A systematic review on multi-criteria group decision-making methods based on weights: Analysis and classification scheme

David Boix-Cots, Francesc Pardo-Bosch, Pablo Pujadas

ABSTRACT

Interest in group decision-making (GDM) has been increasing prominently over the last decade. Access to global databases, sophisticated sensors which can obtain multiple inputs or complex problems requiring opinions from several experts have driven interest in data aggregation. Consequently, the field has been widely studied from several viewpoints and multiple approaches have been proposed. Nevertheless, there is a lack of general framework. Moreover, this problem is exacerbated in the case of experts’ weighting methods, one of the most widely-used techniques to deal with multiple source aggregation. This lack of general classification scheme, or a guide to assist expert knowledge, leads to ambiguity or misreading for readers, who may be overwhelmed by the large amount of unclassified information currently available. To invert this situation, a general GDM framework is presented which divides and classifies all data aggregation techniques, focusing on and expanding the classification of experts’ weighting methods in terms of analysis type by carrying out an in-depth literature review. Results are not only classified but analysed and discussed regarding multiple characteristics, such as MCDMs in which they are applied, type of data used, ideal solutions considered or when they are applied. Furthermore, general requirements supplement this analysis such as initial influence, or component division considerations. As a result, this paper provides not only a general classification scheme and a detailed analysis of experts’ weighting methods but also a road map for researchers working on GDM topics or a guide for experts who use these methods. Furthermore, six significant contributions for future research pathways are provided in the conclusions.

1. Introduction

Decision-making is a key concept in our everyday lives. Albeit unconsciously, people are constantly choosing the most appropriate option from a set of possible alternatives regarding pre-defined requirements, criteria, or indicators. Consequently, multiple experts have developed structured and scientific decision-making methods rather than intuitive methods. These techniques are known as multiple-criteria decision-making (MCDM) methods as they integrate several criteria. They attempt to ensure objective and deliberate analysis to manage challenges successfully [1].

However, constant scientific development and increasingly specific expert knowledge have made problems more complex. Nowadays, a single data source or decision maker (DM) may not be able to consider all relevant aspects of a decision [2] or may not have all the required information. Multiple data sources or decision makers (DMs) could be required for a group decision-making (GDM) process which involves multiple criteria. In these cases, the process is known as multiple-criteria group decision-making (MCGDM) methods. As a consequence, there is a pressing need to explore how to join, incorporate, or consider their opinions to reach a consensus among them and this is constantly discussed among experts [3].

Some 1950s models proposed group negotiation as a solution [4] while others studied the application of election theory [5], which led to the scheme of social choice according to individual values [6]. At the same time, group behavior regarding solution gains and losses was being analysed [7,8]. These studies contributed to the final social choice theory scheme [9].

Besides these psychological aggregation proposals, more mathematical methods arose such as central tendency studies focused on...
finding the statistical value which best fits all DMs [10–12]. Others use the individual DM rankings to aggregate them into a group ranking. The combination techniques include probability distributions [13], Borda heuristic models and their extensions [14,15] or stochastic optimization [16–18].

However, some researchers state that it is impossible to have a homogeneous expert group whose experiences, attitudes, knowledge are the same or similar [19]. In these cases, considering equity between DM opinions could lead to bias or distorted results. Consequently, some methods focus on deriving DM weight to assign differing importance at the opinion aggregation. These methods can be commonly used as multi-source information fusion techniques because the proposed algorithms can consider parallelism between DMs or informatic sources. A huge number of studies to determine the weight of criteria can be found in the related literature, although there are limited studies determining the weight of experts [20,21]. Even so, some schemes have been proposed to divide and classify these methods. For example, they can be clustered in two groups: subjective and objective techniques. In subjective techniques, there is a supervisor or manager who directly assigns different weight to each expert, or the experts directly evaluate themselves. In objective techniques, mathematical approaches are used for this assignment. These approaches can use concepts such as distances between opinions or consistency between others. Another scheme proposed classification of objective methods [19]. This scheme divides methods by similarity of approach: index-based approaches, clustering-based approaches, and integrated approaches.

Nevertheless, these classifications are too broad and general to consider all current methods. In a scientific field where new approaches are constantly being proposed, and use of existing approaches extended by considering uncertainty and multiple data types, the lack of outline to locate the main operation of a particular method can overwhelm experts with information. Furthermore, lack of consensus on the nomenclature can lead to misreading or misunderstanding among inexperienced readers. It is common for multiple MCGDM techniques to use the same terminology, such as consensus index, support, similarity or credibility, to mean the same thing. At the same time, other academic papers may use the same terminology to refer to unrelated processes.

Additionally, an in-depth literature review has highlighted some points on MCGDM approaches that require attention. First, although social choice theory was one of the first methods to be developed, few references are made to it in the articles, even though they focus on the choice of alternatives. Second, the stage at which methods can be applied needs to be analysed, as there are clear differences between methods that do not require accurate information, techniques that need opinions and procedures that need to know the alternatives. Third, a study should explore whether these methods accept the existence of prior influence since it is rare that all decision-makers have equal importance.

This paper aims to provide a novel general GDM classification scheme and go into greater depth on the state-of-the-art concerning all experts’ weight assignment methods. For this purpose, the weight-assigned methods classification is further developed by a literature review. This new classification scheme will allow not only experts and scientists working in the field of MCGDM to know which methods exist but also explain the main features of each technique, such as the implementation phase, the data used and the main processes.

Furthermore, multiple MCGDM critical points are analysed to highlight future research pathways, providing information on gaps in current methods in each proposed category.

The remainder of this paper is divided into the following sections. Section 2 presents the proposed GDM method classification. Section 3 presents the literature review methodology and main reviewed characteristics. Section 4 outlines the literature review. Section 5 reports the results analysis and discussion, explaining future research pathways and their challenges. Conclusions are drawn in the last section.

2. Classification of group decision methods

Although the study focus on DM weighting techniques, the literature review has made it possible to detect and classify multiple types of GDM practices. Consequently, a general GDM methods scheme is presented which classifies the decision makers’ weighting techniques in detail. Future research can be applied to other clusters, although they are not part of this study. Other current schemes are presented in the first part of this section to help readers understand current classification practice.

2.1. Previous GDM classifications

Several classification schemes have been proposed to date. For example, one of the earliest divided methods into process- or content-oriented [22]. Process-oriented approaches are methods intended to generate new ideas to understand and structure the problem. Content-oriented approaches work with the opinions that are presented to aggregate them. This scheme was later improved with the second content-oriented layer [23] shown in Fig. 1, which divides the group into implicit multiple attribute evaluation, explicit multiple attribute evaluation and game-theory approaches.

Implicit multiple attribute evaluation refers to methodologies closely related to social theory techniques. Criteria, indicators, or requirements used for the selection are not explicitly stated by DMs. On the other hand, explicit multiple attribute evaluation clearly demonstrates the criteria and preferences of each DM. MCDM methods and their extensions fall into this category. Finally, game theory approaches are methods which follow traditional game theory. There are two game theory approaches: cooperative and non-cooperative. In GDM processes these methods mainly study coalitions among DMs.

Subsequently, the wide range of explicit multiple attribute evaluation methods made cluster upgrade mandatory. A new group named objective methods for deriving DM weight in GDM was created as part of explicit multiple attribute evaluation methods [19]. As shown in Fig. 2, this group is formed by similarity-based approaches, index-based approaches, clustering-based approaches, integrated approaches, and other approaches.

In this scheme, similarity-based approaches use the distance to weight the DMs. This distance can stem from DM assessments, or a fixed point usually named the ideal solution. Commonly, when a DM has a group-centered assessment it gains importance. Index-based approaches are divided by consensus or consistency approaches. The first group generates an index regarding the concept of “soft consensus” [24], which aims to improve the consensus between DMs allowing discussions or changes. The second group uses DMs’ assessment consistency to weight them. The more consistent a DM, the more weight it receives.

Clustering-based approaches are commonly used with large groups of DMs. These methods usually divide the process into two layers. The first layer consists of an expert division by cluster, and each cluster receives an initial weight. The second layer assigns an individual expert weight inside the cluster. The importance of the final DM is derived from the aggregation of both weights. However, these layers use different weighting methods (i.e., cluster size for first layer and similarity between experts for the second layer) which came from other approaches. Finally, integrated approaches are generated by aggregating two or more previously presented methods, and other approaches refer to any
In GDM, techniques have been developed from multiple fields (i.e., based on analysis of previous proposals and their limitations and inherent weaknesses. The scheme shown in Fig. 1 has an unduly broad approach regarding the large number of proposed methodologies. In GDM, techniques have been developed from multiple fields (i.e., psychology, economy, or politics), which demand a more accurate classification. Subjective or objective division has similar limitations. The scheme shown in Fig. 2 only considers objective methods and provides clusters that can be easily misread by inexperienced readers. As an example, soft consensus can be easily misread as similarity because one of the few differences is how weight evolves. Furthermore, cluster-based approaches are mainly generated by combining other methodologies, similar to integrated approaches.

