short, this indicator should still be considered to be an "inexcusable compliance" one, increasing its weighting by the CONGDE, in consonance with the opinion of the decision-makers. Two of the indicators that are considered to be "relevant" in terms of representing good governance in the field of NPO's economic management present similar importance according to the decision-makers and the CONGDE and besides, both indicators have the lowest weighting for both proposals. Finally, two aspects are given considerably similar importance by the decision-makers and the CON-GDE; BG4.8, regarding no financier contributes more than 50% of total income for the year (Consensus value: 0.036; CONGDE: 0.05), and BG4.9, about the fact that the NPO does not accumulate liquid assets or financial assets in the previous audited year greater than the expense of the current year (consensus value: 0.052; CONGDE: 0.05). In linguistic terms, all the indicators form BG4 present the same values according to the decision-makers and the CONGDE.

5.6 Results from BG5: Human Resources

Table 18 presents the results from each indicator from the block BG5 about the human resources management to represent good governance in the Third Sector.

Firstly, two of the indicators for which the decision-makers give the higher weights are BG5.4, based on a list of organization profiles and the continuous development of its operational team (consensus value: 0.161), and BG5.5, regarding NPO that promotes training and continuous development of the operational team (consensus value: 0.179). On the contrary, the CONGDE gives a weight valued at 0.1 and a degree of importance of "relevant". In contrast, other indicators for which the CONGDE gives higher weights (0.15) but not a higher degree of importance (relevant) are the indicators BG5.7, about the volunteer plan that includes minimum objectives and activities (consensus value: 0.08), and BG5.8, about the percentage of women who are part of the responsible executive structure (consensus value: 0.045). In this sense, the decision-makers do not consider these facts so important and thus, they consider these indicators as two of the lowest weights for the whole block BG5. Our results suggest the need to decrease the importance given to the weights of these indicators. In addition, we highlight the indicator BG5.6, related to the agreement model for volunteering. On the one hand, the CONGDE gives it the lowest value of the entire block (0.05). On the other hand, CONGDE considers it "inexcusable compliance" in terms of the degree of importance. We also highlight how the decision-makers give it a value of something more than double (consensus value: 0.104) than that proposed by CONGDE. Finally, regarding linguistic values, we noticed two differences. First, BG5.6 (consensus value: AI (average important); CONGDE: U (Unimportant)) and second, BG5.8 (consensus value: U (Unimportant); CONGDE: AI (average important)). For the remaining indicators, the consensus value of the decision-makers and the CONGDE present the same linguistic value (AI; average important).

5.7 Results from BG6: Stakeholders

Lastly, Table 19 presents the results from each indicator from the block BG6 regarding the relationship and communication with stakeholders to measure the levels of good governance in the Third Sector.

This block BG6 presents several differences on the values proposed by the CONGDE and by the decision-makers' opinions. The first one, the indicator BG6.1, for which the decision-makers give the highest weight (consensus value: 0.255), is referred to the fact that the partnership policy approved by the governing body defines relationships with entities with which it carries out its

projects. However, the CONGDE does not consider this fact so important (0.1) and its degree of importance is categorized only as "relevant". Therefore, BG6.1 presents a weight of more than double by the decision-makers. The opposite occurs with the indicator BG6.2 about the collaboration agreement model to be signed with local partners and/or local executing entities that contains purpose, rights and obligations, and duration. This indicator is undervalued by the decision-makers (consensus value: 0.052). However, the CONGDE provides a weighting of almost double (0.1). In addition, the degree of importance is of "inexcusable compliance", which differs from the opinion of the decision-makers. Finally, this block shows linguistically the same value (AI, average important) for all the indicators except for BG6.2 (consensus value: U; CONGDE: AI).

