The 18th International Conference on Intelligent Systems and Knowledge Engineering (ISKE 2023)

Table of Contents

◆ Deep Learning for Time Series Forecasting

◆ Logic, Computation and Artificial Intelligence

◆ Explainable Deep Learning Techniques and Applications

◆ Intelligent Manufacturing Enabled Sustainable Development

◆ Artificial Intelligent Methodology and System

◆ Knowledge Engineering and Management

Practical Applications of AI and Knowledge Engineering

A Framework for the Selection of Indicators through Fuzzy Thresholds: A Circular Economy Application

1st Diego García-Zamora
Dept. of Computer Science
University of Jaén
Jaén, Spain
dgzamora@ujaen.es

2nd Álvaro Labella
Dept. of Computer Science
University of Jaén
Jaén, Spain
alabella@ujaen.es

3rd Rosa M. Rodríguez
Dept. of Computer Science
University of Jaén
Jaén, Spain
rmrodrig@ujaen.es

4st Luis Martínez
Dept. of Computer Science
University of Jaén
Jaén, Spain
martin@ujaen.es

***Abstract*—The evaluation of indicators has become increasingly important in various domains, necessitating a comprehensive approach to assess their significance and reliability. This paper introduces a novel framework that addresses the challenges of indicators evaluation by incorporating minimum cost consensus (MCC) and fuzzy thresholds. The framework tackles three key issues: experts' preference modeling, the consensus among experts, and acceptance/rejection conditions. By leveraging linguistic preferences and the MCC model, the framework generates collectively agreed opinions and manages disagreements whereas fuzzy thresholds account for uncertainty and subjectivity, offering a flexible representation of acceptance criteria. Advantages include closer alignment with experts' thinking, consensual indicators' evaluation, and reliable fuzzy acceptance degrees. Furthermore, the proposal is applied to a circular economy context, demonstrating its effectiveness.**

***Index Terms*—Group decision-making, minimum cost consensus, fuzzy threshold, circular economy**

I. Introduction

In recent years, the evaluation of indicators has gained significant importance across various domains, including economics, finance, environmental sciences, and social sciences [1]–[3]. Indicators serve as valuable tools to measure and monitor the performance, progress, and overall health of systems, processes, or entities. However, assessing the significance and reliability of these indicators often poses a challenge, as their evaluation requires consideration of multiple factors, such as consensus among stakeholders, cost implications, and the inherent uncertainty associated with their measurements.

Researchers have addressed the evaluation of indicators in multiple contexts from different approaches, most of them based on expert knowledge. Moreno et al. [1] identified the most relevant good governance nonprofit indicators to evaluate the transparency in nonprofit organizations following a Best-Worst method approach [4]. Nuñez et al. [3] developed a measurement scale based on an initial set of 234 circular economy (CE) indicators, obtained from a literature review and analysis of company reports. In this case, a Delphi approach [5] was used to select the most important CE indicators. In [6], Yang et al. conducted a 2-rounds online survey to collect information and evaluate the importance of the sustainable development goals [7]. However, these approaches demand a lot of attention from experts due to their long duration and complexity, which could result in the attrition of experts and excessive costs.

Regardless of the evaluation mechanism used, three main issues should be kept in mind in the indicators' evaluation process.

1) Experts preference modelling: the way in which experts express their opinions is key to the correct evaluation of indicators. It is quite common to ask experts to evaluate indicators on the basis of numerical assessments [1], [3]. However, it is sometimes very difficult for experts to evaluate the importance of indicators with precise numerical ratings. In that sense, the use of fuzzy linguistic terms such as "important" or "unimportant" may help experts to express their actual opinions [8].
2) Consensus among experts: when the acceptance or rejection of the indicators is based on the opinions of several experts, disagreements may appear [9]. Making decisions ignoring such disagreements may lead to unsatisfactory results, thus, they should be smoothed out to obtain agreed solutions.
3) Acceptance and rejection conditions: determine the thresholds to accept or reject an indicator is key. However, in many cases, simple numerical thresholds may be too rigid and fail to take into account the uncertainty and subjectivity inherent in the evaluation of indicators. A fuzzy definition of the thresholds [10] is able to provide a more flexible and adaptable representation of the acceptance criteria.

On this basis, this paper presents a novel framework that aims to address the challenges of indicators evaluation by proposing a comprehensive approach to evaluating indicators based on minimum cost consensus (MCC) [11] and fuzzy thresholds. First, the experts will be asked to provide their opinions about the different indicators using fuzzy linguistic terms. Then, an MCC model is proposed to estimate the collectively agreed opinions for the group that achieves an