### 2.2. Proposed classification of GDM methods

The proposed GDM methods classification scheme, shown in Fig. 3, is based on analysis of previous proposals and their limitations and shortcomings, found in the current GDM literature. This classification system encompasses all GDM methodologies in six groups, although only the DMs’ weighting methods cluster (the core of this manuscript) is developed. This therefore provides a clear primary guideline for other classification schemes. Developed decision maker’s weighting methods are shaded.

This method in which the DMs weight each other. These practices are developed further, is divided into whether the analysis is carried out on the experts, on their evaluation or if these two methods are merged. In expert analysis methods, expert characteristics, interrelationships, capacities, or coalition power indexes are the basis for the weighting process:

- The first subdivision contains external assessor or interpersonal assessment (EAIA) techniques. In these processes, there is an external assessor (or manager) who assigns weight directly to an expert panel. Furthermore, some techniques propose an interpersonal assessment method in which the DMs weight the other. These practices are commonly known as subjective methods.
- The second subdivision is focused on the study of DMs’ capabilities or the needs they present for the problem (CNA). These techniques generally use a common characteristic to compare the DMs such as experience and expertise, their university degree, or knowledge related to the problem field.
- The third and last subdivision is related to game theory approaches (GT). These methods are commonly used when a previous influence must be considered, and they study the DMs’ ability to form coalitions or to be the main promoter of a decision.

On the other hand, assessment analysis methods only study the opinion, assessment, or position that each DM has taken in the problem. Although there are simple works based on the distance between the opinions and a known statistical value [10–12], the methodologies have been developed to consider criteria multiplicity. This development has led to a wide range of methodologies and consideration of the following typologies:

- Group position methods (GrP) weight DMs in terms of their assessment distances within the group. The main objective is to reduce the distance by modifying the opinion weight to obtain the closest position among them.
- Group centroid (GC) methodologies consider the existence of an “ideal solution”, a real or virtual value which represents the group’s desired assessment. This value is named the group centroid. DMs are weighted regarding the distance between their assessment and this centroid.
- Optimization techniques (O) are generally built on these first two methods. By programming algorithms or using feedback, an iterative process is carried out with the aim of minimizing distances. It is worth mentioning that many of these methods present the goal in a consensus index to be maximized, calculated as the inverse distance between points.
- Quality assessment (QA) techniques weight DMs according to the quality of their opinions, generally compared within the group. This quality is estimated by mathematic values of entropy, certainty, or consistency, among others.

Data-aggregation methods analyze evaluations given by the DMs to be able to add them directly using operators. These operators may be closely related to the GrP, QA and GC of the previous cluster; however, its procedures do not contemplate giving explicit weight to DMs. Therefore, in many cases there are the same type of studies in the calculation of entropy, certainty, and consistency, which could be considered QA if they openly relate this weight to the experts. The same happens in the distance to the mean operators, which are extremely similar to GC, without giving weight to the experts.

Although these operators can be as simple as direct aggregation or AHP-derived results [29], operators for uncertainty using multiple data
types have also been developed, such as spherical fuzzy sets [30,31] or linguistic information [32–34]. These have become more important as they are included in the MCGDM field. Some authors use them as the main link to obtain input data for TOPSIS [35–37], ELECTRE [38,39], or TODIM [40] methods. Among others, there are operators that use the Dempster-Shafer theory of evidence (DST), in which each expert’s evaluation is treated as evidence [41].

Ranking aggregation methods can be subdivided according to their theoretical basis. In a first group, there are set-ordering probability distribution techniques that follow the Thurstone scale and the law of comparative judgment [42] or binary comparisons [13]. Heuristic methods fall into a second group, which instead of seeking optimization, focuses on obtaining a solution by simple and intuitive mathematical approximations. For instance, this division includes Borda methods and their extensions [14], as well as studies of government elections [15] or Markov chains [43,16].

Finally, a third group of stochastic optimization methods can be generated. These techniques synthesize the scenarios given by all DMs based on distance optimization criteria. The ideal final ranking is generally obtained by minimizing the Kendall or Spearman distance between each iterative approach and the DMs’ rankings. Some notable works in this field have been probabilistic approximations using Monte Carlo models [18], maximum probability Bayesian approximations [44] or uncertainty inclusion due to lack of reliability in incomplete rankings [45].

The final two groups are integrated and other methods. Integrated methods are generated by aggregating two or more previously presented methods, and other approaches refer to any that do not belong to any of the given categories.

It is important to emphasize the classification assigned for large group methods which use clustering systems. Commonly, these methods have been assigned to a specific group for them, as they use a double-weighting layer. The first one usually assigns a weight regarding the cluster size. The second one uses one of the previously presented methods. Hence, in this GDM framework large group methods with clustering systems are classified by the second layer usage, or the most appropriate technique.

3. Literature review methodology

This section presents the literature review methodology. A systematic review involves a five-stage structure [46]: (i) formulation of the problem, (ii) determination of the data collection strategy, (iii) evaluation of the retrieved data, (iv) analysis and interpretation of the literature and finally, (v) presentation of the conclusions.

To address the paper’s objective, the main question formulated in this study was (i): Which methods and techniques have been proposed to aggregate opinions and assessments in GDM environment? A data collection strategy was determined to answer this question. This strategy (ii) is shown in the following subsections, and it involves data gathering and a results classification process.

3.1. Gathering process

Articles were identified by the internationally-recognized bibliographic database Web of Science (WoS), which analyses articles from over 12,000 journals worldwide [47]. The use of this database was mainly justified by the depth of its coverage, yielding more outputs than any other database collection. Furthermore, it is a major multidisciplinary database of Thomson Reuters Web of Knowledge.

Having determined the database, two different research methods were used, as shown in Fig. 4. The first one focused on 5 general titles and topics related to GDM. The results obtained were restricted by a minimum of 1 citation (c≥1), given the general nature of the search. The second method centres on a specific search of closely associated DM weighting topics and a manual search in related journals. No minimum citation restriction was applied to the results.

13,717 papers were obtained thought the literature research method. These were later filtered to obtain the final set of articles on which to run the qualitative and quantitative analysis. The filtering process was conducted as follows: (i) discard duplicated articles; (ii) automatic eliminate papers if the title includes the words “operators” and/or “operators”. Operator methods commonly use a direct aggregation without assigning weight to the DMs. (iii) automatic selection of papers if the title includes “consensus”, “importance”, “weights” or “experts”. (iv) title and abstract analysis to detect potential results. (v) working frame analysis to identify DM weighting techniques.

As a result of this process, a final set of 208 references was selected for further analysis and interpretation. Among them, 201 different methodologies were identified and classified using the scheme mentioned above. The remaining 7 selected results concerning already-classified methods.

It is highlighted that article categorization was only based on the DM weighting process which was presented. As multiple references contain
In multiple analyses, the external expert is considered a moderator rather than a weight allocator for DMs. GP convergences refer to the common goal programming systems used by optimization methods. These weighting techniques are usually linked to a distance (deviation) minimization (GP-DO), a consensus index maximization (GP-IC) or a consistency index maximization (GP-CsI). In other cases, there are multiple system methods which use entropy minimization with consistency maximization (GP-ECsI), or entropy and distance minimization (GP-EDO).

It is important to emphasize that multiple authors use the same terminology for different purposes. Consensus indices are generally understood as distances between DMs’ assessments, but also between them and the ideal solution. Therefore, methods that consider the consensus index as a distance between DMs will be contained in GP-DO. Methods that consider this index as the distance to the ideal solution will be presented as GP-IC.

Finally, the last characteristic is the ideal or objective solution consideration. Generally, this solution is considered the mean of the group, either arithmetic or geometric. However, some articles use general uncertainty approximations (RA), the fixed-point theory (FPT) or the Fermat-Torricelli point (FT) to detect optimal consensus points, or group centroids among others.

4. Results and classification

This section presents and classifies the retrieved data and corresponds to the third Cooper stage (1989). The following sub-sections coincide with the previously stated weighting methods for DMs from the proposed classification scheme.

Hence, selected papers are directly categorized according to their group. Furthermore, the aforementioned categorization tables are provided at the end of each sub-section. It is important to mention that the results are ordered by methodology similarities rather than publication year.

4.1. External assessor or interpersonal assessment (EAIA)

In external assessor assessment methods, Keeney and Kirkwood [48] propose the use of a weighted additive social welfare function with all DMs’ opinions to aid the assessor with the weighting allocation. This idea was later developed by Kelemenis et al. [49] to allow multiple criteria. In that case, the assessor can give a different weight to each DM in each of the criteria.