As a final comment on the results of the 6 blocks of indicators, there are many differences that are underscored in the importance of most of the indicators proposed by CONGDE to measure desirable levels of good governance in an NPO. As a general conclusion, we have contributed with this analysis to the research field by proposing an optimal weight for the importance of those aspects as indicative of the extent of good governance measurement in the Third Sector. Considering all the above and observing the discrepancies found between the CONGDE and the decision-makers in many of the indicators, we consider it is necessary to review the importance that the CONGDE proposed for each indicator in comparison with the consensus value that the decision-makers in the non-profit sector grant.

6 Comparison with Existing Studies

There have been some studies focusing on weighting attributes approaches [19, 23, 61] with different features. These approaches will be analyzed and compared with our proposal regarding 3 main aspects related to the resolution of MCGDM problems: the management of experts groups, the processing of consensual solutions and the use of linguistic information to model opinions and results. Table 20 shows a summary of these characteristics for the studied proposals.

1. Rezaei introduced for the first time the BWM in [19]. This new approach was proposed to solve some behavioral errors in similar MCDM methods and reduces the number of pairwise comparisons and inconsistency in experts' preferences elicitation. Although our proposal is mostly based on this one, it does not consider the resolution of MCGDM problems where several experts are involved in the

decision process and thus, the disagreements among them are omitted. Additionally, the classical approach also does not consider the use of linguistic information to model experts' opinions and results.

2. Labella et al. [23] introduced the 2-tuple BWM, which is the BWM approach in which our proposal is based on. Nevertheless, this proposal uses a symmetrically distributed linguistic terms set to model the linguistic results, which is not convenient to represent properly the criteria importance (see Example 1). In addition, although this approach faces the resolution of MCGDM problems, it ignores the consensus among experts.

3. Safarzadeh et al. [61] proposed a MCGDM approach based on the BWM. Again, this proposal is able to deal with group of experts but does not apply any consensus approach to smooth the possible disagreements among them. Furthermore, this proposal models experts' preferences using the classical numerical scale used in BWM and does not provide a linguistic representation of the information.

7 Conclusions

Non-profit good governance has become a necessary mechanism for the supervision of NPOs as it would allow NPOs to better achieve their goals through the most effective, transparent and objective use of their resources, providing an adequate ethical content for their stakeholders. Therefore, the main objective of a good governance mechanism should be the increase and improvement of the accountability practices of NPOs. Several institutions have provided different list of indicators to measure the levels of non-profit good governance that are either not weighted or are weighted without a theoretical or empirical basis. In this paper, it has been proposed the use of a MCGDM approach for modeling the problem of weighting the list of indicators in a consensual way. Our proposal takes as basis the 2-tuple BWM that allows deriving the indicators weights and providing a linguistic representation of them to facilitate their understanding. Moreover, we have extended the 2-tuple BWM to deal with fuzzy unbalanced linguistic scales that represent better the importance of the indicators. Afterwards, we have applied a consensus process based on CMCC models to remove possible disagreements in the indicators weights provided by the decision-makers. To show the validness of our proposal and contribute to improve the measurement of good governance in NPOs, we apply our MCDGM approach to the weighting of several indicators proposed by the CONGDE. Through this

research and analysis of the results obtained, we have found significant differences between the weights of the indicators proposed by CONGDE and those obtained from the decision-makers. As previously stated, the indicators proposed by CONGDE lack empirical validation by decision-makers. In general, our results show that the block of indicators that are considered to be the most important when measuring good governance in NPOs are those from BG2 "Mission, vision and values" (consensus value = 0.296). Aspects such as defining the meaning, focus and reason to exist of the NPO in the long term. On the contrary, those indicators that measure the relationship of the NPO with its stakeholders (BG6) are the least valued (consensus value: 0.096). Analyzing such differences by blocks, the greater divergence comes from indicators BG1.2 (consensus value by decision-makers is 0.079 lower than the CONGDE value), BG2.1 (0.112 greater), BG3.3 (0.099 greater), BG4.3 (0.114 greater), BG5.8 (0.105 lower) and BG6.1 (0.155 greater). The indicators with the highest value in each block are BG1.5 that represents the fact that "80% of the members, minimum, attends at least 50% of the meetings" (consensus value: 0.176); BG2.1 "Mission formulation approved by the governing body" (consensus value: 0.266); BG3.8 "Document that reflects the policy, system or procedure for monitoring and evaluating the organization's own activity projects and programs, directly linked to the mission's fulfillment" (consensus value: 0.156); BG4.3 "Annual budget settlement executed, reviewed and approved by the governing body" (consensus value: 0.264); BG5.5 "Promotes training and continuous development of the operational team" (consensus value: 0.179); and BG6.1 "Partnership policy approved by the governing body that defines relationships with entities with which it carries out its projects" (consensus value: 0.255).