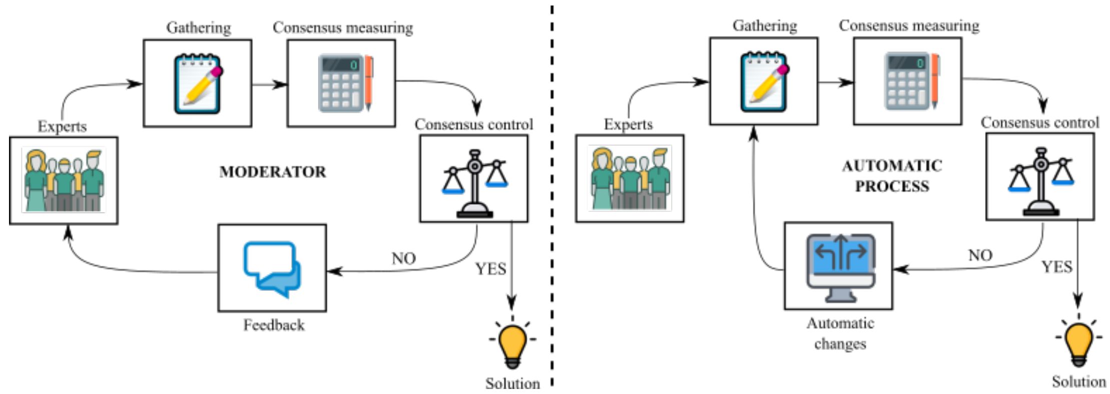

Fig. 1. CRP scheme

acceptable level of agreement among experts. To carry out computations with linguistic information in the proposed MCC model, the 2-tuple linguistic model is used. This model is able to carry out the computing with words processes [12], [13], keeping the interpretability and precision of the results [14]. In addition, we leverage the power of fuzzy thresholds within our framework to account for the inherent uncertainty and imprecision often associated with indicator measurements. To do so, we define a fuzzy threshold to determine the acceptance degree for each indicator that accounts for both the collective opinion on the importance of the indicator and the similarity between the experts' original opinions and the corresponding adjusted preferences obtained from the MCC model.

To sum up, the main advantages of this proposal are:

- Linguistic preferences: experts' opinions are modeled by linguistic information, closer to their common way of thinking.
- Consensual indicators' evaluation: disagreements among experts are managed to obtain a final selection of indicators that satisfy all the members of the group.
- Fuzzy acceptance degrees: fuzzy acceptance thresholds provide a more flexible and adaptable representation of the acceptance criteria and, consequently, a more reliable evaluation.

The remainder contribution is organized as follows: Section II reviews some notions related to the proposal. In Section III, a fuzzy framework to evaluate indicators is proposed. Section IV applies the proposed framework in the selection of CE indicators and compares it with a similar approach. Finally, Section V draws conclusions and future works.

II. Preliminaries

This section revises some basic concepts regarding, group decision-making (GDM), consensus reaching process (CRP), and the 2-tuple linguistic model.

A. Group Decision-Making and Consensus Reaching Processes

GDM involves a collective effort to generate, evaluate, and select the best course of action from a range of options [15]. It capitalizes on the diverse knowledge, expertise, and perspectives of individual group members to enhance the quality of decisions. By pooling together various viewpoints, GDM aims to achieve better outcomes, improve problem-solving capabilities, and foster a sense of shared responsibility.

While GDM offers numerous benefits, it is not without challenges that can impede the decision-making process, lead to prolonged discussions, and hinder the achievement of consensus. Addressing these challenges is crucial for effective GDM, and this work focuses on the management of conflicting opinions in GDM through CRPs.

CRPs are designed to guide groups towards a shared agreement that reflects the collective opinion and maximizes group satisfaction. These processes aim to achieve a balance between individual preferences and the overall group objectives, performing several steps (see Fig. 1). Two main CRPs approaches have been studied in the specialized literature [9]:

1) CRPs with feedback process: these CRPs are usually guided by a moderator who identifies the disagreements among the DMs and ask DMs to change their initial views for bringing their opinions closer and reaching a higher level of agreement [16], [17].
2) Automatic CRPs: these CRPs are automatically supervised. In this case, the DMs are not asked if they want to change or not their initial preferences to increase the level of consensus in the group, but they are changed automatically [11], [18]. This contribution is focused on this kind of CRPs.

A CRP will end either when a desired threshold of consensus (μ_0) is reached, which will be defined a priori, or when a maximum number of rounds of discussion, also defined a priori, have been carried out.

B. 2-tuple linguistic model

The 2-tuple linguistic model [14] was developed to address the limitations of classical linguistic computational approaches. It employs a continuous fuzzy representation of linguistic information and a computational model capable of performing symbolic precise computations, thereby avoiding approximations and achieving accurate linguistic results based on the computing with words scheme [12], [13].

A 2-tuple linguistic value is represented as $(s_i, \alpha) \in \overline{S}$, where s_i is a linguistic term from a predefined set $S = \{s_0, s_1, \ldots, s_g\}$ (with g being a fixed even number), and α denotes the symbolic translation. The symbolic translation α is a numerical value that indicates the shift in the fuzzy membership function of s_i. It is important to note that the possible values for α in the 2-tuple linguistic value $(s_i, \alpha) \in \overline{S}$ are limited to the interval $[-0.5, 0.5[$.

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

One notable feature of 2-tuple linguistic expressions is their ability to be translated into a numerical value x within the range of $[0, g]$. This translation facilitates the computational processes by simplifying the calculations involved.