Similar direct weighting approaches, in which the external assessor directly evaluates DMs, are proposed by Palomares and Martinez [50] and Tabatabei et al. [51], that name the external assessor as a manager or moderator. Hafezalkotob and Hafezalkotob [52] recommend that this manager assigns weight regarding DM talent, experience, and knowledge, like in a CNA methodology. Wang et al. [53] present a clustering method in which a second layer is assigned by the external assessor.

Other authors have developed tools to make the assessor’s work...
easier. For example, Wu et al. [54] develop a linear programming system to decide the DMs’ weight based on a comparison. Other authors use the Analytic Hierarchy Process (AHP) method in uncertainty to assign weight in personnel selection [55,56], or propose that the project leader uses the best-worst method (BWM) among DMs [57]. Gupta et al. [58] propose a voting system for a panel of managers who weight a panel of DMs on intuitionistic fuzzy information. Yang et al. [59] transfer this voting process to a single external assessor worst-best pairwise comparison system. Finally, Sun and Huang [60] propose a gray correlation adjustment tool to be applied once the assessor has decided on the weight. Its main objective is to minimize the negative effects that these may have on each expert’s objective.

Regarding interpersonal assessment methods, Bodily [61] presents a commitment model using Markov chains. This approach is based on how experts delegate power to external utility functions until a valid conclusion is reached for the whole team. However, experts’ scoring methods are more common than delegation approaches.

Ramanathan and Ganesh [62] use the original AHP to allow DMs to assess each other based on problem contexts. Van der Honert [63] uses multiplicative AHP (MHP) and SMART techniques to develop REMEMBRANT, an interpersonal evaluation software tool. Other models consider that score as a direct degree of confidence between experts [64, 65]. Liu et al. [66] use that degree of confidence assignation with a self-confidence index to allow a self-evaluation for each expert. However, this level of confidence among DMs can be analysed through social graph theories. Tian et al. [67] weight DMs regarding the centrality and connections of each DM’s opinion, considering closeness as confidence.

Finally, some EAIA methods introduce a feedback system. Dong et al. [68] propose an interpersonal multi-criteria evaluation concerning professional skills, cooperation, or fairness, among others. If the level of consensus is not reached, feedback to DMs is required so that they can modify their opinions.

4.2. Capabilities and needs analysis (CNA)

The only study that considers assessments as needs is Brock’s approach [69]. This author adopts a Nash theory approach to use Pareto’s limits and frontiers to assign a fair weight to each DM.

On the other hand, many studies have used analysis of expert capabilities to assign importance within the group. Herowati et al. [70,71] propose a DM discrimination and inconsistency capability analysis through repetitive evaluations. Discrimination is also studied by Cheng et al. [72], who analyze the incomplete responses of DMs to assign them a degree of experience according to their preference matrices. Chakhar et al. [73] use the DMs’ accuracy capability by comparing it against the team’s pre-ordered groups of preference. Chinnua et al. [74] assign the weight considering the DMs’ experience and knowledge regarding alternatives such as function, manufacturing, environment, economy, and society.

This experience has been widely applied in capability studies. Ivlev et al. [75] present a multi-attribute function considering each DM’s years of experience and education. Liu et al. [76] continue along this line by scoring DMs based on their professional qualification and work experience. Borissova [77] uses a similar function considering experience and level of competence in the problem context. Elbarkouy and Faye [78] send each DM a form to fill in that considers their years and diversity of work experience, position and time in the company, plus enthusiasm and willingness to participate. DMs are weighted by comparing the group forms. Sellak et al. [79] propose a degree of experience comprising the evaluation hesitation and DMs’ interests and preferences presented in the preference matrices. Bai et al. [80] develop an online database text algorithm to define DMs’ ability and experience based on their internet profiles, literature, and scientific achievements.

Other approaches are based on personal bias, self-confidence, and roles. For example, Chatterjee and Bhattacharyya [81] propose a probabilistic DM approach index based on three factors; capacity for precision regarding the majority opinion, subjective biases shown and the difficulty of the problem to be solved. Slevin et al. [82] propose a form to state the self-confidence of each DM in their respective assessments. Finally, Liao et al. [83] develop a role study for a large group of experts. Once divided into clusters, if one DM has different characteristics from the majority, it obtains more weight to balance the decision from different points of view.

4.3. Game theory approaches (GT)

Game theory methods start from a previously-given influence to analyze the participant’s power in a decision [84]. That power index (PI), understood as weight in this study, is considered the ability of a DM to be a decision promoter by coalition capability. GTs are usually divided into cooperative or non-cooperative or complete and incomplete depending on the exchange and type of information. However, the main characteristic in this study is how PIs are calculated. Hence, the main difference between methods is related to the consideration of winning coalitions.

Shapley and Shubik [85] show that all coalitions are possible, and PI is given by the number of times a DM generates a winning coalition, known as a pivot. Banzhaf [86] states that only winning coalitions should be used, and critical DMs should be studied. DMs are considered critical when a modification of their votes make the coalition lose its winning status. Deegan and Packel [87] introduce the minimum winning coalitions, which do not contain any unnecessary DMs. Felsenthal [88] modifies the minimum winning coalitions for minimum size winning coalitions as a winning DM aims to deal with the minimum number of possible participants in their coalition.

4.4. Group position methods (GrP)

Position among assessments can be considered a negative consensus effect. For example, some studies reduce DMs’ weight if their evaluations are further from the others [89], calling it ineffectiveness by means of thinking errors [90]. On the contrary, Xu and Zhou [91] use the maximum deviation method to reduce DMs’ weight. These authors consider that similar (closer) evaluations are ineffective to differentiate alternatives.

Nevertheless, assessment distances are generally considered a positive characteristic like compatibility [92] or correlation [93]. From these perspectives, the closest DMs receive the highest weight. To handle uncertain distances, Wu et al. [94] propose the use of a probability distribution with geodesic distance to approach the Manhattan distance. Lu et al. [95] incorporate the Dempster-Shafer theory (DST), in which assessments are considered evidence, into a probability function. Whang et al. [96] use the same theory in SLNI. Other authors [97,98] have used a similar DST procedure with Bayesian approaches [99]. This latest technique has been modied for use of TrGFI [100] or DHHFLI and q-ROFI [101,102].

Other GrP methods consider distance among assessments such as similarity [103–105], opinion transition facility among DMs [106,107], consensus degree index [108], or amount of support received [109,110]. Zang et al. [111] use the preference matrix distances to assign DMs’ weight and introduce the results to ELECTRE. Liu and Cheng [112] use a similar P-MNI measurement to obtain the input values to MABAC. Chen and Zou [113] include a reassurance factor based on the similarity of evaluations between experts to reinforce their shared positions.

However, other approaches are proposed with viewpoints beyond negative or positive effects. Gupta et al. [114] consider the distance as an advantage or a disadvantage. The more address matches it has with other assessments, the higher advantage score an assessment will receive. French [115] studies the assessment modification in a discussion to assign influence to each DM, measured by the systematic position variations throughout a debate. This is translated into weight regarding the distance covered in each assessment. Subsequently, Pérez [116]...
expands this study to include social interaction graphs.

Distance variation throughout a debate has been also studied by other authors. Xue et al. [117] analyze DMs who have approached a third-party position to increase their weight, as their confidence has been reinforced. Fu et al. [118] propose that the experts’ weight is given according to their reliability in a discussion process. In this approach, reliability is calculated according to the variation of the distances between assessments. Finally, Zhou et al. [119] use an approach which considers not only how DMs approach their position, but their persistence in defending their opinion.

4.5. Group centroid methods (GC)

Some authors have even studied the real centroid of a group assessment [12], where average is the most common option. Average deviation [120,121] and direct distance [122–124] are easy ways to consider a fair decision.

Nevertheless, average is also used as an indirect measurement. Ye [125] uses a scoring function to rate the differences between DM assessments and the average set, comparing them to assign the weight. Gitinavard et al. [126] use preference matrix deviations regarding the aggregate mean to obtain the least deviated matrix. This matrix is used to assign weight regarding original DMs matrices.

Other authors appoint this distance as a ‘similarity degree’. Chai et al. [127] use the similarity to assign the weight by comparison among DMs. Ye [128] uses a similar formulation to assign weight degree to DMs and criteria at the same time. This degree has been widely used to obtain TODIM input data regarding multiple information as TFI [129], P-HFI [130] or PLI [131] to adapt MCGDM to an existing MCDM.

These studies propose a double step; first, DMs receive an assigned weight and subsequently, the MCDM is applied. However, some studies have directly proposed a MCDM adaptation. Yue was inspired by Shih et al. [132] studies to introduce his first distance approaches [133], which led to his well-known extended TOPSIS or ETOPSIS [2]. Using a modified TOPSIS technique, this author develops a methodology based on closeness coefficients to adapt this MCDM to a MCGDM. Lately, this methodology has been adapted to II [134–136] and IFI [137], which have been used in other studies [138,139]. Other authors have extended Yue’s work modifying input data to OFI [140], PFI [141], MGLI [142] or IFI with Hamming distance [143], proposing new closeness coefficients [144–146] or with the use of fuzzy set theory approaches instead of the arithmetic average [147] for the ideal solution.

