Information Fusion 121 (2025) 103157

Contents lists available at ScienceDirect

Information Fusion

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

Divide and conquer? A combination of judgments method for comparing DSSs. Pairwise comparison vs. holistic paradigms

Carlos Sáenz-Royo [a,*], Francisco Chiclana [b]

[a] University of Zaragoza, Department of Business Organization and Management, C/Pedro Cerbuna, 12, 50009 Zaragoza, Spain
[b] School of Computer Science and Informatics, De Montfort University, Leicester, UK

ARTICLE INFO

Keywords:
DSS
Holistic evaluation
Pairwise comparison
IBR

ABSTRACT

Despite the prevalence of Decision Support Systems (DSSs) in the field of decision-making, there is a paucity of research dedicated to the evaluation and comparison of these systems. This paper put forward a novel approach to symbolically encoding a DSS, which enables the generalization of comparisons between DSSs for any distribution of performances of the alternatives. The only hypothesis required in the proposed methodology is that the probability of choosing each alternative is proportional to its latent performance. The approach developed is demonstrated with its application to compare two paradigms commonly employed in DSS: holistic versus pairwise. Using a set of three alternatives, the present study provides mathematical proof that a DSS based on the pairwise comparison paradigm achieves higher expected performance than a DSS based on the holistic evaluation paradigm. This result challenges the emerging preference for holistic evaluation of alternatives and suggests that this result may apply to any number of alternatives.

1. Introduction

A Decision Support Systems (DSSs) is used to describe an integrated system of algorithms to collect, structure, and process information on a decision problem in an attempt to discover the performance of a set of (feasible decision) alternatives [1–4]. Specifically, this paper focuses on the priority vector of the set of alternatives generated by DSSs with information on the alternatives represented by the expert's judgments regarding their performance based on the expert's comparisons between alternatives. Currently, there is no consensus on the key elements that determine the quality and effectiveness of these DSSs. Furthermore, the comparison of these systems is an understudied area of research.

Unstructured decisions seek to respond to dynamic problems pervaded by uncertainty and, therefore, they lack prior systematized information, making it difficult to formalize the relationships between variables [5] and becoming expert's judgments to be the crucial source of information [6–10]. Since many DSSs are based on expert judgments as information, an important relevant research question to address concerns their comparative analysis "a priori" (before soliciting expert input) to determine the optimal DSS for addressing the specific problem at hand.

To reveal which DSS exploits best the decision information, a comparison of different DSSs based on their expected performance with respect to different distributions of the latent performances of the alternatives is carried out. Latent performance refers to the real but unobservable value of each alternative, which exists independently of the experts' judgments. This value represents the objective utility or desirability of each alternative under the criteria considered.

The comparison of different DSSs based on experts' judgments is a challenging task to model due to the numerous existent decision-making approaches and their diverse characteristics. In addition, the performances of the alternatives may also be difficult to determine a priori and, sometimes, even a posteriori (after the decision has been made), especially when it is only possible to quantify the performance of the course of action but not of all the alternatives [11]. Thus, to reveal which DSS exploits best the decision information, an a priory approach to compare DSSs is based on their expected performance with respect to different distributions of the latent performances of the alternatives, while a posteriori performance measures used in existent research reported in the literature include the level of consensus [12–15], the level of expert consistency [16,17], and sensitivity analyzes [18,19], among others.

In any case, it is challenging to compare the performance of DSSs that are difficult to quantify. Traditionally, it has been assumed that the

* Corresponding author.
E-mail address: csaenz@unizar.es (C. Sáenz-Royo).

https://doi.org/10.1016/j.inffus.2025.103157
Received 11 June 2024; Received in revised form 23 March 2025; Accepted 25 March 2025
Available online 26 March 2025

higher the level of information collected by a DSS, the higher its performance would be. Most comparative studies between DSSs informed by experts' judgments are descriptive in nature and focus on analyzing differences with regard to their underlying methodology. Consequently, the findings are often difficult to generalize. Within this category of comparative studies, we can cite the empirical work by Opricovic and Tzeng [20] on how the application of different techniques can yield disparate conclusions; Wallenius et al. [21] analysis of the popularity of decision-making techniques based on the number of times they had been used in the literature as a proxy to derive a ranking of "quality;" Triantaphyllou [22] comprehensive comparison of decision-making techniques based on a sensitivity analysis of decision outcomes with regard to various parameters; Belton [23] "a priori" theoretical and "a posteriori" practical mathematical analysis of how different decision-making techniques model expert judgments. There is, though, a paucity of studies that propose a priori comparisons between DSSs. The most notable study is that of Zanakis et al. [24], in which scenarios with different numbers of alternatives, criteria, and distributions are simulated to analyze the degree of similarity between the compared decision-making techniques. In general, to the best of our knowledge, the evaluation of DSSs has been limited to sensitivity analysis, consensus levels, and some simulation analysis. No study has specified the superiority of a DSS but only differences between the compared DSSs as reviewed above.

This paper proposes an approach based on Simon's conceptual framework of bounded rationality[5,25,26]. This framework hypothesizes that there is a unique correct priority vector of the alternatives and that different priority vectors result from the inability of a DSS to manage all the available decision information. This framework diverges from the traditional framework for DSS based on expert judgments, which posits a non-deterministic (utility or other) or unknown performance of the alternatives [27], thereby impeding the ability to distinguish between errors of judgment or differences in preferences. The absence of a clearly defined objective necessitates a comprehensive examination of each problem [28] and a degree of flexibility on the part of the DSS to adapt to the specific circumstances of the problem [29]. This traditional framework acknowledges the challenges associated with aggregating and comparing judgments, and the contribution of a DSS must be substantiated through indirect measures, such as the level of agreement, consistency, or resilience of the result (an analysis of the different aspects to consider can be obtained from [30,31]).