Bearing in mind the previous paragraph, the main findings of this proposal can be summarized as follows:

- The 2-tuple BWM has been significantly improved. The classical proposal uses a symmetrically distributed fuzzy linguistic term set to represent linguistically the experts' weights. However, both the syntax and semantic of such a linguistic term set is not suitable to represent appropriately the importance of the criteria. We have defined a fuzzy unbalanced linguistic terms set based on the number of elements to compare whose syntax and semantic allow to obtain more comprehensive linguistic results.

- The BWM based-approaches do not consider disagreement among the experts. For this reason, we have incorporated a consensus approach based on a CMCC model to derive the collective consensual criteria weights obtained from the 2-tuple BWM.

- We have applied our proposal to a real case study focuses on measuring the good governance in NPO. The analysis of the results has allowed us to build a battery of indicators that is appropriately validated by the opinion of NPO accountability decision-makers, giving an optimal weight to each indicator. Therefore, our study contributes not only to previous literature on good governance of accountability for NPOs by proposing a rigorously mechanism based on a set of indicators, but it also contributes to regulators and professionals. This mechanism is supported by the need to adhere to the standards of ethics and honesty that have been consolidated in the non-profit sector as relevant tools to cultivate the image of the NPOs.

Appendix A: Description of the Indicators

The indicators to be weighted proposed by the CONGDE for measuring good practices in NGO are described below:

- BG1: Governing bodies

- BG1.1: Minimum of 5 members in the governing body.
- BG1.2: The proportion of women in the governing body is equal to or greater than 40%.
- BG1.3: The maximum proportion of members of the governing body who maintain a kinship relationship up to the second degree or who are de facto couples is 20%.
- BG1.4: Minimum of 2 meetings a year.
- BG1.5: 80% of the members of the governing body attend at least 50% of the meetings.
- BG1.6: No more than 40% of members of the government body receive remuneration for other positions.
- BG1.7: Election and replacement of positions in associations, maximum every 5 years.
- BG1.8: Minimum 60% of the members of the governing body have a continuous cumulative maximum term of < 8 years.

– BG1.9: No member of the governing body has a lifetime allowance.

- BG2: Mission, vision and values

– BG2.1: Mission formulation approved by the governing body.
– BG2.2: Mission will be reviewed at least every 10 years.
– BG2.3: Vision formulation approved by the governing body.
– BG2.4: Review of the mission every time the strategic plan is updated.
– BG2.5: Values formulation approved by the governing body.
– BG2.6: Values are reviewed at least every 10 years.
– BG2.7: The definition / revision of the Mission, Vision, Values refer to the Code of Conduct of the Coordinator and / or the III Strategic Plan of the Third Sector of Social Action.
– BG2.8: There is a own Code of Conduct on abuse, harassment, fraud and illegal conduct.

- BG3: Planning and evaluation

– BG3.1: Long term strategic planning approved by the governing body (maximum 10 years).
– BG3.2: The development of strategic planning ensures that it is defined in a participatory way.
– BG3.3: Strategic planning explicitly includes references to the mission, vision and values and ethical recommendations of the Third Sector.
– BG3.4: Strategic planning includes an estimate of the private and public income to be achieved annually during the period of application of the plan.
– BG3.5: Strategic planning is specified in periodic operational schedules approved by the governing body.
– BG3.6: Governing body monitors and evaluates strategic planning.
– BG3.7: Governing body monitors and evaluates operational schedules.
– BG3.8: Document that reflects the policy, system or procedure for monitoring and evaluating the organization's own activity projects and programs, directly linked to the mission's fulfillment.
– BG3.9: The Assembly or Board of Trustees is informed of the organization of the number,

typology and measures adopted for cases analyzed by the monitoring body of the complaints channel.