Other TOPSIS-related works include a similar Yue’s approach by Jiang y Wang [148], using ITRFII and the closeness coefficient, criteria disagreement between DMs [149], or HFI [150] and SCI adaptations [151].

Beyond existing MCDM adaptations, some novel MCGDM methods have been developed. Yue [152] presents a new method based on vectorial projections between experts and the ideal solution. Later, he extends this method to IFI [153]. As in ETOPSIS, other authors have worked with this method and have proposed new adjustments or extensions. Yang and Du [154] propose aggregating cosine for expert direction comparisons. Xu and Liu [155] and Sun [156] extend the method to II and IFI respectively. Liao et al. adapt the method to IFI [157] considering the Euclidian and phylogenical distances, used by Luo et al. [158] to obtain MULTIMOORA input data.

Boix-Cots et al. [159] develop HIVES, a novel methodology which uses social theory constraints and statistical mathematics to assign DM weight. This technique uses the social ideal consensus point (SICP) as the ideal solution, which corresponds to the arithmetic mean of the second and third data quartiles. The value that this obtains ensures that bias is avoided and equals the optimal value for the majority. This method can manage previously assigned influence on each DM or source and maximizes the response (or DM satisfaction), as it analyses MCGDM problems criterion by criterion.

Finally, Gong et al. [160] propose to give the weight by obtaining two factors based on the ranking analysis method, which studies the preference position of each expert with respect to their aggregate mean.

4.6. Optimization methods (O)

There are methods based on optimizing these GrP and GC distances. In these cases, the weight takes on a dynamic form until the distance has been minimized or a consensus index threshold is reached. Nevertheless, the consensus index is usually distance-based.

Among the GrP DMs distance or deviation optimization systems, some studies use programmed minimization models under the pretext of decreasing discordance and increasing consensus [161–165], or to increase the support index, generated by minimizing the distance among DMs [166]. On these, some authors have proposed approximation systems for the same purposes [167]. Other minimization models are focused on the quadratic distance between DMs’ preference matrices [168,169].

There are techniques that focus analysis on the assessed alternatives. Meng et al. develop a quadratic programming model that maximizes the consensus index for each evaluated alternative [170]. Regarding rankings, Li et al. program a distance minimization system [171] and Zhang and Guo program a variance minimization system [172].

Some methods introduce feedback. For instance, Reagan et al. propose that DMs get a degree of respect according to the distance between them [173]. If the level of consensus with the final solution is not satisfactory, the opinion can be iteratively modified. Similar models are developed by Jin, Dong, and Cooper to minimize distances and improve consensus [174,175] through voluntary assessment modification.

Other optimization systems try to decrease the distance between DMs and the aggregated solution. In these cases, the consensus index uses this distance as a basis instead of the distances between DMs, as in GC. The clearest examples are programming dynamic weight in TOPSIS to minimize the distance to the ideal solution [176] in HFI and the minimization of distances to the aggregated solution through programming [177].

However, some studies consider that the aggregated matrix is calculated as the mean of the results obtained in each iteration, so they try to reduce their distances [178–180] or deviations [181]. Others consider the mean as the starting point [182] to calculate the initial weight, using iterative aggregation to calculate the optimization objective matrix [183,184]. Zhang and Xu consider that instead of the direct distance between matrices, the degrees of jointness and interaction should be used to perform the computation [185,186] according to the dominance of alternatives.

These differences between data are also known as discordance. Xu [187] develops two linear programming systems to minimize it by reducing the number of adjustments suffered by the DMs due to the weight, a technique that has been applied to obtain TOPSIS data [188]. TOPSIS has also been used in conjunction with programming models to minimize both distances; among DMs’ assessments and between them and the aggregate solution [189].

Other authors use attribute values as a prior step. For example, Yu and Lai propose that DM weight is decided once the attributes are aggregated using an operator [190], to minimize the quadratic distance and discordance. One key point of this study is that the level of consensus is used to indicate whether the aggregation method should be varied. Chen et al. use this model as a basis for developing a minimum-maximum attribute weight optimization method [191]. Taking this operator use idea, Li et al. develop nonlinear programming in MATLAB for the minimization of discordances in a DM cluster [192] in an IVIF uncertainty environment. Among other characteristics, Wan et al. use the degrees of fuzzy membership for each assessment component to optimize distances by minimizing discordance [193].

Some studies propose that these methods should be accompanied by a feedback and re-evaluation system [194–196]. If the consensus indices do not exceed the thresholds, the DMs must modify their assessment at
the marked critical points, to iterate the process.

There are also techniques focused on large groups of DMs. Tan et al. propose that weight should be assigned according to consensus and consistency. Taking the mean as a reference, consensus is calculated as the difference between distances and consistency as the difference of the deviations in the linguistic evaluation [197]. If the minimums are not reached, feedback is sent to the most biased DMs. Rodríguez et al. and Tang et al. [198,199] propose that DM weight depends on both group size and the displayed cohesion, calculated analogously to the similarity with the cluster centroid. This centroid is initially calculated as a random preferred alternative. If a minimum consensus is not reached, individual and group feedback is generated so the opinions are modified.

Finally, there are methods with uncommon consensus- and ranking-based methods. In the former, Yang and He use the fixed-point theorem for the optimization calculation of DM weight [200]. According to this theorem, a solution will converge when the iterative change of DM weight is imperceptible. In the latter, Ben-Arieh and Chen present a consensus level based on the distance between the final aggregate ranking and DM rankings [201]. However, weight is modified according to its contribution to the group, measured as the level of consensus reached in the group without its participation.

4.7. Quality assessment methods (QA)

Quality assessment methods use comparison indices not based on distances, but on the properties of the assessments themselves to weight DMs. These indices can be optimized as Optimization methods are with distances. One of the most-frequently used is the consistency index developed by Saaty, used directly by several studies in this review in multiple data types such as CI [202], ZTLI [203], or IFI [204].

The method proposed by Toloie-Eshlaghy and Farokhi is similar, based on the iterations needed to obtain the convergence vector [205]. Wu et al. [206] extend these studies by introducing an amplification parameter to the consistency of evaluations to differentiate experts with similar opinions. The optimization has also been introduced in consistency maximization. Xu and Cai [207] program an algorithm that maximizes group consistency by minimizing the weight of low-consistency evaluations. Liu et al. [208] present an optimization system adapted to the uncertainty arising from the alternatives themselves in autonomous bootstrapping systems. In this case, he uses an entropy and weight modification system capable of increasing DM consistency.

The Shannon entropy, understood as the lack of information contained in each of the evaluations, has also been a widely-used index [209–213] to assign weight to DMs. Cheng et al. [214] use a relative entropy minimization algorithm to deal with incomplete information, which makes it possible to find and minimize entropy distances. Cai and Balaman [215] use the entropy to find the DMs’ degree of divergence, subsequently using ELECTRE and VIKOR. Li et al. [216] use a scoring factor instead of the degree of divergence between assessments to relate entropy to the weight, used to introduce TODIM. This entropy has been used in conjunction with the hyper entropy, considering the result as DMs’ personalities [217] in a LI, P-LI, and LHI study.

Multiple studies have proposed other quality assessment indices. For example, Jin et al. [218] use a Sugeno integral to analyze the information linked to evaluations to find its degree of certainty. Liao et al. [219] develop a scoring formulation to calculate the quality of fuzzy evaluation, named degree of hesitation in the expert’s assessment. Wang et al. [220] interrelate the multiplicative preference matrices with the characteristic evaluation matrices using a quadratic logarithm to obtain a degree of information reliability. To solve the interrelated system, he presents a difference minimization algorithm that increases the degree of reliability.

Finally, there are cluster methods which use a double quality assessment. Ma et al. [221] use a combination of familiarity and cognitive reliability related to consistency. An optimization model to find the minimum entropy of the group is used for the familiarity calculation. Then reliability is added, calculated as cluster consistency. Zhou et al. [222] use entropy and consistency of the cluster members’ evaluations, added to the cluster size, to obtain DM weight.

4.8. Integrated methods

The latter group corresponds to methods that incorporate two or more of the above-mentioned techniques. One of the most used combines DM distances and quality assessments. For instance, Qi et al. [223] use a quality assessment study with a degree of uncertainty, subsequently modified by a degree of divergence between evaluations to obtain the final DM weight. Liu et al. [224] use a similar idea to calculate large groups of DMs, which are weighted by a double-layer system by entropy and minimum variance method. Zhang and Chen [225,226] use the consistency and proximity of assessments. Wang et al. [227] use the similarity degree among DM rankings, modified by a degree of support for when these rankings differ in validity.