This paper put forward a novel methodology for the a priori comparative analysis of DSSs, wherein the probability of selecting each alternative is used as the primary metric. The proposed approach represents a shift in research focus, and it necessitates a novel method of encoding DSSs. The proposed modeling approach is based on the regularities of human judgments and the computation of performance distributions of the alternatives, and it is designed to contribute to the systematization and generalization of DSS analyzes beyond the scope of case studies because it does not necessitate the establishment of a specific scenario and computes the DSS's expected performance value by appropriately substituting the latent performances of the alternatives. Therefore, this paper research contribution addresses the a priori comparative analysis of DSSs issue with a fully symbolic modeling approach of a DSS that elucidates the configuration of the priority vector of alternatives and ultimately facilitates a more comprehensive understanding of the design of DSSs. This constitutes a significant contribution to the existing literature on a posteriori comparison approaches (for a set of given judgments) and the shortage of a priori comparison approaches (simulations with a distribution of performance of alternatives to which a known distribution error is added). Moreover, the proposed methodology is of interest to company managers and directors because it provides an understanding of the potential profit percentage gains that could be achieved through the implementation of specific DSSs. For researchers in this field, the proposed method offers a means of justifying the development of new DSSs through a comparison of their efficiency across a range of potential scenarios.

One of the techniques proposed to improve decision performance is based on the principle of "simplification," which suggests that complex problems can be broken down into smaller, more manageable problems whose solutions are subsequently aggregated methodologically to obtain the solution to the original complex problem [32]. On the one hand, making simple judgments has been demonstrated to enhance several psychometric properties, including convergent validity and temporal stability, thereby facilitating greater precision in judgments [33–36]. On the other hand, the statistical properties of DSSs have been shown to exhibit lower variances in the final decision [37,38]. The most commonly used DSSs, such as the Analytic Hierarchy Process (AHP) [39], ELECTRE [40], and PROMETHEE [41], are based on the principle of simplification applied to the comparison of the set of alternatives through their pairwise comparisons. Since a significant drawback of this approach is the high number of judgments needed from experts and its corresponding computational resources from the DSS [42], other scholars argue that a comprehensive assessment of potential options is essential, citing several reasons for this approach: (i) it is a more accurate method than making partial pairwise comparisons of alternatives [43,44]; (ii) it reduces the cognitive load required of experts when making a large number of pairwise comparisons of alternatives [45]; (iii) it captures interactions and synergies in the decision that may be lost when making pairwise comparisons of alternatives [46]; (iv) it circumvents the scaling and normalization issues associated with aggregating simple pairwise comparison judgments [47]; (v) it is a more "natural" means of expressing preferences in a direct and comprehensive manner [48]; (vi) it accelerates the decision-making process by eliminating the need for extensive elicitation of preferences or consistency [42], which can have significant in expected performance [49,50]; (vii) it permits the incorporation of emotional responses, intuition, and other intangible factors into the evaluation process [51]; (viii) it is more resistant to change over time [52]; and (ix) experts are more likely to be invested in decisions made through a holistic approach [53]. One of the disadvantages of this approach is that a single judgment does not provide insight into the reasoning that led the expert to express their opinion. This lack of transparency makes it challenging to audit the decision-making process.

The majority of studies assume the superiority of the pairwise comparison paradigm [54–60], although this assumption has not been mathematically demonstrated. This may be because the comparison has not been adequately modeled. This paper investigates the possibility of demonstrating that a DSS based on a pairwise comparison of alternatives has a higher expected performance value than a DSS based on a holistic evaluation of alternatives in all possible scenarios. This is done in the context of experts' judgments that show regularities in the probabilities of success and failure in the case of three alternatives. However, the proposed model allows the inequality to be computed for any number of alternatives, with no limitation other than the computational capacity of the mathematical software used for resolution.

The rest of the paper is organized as follows. Section 2 presents a new proposal for modeling DSSs to allow for a general comparison approach to DSSs. Section 3 analyzes the representations of regularities in experts' judgments, linking this concept with Simon's concept of bounded rationality [61]. This Section proposes an approach to represent both the holistic evaluation of alternatives and the pairwise comparison of alternatives. Section 4 demonstrates the superiority, regarding the expected performance value, of the pairwise comparison paradigm, i.e. it is proved that the pairwise comparison of alternatives paradigm has a higher expected performance value than the holistic comparison of alternatives paradigm. Section 5 discusses some results of the proposed model and suggests potential future research avenues among the multiple possibilities provided by the new modeling. Section 6 presents the main conclusions of this research study.

2. Model

DSSs based on experts' judgments are particularly valuable for problems that are challenging to quantify and for which no prior information is available. Generalized modeling of judgments is a complex undertaking due to the considerable number of DSS approaches. Having said this, the objective of the proposed modeling herein is to obtain all possible *combinations of judgments* required by a DSS to compute the probability of choosing each of the feasible alternatives. Furthermore, we aim to work with this probability function symbolically, for which it is required only that the probability of choosing each alternative is related to its performance value via an increasing function (elaborated in Section 3). Without loss of generality, it is assumed that there is a set of different (positive) latent performance values $V = \{V_1, ..., V_n\}$ associated with a set of n alternatives $A = \{A_1, ..., A_n\}$. The assumption of having different latent performance values is based on the retention of only one of any two alternatives with the same latent performance value.

A DSS based on judgments can be viewed as a function with input variables the expert's judgments and precision values ($g_{ix}(r)$ represents the judgment on how alternative A_i performs with respect to x, with a precision value $r \in R$), and with output variable a priority vector (ranking) of the alternatives (s):

$$s = DSS(g_{1x}(r), \dots,\ g_{nx}(r)). \tag{1}$$

The number of judgments (M) for a "*combination of judgments*" denoted $G = (g_{1x}(r),\dots,\ g_{nx}(r))$, with $|G| = M$, and the precision scale (R) are determined by the considered DSS. The "*combination of judgments*" depends on the comparison of alternatives paradigm used by the DSS. Indeed, a DSS based on the pairwise comparison of alternatives paradigm requires $M = n(n-1)/2$ judgments $g_{ij}(r)$ $(i < j)$ representing the expert's intensity of the preference of alternative A_i over alternative A_j as modeled with precision $r \in R$; a DSS based on the holistic comparison of alternatives paradigm does not require intensity of precision because the expert requires to provide just a single judgment identifying which of the alternatives is the best, i.e. a combination of judgment is represented as a unitary vector $g_{iA} = (0,\ 0, \dots,\ 1, \dots,\ 0)$. In the pairwise comparison of alternatives paradigm, the precision $r \in R$ refers to the quantification of the intensity of the preference when comparing two alternatives; different scales are possible, with the simplest scale of preference having cardinality 2, $R = \{0, 1\}$, used to represent "preferred" ($r = 1$) vs. "less preferred" ($r = 0$); a finer scale of cardinality 4, $R = \{4, 2, 1/2, 1/4\}$, may be used to represent intensity of preference as follows: "very preferred" ($r = 4$); "preferred" ($r = 2$); "less preferred" ($r = 1/2$); "much less preferred" ($r = 1/4$).

