Expert Systems With Applications 171 (2021) 114597

Contents lists available at ScienceDirect

Expert Systems With Applications

journal homepage: www.elsevier.com/locate/eswa

A decision support system for detecting and handling biased decision-makers in multi criteria group decision-making problems

Meysam Rabiee [a,*], Babak Aslani [b], Jafar Rezaei [c]

[a] *Lundquist College of Business, University of Oregon, USA*
[b] *School of Industrial and Systems Engineering, University of Oklahoma, Norman, Oklahoma, 73019, USA*
[c] *Faculty of Technology, Policy and Management, Delft University of Technology, The Netherlands*

ARTICLE INFO

Keywords:
Group decision-making (GDM)
Biased decision-maker
Performance matrix
Multi-criteria decision-making

ABSTRACT

Detecting and handling biased decision-makers in the group decision-making process is overlooked in the literature. This paper aims to develop an anti-biased statistical approach, including extreme, moderate, and soft versions, as a decision support system for group decision-making (GDM) to detect and handle the bias. The extreme version starts with eliminating the biased decision-makers (DMs). For this purpose, the DMs with a lower Biasedness Index value than a predefined threshold are removed from the process. Next, it continues with a procedure to mitigate the effect of partially biased DMs by assigning different weights to DMs with respect to their biasedness level. To do so, two ratios for the remaining DMs are calculated: (i) Overlap Ratio, which shows the relative value of overlap between the confidence interval (CI) of each DM and the maximum possible overlap value. (ii) Relative confidence interval CI which reflects the relative value of CI for each DM compared to the confidence interval CI of all DMs. The final step is assigning weight to each DM, considering the two values Overlap Ratio and Relative confidence interval. DMs with closer opinions to the aggregated opinion of all DMs, or those with an adequate level of discrimination in their judgments gain more weight. The framework addresses and prescribes possible actions for all possible cases in GDM including without outliers, cases with partial outliers, and extreme cases with complete disagreement among DMs, or when none of the DMs show an adequate level of discrimination power. The moderate version preassigns a minimum weight to all unbiased DMs and then follows the weighting step for the remaining total weight. However, the soft version follows the preassignmnet of weights to all DMs in the initial pool, meaning there is no elimination in this setting. The proposed approach is tested for several scenarios with different sizes. Four performance measures are introduced to evaluate the effectiveness of the proposed method. The resulted performance measures show the reliability of the outcomes.

1. Introduction

Decision-making, defined as processing the information to determine the best alternative from among a set of possible alternatives, plays a crucial role in personal life and professional interactions in business, industry, and social contexts (Kabak & Ervural, 2017). As the nature of problems is becoming more complex, considering several criteria in decision-making processes is a necessary action. Therefore, Multi-Criteria Decision-Making (MCDM) is making its way to different disciplines as a useful tool for assessing alternatives based on various norms. MCDM is a branch of Operations Research dealing with problems that typically relate to evaluate, select, sort, or rank multiple alternatives that typically involve various (conflicting) criteria. MCDM approach can deal with a range of decision problems to help decision-makers achieve (relatively) consistent and robust solutions (Cinelli, Kadzinski, Gonzalez, & Slowinski, 2020).

MCDM problems have been classified in different ways, one of which is based on the number of decision-makers (DM) contributing to the process, which leads to two classes, single decision-maker and group decision-maker. Single DM class is a process in which only a single DM is responsible for defining the problem and assessing the alternatives based on a set of criteria. In the group decision-making (GDM) class, on the other hand, the opinion of several DMs are collected and aggregated to make the final decision. Nowadays, GDM plays a crucial role in decision support systems (DSS) and information systems in the presence of ever-increasing complexity of real-life problems. The intensified complexity

* Corresponding author.
E-mail addresses: meysamr@uoregon.edu (M. Rabiee), b.aslani@ou.edu (B. Aslani), j.rezaei@tudelft.nl (J. Rezaei).

https://doi.org/10.1016/j.eswa.2021.114597
Received 7 June 2020; Received in revised form 8 January 2021; Accepted 8 January 2021
Available online 19 January 2021
0957-4174/

of the decision-making process in a range of contexts such as health, business, management, and politics, has made stakeholders and decision-makers rely on group wisdom instead of a single individual. The outcome of decision-making problems usually affects a range of stakeholders at different levels with different preferences. So, in a GDM process related to a real problem, several DMs with expertise in various fields are present. These DMs are usually looking for a single final decision in the process, though (Bouzarour-Amokrane, Tchangani, & Peres, 2015; Dong, Zhou, & Martínez, 2019; Jin et al., 2017; Liu, Du, & Wang, 2019).

GDM as a response to the ever-growing complexity of modern environments related to decision-making problems has found its way in industry, where technical groups make the decisions on designing products, advancing plans and strategies, and in service sector such as healthcare, in which the critical decisions are made by a group of advisors and experts (Ambrus, Greiner, & Pathak, 2009). The GDM process generally starts with identifying the characteristics of the problem, including alternatives, criteria related to the alternatives, and their importance (weights), as well as DMs and stakeholders. During this stage, compromising among DMs to have an agreement over the mentioned features is highly essential. Following that, a score is assigned to each alternative on each criterion, either by the DMs or other sources. Aggregating the opinions of DMs to rank or sort the alternatives or to select the best option is the final stage of the process. In other words, in GDM problems, separate preferences of DMs should be aggregated in a collective and well-structured way to make the final decision (Liao, Zhang, & Luo, 2018; Xia & Chen, 2015; Xu, Huang, & Li, 2019). GDM process usually includes the opinions of DMs who have different backgrounds and knowledge bases. As a result, interpreting and analyzing the preferences of these DMs is a complex task compared to single DM processes (Wittenbaum, Vaughan, & Strasser, 2002; Yu, Li, & Fei, 2018).

Particular challenges have been recognized in the literature for the GDM process. One possible shortcoming of GDM is creating a diffusion of responsibility resulting in a lack of accountability for the outcomes of the process. Moreover, group decisions can make it easier for members to deny personal responsibility and blame others for the undesired final decision (Palomares, Estrella, Martínez, & Herrera, 2014). GDM involves many complex and conflicting aspects intrinsic to human individuality and human nature. For instance, when a team of experts with various (usually conflicting) personal goals and intentions takes part in the process, having disagreement is a natural expectation (Parreiras, Ekel, Martini, & Palhares, 2010). As another challenge, because of different individual concerns about alternatives, DMs usually adopt heterogeneous preference representation structures to express their preferences (Xu, Du, Chen, & Cai, 2019). Experiments related to GDM in different fields show clear contrasts between the results of the GDM process compared to decisions reached by DMs making choices in isolation. DMs in a real decision-making problem may have various fields of expertise, meaning that each DM can have a dissimilar effect on the final decision (Hafezalkotob & Hafezalkotob, 2017).

GDM can be categorized based on its dynamics to cooperative or noncooperative. Cooperative GDM is significant in engineering, medical, and scientific fields, while non-cooperative GDM is common in economic and political areas. Even in cooperative GDM, reaching a complete agreement among all members of the group on the final solution is nearly impossible since DMs, who are supposed to have similar goals, may have some opinion conflicts, in practice (Dong et al., 2018). Having consistency and consensus are other noteworthy challenges in the GDM process. Consistency is directly related to the credibility of the GDM process. Consensus, on the other hand, means that the majority of DMs accept the final result of the process (Song & Hu, 2019). In the case of political context, the presence of the correlated opinions can pose another challenge to the aggregation of individual opinions. For instance, the opinions of the members associated with a certain political party are expected to be highly correlated. These correlated clusters can be identified and labeled as biased opinions in a GDM aggregation phase since they do not reflect the true opinions of DMs (parliament members or congressmen in this context), but only represent their political attachments. This specific problem is not in the scope of the current study, but has been tackled in the literature (see Akram, Adeel, & Alcantud, 2019; Akram, Bashir, & Garg, 2020; Akram, Yaqoob, Ali, & Chammam, 2020; Gonzalez-Arteaga, Alcantud, & de Andres Calle, 2016; Liu, Shahzadi, & Akram, 2020).

One of the main challenges in the GDM process is having biased DM (s), which can significantly challenge the result of the process in such a way that the stakeholders might not accept the final decision. The biased DM(s) distort the consensus among other DMs. The unbiased DM(s) will not reach an agreement over the results in the presence of biased DM(s). These mentioned challenges show the significance of detecting and managing biased DMs in the GDM process (Kabak & Ervural, 2017; Liao, Wu, Mi, & Herrera, 2019; Mohammadi & Rezaei, 2020; Song & Hu, 2019; Yu et al., 2018). In this study, we consider DMs with inadequate discrimination power in their judgment/evaluation, those who are giving their scores with little variation to the available options in the predefined criteria as biased ones in a GDM process. Further explanation will be provided in the methodology section to clarify the scope of the proposed framework in dealing with this type of biasedness.

The studies related to GDM have handled bias in the process from different points of view. A stream in the literature is mainly focused on identifying sources of bias (see Jones & Roelofsma, 2000), (Xiao & Benbasat (2018),Ceschi, Costantini, Sartori, Weller, & Di Fabio (2019)). These works are mainly concerned with the reason of having bias in the GDM and do not present any solution to resolve or eliminate bias. Adopting consensus reaching process (CRP) is another rich body of knowledge in handling biasedness in GDM (see Bouzarour-Amokrane et al., 2015; Liu, Zhang, Chen, & Yu, 2018; Xu et al., 2019). These studies generally present frameworks to improve the consensus among the DMs by closing their preferences (scores) assigned to the alternatives in each criterion. In some of these frameworks, such as Liu et al., 2018, the DMs have the freedom of changing their opinion after giving their initial opinion. Another research avenue related to bias is identifying the biased DMs (outliers) in the GDM process (see McShane, Nirenburg, & Jarrell (2013)). Even though this problem has been addressed in a range of contexts such as business in Lee (2014), NBA draft season in Ichniowski and Preston (2017), there is a shortage of efforts to present a general framework for identifying and managing outliers in GDM.