Proposition 1: [14] Let $S = \{s_0, \ldots s_g\}$ be a linguistic term set. Then, the function $\Delta_S^{-1} : \overline{S} \to [0, g]$ defined by

$$\Delta_S^{-1}(s_i, \alpha) = i + \alpha,\ \forall\ (s_i, \alpha) \in \overline{S}$$

is a bijection whose inverse $\Delta_S : [0, g] \to \overline{S}$ is given by

$$\Delta_S(x) = (s_{round(x)}, x - round(x))\ \forall\ x \in [0, g],$$

where $round(\cdot)$ is the function that assigns the closest integer number $i \in \{0, \ldots, g\}$.

Remark 1: Note that any linguistic term $s_i \in S$ can be represented as a 2-tuple linguistic value by considering $(s_i, 0) \in \overline{S}$.

III. A FRAMEWORK TO EVALUATE INDICATORS BASED ON MINIMUM COST CONSENSUS AND FUZZY THRESHOLDS

In this section, we present a framework designed to provide a comprehensive methodology for evaluating indicators based on the principles of MCC and fuzzy thresholds. By introducing this framework, we aim to address the inherent challenges in indicator evaluation, including experts agreement, similarity considerations, and the uncertainty associated with measurements. By leveraging MCC and incorporating fuzzy thresholds, we offer a robust framework that enables a more nuanced and context-sensitive evaluation, enhancing the effectiveness of indicators assessment.

To facilitate the elicitation process, here we assume that a group of experts, $E = \{e_1, e_2, \ldots, e_m\}$ provide their opinions over a set of indicators, $I = \{i_1, i_2, \ldots, i_n\}$, using linguistic values. Therefore, let us consider a linguistic term set $S = \{s_0, ..., s_g\}$ where $g \in \mathbb{N}$ is an even number and experts' opinions are modeled through a decision matrix with 2-tuple linguistic values $O \in \mathcal{M}_{m\times n}(\overline{S})$. The 2-tuple linguistic representation allows performing computations with linguistic information by using Prop. 1 and guarantees precise results. Thus, an automatic CRP based on MCC [19] is defined as follows:

$$\begin{aligned} &\min_{\overline{O}, g} \sum_{k=1}^{m} \sum_{i=1}^{n} |\Delta_S^{-1}(\overline{o}_{ki}) - \Delta_S^{-1}(o_{ki})| \\ &s.t. \begin{cases} \Delta_S^{-1}(G_i) = \frac{1}{m} \sum_{k=1}^{m} \Delta_S^{-1}(\overline{o}_{ki}),\ i = 1, 2, \ldots, n. \\ 1 - \frac{1}{mg} \sum_{k=1}^{m} |\Delta_S^{-1}(\overline{o}_{ki}) - \Delta_S^{-1}(G_i)| \geq \mu_0,\ i = 1, 2, \ldots, n. \end{cases} \end{aligned} \tag{MCC}$$

where $\mu_0 \in]0, 1[$ is the consensus threshold, $\overline{O} \in \mathcal{M}_{m\times n}(\overline{S})$ are experts' adjusted opinions and $G \in \overline{S}^n$ represents the group collective opinion computed by using the arithmetic mean.

After applying this model, we obtain a collective preference $G \in \overline{S}^n$ such that for each indicator i, the collective value G_i is accepted by the members of the group up to the consensus level μ_0. Now, for each indicator $i \in \{1, 2, ..., n\}$ we can compute the average similarity between experts' original and modified opinions by using the formula:

$$S_i = 1 - \frac{1}{m \cdot g} \sum_{k=1}^{m} |\Delta_S^{-1}(\overline{o}_{ki}) - \Delta_S^{-1}(o_{ki})|.$$

In order to automatize the indicator selection process, we need to apply some key rules for each indicator i:

- If the collective preference G_i/g is lower than a certain threshold $a \in [0, 1[$, then the indicator i must be rejected. Additionally, the performance of an indicator i is considered excellent if the value G_i/g surpasses a threshold $b \in]0, 1]$.
- If the average similarity S_i is lower than a certain threshold $c \in [0, 1[$, it will be discarded. It will be considered good enough if S_i is greater than a threshold $d \in]0, 1]$.

In this context, it is reasonable to assume that, if the collective preference $\Delta_S^{-1}(G_i)/g$ is within the interval $]a, b[$, the acceptance degree should be close to 0 when $\Delta_S^{-1}(G_i)/g$ is close to a, whereas it should be close to 1 if $\Delta_S^{-1}(G_i)/g$ is close to b. In the same way, when $S_i \in]c, d[$, its acceptance should be close to 0 when S_i is close to c, and close to 1 if S_i is close to d. To model this behavior, given $0 \leq \phi < \psi \leq 1$, we can consider the piecewise linear fuzzy number $Q : [0, 1] \to [0, 1]$ (see Fig. 2) defined by:

$$Q_{\phi,\psi}(x) = \begin{cases} 0 & \text{if } x < \phi \\ \frac{x-\phi}{\psi-\phi} & \text{if } \phi \leq x \leq \psi \\ 1 & \text{if } x > \psi \end{cases} \forall x \in [0, 1].$$