The universe of events (γ) comprises all *combinations of judgments* (K) that an expert can express, $\gamma = \left(G^1, \dots, G^K\right)$, which depends on the granularity of the precision scale (R) and the number of judgments (M). The K values for the pairwise comparison of alternatives paradigm and the holistic comparison of alternatives paradigm are $R^{n(n-1)/2}$, and n, respectively. Unless it is necessary to explicitly state the precision scale, the comparative judgment $g_{ix}(r)$ will be denoted simply as g_{ix}. Since it is assumed that the probability of an expert making a particular judgment is a function (increasing) of the performance of the alternative, then the *combination of judgments* G^k will be assigned the probability value $p_{G^k} = \prod_{i}^{n} p_{g_{ix}^k}$, where $p_{g_{ix}}$ denotes the probability of the comparative judgment $g_{ix}(r)$. From a *combination of judgments,* a priority vector of alternatives is derived. Thus, the probability that the DSS chooses alternative A_i, p_{A_i}, is obtained as the sum of the probability values of all *combinations of judgments* in γ with priority vector having A_i as the best alternative. As it has been described, these probability values of alternatives are obtained symbolically, i.e. without the need to know the performances of the alternatives but the quantification of the expert's judgments on the set of alternatives. These symbolic probability values of alternatives are used to compute the expected performance value of a DSS and, therefore, it will be possible to compare different DSSs as per their expected performance values.

This proposed modeling can be synthesized in 7 steps:

Step 1. Determine the number of judgments and the precision of the DSS comparison paradigm.
Step 2. Assign a probability value to each judgment-precision.
Step 3. List all *combinations of judgments* (universe of events).
Step 4. Compute the symbolic probability value of each *combination of judgments*.
Step 5. Derive the priority vector of each *combination of judgments*.
Step 6. Compute the symbolic probability of choosing each alternative.
Step 7. DSSs are compared using their corresponding symbolic probability values of alternatives.

The following subsections justify the use of an increasing probability function on the performance values of the alternatives and its application to the holistic and pairwise comparison paradigms.

3. Alternative judgment probability and alternative performance values

The work of Luce [62] was pioneering in providing a theoretical justification for modeling choice as a probabilistic process. This modeling is compatible with empirical evidence supporting the possibility of experts showing intransitive and inconsistent preferences, while also manifesting well the limited capacity of experts to discriminate small differences in performance between alternatives. Luce proposes a choice rule where the probability of choosing an alternative is proportional to its performance compared to other available alternatives. This captures the relative nature of choice by showing how its probabilistic theory fits well with experimental data of choice between alternatives.

Luce's proposal serves as a pivotal reference point for the subsequent evolution of probabilistic choice theories. Luce's approach results in the summarization of the information processing of each expert in the form of a probability function for each alternative. This is done as if it were a physical law governing the behavior of a particle. Each expert (decision-maker) expresses their judgment in each relationship in favor of one of the alternatives with a certain level of error. In other words, the decision-maker aims to anticipate the value of the corresponding latent performances of the alternatives through the analysis of information signals.

The logic behind the fact that the probability of choice of the alternatives is proportional to their performances is that the human judgment has less probability of error when the performances of the alternatives have very disparate values than when the performance values are very similar. The expert's mental comparison means that each alternative will have a probability of being chosen proportional to the latent performance, which indicates that human judgment is more likely to get it right than to fail.

To avoid scale effects, the performances of alternatives are normalized $v_i = V_i / \sum_{k=1}^{n} V_k$. Luce proposes the existence of a function that relates the latent performances of the compared alternatives and the probability of choosing each alternative: $p_i = \vartheta_i(v_1,\dots,v_n)$, where p_i is the probability of choosing alternative A_i. There are some theories that model the probabilistic distribution of human choice, which according to their functional form can be grouped into two classes:

Proportional representation: the most prominent theories that use this representation are those of proportional selection theories in evolutionary algorithms [63–65]. These theories, when transferred to the decision framework, maintain the idea that natural selection occurs proportionally to performance. The main algorithms are "Odds" [66,67] and "Fitness proportionate selection" [68,69]. In this group, the probability of choosing alternative A_i has the form:

$$p_i = \frac{v_i}{\sum_{j=1}^{n} v_k} = \frac{\frac{V_i}{\sum_{k=1}^{n} V_k}}{\sum_{j=1}^{n} \frac{V_i}{\sum_{k=1}^{n} V_k}} \tag{2}$$

Representation based on logistic function: The most notable theories that use this representation are: i) the discrete choice theory [70,71], which justifies the logistic functional form because it is the one that fits best the empirical evidence of error distributions of decision makers; ii) the full-range log-odds model [72–75], which establishes a relationship between four constructs (true judgment, error, overconfidence, and response) and the logistic function reproduces the empirical regularities found in error; and iii) the Intentional Bounded Rationality (IBR) methodology [76], which justifies this functional form as a consequence of the human way of thinking which only makes an effort to improve its precision if the balance is positive (greater performance obtained by improving the precision minus the effort required). In this group, the probability of choosing alternative A_i has the form:

$$p_i = \frac{e^{v_i}}{\sum_{j=1}^{n} e^{v_j}} = \frac{1}{1 + \sum_{j\neq i}^{n} e^{v_j - v_i}} = \frac{1}{1 + \sum_{j\neq i}^{n} e^{-(v_i - v_j)}} \tag{3}$$

The IBR theory is worthy of further consideration due to its enhanced capacity for representing a broader range of phenomena with greater generality. Its main hypothesis is that all information is available but human limitations do not allow it to be handled, which leads to errors. Simon's concept of bounded rationality [61] inspires this theoretical framework: when making decisions, experts do the best information processing they can within their limitations (decisions are not purely random). Thus, according to the IBR theory, people's fallibility is assumed not to be the result of noisy information signals [37,38,77,78] but the consequence of bounded but intentional rationality. The more precise the analysis or judgment must be, the higher the level of information that must be processed. Trade-offs between alternatives are expressed in relative terms because it is reasonable to assume that the probability of making a choice error is captured more accurately by the differences between the alternatives' performance values: $v_i - v_j$.