In specific situations, the opinions of biased DMs can be excluded from the final pool of decision-making. For instance, if the aggregated opinion of many DMs is desired (like voting systems), eliminating DMs who represent the minority voices is justified. Likewise, when the collective agreement is acceptable by everyone even those whose opinions have been eliminated, we can rationalize the elimination of DMs in the GDM process. One example is the GDM procedures using agents to collect and aggregate anonymous votes of DMs to find the results with the highest degree of consensus. For these cases, we propose the Extreme Anti-Biased Method (EABM) and Moderate Anti-Biased Method (MABM) version of our proposed method, both of which include elimination and weighting based on the consensus degree of each DM with the remaining pool or their discrimination power. However, in the moderated configuration, a threshold of weight is assigned to each DM to mitigate the biasedness in assigning the weights only based on consensus. On the other hand, if all DMs are somehow related to the final results in terms of interest, like stakeholders or board of directors, the opinions of outliers cannot be eliminated easily (or at all). As the DMs observe the consensus reaching process and they do not consent to be eliminated from the GDM process, a different version of the presented framework is designed for these environments. To handle these cases, we propose the Soft Anti-Biased Method (SABM), in which there is no elimination, but the weights are assigned to all DMs based on a combination of a predefined threshold and consensus-based weights.

The main contribution of this study is to present a general statistical-based framework to detect and handle biased DMs in a GDM process. The diverse versions of the proposed framework are capable of providing viable solutions to all the possible scenarios in the aggregation phase of a GDM process. If there is no outlier in the pool, the framework will assign a relative weight to each DM based on the agreement/similarity level to the total pool. In the case that some of the DMs are biased, the framework can either detect and exclude them from the decision-making process or incorporate the biasedness in assigning the weights to the DMs. Finally, if there is no common ground for all DMs, the agent responsible for the aggregation can warn the DMs to adjust their opinions or even incorporate a new set of DMs to the pool to find some common opinions on which she/he can establish the aggregated opinion. Even though detecting and handing outliers in GDM has high importance in various contexts, such as voting systems, supplier selection, and portfolio optimization, there are not enough works in the literature to tackle the problem.

The remaining of the paper is structured as follows. In Section 2, the literature review, as well as the gap, is presented. The proposed method is explained in detail in Section 3. Section 4 includes the result of applying the proposed approach to a numerical example. The result of the method on a large sample of problems and a detailed analysis is provided. Finally, the conclusion and suggestions for possible future research are provided in Section 5.

2. Literature review

As a main challenge in the GDM process, when multiple DMs are involved, it is necessary to aggregate individual judgments into a single representative judgment for the entire group. The studies in this research stream have tried to use different aggregation methods to achieve a collective decision in a well-structured way.

Ordered Weighted Aggregation (OWA) and its extensions as a popular aggregation operator for GDM is present in several works such as Xu and Jian (2007), Liu (2011), and Jin, Mesiar, and Yager (2018). Other aggregation procedures has been intensively proposed in various contexts, such as individual judgments and priorities in Forman and Peniwati (1998), group calibration process of slightly different preferences in Rokou and Kirytopoulos (2014), Introducing a representative value function as the aggregated opinion in Kadziski et al., (2013), Stochastic Group Decision Support System (GDSS) for the GDM with uncertainty in Mokhtari (2013), and incomplete individual ratings in group resource allocation decisions in Stengel (2013). The main shortcoming of these operators is that it cannot appropriately operate in many real cases in which the arguments being aggregated have different significance. This issue massively challenges the credibility of the combined aggregated decision in GDM.

The Delphi method is a framework based on the results of multiple rounds of questionnaires sent to a group of experts. The anonymous responses are aggregated and shared with the group after each round. The experts can adjust their opinions in subsequent rounds, based on their perception of the group's opinion. As the Delphi method seeks to reach a desirable response through consensus, it has practicality in the aggregation stage of GDM. For instance, a novel Delphi method known as agile Delphi was used as an aggregation approach in Xie, Liu, Chen, Wang, and Chaudhry (2012). They aimed to present a GDM process to rapidly develop and execute the response approach to the unconventional emergency events. Their agile Delphi decision-making platform based on ASP.NET. allowed experts to conduct appropriate alternative portfolio and enables researchers to find a new combination of experts to make the decision-making more efficient. Lack of physical interactions, long response time, and the usefulness of received information from experts are the main drawbacks of this method.

A richer body of knowledge in the literature of GDM contains studies that tried to increase the consistency and consensus in the aggregation phase of the GDM process. Wibowo and Deng (2013) presented an interactive algorithm for consensus building in the group decision-making process. They touched the subjectiveness and the imprecision of the decision-making process by using intuitionistic fuzzy numbers. A decision support system framework was also presented for improving the effectiveness of the consensus building process. Dong and Cooper (2016) proposed a novel consensus reaching model in a dynamic decision environment within the context of the group Analytic Hierarchy Process (AHP). They adopted a Markov chain method to determine the DMs' weights of importance for the aggregation process with respect to the group members' opinion transition probabilities. Their developed model also provided feedback suggestions to adaptively adjust the credibility of each DM in each round so that the dynamic feature of decision-making process is also addressed. Wu and Xu (2016) developed separate consistency and consensus processes to deal with hesitant fuzzy linguistic preference relation (HFLPR) individual rationality and group rationality. They included an easy understandable local revision strategy in their framework. The feedback system was based directly on the consensus degrees to reduce the proximity measure calculations.

Jin et al. (2017) proposed two new GDM methods to improve the multiplicative consistency of linguistic preference relation (LPRs) until they are acceptable, and the priority weight vector of the alternatives is derived from adjusted LPRs. They introduced novel concepts of order consistency, multiplicative consistency, and a consistency index in their proposed approach. Other components of their methodology were two linear optimization models to generate the normalized crisp individual and group weight vectors and two GDM methods to help DMs in obtaining reasonable and reliable results. Liu et al., 2018 proposed an iteration-based consensus building framework for GDM problems with self-confident multiplicative preference relations. They developed an extended logarithmic least squares method to derive the individual and collective priority vectors from the self-confident multiplicative preference relations. They used a two-step feedback adjustment mechanism to assist the decision-makers to improve the consensus level. Dong et al., 2019 introduced a novel framework to achieve a potential consensus considering too polarized opinions. By considering the existing dynamic relationships among the experts, they suggested using Social Network Analysis to reach agreement in this specific class of GDM problems. They claimed that the framework is particularly useful when achieving the agreement through CRPs is impossible. Morente-Molinera, Kou, Samuylov, Urena, and Herrera-Viedma (2019) presented a novel model for experts to carry out GDM processes using free text and alternative pairwise comparisons and two ways of applying consensus measures over the GDM process. In the context of social networks, they adopted Sentiment analysis procedures to analyze free texts and extract the preferences that the experts provide about the alternatives. They also presented two ways of applying consensus measures over the GDM process.

Song and Hu (2019) developed an iterative algorithm to improve the consistency to obtain the probabilistic linguistic preference relation (PLPR) with satisfactory consistency in large-scale GDM problems. They used PLPR in the presence of partial preference information coming from stakeholders in the GDM procedure. A probability computation model was defined by mathematical programming to derive the missing probabilities of PLPR. Song and Hu (2019) developed a GDM model based on obtained multiplicative consistency linguistic preference relations (LPRs) in consideration of the group consensus reaching process. Their proposed mathematical programming method was able to find the incomplete LPR with the highest consistency and increasing inconsistent LPR to complete consistent LPR using multiplicative consistency based on given incomplete HFLPR. Tang, Meng, and Zhang (2019) developed a programming model to judge the consistency of multiplicative linguistic intuitionistic fuzzy preference relations (MLIFPRs), and an approach to derive consistent MLIFPRs. They developed a direct consensus framework based on cumulative prospect theory to solve GDM problems with eight kinds of preference representation structures. Xu et al. (2019) developed a confidence consensus-based model for large-scale group decision making that provides a novel approach to addressing non-cooperative behaviors. They defined new concepts of collective adjustment suggestion and rationality degree. Then, they combined the rationality and non-cooperation of the adjustment information to construct the concept of a confidence level to measures the impartiality and objectivity of the adjustment information and for managing non-cooperative behaviors. Liu et al., 2019 constructed a two-step optimization model to solve for the group consensus ideal scheme and its measure value matrix. They adopted Nash's bargaining idea and maximizing group negotiation satisfaction and minimizing system coordination deviation in their developed framework. They measured the closeness degree of decision-maker information and attribute information by using the distance between the group measure matrices of scheme and consensus ideal scheme.

The main challenge of consensus reaching process frameworks is that since they ignore external factors, such as weather conditions, price fluctuations, alternatives and/or experts' availability, they lose their practicality in many real-life settings. Higher time consuming, more time on constant preferences supervision and the higher complexity with respect to dealing with experts' non-cooperative behaviors are notable patterns of CRPs when they are applied on large-scale GDM problems, too.

Handling bias in GDM process is another section of the literature, in which either the source of bias is identified, or the bias is managed by different approaches. One major source of bias is in the data collection phase of GDM, in which DMs assert their preferences (or scores) of alternatives in the defined criteria. Cognitive bias, which is a systematic pattern of deviation from norm or rationality in judgment, and is very common in the GDM process, is addressed in Stewart (2005). Anchoring bias happens when DMs are biased towards the first piece of information they receive. This initial information (anchor) affects further evaluations/judgments conducted by DM and the adjustments which are later done to the evaluation are not sufficient. This source is analyzed comprehensively in the work of Lieder, Griffiths, Huys, and Goodman (2018) and Rezaei (2020) in the context of multi-criteria decision-making. Jacobi and Hobbs (2007) investigated the splitting bias, which refers to a situation where presenting an attribute in more detail, may increase the weight it receives. For more details on the studies of this stream of literature, we refer to a comprehensive study of Montibeller and Winterfeldt (2015), which provides an overview of different cognitive biases at different stages of the decision-making process. Even though these various types of bias are important in the GDM process, we are not going to address them in our framework.