Some studies propose including optimization in these techniques. Chen et al. [228] use the consistency degree adding a consensus index as a benchmark in an iteration system. This consensus refers to distances between DMs, and if it is not less than the threshold, DMs must modify their evaluations. Zhang et al. [229] propose a method to maximize group evaluation consistency while minimizing DM distances. Qi et al. [230] propose that the weight is given by the quality of the evaluation, which is modified by a minimization algorithm for distances between DMs. There are similar studies which use these double programming methods to minimize the hamming distance between DMs and the aggregated mean and to minimize each DM’s weighted entropy [231]. For high assessment uncertainty, Xu et al. [232] propose a double algorithm to minimize entropy and distances between DMs.

Quality assessment methods are also used with GC techniques. Pang et al. [233] use the uncertainty degree of assessments with their mean and limit matrix distances. Liu et al. [234] modify the attributes using the assessments’ entropy, from which a media value is obtained to compare DMs. Liu et al. [235] use a double layer weight assignment for large groups of DMs. In this case, the first layer uses an entropy index, which is added to the minimum variance model index obtained in the second layer. Li et al. [236] develop two cross-entropy programming models which are simplified to exact equations to calculate experts’ professional experience and expertise using the aggregate group mean as a reference. Even though it can be misread as a CNA method, professional experience is obtained by the entropy deviation and expertise by the direct distance among DM assessments and the media. A similar system is proposed by Pramanik et al. [237], with an ideal matrix composed of minimum and maximum values, regarding cost or benefit indicators, used to apply cross-entropy.

GrP are also used with GC techniques. Davoudabadi et al. [238] apply similarity among DMs to assign a first step weight, which leads to an objective matrix. Later, distance to the objective is applied to obtain the modified weight. Wan et al. [239] suggest adding the DM similarity degree technique to modify the ideal ETOPSIS solution.

A group of interesting integrations mix assessment analysis with expert analysis techniques. In this context, it is common to use GC, Chen et al. [240] aggregate a weight assigned by the problem organizers with a weight obtained by the distance with the geometric mean. Jabeur et al. [241] suggest assigning a weight regarding an interpersonal ranking of the importance, which is modified by a distance optimization system between individual and final rankings. Yang et al. [242] use an interpersonal AHP evaluation to obtain the aggregated matrix, to calculate the DM cosine degree of similarity and obtain VIKOR input data.

Other models use both mean matrices and minimum-maximum matrices. Liu et al. [243] propose to include initial subjective weight to modify ETOPSIS mean matrix. Mohagheghi’s suggestion is similar, with an initial weight according to the importance in their area of expertise to modify TOPSIS values [244].
GrP appears in several studies using expert analysis techniques. Mianabadi and Afshar [245] integrate the evaluation of a group manager with the analysis of DM distances. Xu [246] extends Bodily’s interpersonal evaluation contribution with a measure of linguistic preference deviations that modifies the initial weight according to similarity between DMs. Ölçer and Odabaşı [247] use a subjective comparison matrix which scores the importance of each expert, and the difference between their evaluations. Wan et al. develop a multi-objective programming to minimize distances between DMs and their thrust degree [248]. Bai et al. [249] propose the use of age, level of education, position, and experience to draw an initial weight, which he modifies according to DM distance.

There are proposed approaches which integrate both distance methods. Liu and Li [250] use an initial subjective weight, which is modified according to the distance between DMs. It is subsequently adjusted regarding the distance between the DM and the optimal Fermat-Torricelli center-point distance. Chen et al. [251] use the distance between DMs to obtain the modified Shapley index value, which is compared with the distance to their mean.

Finally, there are expert-expert analysis integration methods. Liu et al. [252] suggest adding an initial weight assigned by an organizer with a degree of self-confidence stated by the DMs themselves. Ren et al. [253] use the degree of professionalism to evaluate experts according to how they give information, adding assessment entropy methods.

5. Analysis and discussion

This section analyses the selected literature and corresponds to the fourth Cooper stage [46]. The analysis contains discussion on a generalized problem, a comparison and interpretation of the articles and their classifications, and an analysis of internal characteristics linked to decision-making.

The pervasive problem that has been found throughout the literature has been the need to adopt a common framework in the methods of assigning weight to DMs. The lack of unanimity in their classification and in the existing categories promotes the existence of ambiguous studies. In this study, the author’s point of view has been prioritized, so the explanations and terminologies used by the authors have been paramount for the classification. Also, the approaches presented in the studies have been determining for the category assignment.

For example, some EAIA methods encourage the external evaluator to give weight according to CNA capabilities. Another ambiguity is found in CNA methods that use QA bases but refer to the DM capacities. In this case, the discrimination and consistency analysis capabilities [70] or the contributions provided [73] could easily be interpreted as the second category.

The major sign of the currently general disagreement are these authors’ terminologies. Beyond the framework, multiple uses and explanations of indices or characteristics that share the same root have been detected. The same data receives different names and explanations, as shown in Fig. 5, whether from subjective or objective sources.

Irrespective of whether the distance is calculated between direct assessments, using the preference matrix or by analysing the rankings, multiple terminologies refer to the same distance or even both. For novice researchers or non-experts in the field this can lead to misunderstandings and misconceptions.

Hence, a simplification beyond the 2.2 section classification should be adopted. Table 11 presents two indices that refer to the various distances being considered. The first is the consensus index among experts (CIE), which brings together all those terminologies that refer to the calculation among DMs. For example, one study can seek to improve the CIE by reducing the distances among experts or differences among
assessments’ characteristics, such as entropy. Analogously, the second proposal is the consensus index to the solution (CIS). This index is used to unify the terminologies that refer to the calculation among each DM and the ideal solution, whether using distances or assessments’ characteristics. Adopting these indices, or their inclusion in subsequent studies, might be of great help to the community.

Regarding the results comparison and interpretation, there has been a notable increase in MCGDM-proposed methods. In Fig. 6A, categorized results are shown by decade, from 1950 until after 2020. Fig. 6B uses that information to demonstrate the category dominance per decade. It is interesting to note that the earliest dates back to 1950, when Kenneth Arrow’s social choice theory (1951) seemed to spur interest in opinion aggregation methods. GT techniques quickly appeared to mainly solve economic power problems [85], and the first GrP aggregations [115] were proposed to solve discussion processes. This dominance was extended during the 1960s, when some authors discussed existing techniques and proposed new GT [86] and GrP [90] approaches.

In the 1970s, 1980s and 1990s, external [48] or interpersonal [61, 62] assessments, novel GT approaches [87], and CNA studies [69,82] were proposed. They were all part of what we consider to be subjective methods.

However, as the 21st century began, a paradigm shift took place. The trend has shifted in favor of more objective methods for several reasons. Although expert analysis assessments may be easier to apply, EAIA methods can be subject to dishonesty or bias in interpersonal or supra-evaluator assessments. In addition, unfairness can also arise in CNA and GT techniques when DMs have information from their competitors.
26

Presenting unrealistic needs, use of certain criteria that favor some DMs in advance or the preconception of coalitions through the delegation of power can lead to a biased consensus. Therefore, assessment analysis methods which ignore DMs and only look at their evaluation have been proposed as possible solutions to these problems. These techniques have grown exponentially since 2000, with major contributions on GrP, O, QA, and GC procedures, and the latter has been particularly noteworthy.

In fact, this development of assessment analysis methods might be directly related to the growing interest shown in MCDM methods. As the use of MCDM in multiple alternative analysis has recently expanded due to its objectivity, adding these methods to group decisions gives the process further objectivity.

Several articles use their proposed methods in conjunction with MCDM. Fig. 7 shows the MCDM used regarding each category. Even though numerous MCDMs have been applied TOPSIS and VIKOR are the most frequently used techniques, while other methods have few appearances. This fact is surprising, considering the large number of MCDM techniques constantly applied in all scientific fields. Regarding categories, the GC wide application is remarkable. This is mainly due to Yue’s work on Extended TOPSIS and Projection, which appear multiple times in the categorization. On the other hand, the lack of GC techniques applied with VIKOR is remarkable.

It is precisely this widespread GC use that leads the discussion to another interesting point. Although multiple possibilities have been proposed as an ideal solution, data shown in Fig. 8 make the general preference clear. This figure considers methods which use an ideal solution and separates the percentages according to which point is considered.

The mean is by far the most used statistical point to assume consensus by 56 out of 67 articles, representing an astonishing 83.58%. On the other hand, few articles use other approximations and percentages are widely distributed. Two articles (2.99%) use the mean affected by a previous weight [243,244] and 4 (5.97%) consider the consensus by majority [81,83,235,246]. RA [147], FTP [200], SICP [159], Max/Min [237], and FT [250] are only used by 1 article each, representing 1.49% per approximation.