The IBR maintains that the ability to anticipate the performance of each alternative depends on the expertise of the individual by adding to Eq. (3) a parameter $\beta(\geq 0)$ associated with the human capital of the expert. Thus, the IBR defines the probability with which the expert chooses an alternative A_i as:

$$p_{\beta i} = p_\beta(A_i) = \frac{e^{\beta v_i}}{\sum_{j=1}^{n} e^{\beta v_j}} = \frac{1}{1 + \sum_{j\neq i}^{n} e^{-\beta(v_i - v_j)}}. \tag{4}$$

Notice, that with value $\beta = 1$ Eq. (4) becomes Eq. (3), so this last model is more general. In the IBR theoretical framework, the error is a consequence of the expert's inability to process all the information signals that perfectly anticipate the performance of every one of the alternatives, which is captured via the parameter β. The parameter β is important because it measures the expert's ability to process information. When the expert knows nothing about a problem, i.e. when an expert does not have the ability to process information or does not know what the relevant information is, then $\beta = 0$, and all alternatives have the same probability of being chosen regardless of their latent performance values. High values of β imply a high capacity to process information (little effort to do so), which increases the probability of choosing the best alternative. When the expert is infallible ($\beta = \infty$), then the expert can judge all comparisons without error. In this theoretical framework, the expert can process as much information as he/she considers appropriate, reducing the error; however, he/she maintains a difficult balance, making the cost of the required effort less than the benefit of reducing error. In the extreme cases $\beta = \infty$ and $\beta = 0$, it does not make sense to use a DSS, and structuring the expert's judgments is irrelevant because the expected performance of all DSSs will be the same.

The IBR framework has been used in the literature to analyze aspects that are difficult to evaluate in empirical work (omission errors, influence relationships, expertise levels, systems reliability, etc.). Some examples are the analysis of the impact of influence-communication networks on group performance [79–81]; the influence of participation in innovation decisions [82,83]; the value of time in decision-making [49]; the theoretical advantages of blockchain technology [84]; and the way to combine expertise, diversity, and number of members in steering committees [8]. In all these theoretical studies the expert is attracted to what is correct to a greater degree than to what is incorrect as a physical law, a property that allows the development of DSS that improves the expected performance value.

This work makes a significant contribution to the field by proposing a novel coding system that aims to establish a general representation encompassing all the approaches described in Eqs. (2), (3), and (4). In this system, the probability of selecting alternative A_i is defined as follows:

$$p_i = \frac{B_i}{\sum_{j=1}^{n} B_j} \tag{5}$$

where B_i can be replaced by any of the representations:

i) $B_i = v_i$ if we accept the representation established by evolutionary theories.
ii) $B_i = e^{v_i}$ if we accept the representation established by the discrete choice theory or the full-range log-odds theory.
iii) $B_i = e^{\beta v_i}$ if we accept the representation established by the IBR theory.

In general, it can be assumed that $B_i = f(V_i)$ with f a real positive increasing function. This change of variable is an important contribution since it allows us to propose a new symbolic model that maintains the generality of the calculations. The alternatives can be denoted in such a way that the subscript i establishes the ranking of their performance without loss of generality. In this way the alternative A_1 will have the highest latent performance, A_2 the second highest and so on up to A_n which will have the lowest performance. This notation allows to keep in mind that depending on the subscript of B its value will be higher or lower and with this property symbolic manipulations of the codings of the studied DSS can be done. Thus, the results obtained in the symbolic manipulations of Section 4 are valid for any of the three functional forms described in the three groups of theories, the only necessary condition being the fact that the probability of choice of each alternative is proportional to its latent performance as shown in Eq. (5).

3.1. Holistic evaluation of alternatives paradigm

The holistic evaluation of alternatives paradigm performs a comparison of all alternatives simultaneously, obtaining the probability of choosing each alternative. This approach aims to represent how an individual faces a series of alternatives among which he/she can only choose one of them. The expert anticipates a normalized value for each of the alternatives $\widehat{V}_i / \sum_{k=1}^{n} \widehat{V}_k$, while their real normalized performances $V_i / \sum_{k=1}^{n} V_k$ are latent. The difference between judgment and reality, denoted as u_i, is a random error:

$$\frac{\widehat{V}_i}{\sum_{k=1}^{n} \widehat{V}_k} = \frac{V_i}{\sum_{k=1}^{n} V_k} + u_i$$

The error is defined for a specific choice situation, and for the information that has been processed in that situation. Thus, the characteristics of u_i will depend on how the expert processes the information. If the expert were able to anticipate with absolute precision the value of all the alternatives, the error value would be zero for all alternatives.

The probability that the expert chooses the alternative A_i compared to the other alternatives (holistic choice) will be:

$$p(A_i)^{\mathrm{h}} = p_i^{\mathrm{h}} = Prob\left(\frac{\widehat{V}_i}{\sum_{k=1}^{n}\widehat{V}_k} > \frac{\widehat{V}_j}{\sum_{k=1}^{n}\widehat{V}_k}\middle|\forall i \neq j\right)$$
$$= Prob\left(\frac{V_i}{\sum_{k=1}^{n}V_k} + u_i > \frac{V_j}{\sum_{k=1}^{n}V_k} + u_j\middle|\forall i \neq j\right)$$
$$= Prob\left(u_j - u_i < \frac{V_i - V_j}{\sum_{k=1}^{n}V_k}\middle|\forall i \neq j\right) \tag{6}$$

Therefore, if the difference between the errors is lower than the difference between the true normalized performances of the alternatives, the best alternative will be chosen. Thus, if we denote the random error by $w_i = (w_{i1}, w_{i2}, \ldots, w_{in})$, where $w_{ij} = u_j - u_i$, and its joint density function is $h(w)$, then the probability that the expert chooses the alternative A_i holistically becomes:

$$p_i^{\mathrm{h}} = \int_w I\left[w_{ij} < \frac{V_i - V_j}{\sum_{k=1}^{n}V_k}\middle|\forall i \neq j\right]h(\mathrm{w})d\mathrm{w} \tag{7}$$

Depending on the theoretical framework chosen, Eq. (7) will adopt the functional form of natural selection models (Eq. (2)), discrete choice models, full-range log-odds models (Eq. (3)), or IBR models (Eq. (4) with individual expertise represented by parameter β).