In the current study, we assume that there is bias in the form of biased DM(s) in the GDM process, and we need to identify and handle it in an organized manner. In this section, we focus on papers trying to detect and eliminate the present bias in the GDM problems. There is a limited number of studies in the literature, in which handling biased DMs in a specific case and context is addressed. For instance, McShane et al. (2013) presented the OntoAgent model for automatic detection of decision-making biases, using clinical medicine as a sample application area. In particular, the intelligent agent was able to follow doctor-patient interaction and warn the doctor if the patient's decisions were being affected by biased reasoning. Bouzarour-Amokrane et al. (2015) proposed a resolution model based on individual bipolar assessment. Each decision-maker evaluates alternatives through selectability and rejectability measures, representing the positive and negative aspects of alternatives. They also included the impact of human behavior (including influence, individualism, fear, caution) on decisional capacity. They adopted consensus building models based on game theory for collective decision problems to achieve a common solution in the GDM process. Dong and Cooper (2016) proposed a novel consensus framework for managing non-cooperative behaviors in GDM. A self-management mechanism was responsible for generating experts' weights dynamically based on multi-attribute mutual evaluation matrices (MMEMs). During the process, the experts could provide and update their MMEMs regarding the experts' performances (e.g., professional skill, cooperation, and fairness) so that their weights were updated. He, Chan, and Jiang (2018) introduced a quantum framework based on quantum probability theory to model the subjectivity in multi-attribute group decision making (MAGDM) when subjective beliefs towards DMs' independence or relations are present. The opinions of DMs are viewed as various wave functions that are occurring at the same time. Then the beliefs will interfere with each other and influence the aggregated result. Classical MAGDM techniques were used for the preparation stage. Then, they constructed a Bayesian network and extended it to a quantum framework. When all the DMs are deemed independent, the quantum framework will degenerate into a classical Bayesian network.

Based on the studies reviewed, a significant gap in the scope of our study, detection, and handling bias in the GDM environment, is

Table 1
Nomenclature of the proposed framework.

Notation	Description
d_i	Decision-maker (DM) $i, i = 1, \ldots, i$
a_j	Alternative j
c_k	Criterion k
s_{ijk}	Score assigned by d_i to a_j with respect to c_k
ψ_{ijk}	The normalized value of s_{ijk}
S_k^{min}	Minimum value of c_k across all decision-makers and alternatives or $min\{s_{ijk}\}$
S_k^{max}	Maximum value of c_k across all decision-makers and alternatives or $max\{s_{ijk}\}$
$\overline{x}_i$	Average of normalized scores of d_i
$\overline{X}$	Average of total normalized scores for all DMs
σ_i	The standard deviation of normalized scores of d_i
σ	The standard deviation of total normalized scores for all DMs
CI_i	Confidence Interval for d_i
UB_i	Upper bound of Confidence Interval for d_i
LB_i	Lower bound of Confidence Interval for d_i
l_i	Length of Confidence Interval for d_i
CI	Confidence Interval for total normalized scores for all DMs
O_{mn}	Overlap among CI_m and $CI_n, m, n \in i$
O_i	Overall overlap for d_i
M_i	Maximum possible value of overlap of d_i with other DMs
α	Significance level for calculating CI's (The range is usually between 0.9 and 0.99)
$\widetilde{O}_i$	Overlap Ratio for d_i
$\widetilde{CI}_i$	Relative CI for d_i
B_i	Biasedness Index for d_i
B	Threshold for detecting biased DM (The range can be different from 0 to $I-1$)
R	Threshold for the number of eliminated DMs in the first step
N	The number of normalized scores of each DM in the GDM process
I	The number of DMs in the initial pool of the GDM process
I'	The number of DMs remained in the pool in the second phase of the framework
λ	The total number scores for all DMs in the second phase of the framework
w_i	Weight assigned to d_i
$\widehat{w}_i$	Initial Weight of d_i
w_m	The minimum predefined weight for MABM and SABM versions
γ	Total predefined weight

identified. The studies addressed bias in the GDM process in the literature are more focused on using various models and techniques to increase the agreement level among DMs, which is an essential aspect of GDM. Although there are some previous works approaching bias detection in individual cases of GDM problems, presenting a general framework to this purpose is a significant gap in the literature. The main contribution of our work is to develop a systematic approach to tackle one crucial challenge in GDM, identifying and handling the outliers (biased) DMs to improve the credibility of the aggregated opinion. To bridge this gap, we proposed a two-stage statistical approach to identify outliers and modify the final weights of DMs in the decision-making process. The proposed framework in this study can be applied to any GDM problem regardless of size and context. Also, this framework can handle extreme situations without any agreement as well as the partial and total agreement among DMs. The details of the suggested framework will be provided in Section 3.

3. Methodology

Before moving to the details of the proposed framework, we need to clarify our definition of a biased DM in this paper. While there might be several interpretations of biasedness, we consider a DM who lacks adequate discrimination power in a GDM process as a biased one. In other words, DMs whose scores are high for all cases, are low for all cases, or are mainly distributed in the middle range of the spectrum of scores, and the confidence interval of their aggregated opinion has very small or no overlap with other DMs are treated as biased DMs in our methodology.

In this section, we first explain EABM version of our proposed method in detail. To save the space, we only then focus on the differences among the moderate and soft versions and the extreme one. The EABM for handling bias in GDM is composed of two phases. First, by forming the normalized integrated matrix of scores, the biased DMs are detected and eliminated from the decision-making process. Then, by following a well-defined procedure based on the relative amount of overlap for each DM by other DMs and the total data, a weight value is assigned to each DM. This section includes the detailed process of these steps of the proposed framework. Table 1 presents the notations used in the following sections to help the readability of the equations.

Please note that since several DMs from the initial pool might be excluded as outliers, we defined two sets for DMs. The first set is the initial pool and contains I DMs and the pool in the second phase consists of I' DMs. Regardless of this change, the initial label of DMs is kept for evaluating the alternatives based on the assigned weights and the associated performance matrix. Please also note that for parameter B, the range can change from 0 to $I-1$ and for α, this variation is usually between 0.9 and 0.99 from a statistical point of view.

3.1. Outlier elimination phase

3.1.1. Gathering data for the decision-making process

Group decision-making process includes a number of decision-makers, $d_i = 1, .., I$, who evaluate several alternatives, $a_j = 1, .., J$, with respect to some criteria, $c_k = 1,..,K$. It is worth mentioning that the score table, which is also called performance matrix, sometimes is available with objective scores, while it is sometimes formed by the evaluations done by the experts in a subjective manner or the combination of both. In this study, we consider the last two cases. The following matrices show the raw data available at the beginning of the process.

$$d_1 \Rightarrow \begin{array}{c} a_1 \\ a_2 \\ \vdots \\ a_J \end{array} \begin{pmatrix} c_1 & c_2 & \cdots & c_K \\ s_{111} & s_{112} & \cdots & s_{11K} \\ s_{121} & s_{122} & \cdots & s_{12K} \\ \vdots & \vdots & \vdots & \vdots \\ s_{1J1} & s_{1J2} & \cdots & s_{1JK} \end{pmatrix}, \ldots, d_I \Rightarrow \begin{array}{c} a_1 \\ a_2 \\ \vdots \\ a_J \end{array} \begin{pmatrix} c_1 & c_2 & \cdots & c_K \\ s_{I11} & s_{I12} & \cdots & s_{I1K} \\ s_{I21} & s_{I22} & \cdots & s_{I2K} \\ \vdots & \vdots & \vdots & \vdots \\ s_{IJ1} & s_{IJ2} & \cdots & s_{IJK} \end{pmatrix}$$

3.1.2. Creating the integrated decision matrix

As the first step, the opinions of decision-makers should be integrated as a comprehensive decision matrix. Matrix P^1 is comprised of the score of each decision-maker to all alternatives in every criterion. Matrix P^1 is a sample decision matrix for a situation with I decision-makers, J alternatives, and K criteria. We assume, in our case, that all DMs participating in the GDM process know a minimum and maximum level for the possible scores in the process and are providing comparable inputs to the performance matrix. Therefore, the integrated performance matrix can be normalized to build the foundation for the next steps of the framework.

$$P^1 = \begin{array}{c} d_1 \\ \vdots \\ d_I \end{array} \begin{pmatrix} s_{111} & s_{112} & \cdots & s_{11K} \\ s_{121} & s_{122} & \cdots & s_{12K} \\ \vdots & \vdots & \vdots & \vdots \\ s_{1J1} & s_{1J2} & \cdots & s_{1JK} \\ & \vdots & & \\ s_{I11} & s_{I12} & \cdots & s_{I1K} \\ s_{I21} & s_{I22} & \cdots & s_{I2K} \\ \vdots & \vdots & \vdots & \vdots \\ s_{IJ1} & s_{IJ2} & \cdots & s_{IJK} \end{pmatrix}$$

3.1.3. Normalizing the integrated decision matrix

In real-world problems, criteria usually have various natures and scales. So, in order to compare these criteria regardless of their units, they should be normalized to obtain dimensionless scores. To normalize the scores several methods are developed in the literature like Max, Max–Min, and logarithmic approaches (Vafaei, Ribeiro, & Camarinha-Matos, 2016). To convert the scores given to various alternatives, the decision matrix should be normalized by using Eq. 1 to 4. Eq. 1 shows the normalization procedure for criterion with positive nature (like profit, quality, and so on). For negative criterion such as cost and pollution, we can use Eq. 2. Eqs. 3 and 4 show that the minimum value and the maximum value are calculated in each column related to the criterion by $min\{s_{ikj}\}$and $max\{s_{ijk}\}$, which are the minimum and maximum value among each column of the original decision matrix (Anojkumar, Ilangkumaran, & Sasirekha, 2014). The final output of this normalization process is the matrix labeled as P^1_{norm}.

$$\psi_{ijk} = \frac{s_{ijk} - S_k^{\min}}{S_k^{\max} - S_k^{\min}}, \forall i = 1, .., I, j = 1, \ldots, J, k = 1, \ldots, K \Rightarrow \text{For positive criterion} \tag{1}$$

$$\psi_{ijk} = \frac{S_k^{\max} - s_{ijk}}{S_k^{\max} - S_k^{\min}}, \forall i = 1, .., I, j = 1, \ldots, J, k = 1, \ldots, K \Rightarrow \text{For negative criterion} \tag{2}$$

$$S_k^{\min} = \min\left(s_{ijk}\right), \forall i = 1, .., I, j = 1, \ldots, J, k = 1, .., K \tag{3}$$

$$S_k^{\max} = \max\left(s_{ijk}\right), \forall i = 1, .., I, j = 1, \ldots, J, k = 1, .., K \tag{4}$$

$$P^1_{norm} = \begin{array}{c} d_1 \\ \vdots \\ d_I \end{array} \begin{pmatrix} \psi_{111} & \psi_{112} & \cdots & \psi_{11K} \\ \psi_{121} & \psi_{122} & \cdots & \psi_{12K} \\ \vdots & \vdots & \vdots & \vdots \\ \psi_{1J1} & \psi_{1J2} & \cdots & \psi_{1JK} \\ & \vdots & & \\ \psi_{I11} & \psi_{I12} & \cdots & \psi_{I1K} \\ \psi_{I21} & \psi_{I22} & \cdots & \psi_{I2K} \\ \vdots & \vdots & \vdots & \vdots \\ \psi_{IJ1} & \psi_{IJ2} & \cdots & \psi_{IJK} \end{pmatrix}$$