The high degree of acceptance of the mean as an ideal solution is striking, even though some works oppose its unquestionable use [12]...
and propose alternatives that improve consensus or present the potential loss of information produced by its use [163]. It should be noted that many methods do not use the mean as a direct solution, but as an intermediate step or as an auxiliary tool to achieve other objectives. For instance, integral methods are used to avoid the loss of information produced by its use [163]. It should be noted that many methods do not use the mean as a direct solution, but as an intermediate step or as an auxiliary tool to achieve other objectives. For instance, integral methods are used to avoid the loss of information produced by its use [163].

However, even though scarce studies present social considerations in the procedures as shown in Fig. 9, many lack their explicit incorporation in the social context. As was noted, these axioms are used to force mathematical methods to contain a social basis that gives them some sociological weight and avoids situations of injustice, promoting equality.

Nevertheless, the assessment analysis results are astonishing. Only 4 out of 132 (3.03%) articles present social considerations, even though they are developed to work with phase 1 of the GDM process. The last comparison and interpretation of results and their classification is presented in the characteristic ‘Social considerations’ of Table 10. The clear value dominance can be observed in most categories, except for GC and Integrated methods. This may be because the main objective of this study has been to classify new methods, which are usually presented with exact values. Subsequently, their complexity increases with new data. As many GC methods use the aggregate mean, several proposals are based on modifying the data used on this value. A clear example is the multiple input data modifications of methods such as ETOPSIS and projection. Integrated methods, on the other hand, generally start from combining existing methods and can therefore move on to other data. It is remarkable that all the GT methods are concentrated in the CI category. The scarce information obtained by the search process should be mentioned, and this would have to be expanded.

The last comparison and interpretation of results and their classification is the point at which these methods are applied in the GDM process. Fig. 11 shows the articles phase application percentages regarding each category. It is noted that phase 0 articles are contained in the CI category. The scarce information obtained by the search process should be mentioned, and this would have to be expanded.
opinion analysis (16 out of 132). Furthermore, 6 out of these 16 articles are related to QA, as the method compares assessment characteristics. Thus, only 5, 1 and 4 phase-1 articles are GrP, GC and O, respectively.

There are many cases where alternatives are not known or may vary. Therefore, methods that might be applied without awareness of the alternatives (phase 0 and 1) are an interesting aspect to develop. Remember that methods using alternatives have been given phase 2 consideration or role contemplation), assessment analysis technique characteristics. Thus, only 5, 1 and 4 phase-1 articles are GrP, GC and O, respectively.

Finally, two internal method characteristics have been analysed due to their close connection with real-world GDM. Previous influence acceptance is one. In actual decision-making, it is not unusual for experts to differ in importance. Examples of this might be different business experience or knowledge. The other is a component division process. Some studies have stated that dividing the previous influence treatment. Finally, in Integrated methods, the Shapley power index is used as the previous influence treatment. In this paper, a comprehensive analysis of GDM methods has been carried out to provide an in-depth state-of-the-art understanding on decision makers’ weight assignment methods. In addition, the knowledge gained through the literature review has made it possible to propose a new classification scheme for all existing and future MCGDM methods. The development of this classification will allow scientists developing new methods to place their approach within a classification group, thus gaining insight into existing advances in the field. Furthermore, for the first time, experts will be able to get a clear picture of the

6. Conclusion

In this paper, a comprehensive analysis of GDM methods has been carried out to provide an in-depth state-of-the-art understanding on decision makers’ weight assignment methods. In addition, the knowledge gained through the literature review has made it possible to propose a new classification scheme for all existing and future MCGDM methods. The development of this classification will allow scientists developing new methods to place their approach within a classification group, thus gaining insight into existing advances in the field.
state of play of MCGDM methods, clearly displaying the main features of their procedures and the data they work with. Furthermore, six issues for future research pathways have been highlighted:

First, the need for a unified working framework has been outlined. Ambiguity, misunderstandings, or simple misreading can arise if multiple, interlocking concepts continue to be used. Hence, two indices have been presented to help the research community, novel researchers, or non-experts in GDM.

Second, the use of the mere mean as direct ideal solution should be avoided. Several authors have questioned the veracity of this objective.

Table 8
Optimization methods cataloging table.

<table>
<thead>
<tr>
<th>Author</th>
<th>Year</th>
<th>Data</th>
<th>Phase</th>
<th>Weight</th>
<th>Convergence</th>
<th>Consensus</th>
</tr>
</thead>
<tbody>
<tr>
<td>Li et al.</td>
<td>2017</td>
<td>HFI</td>
<td>Phase 2</td>
<td>Dynamic</td>
<td>GP-DO</td>
<td>–</td>
</tr>
<tr>
<td>Lin and Wang</td>
<td>2018</td>
<td>IFI</td>
<td>Phase 2</td>
<td>Dynamic</td>
<td>GP-DO</td>
<td>–</td>
</tr>
<tr>
<td>Ma et al.</td>
<td>2020</td>
<td>MGLI</td>
<td>Phase 2</td>
<td>Dynamic</td>
<td>GP-DO</td>
<td>–</td>
</tr>
<tr>
<td>Xu and Wu</td>
<td>2013</td>
<td>2TLI</td>
<td>Phase 2</td>
<td>Dynamic</td>
<td>GP-IC</td>
<td>–</td>
</tr>
<tr>
<td>Xu and Cai</td>
<td>2012</td>
<td>CI</td>
<td>Phase 2</td>
<td>Dynamic</td>
<td>GP-DO</td>
<td>–</td>
</tr>
<tr>
<td>S. Wan et al.</td>
<td>2021</td>
<td>EMO</td>
<td>Phase 2</td>
<td>Dynamic</td>
<td>GP-DO</td>
<td>–</td>
</tr>
<tr>
<td>Abootalebi et al.</td>
<td>2018</td>
<td>CI</td>
<td>Phase 2</td>
<td>Dynamic</td>
<td>GP-DO</td>
<td>–</td>
</tr>
<tr>
<td>Xu et al.</td>
<td>2016</td>
<td>LI</td>
<td>Phase 2</td>
<td>Dynamic</td>
<td>GP-DO</td>
<td>–</td>
</tr>
<tr>
<td>Xu et al.</td>
<td>2017</td>
<td>HFI</td>
<td>Phase 2</td>
<td>Dynamic</td>
<td>GP-DO</td>
<td>–</td>
</tr>
<tr>
<td>Meng et al.</td>
<td>2017</td>
<td>IFI</td>
<td>Phase 2</td>
<td>Dynamic</td>
<td>GP-DO</td>
<td>–</td>
</tr>
<tr>
<td>Li et al.</td>
<td>2015</td>
<td>CI</td>
<td>Phase 2</td>
<td>Dynamic</td>
<td>GP-DO</td>
<td>–</td>
</tr>
<tr>
<td>Zhang and Guo</td>
<td>2016</td>
<td>CI</td>
<td>Phase 2</td>
<td>Dynamic</td>
<td>GP-DO</td>
<td>–</td>
</tr>
<tr>
<td>Regan et al.</td>
<td>2006</td>
<td>CI</td>
<td>Phase 2</td>
<td>Dynamic</td>
<td>GP-IC</td>
<td>–</td>
</tr>
<tr>
<td>Dong y Cooper</td>
<td>2016</td>
<td>CI</td>
<td>Phase 2</td>
<td>Dynamic</td>
<td>GP-DO</td>
<td>–</td>
</tr>
<tr>
<td>Ji et al.</td>
<td>2021</td>
<td>CI/PI</td>
<td>Phase 2</td>
<td>Dynamic</td>
<td>GP-DO</td>
<td>–</td>
</tr>
<tr>
<td>W.-C. Zou et al.</td>
<td>2022</td>
<td>HFI</td>
<td>Phase 2</td>
<td>Dynamic</td>
<td>GP-DO</td>
<td>Mean</td>
</tr>
<tr>
<td>A. H. Liu et al.</td>
<td>2021</td>
<td>LI</td>
<td>Phase 2</td>
<td>Dynamic</td>
<td>GP-DO</td>
<td>Mean</td>
</tr>
<tr>
<td>Meng et al.</td>
<td>2016</td>
<td>I2TLI</td>
<td>Phase 2</td>
<td>Dynamic</td>
<td>GP-DO</td>
<td>Mean</td>
</tr>
<tr>
<td>S. Wan et al.</td>
<td>2022</td>
<td>P-LI</td>
<td>Phase 2</td>
<td>Dynamic</td>
<td>GP-DO</td>
<td>–</td>
</tr>
<tr>
<td>Wan et al.</td>
<td>2016</td>
<td>IFI</td>
<td>Phase 2</td>
<td>Dynamic</td>
<td>GP-IC</td>
<td>Mean</td>
</tr>
<tr>
<td>Zhang et al.</td>
<td>2020</td>
<td>qROI</td>
<td>Phase 1</td>
<td>Dynamic</td>
<td>GP-DO</td>
<td>Mean</td>
</tr>
<tr>
<td>Lee</td>
<td>2002</td>
<td>FI</td>
<td>Phase 1</td>
<td>Dynamic</td>
<td>GP-DO</td>
<td>Mean</td>
</tr>
<tr>
<td>Li and Wang</td>
<td>2019</td>
<td>HFI</td>
<td>Phase 1</td>
<td>Dynamic</td>
<td>GP-DO</td>
<td>–</td>
</tr>
<tr>
<td>Ferreira et al.</td>
<td>2010</td>
<td>LI</td>
<td>Phase 1</td>
<td>Dynamic</td>
<td>GP-DO</td>
<td>–</td>
</tr>
<tr>
<td>Zhang and Xu</td>
<td>2014</td>
<td>IFI</td>
<td>Phase 1</td>
<td>Dynamic</td>
<td>GP-IC</td>
<td>Mean</td>
</tr>
<tr>
<td>Zhang and Xu</td>
<td>2015</td>
<td>IFI</td>
<td>Phase 2</td>
<td>Dynamic</td>
<td>GP-IC</td>
<td>Mean</td>
</tr>
<tr>
<td>Xu</td>
<td>2018</td>
<td>IFI</td>
<td>Phase 2</td>
<td>Dynamic</td>
<td>GP-IC</td>
<td>Mean</td>
</tr>
<tr>
<td>S. Wan et al.</td>
<td>2021</td>
<td>P-LI</td>
<td>Phase 2</td>
<td>Dynamic</td>
<td>GP-DO</td>
<td>–</td>
</tr>
<tr>
<td>Yu and Lai</td>
<td>2011</td>
<td>CI</td>
<td>Phase 2</td>
<td>Dynamic</td>
<td>GP-IC</td>
<td>Mean</td>
</tr>
<tr>
<td>Chen et al.</td>
<td>2015</td>
<td>CI</td>
<td>Phase 2</td>
<td>Dynamic</td>
<td>GP-IC</td>
<td>–</td>
</tr>
<tr>
<td>Li et al.</td>
<td>2017</td>
<td>IFI</td>
<td>Phase 2</td>
<td>Dynamic</td>
<td>GP-DO</td>
<td>Mean</td>
</tr>
<tr>
<td>Dong et al.</td>
<td>2016</td>
<td>CI</td>
<td>Phase 2</td>
<td>Dynamic</td>
<td>GP-IC</td>
<td>–</td>
</tr>
<tr>
<td>Pang et al.</td>
<td>2017</td>
<td>IFI</td>
<td>Phase 2</td>
<td>Dynamic</td>
<td>GP-IC</td>
<td>Mean</td>
</tr>
<tr>
<td>Zhang et al.</td>
<td>2016</td>
<td>IFI</td>
<td>Phase 2</td>
<td>Dynamic</td>
<td>GP-IC</td>
<td>Mean</td>
</tr>
<tr>
<td>Tan et al.</td>
<td>2020</td>
<td>CI</td>
<td>Phase 2</td>
<td>Dynamic</td>
<td>GP-IC</td>
<td>Mean</td>
</tr>
<tr>
<td>Rodríguez et al.</td>
<td>2018</td>
<td>HFI</td>
<td>Phase 2</td>
<td>Dynamic</td>
<td>GP-IC</td>
<td>–</td>
</tr>
<tr>
<td>Tang et al.</td>
<td>2020</td>
<td>CI</td>
<td>Phase 2</td>
<td>Dynamic</td>
<td>GP-IC</td>
<td>–</td>
</tr>
<tr>
<td>Yang and He</td>
<td>2021</td>
<td>IFI</td>
<td>Phase 2</td>
<td>Dynamic</td>
<td>GP-DO</td>
<td>FPT</td>
</tr>
<tr>
<td>Ben-Arieh and Chen</td>
<td>2006</td>
<td>IFI</td>
<td>Phase 2</td>
<td>Dynamic</td>
<td>GP-IC</td>
<td>Mean</td>
</tr>
</tbody>
</table>