3.2. Pairwise comparison of alternatives paradigm

Keeney and Raiffa [85] argue that complex problems can be solved by breaking them down into simpler problems, while Ravinder et al. [38] and Ravinder [37] show statistically that this approach reduces error variance. In a decision problem with a finite set of alternatives, one of the most used simplification approaches to deal with the difficulty of choosing the best alternative is to focus on pairs of alternatives. The use of pairwise comparisons of alternatives has its origins in psychological studies [34,36] which claim that making judgments about two alternatives at a time is easier and more accurate than making simultaneous judgments about all the alternatives. In addition, pairwise comparisons of alternatives allow the consistency (coherence) of judgments to be cross-checked.

With pairwise comparisons of alternatives Eq. (6) becomes:

$$p_{ij} = Prob\left(u_j - u_i < \frac{V_i - V_j}{\sum_{k=1}^{n}V_k}\right) \tag{8}$$

Judgment probability distributions proportional to performance do not establish purely random errors but rather assume the existence of a relationship between decision difficulty and error. Thus, in the IBR framework, when applied to a pair of alternatives being compared, (A_i, A_j), Eq. (4) provides the probability of choosing the first alternative A_i when compared to the second one A_j:

$$p_{ij} = \frac{\mathrm{e}^{\beta v_i}}{\mathrm{e}^{\beta v_i} + \mathrm{e}^{\beta v_j}} = \frac{1}{1 + \mathrm{e}^{\beta(v_j - v_i)}} = \left(1 + \mathrm{e}^{\beta\frac{V_j - V_i}{V_i + V_j}}\right)^{-1} \tag{9}$$

Thus, experts do not need to estimate specific values of the normalized performances of the alternatives, $\widehat{V}_1/\sum_{k=1}^{n}\widehat{V}_k$, $\widehat{V}_2/\sum_{k=1}^{n}\widehat{V}_k$, …, $\widehat{V}_n/\sum_{k=1}^{n}\widehat{V}_k$; it is enough to estimate the difference in performance between two alternatives through a judgment relationship.

One of the simplest DSS based on pairwise comparison is Ordering [42]. Since it is assumed that the performances of alternatives are all different, experts' judgments can only express preference (not indifference). A variable is established for each pairwise judgment ($\widehat{r}_{ij}$): when the expert judges the performance of the alternative A_i greater than that of the alternative A_j (A_i preferred to A_j), it is $\widehat{r}_{ij} = 1 \wedge \widehat{r}_{ji} = 0$; while when the expert judges A_j preferred to A_i, it is $\widehat{r}_{ij} = 0 \wedge \widehat{r}_{ji} = 1$. The final assessment of the alternative A_i will be the sum of its favorable pairwise judgments $\widehat{r}_i = \sum_{j=1;j\neq i}^{n}\widehat{r}_{ij}$. The chosen alternative is the one with the greatest number of favorable pairwise comparisons: $\max_i \widehat{r}_i$.

The IBR Method (IBRM) proposed in [7] details the probabilistic convolution process of choosing each alternative. In the pairwise comparison of a set of three alternatives $\{A_i, A_j, A_k\}$, the probability that alternative A_i is chosen as best is the probability of choosing A_i when compared to A_j multiplied by the probability of choosing A_i when compared to A_k : $p_{ij} \cdot p_{ik}$. If two or more alternatives receive the same number of favorable judgments, then their probability is distributed equally between them.

Notice that IBR Eq. (9) is based on the exponential function and therefore it is a positive and increasing function with respect to the difference $V_i - V_j$. Since (i) $\beta\frac{V_j - V_i}{V_i + V_j} \to -\infty \Rightarrow p_{ij} \to 1$, (ii) $\beta\frac{V_j - V_i}{V_i + V_j} = 0 \Rightarrow p_{ij} = 0.5$, and (iii) $\beta\frac{V_j - V_i}{V_i + V_j} \to +\infty \Rightarrow p_{ij} \to 0$, it is obvious that the equalities and inequalities of the natural selection models of Eq. (2) are respected when the performance values of the alternatives are positive, and in the IBR case too: (i) $\beta\frac{V_j - V_i}{V_i + V_j} \to -\infty$ iff $V_j - V_i < 0 \wedge \beta \to -\infty$; (ii) $\beta\frac{V_j - V_i}{V_i + V_j} = 0$ iff $V_j - V_i = 0 \vee \beta = 0$; (iii) $\beta\frac{V_j - V_i}{V_i + V_j} \to +\infty$ iff $V_j - V_i > 0 \wedge \beta \to +\infty$.

Saaty's AHP [86] is also a very popular DSS based on the pairwise comparison of the alternatives paradigm that has led other researchers to propose decision approaches based on this methodology on areas such as supplier selection [87,88], production capability [89], and evaluating Industry 4.0 technology [90].

4. Holistic DSS vs. pairwise comparison DSS

DSSs based on experts' judgments enhance the efficacy of human decision-making processes and imbue the processing of information with scientific rigor [1,2]. A DSS comprises two constituent elements: (i) the enhancement of deficiencies in human judgment and (ii) the mathematical evaluation of these processes [91].

Some DSSs have been developed with consideration of the characteristics of individual experts, including their level of experience [92] and reputation [93], as well as their judgments, such as the presence of hesitation [94,95] and consistency [39,96]. The mechanisms developed for a DSS aim to maximize its expected performance value.

The diversity and complexity of DSSs make their evaluation and comparison a challenging task. Several approaches have been used to analyze and compare the efficiency of different DSSs: provision of decision problems where different DSSs lead to different outcomes [20]; simulation of scenarios to analyze the divergences of the different DSSs [24]; sensitivity analysis of different DSSs to specific decision making parameters [97]; effect of the method used to represent the scale of verbal priorities expressed by experts in the different DSSs [98]; the analysis of distribution of group error versus the distribution of expert error for different DSSs [99]; the representation of alternatives into attributes to reduce error within each DSSs [37]; the resilience of the algorithms of the different DSSs [100]; the probabilistic convolution of choosing each alternative by the different DSSs [101]. The approach chosen in the present article differs in that the superiority of the pairwise comparison of alternatives paradigm over the holistic evaluation of alternatives paradigm is demonstrated theoretically since it only requires that the probability distribution of the judgments is proportional to the performances of the alternatives.