Now, using the product t-norm [20] and given $0 \leq a < b \leq 1$ and $0 \leq c < d \leq 1$, we can construct the fuzzy threshold $D_{(a,b),(c,d)} : [0, 1] \times [0, 1] \to [0, 1]$ as follows (see Fig. 3):

$$D_{(a,b),(c,d)}(x, y) = Q_{a,b}(x) \cdot Q_{c,d}(y)\ \forall\ x, y \in [0, 1].$$

Therefore, the acceptance degree of the i-th indicator can be computed as:

$$d^i_{(a,b),(c,d)} = D_{(a,b),(c,d)}(\Delta_S^{-1}(G_i)/g, S_i).$$

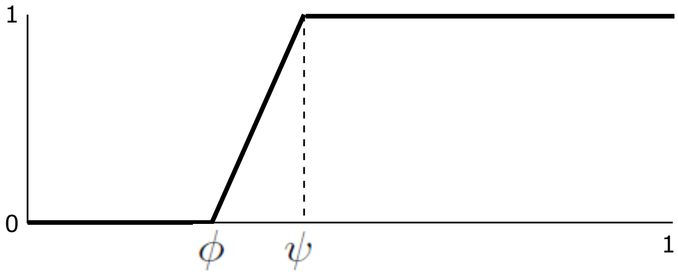

Fig. 2. Membership function of $Q_{\phi,\psi}$

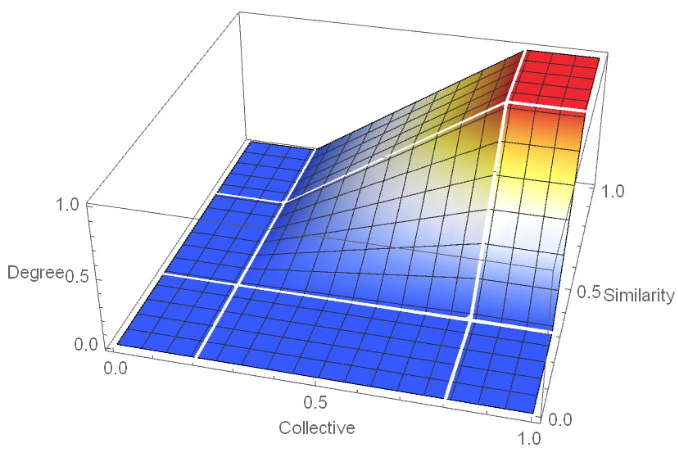

Fig. 3. Membership function of $D_{(0.2,0.8),(0.3,0.7)}$

In that case, if either the collective preference $\Delta_S^{-1}(G_i)/g$ is lower than a or the similarity S_i is under c, the i-th indicator is rejected. Additionally, the higher the collective preference and the similarity, the higher the value of $d^i_{(a,b),(c,d)}$. In the case that simultaneously $\Delta_S^{-1}(G_i)/g \geq b$ and $S_i \geq d$, then $d^i_{(a,b),(c,d)} = 1$.

IV. Case study: Circular Economy indicators evaluation

This section applies the proposed framework to the evaluation of CE indicators and shows its reliability and advantages. To do so, we follow the case study introduced in [3] and compare our results with the ones obtained from the former proposal.

A. Background

The concept of CE has gained significant attention in recent years as a sustainable alternative to the traditional linear economic model [21]. It emphasizes the importance of reducing waste, promoting resource efficiency, and fostering closed-loop systems. While the CE offers numerous environmental and economic benefits, assessing its progress and impact requires the use of suitable indicators [22].

Measuring CE achievements serves multiple purposes [23]. Firstly, it provides a means to monitor the progress of organizations and institutions in transitioning towards circular practices. By quantifying and analyzing key performance indicators, stakeholders can assess the effectiveness of their efforts and identify areas for improvement. Secondly, measuring CE achievements enables benchmarking and comparison among different entities, facilitating knowledge sharing and best practices dissemination. Lastly, it allows policy-makers and decision-makers to evaluate the effectiveness of CE policies and interventions and makes informed decisions based on empirical evidence.

Determining the appropriate dimensions and indicators for measurement is currently a pressing concern for researchers. However, a key challenge in these emerging fields is the absence of pre-validated scales in the existing literature. As a result, researchers must undertake the construction and validation processes using various methods.

Recently, Nuñez et al. [3] developed a measurement scale based on an initial set of 234 indicators obtained from a literature review and analysis of company reports and classified them into 7 categories: energy, 3R's, water management, waste management, materials, emissions, and transition to CE. A Delphi approach [5] was used to select the most important CE indicators based on expert knowledge, resulting in a final set of 54 indicators. The Delphi method is a CRP in which experts participate in multiple rounds of questionnaires or surveys that are administered and facilitated by a moderator. However, this approach has several limitations:

- Long time horizon: it requires a long time to be completed successfully.
- Attrition: participants may drop out because of the long duration of the process.
- Financial cost: the longer the process, the higher the cost.