3.1.4. Reshaping the normalized matrix to a column matrix

To compare the decision-makers in a structured way, the normalized decision matrix should be converted to several columns separately mirroring the scores of each decision-maker to all alternatives in each criterion. This reshaped matrix is shown as matrix P^2.

$$P^2 = \begin{pmatrix} d_1 & d_2 & \cdots & d_I \\ \psi_{111} & \psi_{211} & \cdots & \psi_{I11} \\ \psi_{112} & \psi_{212} & \cdots & \psi_{I21} \\ \vdots & \vdots & \vdots & \vdots \\ \psi_{11K} & \psi_{21K} & \cdots & \psi_{I1K} \\ \vdots & \vdots & \vdots & \vdots \\ \psi_{1J1} & \psi_{2J1} & \cdots & \psi_{IJ1} \\ \vdots & \vdots & \vdots & \vdots \\ \psi_{1JK} & \psi_{2JK} & \cdots & \psi_{IJK} \end{pmatrix}$$

Having this column matrix, a confidence interval (*CI*) for each decision-maker should be calculated. Eqs. 5 and 6 show the required formulas to find the lower bound (*LB*) and the upper bound (*UB*) of these intervals for each column (DM). The $\overline{x}_i$ reflects the average scores of d_i assigned to all the alternatives in all criteria, and σ_i shows the standard deviation of the scores for this DM. N is the total number of scores available in the integrated matrix (such as P^2) for each DM, which is the product of the number of alternatives in the number of criteria ($J \times K$). Since we do not know the exact distribution of the scores, t-student distribution is used for calculating the CIs. Three parameters (Thresholds) should be set before further calculation. The first threshold is the acceptable Biasedness level (B). This threshold is a base for comparing the biasedness level of all DMs.

As the first step to calculate the B_i, the CI of all DMs will be compared together. Then, if that DM has overlap with all other DMs, its B_i will be equal to $I-1$ and if that DM has no overlap with any DM, its B_i will be equal to zero. Basically, B_i is defined as the number of DMs with having an overlap in their CI's with a specific DM. This threshold (B) can get any values between 0 and $I-1$. Choosing a value close to zero means more open to hear extreme voices and closer to I-1 means more restrictive to hear extreme voices from DMs. The B_i value is calculated for one DM in the numerical example in Section 3 to clarify the procedure. The second threshold is α which is the confidence level of statistical tests. Finally, a threshold R should be defined for the ratio of DMs eliminated in the first phase of the framework. This parameter gives the aggregation agent the freedom to send the DMs warning to review and resubmit their scores or (if possible) to consider new DMs in the process to reach an agreement.

$$\overline{x}_i = \frac{\sum_{j=1}^{J}\sum_{k=1}^{K}\psi_{ijk}}{N}, \sigma_i = \sqrt{\frac{1}{N-1}\sum_{j=1}^{J}\sum_{k=1}^{K}\left(\psi_{ijk}-\overline{x}_i\right)^2}, \forall i = 1, .., I. \tag{5}$$

$$CI_i = \left(\overline{x}_i - t(N-1,\alpha)\frac{\sigma_i}{\sqrt{N}}, \overline{x}_i + t(N-1,\alpha)\frac{\sigma_i}{\sqrt{N}}\right), \forall i = 1, .., I. \tag{6}$$

To find the decision-makers who are biased or outliers in comparison with their counterparts, these CI's should be compared in terms of their B. DMs with lower B_i values than the predefined threshold B are excluded from the process as the output of this phase.

3.2. Weight assignment phase

After eliminating the biased DMs in the previous step, an overall comparison should be conducted for the remaining scores in the decision matrix. In this step, a pairwise comparison of CIs for all DMs is conducted to obtain an overlap ratio of DMs. Having these values, the relative CI_i to CI is also calculated to assign a proportionate weight to the remaining DMs. While there is no elimination in this phase, DMs who have more agreement with other DMs and the total opinion or those who show a high discrimination power are considered more influential by having higher weight values. A similar concept is presented in the work

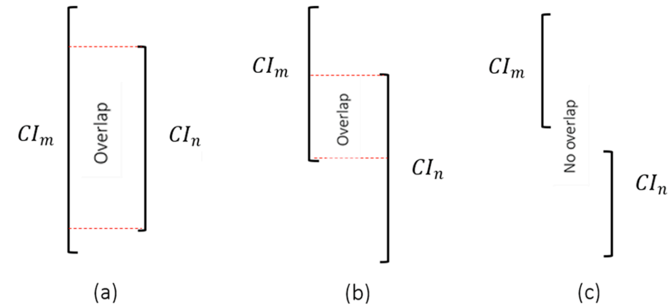

Fig. 1. The graphical representation of overlap among CIs.

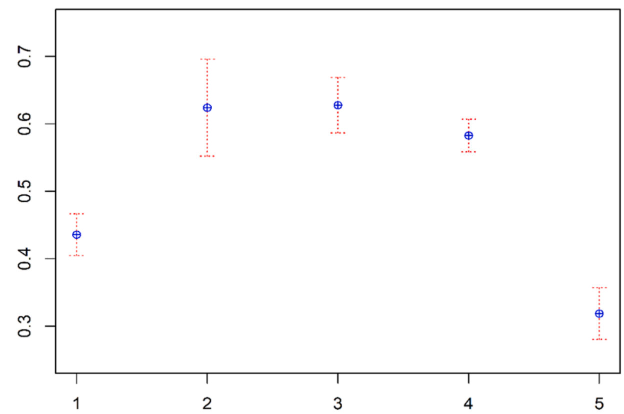

Fig. 2. CIs of DMs.

of Bonner, Baumann, and Dalal (2002), in which the experts with higher expertise receive more weight in the aggregation stage.

3.2.1. Overlap Ratio calculation

As the first step of this phase, the CIs of the remaining DMs in the pool of GDM are compared to one another. For each DM, the overlaps with CIs of other DMs are accumulated to a value known as overlap ratio. Matrix P^3 shows the remaining scores after eliminating the outliers in the first step. Please note that although the new set of unbiased DMs starts from 1 to I', the initial label of DMs is preserved for the final evaluations.

$$P^3 = \begin{array}{c} \\ d_1 \\ d_2 \\ \vdots \\ d_{I'} \end{array} \begin{pmatrix} c_1 & c_2 & \cdots & c_K \\ \psi_{111} & \psi_{122} & \cdots & \psi_{1JK} \\ \psi_{211} & \psi_{222} & \cdots & \psi_{2JK} \\ \vdots & \vdots & \vdots & \vdots \\ \psi_{I'11} & \psi_{I'22} & \cdots & \psi_{I'JK} \end{pmatrix}, \forall i = 1, \dots, I' \in \{\text{unbiased DMs}\}$$

$$\forall \mathrm{m}, n \in \{\text{unbiased DMs}\} \Rightarrow \begin{cases} \text{if } \mathrm{CI}_m \text{has overlap with } \mathrm{CI}_n \Rightarrow \begin{cases} O_m = O_m + O_{mn} \\ O_n = O_n + O_{mn} \end{cases} \\ \text{else} \Rightarrow \begin{cases} O_m = O_m + 0 \\ O_n = O_n + 0 \end{cases} \end{cases}$$

Fig. 1 shows the overlap among CIs in three possible situations: (a) Total overlap, in which one of the CIs is completely overlapped with the other one, (b) Partial overlap, in which only a portion of CIs are

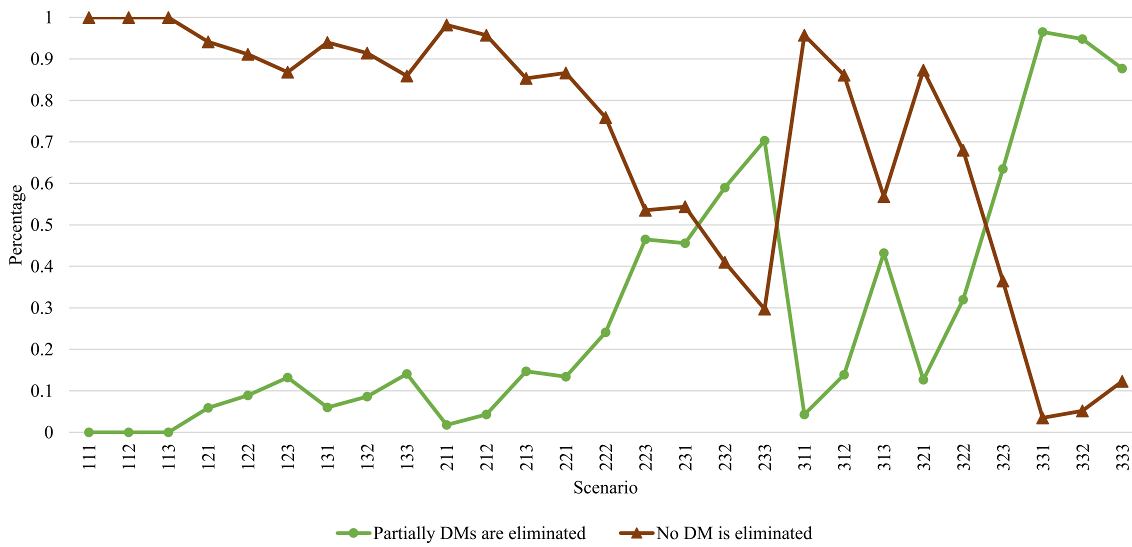

Fig. 3. Trend of elimination in the scenarios.