- BG4: Economic management

– BG4.1: Annual income and expense budget approved by the governing body.
– BG4.2: Government body carries out a monitoring of the minimum budget every 6 months.
– BG4.3: Annual budget settlement executed, reviewed and approved by the governing body.
– BG4.4: An external audit of the annual accounts is carried out if the volume of income exceeds 300,000.
– BG4.5: Written financial investment policy approved by the governing body.
– BG4.6: Long term strategic planning and monitoring explicitly include main criteria for obtaining private and public income.
– BG4.7: Institutional document with traceability criteria for finalist funds received.
– BG4.8: No financier contributes more than 50% of total income for the year.
– BG4.9: It does not accumulate liquid assets or financial assets in the previous audited year greater than the expense of the current year.

- BG5: Human resources

– BG5.1: Policy approved by the governing body of compensation and social benefits criteria.
– BG5.2: Personnel selection and hiring policy approved by the governing body.
– BG5.3: Human resources policies (remunerated and voluntary) approved by the government body make mention of discrimination.
– BG5.4: There is a list of organization profiles and the continuous development of its operational team.
– BG5.5: Promotes training and continuous development of the operational team.
– BG5.6: Incorporation agreement model for volunteering that specifies rights and obligations of both parties, functions, activities, time of dedication and insurance for volunteering.
– BG5.7: Volunteer Plan that includes minimum objectives and activities.
– BG5.8: The percentage of women who are part of the responsible executive structure is equal to or greater than 40%.

– BG5.9: There is a gender policy approved by the governing body in the organization.

- BG6: Stakeholders

– BG6.1: Partnership policy approved by the governing body that defines relationships with entities with which it carries out its projects (commitment to comply with legal regulations).
– BG6.2: Collaboration agreement model to be signed with local partners and/or local executing entities that contains purpose, rights and obligations and duration.
– BG6.3: There is an institutional environmental management policy approved by the governing body.
– BG6.4: Purchase policy approved by the governing body.
– BG6.5: Declaration of principles or collaboration with companies criteria document approved by the governing body that refers to guiding principles on human rights.
– BG6.6: The organization has actively participated in the NGDO State Coordinator, Third Sector Platform and Social Action NGO Platform during the last year.
– BG6.7: Criteria and procedure that develop organizational commitment to respond to complaints, inquiries and requests for information received.
– BG6.8: Own and public complaints channel, regulated procedure and body that ensures its monitoring.