The existent holistic versus decomposition approach [102] is different from the approach proposed in this paper, which is the reason why the term paradigm is used instead. Thus, herein, a holistic paradigm means that the best alternative of a set of feasible alternatives requires a single judgment from an expert, while a pairwise paradigm means that the best alternative of a set of feasible alternatives requires $n(n-1)/2$ judgments from an expert. Differentiating between these two paradigms is particularly useful since most complex DSSs use one of these forms as the basis for their structuring. For example, Condorcet's Jury Theorem [103] enhances the individual holistic decision as described above with a group-decision majority rule. The ranking system proposed by Borda

[104] can be considered a holistic sequential system. DSSs based on the decomposition principle [102] include a holistic (SMARTER, [105]) or a pairwise comparison process (AHP, [106]) to assign criteria importance values. Thus, in this paper, the pairwise comparison is considered a basic decomposition and is justified as a way to facilitate the cognitive process of eliciting preferences that simplifies the decision process [107].

A practical example might be asking 100 expert analysts to evaluate which of three mutual funds will have the highest value at a future date. One can ask if the evaluation be done holistically, simply by giving their judgment on which will be the best, or by requiring the experts to perform a pairwise comparison of the three mutual funds. This same example could be done with tennis players or with other domains. Our work asks which of the two paradigms (holistic vs. pairwise) performs better if the experts have the same probability distribution when making pairwise comparisons as when making holistic comparisons. While an example may be illustrative, it is important to maintain the distinction between the theoretical validity of the model and its empirical application. The main value of the model lies in identifying fundamental regularities in the decision process, specifically whether there is a systematic superiority of the pairwise comparison paradigm over the holistic one or vice versa (see the statement of Lemma 2).

Without loss of generality, it is assumed three alternatives $\{A_1, A_2, A_3\}$ with ordered positive performance values: $V_1 > V_2 > V_3 > 0$. The probability convolution of choosing the alternative A_1 through pairwise comparison with "ordering", p_1^p, consists of two addends: (1) the probability of being preferred in all comparisons: $A_1 > A_2; A_1 > A_3$; (2) the product of multiplying the probabilities of ties resulting from the non-transitive circular arrangements, $A_1 > A_2; A_2 > A_3; A_3 > A_1$ and $A_2 > A_1; A_1 > A_3;\ A_3 > A_2$ (when the expert shows a tie between the three alternatives), and the probability that the alternative is chosen randomly among the alternatives tied (1/3). Thus, we derive the following pairwise comparison probabilities of alternatives:

$$p_1^p = p_{12}{\cdot}p_{13} + \frac{1}{3}p_{12}{\cdot}p_{23}{\cdot}p_{31} + \frac{1}{3}p_{21}{\cdot}p_{32}{\cdot}p_{13} \tag{10}$$

$$p_2^p = p_{21}{\cdot}p_{23} + \frac{1}{3}p_{12}{\cdot}p_{23}{\cdot}p_{31} + \frac{1}{3}p_{21}{\cdot}p_{32}{\cdot}p_{13} \tag{11}$$

$$p_3^p = p_{32}{\cdot}p_{31} + \frac{1}{3}p_{12}{\cdot}p_{23}{\cdot}p_{31} + \frac{1}{3}p_{21}{\cdot}p_{32}{\cdot}p_{13} \tag{12}$$

The subsequent result demonstrates that if the probabilities assigned by experts are proportional to a real positive increasing function of the performance values of the alternatives, as per Eq. (5), then the probabilities associated with the two circular tie orderings previously described are equal.

Lemma 1. If the probability distribution of the judgements of an expert verifies Eq. (5), then the probabilities of the two circular tie orderings are equal:

$$p_{12}{\cdot}p_{23}{\cdot}p_{31} = p_{21}{\cdot}p_{32}{\cdot}p_{13}.$$

Proof. Applying Eq. (5) yields

$$p_{12} = \frac{B_1}{B_1+B_2};\ p_{13} = \frac{B_1}{B_1+B_3};p_{23} = \frac{B_2}{B_2+B_3};\ p_{21} = \frac{B_2}{B_1+B_2};\ p_{31} = \frac{B_3}{B_1+B_3};\ p_{32} = \frac{B_3}{B_2+B_3}.$$

Therefore, it is:

$$p_{12}{\cdot}p_{23}{\cdot}p_{31} = \frac{B_1}{B_1+B_2}\frac{B_2}{B_2+B_3}\frac{B_3}{B_1+B_3} = \frac{B_2}{B_1+B_2}\frac{B_3}{B_2+B_3}\frac{B_1}{B_1+B_3} = p_{21}{\cdot}p_{32}{\cdot}p_{13}\ \blacksquare$$

Thus, when the probability distribution of the judgements of an expert is described by Eqs. (5), (10), (11), and (12) are obtained as the sum of two terms:

$$p_1^p = p_{12}{\cdot}p_{13} + \frac{2}{3}p_{12}{\cdot}p_{23}{\cdot}p_{31}$$

$$p_2^p = p_{21}{\cdot}p_{23} + \frac{2}{3}p_{12}{\cdot}p_{23}{\cdot}p_{31}$$

$$p_3^p = p_{32}{\cdot}p_{31} + \frac{2}{3}p_{12}{\cdot}p_{23}{\cdot}p_{31}$$

The probability of choosing an alternative with the holistic evaluation with probability distribution of the judgements of an expert described by Eq. (5) are

$$p_i^h = \frac{B_i}{B_1+B_2+B_3};\ i \in \{1,2,3\}$$

The following results state that if the probabilities of the judgments of an expert are proportional to a real positive increase function of the performance values of the alternatives, as per Eq. (5), then for the best performing alternative the holistic probability value is lower than the pairwise comparison probability value while for the worst performing alternative, the holistic probability value is higher than the pairwise comparison probability value.