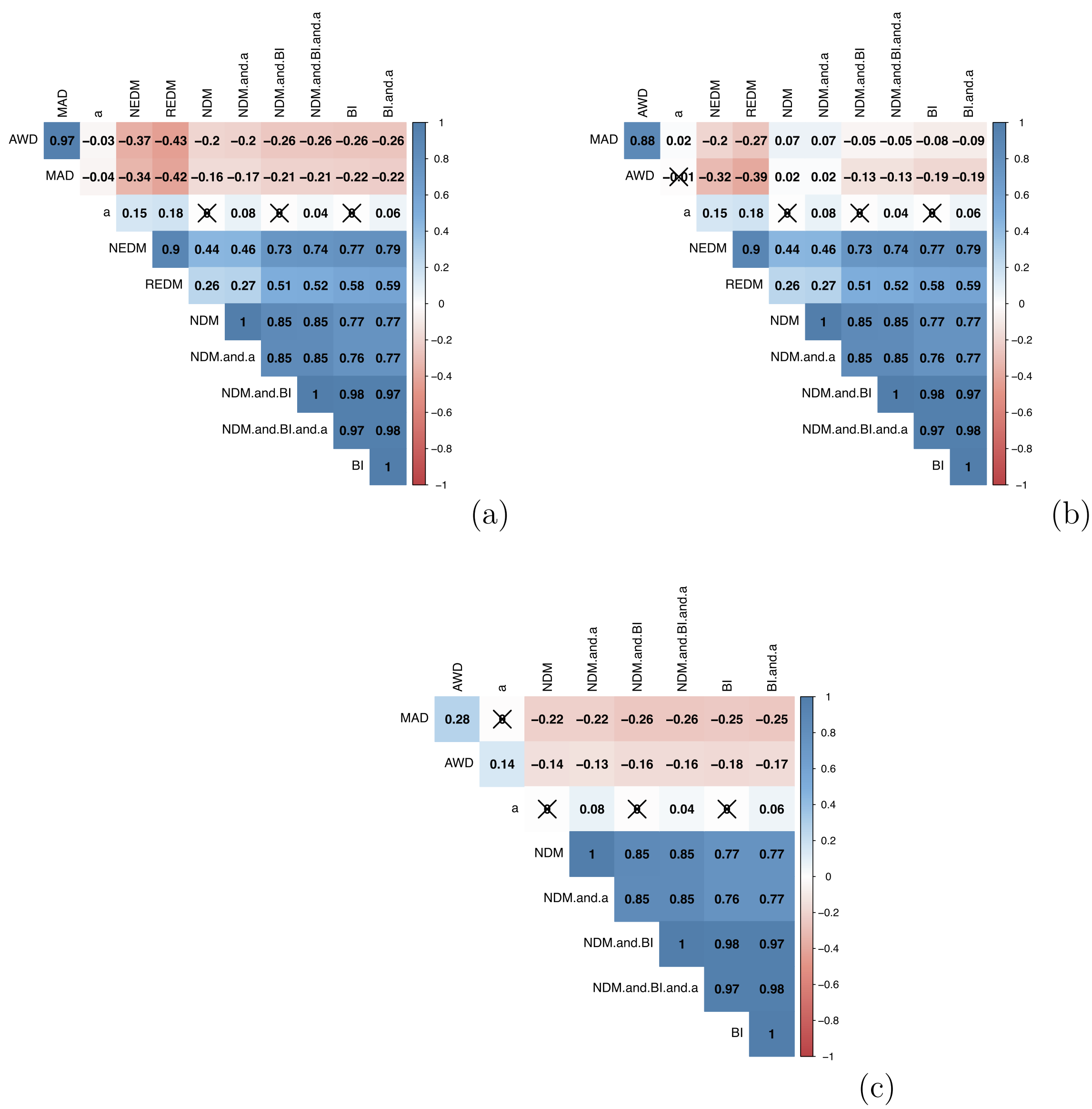

Fig. 4. Correlation analysis result.

Table 2
Applications of the proposed framework.

Version	GDM context
EABM and MABM	The aggregated opinion of many DMs is desired (like voting systems)
	The collective agreement is acceptable by everyone
	An agent collects and aggregates anonymous votes of DMs
SABM	DMs have interest, like stakeholders or board of directors
	DMs observe the consensus reaching process
	DMs do not consent to be eliminated from the GDM process

Table 3
Proposed solution for the possible scenarios in a GDM process.

Scenario	Proposed solution
(i) The pool of DMs does not contain any outlier	The weights of DMs are calculated based on the level of accordance with the total opinion
(ii) The pool of DMs contains outliers	Detect the outliers and handle the biasedness in a systematic approach
(iii) The agreement is unreachable in the GDM	Tune the parameters to reach minimum desirable consensus or give equal weights

overlapped, and (c) No overlap, in which the CIs do not have any overlap whatsoever. The overlap among the CIs of two DMs could show either the DMs have similar evaluation of individual options (which is naturally resulted in similar average and standard deviation) or they have different evaluations but with similar average and similar discrimination power (which is reflected in standard deviation). It is worth mentioning that the reverse cases are not included in the figure to save the brevity of the paper.

The result of this process is a matrix of overlaps such as P^4. The sum of each row will yield the total overlap of each DM with other DMs. It is worth mentioning that the diagonal values of the overlap matrix are the length of CI for the associated DM. However, since we are interested in having the overlap among CIs of different DMs, we reduced the diagonal values from the sum of overlaps for all DMs.

$$P^4 = \begin{array}{c} \\ d_1 \\ d_2 \\ \vdots \\ d_{I'} \end{array} \begin{pmatrix} \mathrm{d}_1 & d_2 & \dots & d_{I'} \\ l_1 & O_{12} & \dots & O_{1I'} \\ O_{21} & l_2 & \dots & O_{2I'} \\ \vdots & \vdots & \vdots & \vdots \\ O_{I'1} & O_{I'2} & \dots & l_{I'} \end{pmatrix}, O_i = \sum_{q=1}^{I'} O_{iq} - l_i, l_i = UB_i - LB_i$$

For each DM, there is an upper bound of CI overlap with others (M_i), which is the product of the length of its CI and the number of other DMs (Eq. 7). Dividing the actual ratio and the maximum possible value in Eq. 8 will give us the overlap ratio for each DM.

$$M_i = (I' - 1) \times l_i \tag{7}$$

$$\widetilde{O}_i = \frac{O_i}{M_i} \tag{8}$$

3.2.2. Relative CI calculation

To have the relative CI of each DM to the total CI, we use Eqs. 5 and 6 to find the CI_i. Then, Eqs. 9 and 10 are used to calculate the total CI based on matrix C input. Note that in Eqs. 11 and 12, the total number of scores in the decision matrix after eliminating the biased ones ($\lambda = N \times I'$) in the first step (matrix P^3) is used to find the CI.

$$\overline{x} = \frac{\sum_{i=1}^{I'} \sum_{j=1}^{J} \sum_{k=1}^{K} \psi_{ijk}}{\lambda}, \sigma = \sqrt{\frac{1}{\lambda - 1} \sum_{i=1}^{I'} \sum_{j=1}^{J} \sum_{k=1}^{K} \left(\psi_{ijk} - \overline{x}\right)^2} \tag{9}$$

$$CI = \left(\overline{x} - t(\lambda - 1, \alpha)\frac{\sigma}{\sqrt{\lambda}}, \overline{x} + t(\lambda - 1, \alpha)\frac{\sigma}{\sqrt{\lambda}}\right) \tag{10}$$

The calculated CI for each DM is then compared with the CI for total data in Eq. 11 to find the relative confidence interval value.

$$\widetilde{CI}_i = \frac{CI_i}{CI} \tag{11}$$

3.2.3. Weight calculation

Having the overlap ratio and the relative CI, we can calculate a baseline value for assigning weights to DMs as it is shown in Eq. 12.

$$w_i = \frac{\widetilde{O}_i \times \widetilde{CI}_i}{\sum_{i=1}^{I'} \left(\widetilde{O}_i \times \widetilde{CI}_i\right)} \tag{12}$$

Eq. 12 shows that a decision-maker has a higher weight (compared to other individual decision-makers) if that decision-maker has more relative agreement or similarity in their discrimination power with the other decision-makers.

3.3. MABM version of the proposed approach

As we mentioned, the MABM version is different from the EABM one in the weighting phase. After eliminating the biased DMs in the elimi-

Table 4
The normalized scores for the numerical example.

		DM1	DM2	DM3	DM4	DM5
Alternative 1	Criterion 1	0.48692	0.57043	0.67077	0.59369	0.36061
	Criterion 2	0.41586	0.75376	0.57771	0.51611	0.34695
	Criterion 3	0.54536	0.30971	0.49205	0.60229	0.21558
Alternative 2	Criterion 1	0.45033	0.67194	0.72409	0.60409	0.35013
	Criterion 2	0.52772	0.59235	0.69956	0.58426	0.28778
	Criterion 3	0.34813	0.83423	0.54178	0.62333	0.17443
Alternative 3	Criterion 1	0.48894	0.61275	0.75015	0.68182	0.16171
	Criterion 2	0.32891	0.72741	0.55521	0.56144	0.42836
	Criterion 3	0.48867	0.71386	0.71333	0.4974	0.30785
Alternative 4	Criterion 1	0.37971	0.43229	0.62814	0.49796	0.39818
	Criterion 2	0.37648	0.79708	0.63033	0.6325	0.26244
	Criterion 3	0.42778	0.53919	0.76235	0.68966	0.34949
Alternative 5	Criterion 1	0.33189	0.27685	0.70551	0.51662	0.20792
	Criterion 2	0.43112	0.56941	0.56176	0.57655	0.30157
	Criterion 3	0.33699	0.77226	0.72425	0.503	0.46047
Alternative 6	Criterion 1	0.54532	0.42602	0.64256	0.59018	0.24999
	Criterion 2	0.52755	0.77619	0.50469	0.63534	0.40827
	Criterion 3	0.40366	0.85662	0.41516	0.5835	0.46604

Table 5
The overlap ratio table.

	d2	d3	d4	$\sum O_{mn}$	O_i	$\widetilde{O}_i$
d2	0.0935	0.08216	0.04844	0.224101	0.130601	0.454877
d3	0.08216	0.10595	0.02032	0.20843	0.102479	0.623649
d4	0.04844	0.02032	0.00365	0.07241	0.068758	0.709726

Table 6
The relative CI.

	d2	d3	d4
CI_i	0.14356	0.08216	0.04844
$\widetilde{CI}_i$	2.06408	1.18133	0.69648

Table 7
The weight calculations.

	d2	d3	d4	Sum
w_i	0.43269	0.33952	0.2278	1

nation phase, a combinatory approach is used to assign weights. First, a minimum weight is assigned to all DMs. Eq. 13 shows this predefined minimum value, in which I' is the number of remaining DMs in the pool of GDM and γ parameter controls the total share of this part of weight in the total weight. Then the other part of weight is distributed among DMs based on the weighting scheme described in Section 3.2.3. Eq. 14 shows the method to compute the final weight for DMs.

$$w_i = \frac{\gamma}{I'} + \left(1 - \gamma\right) w_i \tag{13}$$

3.4. SABM version of the proposed approach

The SABM version skips the elimination phase, and only follows the logic of MABM version for all DMs in the initial pool. The rationale of this version is the fact that sometimes excluding DMs from the GDM pool is impractical, so that we need to tolerate biasedness in a logical way. Therefore, the part of the final weight that comes from the overlap of each DM with other DMs reflect the biasedness in these situations. Again, the minimum predefined weight is the same for this version as the SABM.