References

1. Salamon, L.. M., Sokolowski, S.. W.: Beyond nonprofits: reconceptualizing the third sector. VOLUNTAS: Int. J. Volunt. Nonprofit Organ. **27**(4), 1515–1545 (2016)
2. Teasdale, S.: What's in a name? making sense of social enterprise discourses. Public Policy Adm. **27**(2), 99–119 (2012)
3. Cornforth, C.: Nonprofit governance research: limitations of the focus on boards and suggestions for new directions. Nonprofit Volunt. Sect. Q. **41**(6), 1116–1135 (2012)
4. Schatteman, A.: Nonprofit accountability: to whom and for what? an introduction to the special issue. Int. Rev. Public Adm. **18**(3), 1–6 (2013)
5. Vakil, A.C.: Confronting the classification problem: toward a taxonomy of NGOs. World Dev. **25**(12), 2057–2070 (1997)
6. Wango: Code of ethics & conduct for NGOs, Tech. rep. (2004)
7. Bies, A.L.: Evolution of nonprofit self-regulation in Europe. Nonprofit Volunt. Sect. Q. **39**(6), 1057–1086 (2010)
8. Verbruggen, S., Christiaens, J., Milis, K.: Can resource dependence and coercive isomorphism explain nonprofit organizations' compliance with reporting standards? Nonprofit Volunt. Sect. Q. **40**(1), 5–32 (2011)
9. Benjamin, L.M.: Account space: how accountability requirements shape nonprofit practice. Nonprofit Volunt. Sect. Q. **37**(2), 201–223 (2008)
10. Benjamin, L.M.: The potential of outcome measurement for strengthening nonprofits' accountability to beneficiaries. Nonprofit Volunt. Sect. Q. **42**(6), 1224–1244 (2013)
11. Costa, E., Goulart da Silva, G.: Nonprofit accountability: the viewpoint of the primary stakeholders. Financ. Account. Manage. **35**(1), 37–54 (2019)
12. Boomsma, R., O'Dwyer, B.: Constituting the governable NGO: the correlation between conduct and counter-conduct in the evolution of funder-ngo accountability relations. Account. Organ. Soc. **72**, 1–20 (2019)
13. Byers, T., Anagnostopoulos, C., Brooke-Holmes, G.: Understanding control in nonprofit organisations: moving governance research forward? Corp. Gov. **15**(1), 134–145 (2015)
14. Speckbacher, G.: Nonprofit versus corporate governance: an economic approach. Nonprofit Manage. Leadersh. **18**(3), 295–320 (2008)
15. Moreno-Albarracin, A.L., Liceran-Gutierrez, A., Ortega-Rodriguez, C., Labella, A., Rodriguez, R.M.: Measuring what is not seen-transparency and good governance nonprofit indicators to overcome the limitations of accounting models. Sustainability **12**(18), 7275 (2020)
16. Hourali, S., Zahedi, M., Fateh, M.: Coreference resolution using neural mcdm and fuzzy weighting technique. Int. J. Comput. Intell. Syst. **13**, 56–65 (2020)
17. Mi, X., Liao, H.: An integrated approach to multiple criteria decision making based on the average solution and normalized weights of criteria deduced by the hesitant fuzzy best worst method. Comput. Ind. Eng. **133**, 83–94 (2019)
18. Zhang, Q., Chen, J., Chong, P.: Decision consolidation: criteria weight determination using multiple preference formats. Decis. Support Syst. **38**(2), 247–258 (2004)

19. Rezaei, J.: Best–Worst multi-criteria decision-making method. Omega **53**, 49–57 (2015)

20. Eklund, P., Rusinowska, A., de Swart, H.: Consensus reaching in committees. Decis. Support **178**(1), 185–193 (2007)

21. Labella, Á., Liu, Y., Rodríguez, R.M., Martínez, L.: Analyzing the performance of classical consensus models in large scale group decision making: a comparative study. Appl. Soft Comput. **67**, 677–690 (2018)

22. Rodríguez, R.M., Labella, Á., Tré, G.D., Martínez, L.: A large scale consensus reaching process managing group hesitation. Knowl-Based Syst. **159**, 86–97 (2018)

23. Labella, Á., Dutta, B., Martínez, L.: An optimal Best–Worst prioritization method under a 2-tuple linguistic environment in decision making. Comput. Ind. Eng. **155**, 1071411 (2021)

24. Labella, Á., Liu, H., Rodríguez, R.M., Martínez, L.: A cost consensus metric for consensus reaching processes based on a comprehensive minimum cost model. Eur. J. Oper. Res. **281**(2), 316–331 (2020)

25. Martínez, L., Rodríguez, R. M., Herrera, F.: The 2-tuple linguistic model. Springer, Cham (2015)

26. Herrera, F., Herrera-Viedma, E., Martínez, L.: A fuzzy linguistic methodology to deal with unbalanced linguistic term sets. IEEE Tran. Fuzzy Syst. **16**(2), 354–370 (2008)