Lemma 2. If the probability distribution of the judgements of an expert verifies Eq. (5), then

$$p_1^h < p_1^p \wedge p_3^h > p_3^p.$$

Proof. Notice that

$$p_1^p - p_1^h = \frac{B_1}{B_1+B_2}\frac{B_1}{B_1+B_3} + \frac{2}{3}\frac{B_1}{B_1+B_2}\frac{B_2}{B_2+B_3}\frac{B_3}{B_1+B_3} - \frac{B_1}{B_1+B_2+B_3}$$

and $B_1 > B_2 > B_3 > 0$ because $V_1 > V_2 > V_3 > 0$. Algebraic manipulation of the right-hand side yields:

$$p_1^p - p_1^h = \frac{B_1B_2B_3(2B_1-B_2-B_3)}{(B_1+B_2)(B_1+B_3)(B_2+B_3)(B_1+B_2+B_3)}$$

Since $2B_1 - B_2 - B_3 > 0$, we conclude that $p_1^p - p_1^h > 0$ and the inequality $p_1^h < p_1^p$ is true.

Similarly,

$$p_3^p - p_3^h = \frac{-B_1B_2B_3(B_1+B_2-2B_3)}{(B_1+B_2)(B_1+B_3)(B_2+B_3)(B_1+B_2+B_3)}$$

Since $B_1 + B_2 - 2B_3 > 0$, we conclude that $p_3^p - p_3^h < 0$ and the inequality $p_3^p < p_3^h$ is true.$\blacksquare$

The following results prove that for any two DSSs that verify the above property, the DSS having a higher probability of choosing the best alternative and a lower probability of choosing the worst alternative will be more efficient in the sense of having a higher expected performance value.

Lemma 3. Let S_1 and S_2 be two different DSSs and $\{V_1, V_2, V_3\}$ the latent performance values of three alternatives $\{A_1, A_2, A_3\}$ such that $V_1 > V_2 > V_3 > 0$. If the probability distribution of the judgements of an expert verifies Eq. (5), then:

$$p_1^{S_1} < p_1^{S_2} \wedge p_3^{S_1} > p_3^{S_2} \Rightarrow V_1p_1^{S_1} + V_2p_2^{S_1} + V_3p_3^{S_1} < V_1p_1^{S_2} + V_2p_2^{S_2} + V_3p_3^{S_2}$$

Proof. On the one hand, since $p_1^{S_1} < p_1^{S_2} \wedge p_3^{S_1} > p_3^{S_2}$, there exist $0 < x \le 1 - p_1^{S_1}$ and $0 < y \le p_3^{S_1}$ such that

$$p_1^{S_2} = p_1^{S_1} + x \wedge\ p_3^{S_2} = p_3^{S_1} - y.$$

On the other hand, it is

$$p_2^{S_2} = 1 - p_1^{S_2} - p_3^{S_2} = 1 - p_1^{S_1} - x - p_3^{S_1} + y = \left(1 - p_1^{S_1} - p_3^{S_1}\right) + y - x = p_2^{S_1} + y - x.$$

Thus, it is

$$V_1 p_1^{S_1} + V_2 p_2^{S_1} + V_3 p_3^{S_1} = V_1 p_1^{S_2} + V_2 p_2^{S_2} + V_3 p_3^{S_2} + x(V_2 - V_1) + y(V_3 - V_2).$$

The conditions $V_1 > V_2 > V_3$, and $x,\ y > 0$, imply that $x(V_2 - V_1) + y(V_3 - V_2) < 0$, which is equivalent to

$$V_1 p_1^{S_1} + V_2 p_2^{S_1} + V_3 p_3^{S_1} < V_1 p_1^{S_2} + V_2 p_2^{S_2} + V_3 p_3^{S_2}.$$

This completes the proof.■

Notice that the joint application of the above results leads to the following conclusion: The pairwise comparison paradigm returns a higher expected value than the holistic paradigm for the case of three alternatives. The following corollary of the above results is derived.

Corollary. Under the assumptions of Lemma 1, the expected performance value of a pairwise comparison DSS is higher than the expected performance value of the holistic DSS no matter the probability values of choosing the middle alternative with any of the two DSSs.

Notice that

$$p_2^p - p_2^h = \frac{-B_1B_2B_3(B_1 - 2B_2 + B_3)}{(B_1 + B_2)(B_1 + B_3)(B_2 + B_3)(B_1 + B_2 + B_3)}.$$

Since $B_1 - 2B_2 + B_3$ can take on any sign, we have that the following relation between the holistic probability value and the pairwise comparison probability value of choosing the middle alternative:

(i) if $B_2 = \frac{B_3+B_1}{2}$, it is $p_2^p = p_2^h$;
(ii) if $B_2 > \frac{B_3+B_1}{2}$, it is $p_2^p > p_2^h$;
(iii) if $B_2 < \frac{B_3+B_1}{2}$, it is $p_2^p < p_2^h$.

Under the assumptions of Lemma 1; Lemma 2 is applicable, and we conclude that

$$p_1^p = p_1^h + \frac{-B_1B_2B_3(B_2 + B_3 - 2B_1)}{(B_1 + B_2)(B_1 + B_3)(B_2 + B_3)(B_1 + B_2 + B_3)};$$

$$p_2^p = p_2^h + \frac{-B_1B_2B_3(B_3 + B_1 - 2B_2)}{(B_1 + B_2)(B_1 + B_3)(B_2 + B_3)(B_1 + B_2 + B_3)}$$

$$p_3^p = p_3^h + \frac{-B_1B_2B_3(B_2 + B_1 - 2B_3)}{(B_1 + B_2)(B_1 + B_3)(B_2 + B_3)(B_1 + B_2 + B_3)}.$$

From Lemma 3 above, this would give the following expected performance gain of the pairwise paradigm versus the holistic parading:

$$V_1\frac{-B_1B_2B_3(B_2 + B_3 - 2B_1)}{(B_1 + B_2)(B_1 + B_3)(B_2 + B_3)(B_1 + B_2 + B_3)} + V_2\frac{-B_1B_2B_3(B_3 + B_1 - 2B_2)}{(B_1 + B_2)(B_1 + B_3)(B_2 + B_3)(B_1 + B_2 + B_3)} + V_3\frac{-B_1B_2B_3(B_2 + B_1 - 2B_3)}{(B_1 + B_2)(B_1 + B_3)(B_2 + B_3)(B_1 + B_2 + B_3)}$$

5. Discussion

The novelty of the comparison approach proposed in this paper is that it differs from existing approaches in the literature. The proposed modeling in this paper studies initially the judgments required by the DSS (coding step). Rather than establishing scenarios, all *combinations of possible judgments* are obtained (space of events) to symbolically determine the probability of each alternative generated by the DSS.