3.5. The pseudo-code of the proposed approach

As the final stage of this section, we present the pseudo-code for EABM version of the proposed framework. Since the SABM and MABM are slightly different from this version, we did not include the pseudo-codes for these variants.

3.6. Application and possible scenarios

Based on the nature of GDM problems, different versions of the proposed framework in this study can be applied on certain categories of problems. Table 2 shows the applicability of our framework based on the specific context of a GDM process. Table 3 also summarizes the proposed solution of the framework in all possible scenarios in terms of having outliers in the pool of DMs.

Table 8
The weight calculations for moderate version.

	d2	d3	d4	Sum
w_i	0.3830	0.3364	0.2806	1

Table 9
The weight calculations for soft version.

	d1	d2	d3	d4	d5	Sum
w_i	0.1	0.3163	0.2698	0.2139	0.1	1

Table 10
The levels for scenarios.

Value	Level 1	Level 2	Level 3
I	3	5	10
B	$\frac{I}{3}$	$\frac{I}{2}$	$I-1$
α	0.9	0.95	0.99

4. Computational results

This part includes an explanatory example to clarify the designed steps in the process, along with the analysis of the proposed method on a categorized series of test problems.

4.1. Numerical example

In this section, one example of a GDM problem with 5 DMs, 6 alternatives, and 3 criteria is considered to explain the steps of the proposed approach in detail. Table 4 shows the normalized scored assigned to the alternatives in each criterion sorted based on 5 present DMs in the process.

4.2. EABM version

As the first step, we need to calculate the CIs for all DMs, which is presented in Fig. 2.

Since d_1 and d_5 do not have any overlap with other DMs, they will be excluded from the pool of decision-making as outliers. This elimination is done by comparing the CIs of the DMs. For example, the CI of opinions for d_1 does not have any overlap with other DMs. Likewise, the CI of d_5 is outside the range of CIs for other DMs. On the other hand, d_2, d_3, and d_4 have overlap with two other DMs. By considering a threshold $B = 2, d_1$ and d_5 are eliminated since they do not have any overlap with other DMs. Then, Table 5 shows the overlap ratio for the remaining DMs. Please note that diagonal values are the length of CI for each DM (Refer to Section 3.2.1).

For instance, the overlap between d_2 and d_3 shows that the opinions

Table 11
The results of the regression analysis.

Measure	MAD		AWD		NEDM		REDM	
	Coeff	P-value	Coeff	P-value	Coeff	P-value	Coeff	P-value
NDM	−0.896***	0.000	−0.449	0.056	−0.833***	0.000	0.582 **	0.001
B	2.682***	0.000	1.830 ***	0.000	−1.652***	0.000	−1.873***	0.000
α	0.051***	0.000	0.044**	0.001	−0.017*	0.024	−0.095 ***	0.000
NDM and α	0.851***	0.000	0.467*	0.047	0.300*	0.024	−0.935***	0.000
B and α	−3.078***	0.000	−2.057***	0.000	2.189***	0.000	3.153***	0.000
NDM and B and α	0.215	0.000	-.053	0.178	0.670***	0.000	−0.4373***	0.000

Coeff = Coefficients & Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' '.

and discrimination power of these two DMs are more similar compared to other combinations. d_3 and d_4 have the least overlap in this specific case, though. Column $\sum O_{mn}$ is simply calculated by summation over each row, and column O_i is obtained by reducing the diagonal values form the summations. Finally, column $\widetilde{O}_i$ is derived by dividing the previous column over the maximum possible overlap value for each DM. This column shows that d_4 has the highest ratio, meaning that this DM has 70 percent of potential overlap with other DMs. In contrast, d_2 has the lowest ratio, which means its CI has a moderate overlap with others.

The next step is to calculate the relative CI for each DM. The result of this step is provided in Table 6.

The row CI_i reflects the length of CI for each DM. By dividing this row on the CI total, the next row of relative CI will be obtained. For example, the relative value for d_2 mirrors the fact that its CI is much wider than the CI of total opinions. Among the DMs in the pool, d_4 has the tightest CI compared to the CI total.

As the final step, the multiplication of overlap ratio and relative CI should be divided on the sum of these productions calculated to have the final weights of DMs. The result of this step is provided in Table 7. Having the weights, we can aggregate the opinions of the DMs in the pool. The weighted sum of scores for each alternative will provide the final ranking of the GDM problem.

4.3. MABM version

We need to first find the minimum predefined weight for the remaining DMs to obtain the results for this numerical example for the MABM version. In this example, we consider the γ parameter equal to 0.5 so that the total predefined comprises 50 percent of the total weight. We use the following formula to find this predefined weight for each DM:

$$w_m = \frac{0.5}{3} = 0.1667 \tag{14}$$

Then, we multiply the weight calculated in Table 7 by 0.5 to normalize this segment of the final weight. As the final step, we add the normalized weights to the minimum predefined weight to find the weights in this version. For example, for d_4 the weight is:

$$W_{d_4} = 0.1667 + (0.5) * (0.2278) = 0.2806 \tag{15}$$

The final weights are shown in Table 8.

As we can see from Table 8, the weights assigned to the three DMs are slightly different from those in Table 7, however the order of importance of the DMs remains the same.

4.4. SABM version

Finally, since we do not have any elimination in the SABM version, the threshold weight will be acquired from the following equation:

$$w_m = \frac{0.5}{5} = 0.1 \tag{16}$$

Then, the other part of weight is based on the overlaps between the DMs like Table 5. having these inputs, we can calculate the weights for all DMs. For instance, for DM2 the weight is calculated as:

$$W_{d_2} = 0.1 + (0.5) * (0.43269) = 0.3163 \tag{17}$$

Table 9 shows the final weights for all DMs in the poll in the soft version.

As we can see, the so-called biased DMs of the two other versions are not eliminated and get the minimum assigned weight of 0.1. The other three DMs receive different weights than those they got in other two versions, however the order of importance among the three DMS remains the same. It is important to note that as we technically assign the additional weights only to the unbiased DMs, they always gain more weight than the biased ones.

4.5. Test Scenarios

To test the performance of the proposed approach, we first identified the factors affecting the output of the process, including the number of DMs, the significance level for the statistical calculation of CI, and the B threshold. Table 8 shows the considered levels for the mentioned variables.

A total number of 27 combinations is considered as the benchmark problems for the proposed approach. In each category, 1,000 random instances are generated. The codes of the algorithm are written in MATLAB© software package and are executed on a PC with Intel® Xenon® Bronze 3160 1.70 GHz CPU with 16 GB RAM. On average, the computational time for each problem was about 22 min, resulting in 10-h time to obtain the results for all the cases.

4.6. Performance measures

To compare and analyze the performance of our approach, we defined four performance measures as follows:

4.6.1. Mean Absolute Deviation (MAD)

This indicator mirrors the average difference among the (normalized) weights of the DMs present in the decision-making process and their initial weight (which is considered equal among the DMs) before applying the method. Note that for the biased DMs, the final weights are considered zero. Eq. 18 captures this concept, in which w_i and $\widehat{w}_i = \frac{1}{I}$are the final normalized and the initial weights for the d_i, respectively.

$$MAD = \frac{\sum_{i=1}^{I} \left| w_i - \widehat{w}_i \right|}{I} \tag{18}$$

4.6.2. Average weight deviation (AWD)

This measure captures the deviation of weights assigned to the DMs remained in the GDM pool compared to the equally distributed weights scheme. Eq. 19 shows the formula for this performance indicator:

$$AWD = \frac{\sum_{i=1}^{I'} \left| w_i - \left(\frac{1}{I'}\right) \right|}{I'}, \forall i \in \{\text{Unbiased DMs}\} \tag{19}$$

4.6.3. Number of Eliminated Decision-Makers (NEDM)

As the number of DMs plays a significant role in the decision-making problems, we record the number of eliminated (biased) DMs as a result of applying our method on the GDM procedure. This measure can be calculated by subtracting the number of DMs in the second phased (I') and the beginning of the process (I) reflected in Eq. 18. Please note that this measure is not applicable for the soft version as there is no elimination in this setting.

$$NEDM = I - I' \tag{20}$$

4.6.4. Ratio of Eliminated Decision-Makers (REDM)

This performance measure captures the ratio of the eliminated DMs in the proposed framework to the total number of DMs in the GDM process. Eq. 19 shows the formulation for this indicator. Again, please note that this measure is not applicable for the soft version.

$$REDM = \frac{NEDM}{I} \tag{21}$$

4.7. Output analysis

This section includes the detailed results of the devised scenarios, the trend of the defined performance measures, and the noteworthy patterns. The analysis is performed in three parts. First, the trends of various elimination situations in the generated problems are presented to discuss the patterns and the prescription of the framework to handle the process. Then, to assess the relationship among variables, their interaction, and performance measures, a Pearson correlation analysis is executed. Finally, to evaluate the effect of the variables and their interactions on the performance measures, a multivariate linear regression analysis is designed. The rationale for these separated analyses is to assess two different aspects of the problem. By doing the correlation analysis, we want to investigate the whole possible combinations among the parameters and performance measures in a holistic approach. On the other hand, in the regression analysis, we want to evaluate the impact of variables and their interactions on each performance measure separately.

4.7.1. Elimination analysis

Fig. 3 visualizes the output of the designed scenarios in terms of elimination of DMs from the GDM pool. Please note that the scenarios are referenced by the levels presented in Table 10. For example, 111 scenario is the design with 3 DMs, $B = \frac{I}{3}$ and $\alpha = 0.9$ values. This figure demonstrates two trends in the test scenarios; (1) The share of problems in each case that none of the DMs are eliminated in the proposed framework (the first step) and (2) The share of problems in which a part of DMs (from 1 DM to I-1) are excluded from the final pool. This analysis is only plausible for the EABM and MABM versions of the proposed method.