27. CONGDE, Herramienta de transparencia y buen gobierno, Tech. rep., Plataforma de ONG de Acción Social y Coordinadora de ONGD-España (2019)

28. de España, G.: Law 5/2011, of March 29, on social economy, Tech. rep. (2011)

29. Zadeh, L.A.: The concept of a linguistic variable and its application to approximate reasoning, part i, ii, iii. Inf. Sci. (1975). https://doi.org/10.1016/0020-0255(75)90017-1

30. Yager, R.R.: Approximate reasoning as a basis for computing with words. In: Computing with words and information/intelligent systems 2: applications, pp. 50–77. Physica Verlag, Heidelberg (1999)

31. Tai, W., Chen, C.: A new evaluation model for intellectual capital based on computing with linguistic variable. Expert Syst. Appl. **36**(2), 3483–3488 (2009)

32. Cao, M., Wu, J., Chiclana, F., Ureña, R., Herrera-Viedma, E.: A personalized consensus feedback mechanism based on maximum harmony degree. IEEE Trans. Syst. Man Cybern.: Syst. (2021). https://doi.org/10.1109/TSMC.2019.2960052

33. Gu, Y., Hao, T., Cheng, D., Wang, J., Cheng, F.: Consensus model with double feedback mechanism based on dynamic trust relationship in social network group decision-making. Int. J. Comput. Intell. Syst. **14**, 491–502 (2021)

34. Herrera-Viedma, E., Cabrerizo, F., Kacprzyk, J., Pedrycz, W.: A review of soft consensus models in a fuzzy environment. Inf. Fusion **17**, 4–13 (2014)

35. Palomares, I., Estrella, F., Martínez, L., Herrera, F.: Consensus under a fuzzy context: taxonomy, analysis framework AFRYCA and experimental case of study. Inf. Fusion **20**, 252–271 (2014)

36. Zhang, Z., Gao, Y., Li, Z.: Consensus reaching for social network group decision making by considering leadership and bounded confidence. Knowl.-Based Syst. **204**, 106240 (2020)

37. Ben-Arieh, D., Chen, Z.: Linguistic labels aggregation and consensus measure for autocratic decision-making using group recommendations. IEEE Trans. Syst. Man Cybern. A **36**(1), 558–568 (2006)

38. Zhang, G., Dong, Y., Xu, Y., Li, H.: Minimum-cost consensus models under aggregation operators. IEEE Trans. Syst. Man Cybern. A **41**(6), 1253–1261 (2011)

39. Wu, Z., Xu, J.: A consistency and consensus based decision support model for group decision making with multiplicative preference relations. Decis. Support Syst. **52**(3), 757–767 (2012)

40. Ben-Arieh, D., Easton, T.: Multi-criteria group consensus under linear cost opinion elasticity. Decis. Support Syst. **43**(3), 713–721 (2007)

41. Gong, Z.W., Zhang, H.H., Forrest, J., Li, L.S., Xu, X.X.: Two consensus models based on the minimum cost and maximum return regarding either all individuals or one individual. Eur. J. Oper. Res. **240**(1), 183–192 (2015)

42. Li, Y., Zhang, H.J., Dong, Y.C.: The interactive consensus reaching process with the minimum and uncertain cost in group decision making. Appl. Soft Comput. **60**, 202–212 (2017)

43. Zhang, N., Gong, Z.W., Chiclana, F.: Minimum cost consensus models based on random opinions. Expert Syst. Appl. **89**, 149–159 (2017)

44. Tacon, R., Walters, G., Cornforth, C.: Accountability in nonprofit governance: a process-based study. Nonprofit Volunt. Sect. Q. **46**(4), 685–704 (2017)

45. Hideto Dato, M., Hudong, M., Mersland, R.: Board governance: does ownership matter? Ann. Public Coop. Econ. **91**(1), 5–28 (2020)

46. Jepson, P.: Governance and accountability of environmental NGOs. Environ. Sci. Policy **8**(5), 515–524 (2005)

47. AbouAssi, K.: Testing resource dependency as a motivator for NGO self-regulation: suggestive evidence from the global south. Nonprofit Volunt. Sect. Q. **44**(6), 1255–1273 (2015)