Table 9
Quality Assessment methods cataloging table.

<table>
<thead>
<tr>
<th>Author</th>
<th>Year</th>
<th>Data</th>
<th>Phase</th>
<th>Weight</th>
<th>Convergence</th>
<th>Consensus</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wu et al.</td>
<td>2018</td>
<td>CI</td>
<td>Phase 1</td>
<td>Static</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Cabrerizo et al.</td>
<td>2010</td>
<td>2TLI</td>
<td>Phase 2</td>
<td>Static</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>S. Wan et al.</td>
<td>2020</td>
<td>IFI</td>
<td>Phase 2</td>
<td>Static</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Toloie-Eshlaghy and Farokhi</td>
<td>2011</td>
<td>CI</td>
<td>Phase 1</td>
<td>Static</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Wu et al.</td>
<td>2012</td>
<td>CI</td>
<td>Phase 1</td>
<td>Static</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Xu and Cai</td>
<td>2014</td>
<td>IFI</td>
<td>Phase 2</td>
<td>Dynamic</td>
<td>GP-CdI</td>
<td>–</td>
</tr>
<tr>
<td>Liu et al.</td>
<td>2013</td>
<td>IFI</td>
<td>Phase 1</td>
<td>Dynamic</td>
<td>GP-CdI</td>
<td>–</td>
</tr>
<tr>
<td>Hashemi et al.</td>
<td>2018</td>
<td>IFI</td>
<td>Phase 1</td>
<td>Static</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>RazaviFossi and Samani</td>
<td>2019</td>
<td>IFI</td>
<td>Phase 1</td>
<td>Static</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Wu et al.</td>
<td>2020</td>
<td>LI</td>
<td>Phase 2</td>
<td>Static</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Ye Jun</td>
<td>2013</td>
<td>IFI</td>
<td>Phase 2</td>
<td>Static</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Yue Chuan</td>
<td>2017</td>
<td>CI</td>
<td>Phase 2</td>
<td>Static</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Cheng et al.</td>
<td>2017</td>
<td>LI</td>
<td>Phase 2</td>
<td>Dynamic</td>
<td>GP-EDO</td>
<td>–</td>
</tr>
<tr>
<td>Cali and Balaman</td>
<td>2019</td>
<td>IFI</td>
<td>Phase 2</td>
<td>Static</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Li et al.</td>
<td>2015</td>
<td>IFI</td>
<td>Phase 2</td>
<td>Static</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Mao et al.</td>
<td>2022</td>
<td>LI / P-LI / LHI</td>
<td>Phase 2</td>
<td>Static</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Jin et al.</td>
<td>2018</td>
<td>IFI</td>
<td>Phase 1</td>
<td>Static</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Liao et al.</td>
<td>2019</td>
<td>IFI</td>
<td>Phase 2</td>
<td>Static</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Wang et al.</td>
<td>2013</td>
<td>IFI</td>
<td>Phase 2</td>
<td>Dynamic</td>
<td>GP-CdI</td>
<td>–</td>
</tr>
<tr>
<td>Ma et al.</td>
<td>2019</td>
<td>LI</td>
<td>Phase 2</td>
<td>Dynamic</td>
<td>GP-CdI</td>
<td>–</td>
</tr>
<tr>
<td>Zhou et al.</td>
<td>2011</td>
<td>CI</td>
<td>Phase 2</td>
<td>Dynamic</td>
<td>GP-CdI</td>
<td>–</td>
</tr>
</tbody>
</table>


consider that the alternatives will be known. In a world where possibilities are growing for equal opportunities. However, most methods do not consider them. Group of experts, it seems obvious that there should be social guidelines from sociological theories, such as social choice theory. Since statistic and have analysed other possibilities. Others have applied novel procedures or proposed the use of other, more suitable geometric points. Fourth, novel methods must also promote approaches in the absence of concepts and combinations. It is precisely the increasing complexity of today’s world problems which leads to considering multiple experts or data sources and increasingly more criteria. These experts or sources may not have the same weight, and some studies show an improvement in the decision when analysing the criteria separately.

Table 10
Integrated methods cataloging table. The methods characteristic shows the techniques used for each item.