In the next section, we will replace the Delphi approach used in Nuñez et al. approach [3] by the proposed fuzzy framework based on MCC and show how it can overcome all the previous limitations and obtain very similar results regarding the former. However, for sake of space, we will focus only on the evaluation of 7 indicators belonging to the category water management:

i_1: Use of purified rainwater.
i_2: Environmental chemicals used in the process of treating water and sewage.
i_3: Fresh water consumption.
i_4: Industrial water reuse ratio.
i_5: Industrial, domestic, and ballast effluents.
i_6: Recycled and reused water.
i_7: Water consumption per unit industrial production value.

B. Resolution

Firstly, 31 experts provide their preferences by using linguistic terms. The fuzzy linguistic term set used is defined as $S = \{s_0 : Very\ unimportant, s_1 : Unimportant, s_2 : Fair, s_3 : Important, s_4 : Very\ important\}$ and depicted in Fig. 4. Table I shows the experts' preferences. Notice, the linguistic assessments have been transformed into 2-tuple linguistic values with symbolic translation equal to 0 (see Remark 1)

TABLE I
EXPERTS' PREFERENCES

Expert	i_1	i_2	i_3	i_4	i_5	i_6	i_7
e_1	$(s_3, 0.0)$	$(s_3, 0.0)$	$(s_4, 0.0)$	$(s_4, 0.0)$	$(s_3, 0.0)$	$(s_4, 0.0)$	$(s_3, 0.0)$
e_2	$(s_3, 0.0)$	$(s_4, 0.0)$	$(s_4, 0.0)$	$(s_4, 0.0)$	$(s_4, 0.0)$	$(s_4, 0.0)$	$(s_4, 0.0)$
e_3	$(s_3, 0.0)$	$(s_4, 0.0)$	$(s_4, 0.0)$	$(s_2, 0.0)$	$(s_2, 0.0)$	$(s_4, 0.0)$	$(s_3, 0.0)$
e_4	$(s_1, 0.0)$	$(s_1, 0.0)$	$(s_2, 0.0)$	$(s_3, 0.0)$	$(s_4, 0.0)$	$(s_4, 0.0)$	$(s_2, 0.0)$
e_5	$(s_3, 0.0)$	$(s_4, 0.0)$	$(s_4, 0.0)$	$(s_3, 0.0)$	$(s_3, 0.0)$	$(s_4, 0.0)$	$(s_2, 0.0)$
e_6	$(s_2, 0.0)$	$(s_3, 0.0)$	$(s_3, 0.0)$	$(s_2, 0.0)$	$(s_2, 0.0)$	$(s_3, 0.0)$	$(s_3, 0.0)$
e_7	$(s_2, 0.0)$	$(s_2, 0.0)$	$(s_3, 0.0)$	$(s_2, 0.0)$	$(s_2, 0.0)$	$(s_3, 0.0)$	$(s_3, 0.0)$
e_8	$(s_3, 0.0)$	$(s_3, 0.0)$	$(s_4, 0.0)$	$(s_4, 0.0)$	$(s_4, 0.0)$	$(s_4, 0.0)$	$(s_4, 0.0)$
e_9	$(s_1, 0.0)$	$(s_4, 0.0)$	$(s_4, 0.0)$	$(s_4, 0.0)$	$(s_3, 0.0)$	$(s_4, 0.0)$	$(s_3, 0.0)$
e_{10}	$(s_4, 0.0)$	$(s_4, 0.0)$	$(s_4, 0.0)$	$(s_4, 0.0)$	$(s_4, 0.0)$	$(s_4, 0.0)$	$(s_4, 0.0)$
e_{11}	$(s_3, 0.0)$	$(s_3, 0.0)$	$(s_4, 0.0)$	$(s_4, 0.0)$	$(s_4, 0.0)$	$(s_4, 0.0)$	$(s_4, 0.0)$
e_{12}	$(s_4, 0.0)$	$(s_4, 0.0)$	$(s_4, 0.0)$	$(s_4, 0.0)$	$(s_4, 0.0)$	$(s_4, 0.0)$	$(s_4, 0.0)$
e_{13}	$(s_3, 0.0)$	$(s_2, 0.0)$	$(s_4, 0.0)$	$(s_3, 0.0)$	$(s_2, 0.0)$	$(s_4, 0.0)$	$(s_3, 0.0)$
e_{14}	$(s_3, 0.0)$	$(s_4, 0.0)$	$(s_4, 0.0)$	$(s_4, 0.0)$	$(s_3, 0.0)$	$(s_4, 0.0)$	$(s_4, 0.0)$
e_{15}	$(s_4, 0.0)$	$(s_4, 0.0)$	$(s_4, 0.0)$	$(s_4, 0.0)$	$(s_4, 0.0)$	$(s_4, 0.0)$	$(s_4, 0.0)$
e_{16}	$(s_4, 0.0)$	$(s_3, 0.0)$	$(s_4, 0.0)$	$(s_4, 0.0)$	$(s_4, 0.0)$	$(s_4, 0.0)$	$(s_3, 0.0)$
e_{17}	$(s_3, 0.0)$	$(s_4, 0.0)$	$(s_4, 0.0)$	$(s_3, 0.0)$	$(s_3, 0.0)$	$(s_4, 0.0)$	$(s_4, 0.0)$
e_{18}	$(s_2, 0.0)$	$(s_3, 0.0)$	$(s_4, 0.0)$	$(s_4, 0.0)$	$(s_3, 0.0)$	$(s_4, 0.0)$	$(s_4, 0.0)$
e_{19}	$(s_3, 0.0)$	$(s_3, 0.0)$	$(s_3, 0.0)$	$(s_3, 0.0)$	$(s_3, 0.0)$	$(s_4, 0.0)$	$(s_3, 0.0)$
e_{20}	$(s_2, 0.0)$	$(s_2, 0.0)$	$(s_4, 0.0)$	$(s_3, 0.0)$	$(s_2, 0.0)$	$(s_3, 0.0)$	$(s_4, 0.0)$
e_{21}	$(s_3, 0.0)$	$(s_3, 0.0)$	$(s_4, 0.0)$	$(s_3, 0.0)$	$(s_2, 0.0)$	$(s_3, 0.0)$	$(s_4, 0.0)$
e_{22}	$(s_4, 0.0)$	$(s_3, 0.0)$	$(s_4, 0.0)$	$(s_4, 0.0)$	$(s_4, 0.0)$	$(s_4, 0.0)$	$(s_4, 0.0)$
e_{23}	$(s_5, 0.0)$	$(s_3, 0.0)$	$(s_2, 0.0)$	$(s_4, 0.0)$	$(s_3, 0.0)$	$(s_4, 0.0)$	$(s_2, 0.0)$
e_{24}	$(s_4, 0.0)$	$(s_4, 0.0)$	$(s_3, 0.0)$	$(s_4, 0.0)$	$(s_4, 0.0)$	$(s_4, 0.0)$	$(s_3, 0.0)$
e_{25}	$(s_2, 0.0)$	$(s_2, 0.0)$	$(s_2, 0.0)$	$(s_2, 0.0)$	$(s_2, 0.0)$	$(s_2, 0.0)$	$(s_2, 0.0)$
e_{26}	$(s_3, 0.0)$	$(s_3, 0.0)$	$(s_3, 0.0)$	$(s_3, 0.0)$	$(s_3, 0.0)$	$(s_3, 0.0)$	$(s_2, 0.0)$
e_{27}	$(s_4, 0.0)$	$(s_3, 0.0)$	$(s_4, 0.0)$	$(s_4, 0.0)$	$(s_3, 0.0)$	$(s_4, 0.0)$	$(s_2, 0.0)$
e_{28}	$(s_4, 0.0)$	$(s_4, 0.0)$	$(s_4, 0.0)$	$(s_4, 0.0)$	$(s_4, 0.0)$	$(s_4, 0.0)$	$(s_4, 0.0)$
e_{29}	$(s_3, 0.0)$	$(s_4, 0.0)$	$(s_4, 0.0)$	$(s_4, 0.0)$	$(s_3, 0.0)$	$(s_3, 0.0)$	$(s_3, 0.0)$
e_{30}	$(s_1, 0.0)$	$(s_1, 0.0)$	$(s_2, 0.0)$	$(s_4, 0.0)$	$(s_4, 0.0)$	$(s_3, 0.0)$	$(s_3, 0.0)$
e_{31}	$(s_4, 0.0)$	$(s_3, 0.0)$	$(s_0, 0.0)$	$(s_4, 0.0)$	$(s_0, 0.0)$	$(s_4, 0.0)$	$(s_2, 0.0)$