48. Reid, W., Turbide, J.: Board/staff relationships in a growth crisis: implications for nonprofit governance. Nonprofit Volunt. Sect. Q. **41**(1), 82–99 (2012)

49. Keating, V.C., Thrandardottir, E.: NGOs, trust, and the accountability agenda. Br. J. Politics Int. Relat. **19**(1), 134–151 (2017)

50. Greenlee, J., Fischer, M., Gordon, T., Keating, E.: An investigation of fraud in nonprofit organizations: occurrences and deterrents. Nonprofit Volunt. Sect. Q. **36**(4), 676–694 (2007)

51. Murtaza, N.: Putting the lasts first: the case for community-focused and peer-managed NGO accountability mechanisms. VOLUNTAS: Int. J. Volunt. Nonprofit Organ. **23**(1), 109–125 (2012)

52. Abou Assi, K., Trent, D. L.: NGO accountability from an ngo perspective: Their perceptions, strategies, and practices. Program on governance and local development working paper (2015)

53. Hortsch, D.: The paradox of partnership: amnesty international, responsible advocacy, and NGO accountability. Columbia Hum. Right Law Rev. **42**, 119–155 (2010)

54. Rocha Valencia, L.. A., Queiruga, D., Gonzlez-Benito, J.: Relationship between transparency and efficiency in the allocation of funds in nongovernmental development organizations. VOLUNTAS: Int. J. Volunt. Nonprofit Organ. **26**(6), 2517–2535 (2015)

55. Hyndman, N., McDonnell, P.: Governance and charities: an exploration of key themes and the development of a research agenda. Financ. Account. Manage. **25**(1), 5–31 (2009)

56. Wellens, L., Jegers, M.: Effective governance in nonprofit organizations: a literature based multiple stakeholder approach. Eur. Manage. J. **32**(2), 223–243 (2014)

57. Coule, T.M.: Nonprofit governance and accountability: broadening the theoretical perspective. Nonprofit Volunt. Sect. Q. **44**(1), 75–97 (2015)

58. Torra, V.: Aggregation of linguistic labels when semantics is based on antonyms. Int. J. Intell. Syst. **16**(4), 513–524 (2001)

59. Ruspini, E.H.: A new approach to clustering. Inf. Control **15**(1), 22–32 (1969)

60. Cheng, C.: A new approach for ranking fuzzy numbers by distance method. Fuzzy Sets Syst. **95**(3), 307–317 (1998)

61. Safarzadeh, S., Khansefid, S., Rasti-Barzoki, M.: A group multi-criteria decision-making based on Best–Worst method. Comput. Ind. Eng. **126**, 111–121 (2018)

Ana Licerán-Gutiérrez is a PhD. In Economics and Business Administration and a Bachelor in Law. She works in the Department of Financial Economics and Accounting at the University of Jaén (Spain) as a lecturer and researcher. She is currently researching in different lines about accounting: earnings quality and its effect on cost of financing (this is the line of the Doctoral Thesis, currently extending research), measurement of transparency and good governance of Nonprofit Organizations (line of a PhD student whose thesis I am co-directing), and International Accounting Standards. She participates in innovative teaching projects on active learning and gamification. She has received two awards for best conference papers in international workshops, an award for best doctoral research paper as thesis advisor, and two awards for best bachelor marks. She is a member of the European Accounting Association (EAA), ASEPUC (Accounting University Lecturers), and AECA (Accounting and Business Administration).

Cristina Ortega-Rodríguez has work 6 years as a professor in the Department of Financial Economics and Accounting at the University of Jaén, teaching subjects in Financial Accounting and Financial Statements. She has taught during two years Financial subject in Paris Université ParisEst Créteil (UPEC). She has worked for several years in the private sector in companies such as PriceWaterhouse Coopers. Her research focuses on the transparency and good governance of the non-profit organizations. She has published several articles in journals indexed by the SSCI (Journals Citation Reports) and more than 30 contributions in International Congresses related to her areas.