The numbers of the scenarios are based on the levels of the parameters is Table 6. The share of problems with no elimination confronted a descending trend in general, meaning that as the number of DMs, B, and α rises, the possibility for the elimination of completely biased DMs goes up. In contrast, for partial elimination, the share grows as the parameters increase in the testbed. This trend will be explained in the regression analysis section. Finally, the share of problems in which there is no reachable consensus among the DMs in the first step is relatively constant with certain spikes. This share increases for test cases with higher levels (2 and 3) for B and α parameters. However, we included these problems in the class of problems where no DM is eliminated. In these cases, as reaching an agreement is impossible, there are some options. For one thing, the agent can tune the parameters (mainly B and α) to the point that a minimum desirable consensus is achievable. This minimum consensus is different in various context, like more than 50 percent in some political systems and even the biggest minority in other cases. So, this parameter is entirely in the control of the aggregator agent. Another option is considering an equally distributed weight for all DMs. As there is not any overlaps between the CIs of any two DMs, this strategy assures to keep all the DMs in the pool of GDM. Since we generated random cases in this study and there is no aggregator agent to tune the parameters, we considered the equal weight scenario for this specific situation. However, in real-life applications, both of the proposed solutions can be implemented when there is no consensus in the first phase.

4.7.2. Correlation analysis

Fig. 4 shows the results of the correlation analysis for the output of generated problems for the EABM (a), MABM (b), and the SABM (c) versions of the proposed algorithm. The numbers in the squares represent the correlation among variables and the crossed squares are the insignificant correlations.

There are some interesting correlations between the parameters and measures. For example, there is a strong correlation among NEDM measure and the interactions of NDM and B and NDM, B and α. On the other hand, a relatively meaningful negative correlation exists among MAD and AWD measures and NEDM and REDM measures, meaning they display opposite behaviors in the generated examples. Parameter α has weak correlations with all other ones, though. The patterns are almost similar in all versions, and the only difference is that the correlation between α and the MAD performance measure is insignificant in this configuration and the correlation between AWD and MAD is lower compared to other cases.

4.7.3. Regression analysis

Table 11 presents the results of a multivariate linear regression conducted on the output of the generated scenarios for the extreme version. To save the space, we present the tables for other versions in the appendix (a) for MABM and (b) for SABM.

Analyzing the patterns emerged in Table 11 will lead to the following insights:

(i) B has the highest impact on the weight-related measures, as higher value (level) forces the method to exclude more DMs. Therefore, the changes in the weights are more noticeable.

Algorithm 1: Pseudo-code of the proposed approach

Begin

/* GDM initialization and parameter setting/

Gather the opinions of DMs about all alternatives in each criterion;

Set the B and α value for CI; Combine the data in an integrated decision matrix;

Begin normalization Determine the nature of the criteria (positive or negative);

Find the Min and Max values for each criterion; Calculate the normalized values;

End

/* End of initialization/

/* Column elimination of biased DMs/

Transform the data to a column matrix; Calculate the CIs and B_i for each DM;

If B_i is lower than the B

Eliminate d_i as a biased one;

Else

Keep d_i in the decision-making process;

End if

/* End of column elimination /

If No consensus is reachable, or the number of the eliminated DMs is higher than R

Tune the B and α parameters to have a minimum consensus or give the equal weights to all DMs

End if

/* Weight assignment phase/

Calculate the CIs for the remaining DMs and the total data;

for $DM\ m = 1\ to\ I'$ **do**

  for $DM\ n = 1\ to\ I'$ **do**

    If CI_m has overlap with CI_n

      Calculate the overlap O_{mn} and add to the total overlap of O_m and O_n;

    Else

      The total overlap values will remain unchanged;

    End if

    End

    End

  end

end

for $DM\ i = 1\ to\ I'$ **do**

  Calculate the overlap ratio $\widetilde{O}_i$;

  Calculate the proportional value of CI_i to the total CI;

  Calculate the w_i by normalizing the product of $\widetilde{O}_i$ and $\widetilde{CI}_i$;

end

/* End Weight assignment phase /

/*Calculating performance measures/

Calculate MAD, AWD, NEDM, and REDM measures;

/*End of calculating the performance measures/

End

Conversely, this parameter also has negative coefficients for NEDM and REDM measures.

(ii) In the performance related to weights, MAD and AWD, parameter α has a high coefficient in the estimated models. This can be justified by the fact that by increasing α, the CIs become wider. As a result, the possible overlap among DMs increases, resulting in more fluctuations in their weights compared to the beginning of the process. However, the coefficient for α for two other performance measures, NEDM and REDM, is negative. In other words, with wider CIs, the number of eliminated DMs and the ratio to the total number of DMs decreases.

(iii) The NDM parameter has a considerable negative impact on weight related measures and a strong positive one on the NEDM measure. The effect of NDM on REDM measure is positive and significant, though.

(IV) Interestingly, the interaction of (B and α) has the highest negative effect on all the performance measures related to change of weights and the second highest positive one on the elimination-related measures. This intensified effect can be associated with the synergy of these parameters to convert a problem to an extreme one.

(V) The other interactions, (NDM and α) and (NDM and B and α) show conflicting behavior in terms of the direction of coefficients, yet all of the relations are statistically significant.

5. Conclusion and future research

This paper presents a structured framework to deal with biased decision-makers in the group decision-making process. Our focus was on aggregating the evaluations of a set of alternatives with respect to a set of criteria provided by a set of independent DMs. It starts with determining the entirely biased DMs and removes them from the decision-making pool. Then, it continues with remained DMs and assigns a weight to each DM based on its consensus compared with other DMs. The proposed algorithm is developed using two ratios, which are called Overlap Ratio and Relative Confidence Interval. The Overlap Ratio indicates the relative value of overlap between CI of each DM and other decision-makers by making the pairwise comparison of their confidence intervals. The Relative Confidence Interval shows the relative joint range of CI for each DM compared to the total CI of the remained decision-makers.

The proposed method is one of the first attempts in the context of group decision making where one is concerned with biased decision-makers, tries to detect the sources of biases systematically, and finally to handle the partially biased DMs. One of the main applications of our proposed methods is where there is an agent who is responsible for aggregating the DMs' opinion, and DMs do not have information either about aggregation procedure or the opinion of other DMs. In this setting, the biased DMs can be excluded from the final decision-making pool. However, there are two possible scenarios in case DMs do have information about the aggregation procedure. Either collective agreement is acceptable by all DMs, or it is not. The proposed framework is not applicable in the latter case. Another example can happen when two DMs have complete overlap in their confidence interval which could lead to equal weights despite the fact that they do not have full agreement in terms of scores they assign to the options. Finally, our developed method can be used in the initial phase of any group decision-making procedure, falling into the mentioned categories. The proposed framework also provides satisfactory flexibility in dealing with all possible situations in a GDM problem. If the pool of DMs does not contain any outlier, the weights of DMs are calculated based on the level of accordance with the total opinion. However, the framework can detect the outliers and handle the biasedness in a systematic approach in other cases. In the presence of biasedness, the framework provides different prescription for the cases. The biased DMs can be excluded (if possible) from the GDM problem (EABM version) or they can be kept but assign their partial weights based on a combination of a minimum predefined weight and their disagreement level from other DMs (SABM version). As a middle way, the MABM version excludes the biased DMs, but assign weights to the remaining DMs based on a combination of minimum weight and relative weight based on the level of agreement. Finally, if an agreement is unreachable in the GDM, the framework suggests revising the opinions or having new sources for the evaluation of alternatives based on the criteria. As a result, we can claim that the proposed framework and its variants cover adequately all possible scenarios in a GDM setting from strict exclusion of biased DMs to only reflecting biasedness in the weighting phase.

Several paths can be followed as fruitful research directions for expanding the current study. This research tried to detect the outlier DMs in total. That is to say, we consider a DM biased if they give similar scores to all alternatives. These DMs have small deviations in their scoring which implies their confidence intervals are very concentrated towards the mean value and have no overlap with the confidence intervals of others (so no agreement). We call these DMs with inadequate discrimination power as biased DMs. However, we aim to develop another comprehensive method to detect biased DMs towards particular alternatives. In that regard, a DM who is treated as a non-biased DM is this study might be found biased towards some alternatives. We could then think of mirroring this biasedness by either ignoring their opinions for those alternatives or assigning relatively less weight to them. Another direction could be applying our proposed method is real-world group decision-making problems. This area of research requires more attention, particularly developing new algorithms for individual cases. In specific, in the political context, the decisions are highly correlated (similar opinions of voters belong to a political party for example). Even though this situation is different from a routine GDM problem, the proposed methodology in this study can be extended to incorporate the correlated decisions and handle these inherently different outliers as they adversely affect the output of the aggregation phase.

CRediT authorship contribution statement

Meysam Rabiee: Conceptualization, Methodology, Writing - review & editing. **Babak Aslani:** Conceptualization, Methodology, Writing - original draft, Writing - review & editing, Software, Formal analysis. **Jafar Rezaei:** Conceptualization, Methodology, Conceptualization, Methodology, Writing - review & editing, Supervision.

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.

Appendix A

Table 12
The results of the regression analysis for moderate version.

Measure	MAD		AWD		NEDM		REDM	
	Coeff	P-value	Coeff	P-value	Coeff	P-value	Coeff	P-value
NDM	−0.649**	0.005	−1.092***	0.000	−0.833***	0.000	0.582***	0.001
B	2.083***	0.000	2.451***	0.000	−1.652***	0.000	−1.873***	0.000
α	0.068***	0.000	0.067***	0.000	−0.017*	0.024	0.095***	0.000
NDM and α	0.924***	0.000	1.364***	0.000	0.300 *	0.024	−0.935***	0.000
B and α	−2.978***	0.000	−2.947***	0.000	2.189***	0.000	3.153***	0.000
NDM and *B* and α	0.507***	0.000	0.206***	0.000	0.670***	0.000	−0.437***	0.000

Coeff = Coefficients & Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' '.

Appendix B

Table 13
The results of the regression analysis for soft version.

Measure	MAD		AWD	
	Coeff	P-value	Coeff	P-value
NDM	−0.323***	0.000	−0.462***	0.000
B	1.762***	0.000	1.913***	0.000
α	0.184***	0.000	0.046***	0.000
NDM and *B*	0.547***	0.000	1.141***	0.000
NDM and α	−0.118***	0.000	0.013	0.326
B and α	−0.455***	0.000	0.060.	0.061
NDM and *B* and α	−0.169	0.281	−1.470***	0.000

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' '.