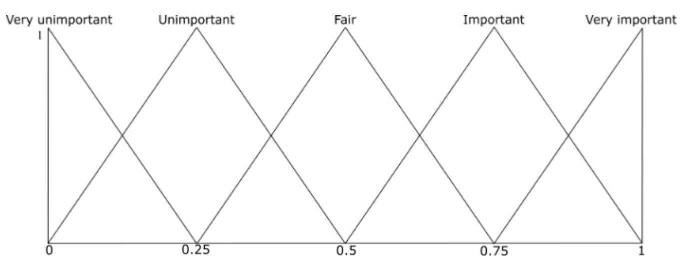

Fig. 4. Fuzzy linguistic terms set for assessments.

To smooth out the possible disagreements among experts, we apply the MCC model. This model changes automatically the initial experts' preferences as less as possible to achieve an acceptable level of agreement. In this experiment, it has been fixed $\mu_0 = 85$.

After obtaining the consensual opinions, the CE indicators are evaluated. The acceptance conditions will be determined by two aspects: (i) the average importance value obtained from the experts' preferences for each indicator, i.e., if the experts evaluate a specific indicator with a low importance, then the indicator will be directly refused, (ii) the cost of modifying the experts' preferences to achieve the consensus, i.e, if achieving a consensus for a specific indicator implies to change a lot the initial experts' preferences and, thus, the resulting cost of the MCC is quite high, we will refuse the indicator considering that an excessive change over the experts' preferences is not realistic (in real situations experts are not receptive to modify their opinions too much). Keeping in mind this, the next step consists of defining the thresholds and the rules of acceptance as follows:

- R1 If the collective value is lower than $a = 0.6$, then the indicator i must be rejected. On the contrary, if the collective performance of i surpasses the threshold $b = 0.9$, the indicator is considered excellent.
- R2 If the average similarity of the expert' opinions over an indicator i is lower than $c = 0.98$, i must be rejected. On the contrary, the indicator i will be considered good enough is its average similarity is equal to the threshold $d = 1$.