<table>
<thead>
<tr>
<th>Author</th>
<th>Year</th>
<th>Data</th>
<th>Phase</th>
<th>Weight</th>
<th>Convergence</th>
<th>Consensus</th>
<th>Methods</th>
</tr>
</thead>
<tbody>
<tr>
<td>Qi et al.</td>
<td>2015</td>
<td>IFI</td>
<td>Phase 2</td>
<td>Static</td>
<td>–</td>
<td>–</td>
<td>GrP + QA</td>
</tr>
<tr>
<td>Liu et al.</td>
<td>2015</td>
<td>2TLI</td>
<td>Phase 2</td>
<td>Static</td>
<td>–</td>
<td>–</td>
<td>GrP + QA</td>
</tr>
<tr>
<td>Zhang and Chen</td>
<td>2019</td>
<td>IFI</td>
<td>Phase 2</td>
<td>Static</td>
<td>–</td>
<td>–</td>
<td>GrP + QA</td>
</tr>
<tr>
<td>Wang et al.</td>
<td>2015</td>
<td>CI</td>
<td>Phase 2</td>
<td>Static</td>
<td>–</td>
<td>–</td>
<td>GrP + QA</td>
</tr>
<tr>
<td>Chen et al.</td>
<td>2015</td>
<td>IFI</td>
<td>Phase 2</td>
<td>Dynamic</td>
<td>GP-DO</td>
<td>O</td>
<td>GrP + QA</td>
</tr>
<tr>
<td>Z. Chen et al.</td>
<td>2023</td>
<td>IFI</td>
<td>Phase 2</td>
<td>Dynamic</td>
<td>GP-DO + ECi</td>
<td>–</td>
<td>O + QA</td>
</tr>
<tr>
<td>Xu et al.</td>
<td>2020</td>
<td>P-LI</td>
<td>Phase 2</td>
<td>Dynamic</td>
<td>GP-EDO</td>
<td>–</td>
<td>O + GrP</td>
</tr>
<tr>
<td>Fang et al.</td>
<td>2013</td>
<td>IULI</td>
<td>Phase 2</td>
<td>Static</td>
<td>–</td>
<td>Mean</td>
<td>GC + QA</td>
</tr>
<tr>
<td>Liu et al.</td>
<td>2017</td>
<td>IFI</td>
<td>Phase 2</td>
<td>Static</td>
<td>–</td>
<td>Mean</td>
<td>GC + QA</td>
</tr>
<tr>
<td>Li et al.</td>
<td>2020</td>
<td>IFI</td>
<td>Phase 2</td>
<td>Static</td>
<td>–</td>
<td>Mean</td>
<td>GC + QA</td>
</tr>
<tr>
<td>Pramanik et al.</td>
<td>2018</td>
<td>SNI</td>
<td>Phase 2</td>
<td>Static</td>
<td>–</td>
<td>Max/Min</td>
<td>GC + QA</td>
</tr>
<tr>
<td>Davoudabadi et al.</td>
<td>2020</td>
<td>IFI</td>
<td>Phase 2</td>
<td>Static</td>
<td>–</td>
<td>–</td>
<td>GrP + GC</td>
</tr>
<tr>
<td>Wan et al.</td>
<td>2015</td>
<td>IFI</td>
<td>Phase 2</td>
<td>Static</td>
<td>–</td>
<td>Mean</td>
<td>GrP + GC</td>
</tr>
<tr>
<td>Chen et al.</td>
<td>2018</td>
<td>CI</td>
<td>Phase 1</td>
<td>Static</td>
<td>–</td>
<td>Mean</td>
<td>EAIA + GC</td>
</tr>
<tr>
<td>Jabeur et al.</td>
<td>2004</td>
<td>CI</td>
<td>Phase 2</td>
<td>Dynamic</td>
<td>GP-IC</td>
<td>–</td>
<td>EAIA + O</td>
</tr>
<tr>
<td>Yang et al.</td>
<td>2019</td>
<td>P-HFLI</td>
<td>Phase 1</td>
<td>Static</td>
<td>–</td>
<td>–</td>
<td>EAIA + GC</td>
</tr>
<tr>
<td>Liu et al.</td>
<td>2016</td>
<td>CI</td>
<td>Phase 2</td>
<td>Static</td>
<td>–</td>
<td>Mean*</td>
<td>EAIA + GC</td>
</tr>
<tr>
<td>Mohagheghi et al.</td>
<td>2017</td>
<td>IFI</td>
<td>Phase 2</td>
<td>Static</td>
<td>–</td>
<td>Mean*</td>
<td>CNA + GC</td>
</tr>
<tr>
<td>Mianabadi and Afshar</td>
<td>2008</td>
<td>FI</td>
<td>Phase 1</td>
<td>Static</td>
<td>–</td>
<td>–</td>
<td>EAIA + GrP</td>
</tr>
<tr>
<td>Xu</td>
<td>2008</td>
<td>Li</td>
<td>Phase 2</td>
<td>Static</td>
<td>–</td>
<td>Majority</td>
<td>EAIA + GrP</td>
</tr>
<tr>
<td>Oļcer and Odabaşi</td>
<td>2005</td>
<td>TrFI</td>
<td>Phase 1</td>
<td>Dynamic</td>
<td>GP-DO + IC</td>
<td>–</td>
<td>EAIA + GrP</td>
</tr>
<tr>
<td>S. Wang et al.</td>
<td>2022</td>
<td>LI</td>
<td>Phase 2</td>
<td>Static</td>
<td>–</td>
<td>–</td>
<td>CNA + GrP</td>
</tr>
<tr>
<td>Bai et al.</td>
<td>2020</td>
<td>TrFI</td>
<td>Phase 1</td>
<td>Static</td>
<td>–</td>
<td>FT</td>
<td>EAIA + GrP + GC</td>
</tr>
<tr>
<td>Liu and Li</td>
<td>2015</td>
<td>II</td>
<td>Phase 2</td>
<td>Static</td>
<td>–</td>
<td>–</td>
<td>GrP + GT + GC</td>
</tr>
<tr>
<td>Chen et al.</td>
<td>2021</td>
<td>q-ROFI</td>
<td>Phase 2</td>
<td>Static</td>
<td>–</td>
<td>Mean</td>
<td>CNA + EAIA</td>
</tr>
<tr>
<td>Liu et al.</td>
<td>2019</td>
<td>LI</td>
<td>Phase 2</td>
<td>Static</td>
<td>–</td>
<td>–</td>
<td>CNA + QA</td>
</tr>
<tr>
<td>Ren et al.</td>
<td>2018</td>
<td>DIFI</td>
<td>Phase 2</td>
<td>Static</td>
<td>–</td>
<td>–</td>
<td>CNA + QA</td>
</tr>
</tbody>
</table>

*The mean has been modified by previously assigned weight.

Table 11
Index proposal.

<table>
<thead>
<tr>
<th>Acronym</th>
<th>Index</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>CIE</td>
<td>Consensus Index among Experts</td>
<td>Distance among experts</td>
</tr>
<tr>
<td>CIS</td>
<td>Consensus Index to Solution</td>
<td>Distance among the ideal solution and experts</td>
</tr>
</tbody>
</table>

Fig. 12. Percentage of articles which state previous influence or component division by category.

statistic and have analysed other possibilities. Others have applied novel procedures or proposed the use of other, more suitable geometric points.

Third, it was found to be necessary to consider concepts and constraints from sociological theories, such as social choice theory. Since this is a process of selecting alternatives or possibilities from among a group of experts, it seems obvious that there should be social guidelines for equal opportunities. However, most methods do not consider them.

Fourth, novel methods must also promote approaches in the absence of alternatives knowledge. In a world where possibilities are growing and alternatives might have multiple modifications, it is not realistic to consider that the alternatives will be known.

Fifth, there is a need to consider including initial influence and separation of components. It is precisely the increasing complexity of today’s world problems which leads to considering multiple experts or data sources and increasingly more criteria. These experts or sources may not have the same weight, and some studies show an improvement in the decision when analysing the criteria separately.

Finally, MCDM techniques and data types being used in GDM methods have been analysed to identify where future research might be focused. Providing a wide range of possibilities for experts to decide which method is more suitable will improve the use rate for these techniques and can open up new avenues of research by detecting which methods need to be adapted to certain types of data.

Nevertheless, the limitations of the present study should be mentioned. Firstly, although more than 13,500 papers have been analysed, many others may have yet to be included in the study. For example, contributions in the field are diluted within much more extensive studies, which treat this aggregation as secondary and focus on other results, such as the criteria and rankings obtained. Another limitation is the length of the manuscript, which makes it challenging to increase the number of comparisons between algorithms.

In future research, other groups of the proposed classification scheme, such as trading techniques or data aggregators, should be developed, expanding and complementing it. Additionally, future studies should analyze the algorithms used by each technique, comparing them with other approaches in the same classification group.

CRediT authorship contribution statement

David Boix-Cots: Conceptualization, Data curation, Formal analysis, Writing – original draft. Francesc Pardo-Bosch: Resources, Supervision, Validation. Pablo Pujadas: Resources, Supervision, Validation.

Declaration of Competing Interest

The authors declare that they have no known competing financial
interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability
No data was used for the research described in the article.

Acknowledgments
The first author acknowledges support from the Spanish Ministry of Universities [grant number FPU18/01471]. The second and third authors wish to recognize their support from the Serra Hunter program. Finally, this work was supported by the Catalan agency AGAUR through its research group support program (2017SGR00227). This research is part of the R&D project I+Q-EDU, reference no. PID2020-117366RB-I00, funded by MCIN/AEI/10.13039/501100011033.

References


D. Boix-Cots et al.

Information Fusion 96 (2023) 16–36

32