References

Akram, M., Adeel, A., & Alcantud, J. C. R. (2019). Group decision-making methods based on hesitant N-soft sets. *Expert Systems with Applications, 115*, 95–105.
Akram, M., Bashir, A., & Garg, H. (2020). Decision-making model under complex picture fuzzy Hamacher aggregation operators. *Computational and Applied Mathematics, 39* (3), 1–38.
Akram, M., Yaqoob, N., Ali, G., & Chammam, W. (2020). Extensions of Dombi Aggregation Operators for Decision Making under-Polar Fuzzy Information. Journal of Mathematics, 2020.
Ambrus, A., Greiner, B., & Pathak, P. (2009). Group versus individual decision-making: Is there a shift. Institute for Advanced Study. *School of Social Science Economics Working Paper, 91*.
Anojkumar, L., Ilangkumaran, M., & Sasirekha, V. (2014). Comparative analysis of MCDM methods for pipe material selection in sugar industry. *Expert Systems with Applications, 41*(6), 2964–2980.
Bonner, B. L., Baumann, M. R., & Dalal, R. S. (2002). The effects of member expertise on group decision-making and performance. *Organizational Behavior and Human Decision Processes, 88*(2), 719–736.
Bouzarour-Amokrane, Y., Tchangani, A., & Peres, F. (2015). A bipolar consensus approach for group decision making problems. *Expert Systems with Applications, 42* (3), 1759–1772.
Ceschi, A., Costantini, A., Sartori, R., Weller, J., & Di Fabio, A. (2019). Dimensions of decision-making: an evidence-based classification of heuristics and biases. *Personality and Individual Differences, 146*, 188–200.
Cinelli, M., Kadzinski, M., Gonzalez, M., & Slowinski, R. (2020). How to Support the Application of Multiple Criteria Decision Analysis? Let Us Start with a Comprehensive Taxonomy. *Omega, 102261*.
Dong, Q., & Cooper, O. (2016). A peer-to-peer dynamic adaptive consensus reaching model for the group AHP decision making. *European Journal of Operational Research, 250*(2), 521–530.
Dong, Q., Zhou, X., & Martínez, L. (2019). A hybrid group decision making framework for achieving agreed solutions based on stable opinions. *Information Sciences, 490*, 227–243.
Dong, Y., Zha, Q., Zhang, H., Kou, G., Fujita, H., Chiclana, F., & Herrera-Viedma, E. (2018). Consensus reaching in social network group decision making: Research paradigms and challenges. *Knowledge-Based Systems, 162*, 3–13.
Forman, E., & Peniwati, K. (1998). Aggregating individual judgments and priorities with the analytic hierarchy process. *European Journal of Operational Research, 108*(1), 165–169.
Gonzalez-Arteaga, T., Alcantud, J. C. R., & de Andres Calle, R. (2016). A new consensus ranking approach for correlated ordinal information based on Mahalanobis distance. *Information Sciences, 372*, 546–564.
Hafezalkotob, A., & Hafezalkotob, A. (2017). A novel approach for combination of individual and group decisions based on fuzzy best-worst method. *Applied Soft Computing, 59*, 316–325.
He, Z., Chan, F. T., & Jiang, W. (2018). A quantum framework for modelling subjectivity in multi-attribute group decision making. *Computers & Industrial Engineering, 124*, 560–572.
Ichniowski, C., & Preston, A. (2017). Does March Madness lead to irrational exuberance in the NBA draft? High-value employee selection decisions and decision-making bias. *Journal of Economic Behavior & Organization, 142*, 105–119.
Jacobi, S. K., & Hobbs, B. F. (2007). Quantifying and mitigating the splitting bias and other value tree-induced weighting biases. *Decision Analysis, 4*(4), 194–210.
Jin, F., Ni, Z., Pei, L., Chen, H., Tao, Z., Zhu, X., & Ni, L. (2017). Approaches to group decision making with linguistic preference relations based on multiplicative consistency. *Computers & Industrial Engineering, 114*, 69–79.
Jin, L., Mesiar, R., & Yager, R. R. (2018). The paradigm of induced ordered weighted averaging aggregation process with application in uncertain linguistic evaluation. *Granular Computing*, 1–7.
Jones, P. E., & Roelofsma, P. H. (2000). The potential for social contextual and group biases in team decision-making: Biases, conditions and psychological mechanisms. *Ergonomics, 43*(8), 1129–1152.
Kabak, Ö., & Ervural, B. (2017). Multiple attribute group decision making: A generic conceptual framework and a classification scheme. *Knowledge-Based Systems, 123*, 13–30.
Lee, Y.-C. (2014). Impacts of decision-making biases on eWOM retrust and risk-reducing strategies. *Computers in Human Behavior, 40*, 101–110.
Liao, H., Wu, X., Mi, X., & Herrera, F. (2019). An integrated method for cognitive complex multiple experts multiple criteria decision making based on ELECTRE III with weighted Borda rule. *Omega*.
Liao, H., Zhang, C., & Luo, L. (2018). A multiple attribute group decision making method based on two novel intuitionistic multiplicative distance measures. *Information Sciences, 467*, 766–783.
Lieder, F., Griffiths, T. L., Huys, Q. J., & Goodman, N. D. (2018). The anchoring bias reflects rational use of cognitive resources. *Psychonomic Bulletin & Review, 25*(1), 322–349.
Liu, P. (2011). A weighted aggregation operators multi-attribute group decision-making method based on interval-valued trapezoidal fuzzy numbers. *Expert Systems with Applications, 38*(1), 1053–1060.
Liu, P., Shahzadi, G., & Akram, M. (2020). Specific Types of q-rung Picture Fuzzy Yager Aggregation Operators for Decision-Making. *International Journal of Computational Intelligence Systems, 13*(1), 1072–1091.
Liu, W., Zhang, H., Chen, X., & Yu, S. (2018). Managing consensus and self-confidence in multiplicative preference relations in group decision making. *Knowledge-Based Systems, 162*, 62–73.
Liu, Y., Du, J.-L., & Wang, Y.-H. (2019). An improved grey group decision-making approach. *Applied Soft Computing, 76*, 78–88.
McShane, M., Nirenburg, S., & Jarrell, B. (2013). Modeling decision-making biases. *Biologically Inspired Cognitive Architectures, 3*, 39–50.
Mohammadi, M., & Rezaei, J. (2020). Bayesian Best-Worst Method: A Probabilistic Group Decision Making Model. *Omega, 96*, Article 102075.
Mokhtari, S. (2013). Developing a Group Decision Support System (GDSS) For Decision Making Under Uncertainty.
Von Montibeller, G., & Winterfeldt, D. (2015). Cognitive and motivational biases in decision and risk analysis. *Risk Analysis, 35*(7), 1230–1251.
Morente-Molinera, J. A., Kou, G., Samuylov, K., Urena, R., & Herrera-Viedma, E. (2019). Carrying out consensual Group Decision Making processes under social networks using sentiment analysis over comparative expressions. *Knowledge-Based Systems, 165*, 335–345.
Palomares, I., Estrella, F. J., Martínez, L., & Herrera, F. (2014). Consensus under a fuzzy context: Taxonomy, analysis framework AFRYCA and experimental case of study. *Information Fusion, 20*, 252–271.
Parreiras, R. O., Ekel, P. Y., Martini, J. S. C., & Palhares, R. M. (2010). A flexible consensus scheme for multicriteria group decision making under linguistic assessments. *Information Sciences, 180*(7), 1075–1089.
Rezaei, J. (2020). Anchoring bias in eliciting attribute weights and values in multi-attribute decision-making. *Journal of Decision Systems*. in press.
Rokou, E., & Kirytopoulos, K. (2014). A calibrated group decision process. *Group Decision and Negotiation, 23*(6), 1369–1384.

Song, Y. & Hu, J. (2019). Large-scale group decision making with multiple stakeholders based on probabilistic linguistic preference relation. Applied Soft Computing, 80, 712-722.

Stengel, D. N. (2013). Aggregating incomplete individual ratings in group resource allocation decisions. *Group Decision and Negotiation, 22*(2), 235–258.

Stewart, T. (2005). Goal programming and cognitive biases in decision-making. *Journal of the Operational Research Society, 56*(10), 1166–1175.

Tang, J., Meng, F., & Zhang, Y. (2019). Programming model-based group decision making with multiplicative linguistic intuitionistic fuzzy preference relations. *Computers & Industrial Engineering*.

Vafaei, N., Ribeiro, R. A., & Camarinha-Matos, L. M. (2016). Normalization techniques for multi-criteria decision making: analytical hierarchy process case study. In *Paper presented at the doctoral conference on computing, electrical and industrial systems*.

Wibowo, S., & Deng, H. (2013). Consensus-based decision support for multicriteria group decision making. *Computers & Industrial Engineering, 66*(4), 625–633.

Wittenbaum, G. M., Vaughan, S. I., & Strasser, G. (2002). Coordination in task-performing groups. In *Theory and research on small groups* (pp. 177–204). Boston, MA: Springer.

Wu, Z. & Xu, J. (2016). Managing consistency and consensus in group decision making with hesitant fuzzy linguistic preference relations. Omega, 65, 28-40.

Xia, M., & Chen, J. (2015). Multi-criteria group decision making based on bilateral agreements. *European Journal of Operational Research, 240*(3), 756–764.

Xiao, B., & Benbasat, I. (2018). An empirical examination of the influence of biased personalized product recommendations on consumers' decision making outcomes. *Decision Support Systems, 110*, 46–57.

Xie, K., Liu, J., Chen, G., Wang, P., & Chaudhry, S. S. (2012). Group decision-making in an unconventional emergency situation using agile Delphi approach. *Information Technology and Management, 13*(4), 351–361.

Xu, W., Huang, S., & Li, J. (2019). A novel consensus reaching framework for heterogeneous group decision making based on cumulative prospect theory. *Computers & Industrial Engineering, 128*, 325–335.

Xu, X.-H., Du, Z.-J., Chen, X.-H., & Cai, C.-G. (2019). Confidence consensus-based model for large-scale group decision making: A novel approach to managing non-cooperative behaviors. *Information Sciences, 477*, 410–427.

Xu, Z.-S., & Jian, C. (2007). Approach to group decision making based on interval-valued intuitionistic judgment matrices. *Systems Engineering-Theory & Practice, 27*(4), 126–133.

Yu, G.-F., Li, D.-F., & Fei, W. (2018). A novel method for heterogeneous multi-attribute group decision making with preference deviation. *Computers & Industrial Engineering, 124*, 58–64.