Remark 2: Notice, the values of the thresholds have not been assigned randomly, but they have been set according to the acceptance conditions established in [3].

Taking account the previous rules, the consensual collective preference is shown in Table II together its corresponding 2-tuple linguistic value. In this case, no indicator presents a low value according to the threshold $a = 0.6$, but one of them is considered with an excellent performance (greater than $b = 0.9$), which is i_6.

In addition, the average similarities for each indicator (measured from the cost of modifying experts' preferences) between modified and original preferences are shown in Table III. Note that the lowest similarity values according to the threshold $c = 0.98$ refer to the indicators i_1, i_3 and i_5.

TABLE II
COLLECTIVE PREFERENCE

Indicator	i_1	i_2	i_3	i_4	i_5	i_6	i_7
Collective	0.75	0.767	0.889	0.878	0.75	0.928	0.782
2-tuple	$(s_3, 0.0)$	$(s_3, 0.068)$	$(s_4, -0.443)$	$(s_4, -0.489)$	$(s_3, 0.0)$	$(s_4, -0.29)$	$(s_{3.0}, 0.128)$

TABLE III
AVERAGE OPINIONS' SIMILARITY FOR EACH INDICATOR

Indicator	i_1	i_2	i_3	i_4	i_5	i_6	i_7
Similarity	0.972	0.984	0.973	0.993	0.972	0.999	0.983

TABLE IV
CONSENSUS DEGREES

Indicator	i_1	i_2	i_3	i_4	i_5	i_6	i_7
Initial	0.823	0.831	0.814	0.842	0.819	0.892	0.831
Modified	0.85	0.85	0.85	0.85	0.85	0.89	0.85

The consensus degrees achieved by the experts for each indicator have been shown in Table IV. Notice all of them satisfy the consensus threshold condition $\mu_0 = 0.85$.

Finally, Table V shows the decision about the acceptance and rejection of the CE indicators. The indicators i_1, i_3 and i_5 have been rejected according to the predefined rules, and the remainder are accepted with different degrees of acceptance, in which i_6 highlights with the maximum degree.

Note that the indicators that have been rejected using the proposed framework are exactly the same as those that were rejected in [3]. However, while this approach took months to perform the indicators' selection, our approach obtains the same results in just a few seconds, demonstrating its potential.

V. CONCLUSIONS

In this contribution, we have presented a framework for the evaluation of indicators, addressing the challenges associated with their assessment. Such a framework introduces several key advantages:

1) it models experts' opinions using fuzzy linguistic terms, allowing them to express their preferences in a manner closer to their common way of thinking. This linguistic preference modeling enhances the accuracy and reliability of the evaluation process.
2) it facilitates the achievement of consensus among experts. By using the MCC model, we estimate the collectively agreed opinion for the group, ensuring that the final selection of indicators satisfies all members and effectively addresses disagreements.
3) Fuzzy acceptance degrees through the use of fuzzy thresholds are included. Considering the uncertainty and subjectivity inherent in indicator evaluation, this enables a more flexible and adaptable evaluation of the acceptance criteria.

To demonstrate the effectiveness of this framework, it has been applied to the selection of CE indicators and afterwards the results have been compared with a similar approach which uses Delphy method. The results showcased the advantages of the proposed method, highlighting the importance of the use of linguistic preferences, consensus-building, and fuzzy acceptance degrees in indicator evaluation.

TABLE V
FUZZY ACCEPTANCE DEGREES

Indicator	i_1	i_2	i_3	i_4	i_5	i_6	i_7
Degree	0.0	0.135	0.0	0.617	0.0	1.0	0.108
Acceptance	R	A*	R	A	R	A	R*

In future works, we will explore different aggregation techniques to improve the MCC approach. Moreover, investigating advanced linguistic models and incorporating semantic analysis methods could enhance the representation and interpretation of experts' opinions. Additionally, the application of the proposed framework to other domains and contexts would provide valuable insights into its generalizability and effectiveness.

ACKNOWLEDGMENT

This work is partially supported by the Andalusian Excellence Research Program with the Research Project ProyExcel_00257, by the Spanish Ministry of Economy and Competitiveness through the Postdoctoral fellow Ramón y Cajal (RYC-2017-21978), by the Spanish Ministry of Science, Innovation and Universities through a Formación de Profesorado Universitario (FPU2019/01203) grant and by the Junta de Andalucía, Andalusian Plan for Research, Development, and Innovation (POSTDOC 21-00461).

REFERENCES

[1] A. L. Moreno-Albarracín, A. Licerán-Gutierrez, C. Ortega-Rodríguez, Á. Labella, and R. M. Rodríguez, "Measuring what is not seen—transparency and good governance nonprofit indicators to overcome the limitations of accounting models," *Sustainability*, vol. 12, no. 18, p. 7275, 2020.