Antonio Luis Moreno Albarracín has work 25 years as an associate professor in the Department of Financial Economics and Accounting at the University of Jaén, teaching subjects in Financial Accounting and Management Accounting for Public Entities. He has held positions of high responsibility in the non-profit sector (National Organization of the Spanish Blind (ONCE), among others). His research is focused on the non-profit sector. He has published several papers in journals indexed by the SSCI and more than 50 contributions in Inter / national Conferences related to his areas.
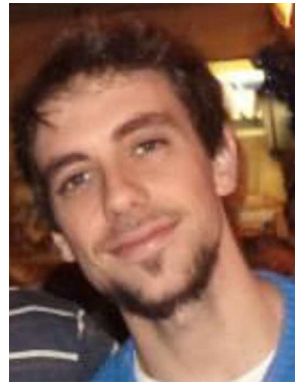
Álvaro Labella received his B.sc diploma from University of Jaén, Spain, in 2014 and holds his PhD in Computer Science from the same university in 2021. Currently, he is a postdoctoral researcher in the SIN-BAD2 research group at the University of Jaén and works in the area of decision making, computing with words and software application development. During the development of his doctoral thesis, he has published 20 articles in international journals, as well as 17 contributions in international conferences and 4 in National Conferences. During his doctoral studies he stayed at the National Technical University of Athens (NTUA) in Athens, Greece and at the Portsmouth Business School in Portsmouth, UK. He was awarded by the University of Jaén with the Ada Lovelace Award in the Best Application modality, in 2019 he received the Best Student Paper Award at the International Conference on Intelligent and Fuzzy Systems in Istanbul, Turkey and the Best Paper Award at the International Conference on Intelligent Systems and Knowledge Engineering in Dalian, China.

Rosa M. Rodríguez received her B.Sc, M.Sc. and PhD degrees in Computer Science, all of them from the University of Jaén in 2006, 2008 and 2013 respectively. She got the extraordinary award of PhD (bestowed in 2017). Currently, she has a Posdoctoral contract (Ramón y Cajal) at the University of Jaén, Spain. Her research interests include linguistic preference modelling, decision making, decision support systems and fuzzy logic based systems. She has more than 40 papers in journals indexed by the SCI. She is Editorial Assistant in International Journal of Computational Intelligence Systems and Associate Editor in IEEE Transactions on Fuzzy Systems, International Journal of Fuzzy Systems and Journal of Intelligent and Fuzzy Systems. She received the IEEE Transactions on Fuzzy Systems Outstanding Paper Award 2012 (bestowed in 2015). Finally highlight that It has been classified as Highly Cited Researcher 2019, 2020 and 2021 according to the Essential Science Indicators of the Web of Science.

Luis Martínez is currently a Full Professor with the Computer Science Department, University of Jaén, Jaén, Spain. He is also Visiting Professor in University of Technology Sydney, University of Portsmouth (Isambard Kingdom Brunel Fellowship Scheme), and in the Wuhan University of Technology (Chutian Scholar). He has been main researcher in 16 R&D projects, also has published more than 190 papers in journals indexed by the SCI and more than 200 contributions in Inter/national Conferences related to his areas. His current research interests include multicriteria decision making, fuzzy logic-based systems, computing with words and recommender systems. He was a recipient o f the IEEE Transactions on fuzzy systems Outstanding Paper Award 2008 and 2012 (bestowed in 2011 and 2015, respectively). He is a Co-Editor-in-Chief o f the International Journal of Computational Intelligence Systems and an Associate Editor of the journals, including the Information Sciences, Knowledge Based Systems, Information Fusion. He is IFSA Fellow 2021, senior member of IEEE and of the European Society for Fuzzy Logic and Technology. Eventually, he has been appointed as Highly Cited Researcher 2017–2021 in Computer sciences.