[2] R. Dhiman, "Identifying the key indicators of financial stability and financial development: a review of financial service sector," *Asian Journal of Management Science and Applications*, vol. 3, no. 4, pp. 302–320, 2018.

[3] P. Nuñez-Cacho, J. Górecki, V. Molina-Moreno, and F. A. Corpas-Iglesias, "What gets measured, gets done: Development of a circular economy measurement scale for building industry," *Sustainability*, vol. 10, no. 7, p. 2340, 2018.

[4] J. Rezaei, "Best-worst multi-criteria decision-making method," *Omega*, vol. 53, pp. 49–57, 2015.

[5] N. Dalkey and O. Helmer, "An experimental application of the delphi method to the use of experts," *Management science*, vol. 9, no. 3, pp. 458–467, 1963.

[6] S. Yang, W. Zhao, Y. Liu, F. Cherubini, B. Fu, and P. Pereira, "Prioritizing sustainable development goals and linking them to ecosystem services: A global expert's knowledge evaluation," *Geography and Sustainability*, vol. 1, no. 4, pp. 321–330, 2020.

[7] O. Cf, "Transforming our world: the 2030 agenda for sustainable development," *United Nations: New York, NY, USA*, 2015.

[8] L. A. Zadeh, "The concept of a linguistic variable and its application to approximate reasoning, part i, ii, iii," *Information Sciences*, vol. 8,8,9, pp. 199 – 249, 301–357, 43–80, 1975.

[9] D. García-Zamora, Á. Labella, W. Ding, R. M. Rodríguez, and L. Martínez, "Large-scale group decision making: a systematic review and a critical analysis," *IEEE/CAA J. Autom. Sinica*, vol. 9, no. 6, p. 949–966, 2022.

[10] L. Zadeh, "Fuzzy logic = computing with words," *IEEE Transactions on Fuzzy Systems*, vol. 4, no. 2, pp. 103–111, 1996.

[11] D. Ben-Arieh and T. Easton, "Multi-criteria group consensus under linear cost opinion elasticity," *Decision Support Systems*, vol. 43, no. 3, pp. 713–721, 2007, integrated Decision Support.

[12] J. M. Mendel, "Computing with words and its relationships with fuzzistics," *Information Sciences*, vol. 177, no. 4, pp. 988–1006, 2007.

[13] R. R. Yager, "On the retranslation process in Zadeh's paradigm of computing with words," *IEEE Transactions on Systems, Man and Cybernetics, Part B (Cybernetics)*, vol. 34, no. 2, pp. 1184–1195, 2004.

[14] L. Martínez, R. M. Rodríguez, and F. Herrera, *The 2-tuple Linguistic Model*. Springer International Publishing, 2015.

[15] R. S. Tindale, T. Kameda, and V. B. Hinsz, "Group decision making," *Sage handbook of social psychology*, pp. 381–403, 2003.

[16] F. Chiclana, F. Mata, L. Martinez, E. Herrera-Viedma, and S. Alonso, "Integration of a consistency control module within a consensus model," *International Journal of Uncertainty, Fuzziness and Knowlege-Based Systems*, vol. 16, no. SUPPL. 1, pp. 35–53, 2008.

[17] E. Herrera-Viedma, F. Herrera, and F. Chiclana, "A consensus model for multiperson decision making with different preference structures," *Systems, Man and Cybernetics, Part A: Systems and Humans, IEEE Transactions on*, vol. 32, pp. 394 – 402, 2002.

[18] D. Cheng, Z. Zhou, F. Cheng, Y. Zhou, and Y. Xie, "Modeling the minimum cost consensus problem in an asymmetric costs context," *European Journal of Operational Research*, vol. 270, no. 3, pp. 1122–1137, 2018.

[19] D. García-Zamora, B. Dutta, Á. Labella, and L. Martínez, "A fuzzy-set based formulation for minimum cost consensus models," *Computers & Industrial Engineering*, vol. 181, p. 109295, Jul. 2023. [Online]. Available: https://doi.org/10.1016/j.cie.2023.109295

[20] M. Eugenia Cornejo, D. Lobo, and J. Medina, "Bipolar fuzzy relation equations systems based on the product t-norm," *MATHEMATICAL METHODS IN THE APPLIED SCIENCES*, vol. 42, no. 17, SI, pp. 5779–5793, NOV 30 2019.

[21] J. Górecki, P. Núñez-Cacho, F. A. Corpas-Iglesias, and V. Molina, "How to convince players in construction market? strategies for effective implementation of circular economy in construction sector," *Cogent engineering*, vol. 6, no. 1, p. 1690760, 2019.

[22] R. M. Rodríguez, Á. Labella, P. Nunez-Cacho, V. Molina-Moreno, and L. Martínez, "A comprehensive minimum cost consensus model for large scale group decision making for circular economy measurement," *Technological Forecasting and Social Change*, vol. 175, p. 121391, 2022.

[23] A. Mayer, W. Haas, D. Wiedenhofer, F. Krausmann, P. Nuss, and G. A. Blengini, "Measuring progress towards a circular economy: a monitoring framework for economy-wide material loop closing in the eu28," *Journal of industrial ecology*, vol. 23, no. 1, pp. 62–76, 2019.