This modeling approach enables the determination of the superiority (if it exists) of one DSS over another and the quantification of the difference in any scenario (distribution of latent performances) by simply imputing specific values into the mathematical expressions obtained. The results obtained are important because they are general in nature:

(i) They are not the product of a simulation for a set of specific values given for the performance values of the alternatives.
(ii) They do not depend on the specific probability distribution of random errors.
(iii) They require only that the probability distribution of the judgements of an expert about the alternatives are proportional to a real positive increasing function of the performance values of the alternatives, an assumption that is widely supported by empirical data, as demonstrated by the numerous references cited in Section 3 of [76].
(iv) An a priori analysis (before the experts make their judgments) is provided that is useful for managers, because it allows them to choose a DSS according to their problem to solve, and for researchers because it allows them to demonstrate the gains of modifications in the DSS of their interest.
(v) A general analysis of the contributions to accuracy of the different parts of the DSS can be carried out.

A promising research direction would be to find out how the distribution of latent performances (mean and variance) increases the expected performance of the pairwise comparison paradigm. In the case of the IBR methodology, the additional parameter representing the expertise level of the expert (β) may also affect the performance of the DSS. Moreover, as per the proposed modeling in this paper, increasing the level of expertise of the expert will increase the difference in expected performance between the holistic paradigm and the pairwise paradigm because the probability of choosing the best alternative increases exponentially with respect to β. This result is relevant for managers because it allows them to analyze the price that they would be willing to pay for each expert based on their level of expertise.

Inconsistency management is one of the most studied aspects of a DSS [108,109]. The proposed modeling allows a general evaluation of the contribution of ordinal inconsistency to the expected performance of a DSS. The pairwise probability value consists of two addends, the first one representing the probability that alternative A_1 is chosen $\left(\frac{B_1}{B_1+B_2}\frac{B_1}{B_1+B_3}\right)$, while the second addend $\left(\frac{2}{3}\frac{B_1}{B_1+B_2}\frac{B_2}{B_2+B_3}\frac{B_3}{B_1+B_3}\right)$ represents a random choice of alternative when the expert shows inconsistent judgments (circular preference, ordinal inconsistency). The establishment of alternative systems to the random distribution of this probability can be quantified in the proposed modeling.

We conjecture that the regularities of the proportional probabilities found for three alternatives are valid for the general case $n > 3$. This conjecture is based on a total of more than a million simulations carried out for $n = 4,\ 5,\ 6$ and 7 with different normalized performance distributions, which suggest that the superiority of the pairwise paradigm over the holistic paradigm. Summarizing, the modeling proposed in this paper contributes to the systematization and generalization of DSS analysis, which will have important implications for its structuring.

6. Conclusions

The evaluation of DSSs quantifies their contribution based on the expected performance value of their outcomes. The probabilistic convolutions that DSS assigns to each alternative is an unexplored line of research worth developing that can provide important advances in all algorithms that use as key information the judgments of experts. This paper proposes a modeling that overcomes the traditional barriers of DSS analysis, which is limited to a given set of judgments (a particular case) or a simulation of scenarios with different performances of the alternatives. Our proposal shows how to model DSS algorithms to obtain a priori information about their expected performance, with the only necessary condition imposed being that the probability of choosing each alternative is proportional to its latent performance, a regularity of human judgments that has been corroborated in numerous works and modeling.

This a priori vision is significantly important to managers because it enables them to ascertain the optimal DSS for addressing a specific problem at hand. For scientists, it facilitates a more profound examination of the potential of the proposed DSS models and the role of their constituent components.

To illustrate the proposed modeling, two fundamental paradigms frequently utilized in the literature have been contrasted: the holistic evaluation of alternatives and the pairwise comparison of alternatives. The latter has been implemented in well-known DSSs: AHP, ELECTRE, PROMETHEE, while the former has been applied in simpler DSSs such as the Borda rule [104] or the SMARTER [105]. In the existing literature, there are arguments for requesting that experts provide holistic judgments: (i) they allow to establish a direct link between "preferences and the recommendation delivered;" (ii) they are generally carried out in a "single space" while pairwise comparisons are made in "different spaces"; (iii) experts feel more comfortable making holistic judgments about a set of alternatives [110]; (iv) increasing the number of judgments can generate overload, deteriorating their quality as evidenced in [111–114]. This paper addresses the mentioned doubts about these two paradigms because it demonstrates that when the empirical regularities found in human judgment are met, the pairwise comparison paradigm results in superior expected performance values than the holistic paradigm for any performance distribution of three alternatives and for any level of experience of the individuals involved.

The comparison of the two paradigms with three alternatives has allowed to show the philosophy of our modeling with manual calculations. Although the model has been illustrated for three alternatives, our modeling allows it to be evaluated symbolically for any number of alternatives n without any limitation other than the computational capacity of the mathematical software used. We conjecture, based on our simulations for $n = 4, 5, 6$ and 7, that the superiority of the pairwise model over the holistic one holds for any n.

The proposed modeling methodology enables the examination of how the distribution of latent performances (mean and variance) influences the maximum performance of each model, thereby determining the sensitivity of consistency or the sensitivity to the expertise of the individuals. Our approach allows for the determination (quantification) of the superiority (if it exists) of one DSS over another, thereby contributing to the systematization and generalization of DSS analysis.

In general, the majority of DSSs are based on one of two paradigms. However, there are a few that propose mixed methods approach, such as the Best-Worst Multi-Criteria (BWM) [45]. The BWM model introduces a second stage wherein the expert must identify the best and the worst criteria, in a holistic way, and then perform the pairwise comparison with respect to these two. The BWM model justifies the use of a holistic evaluation approach because the gains in expected performance value do not outweigh the resource needs by using only a pairwise comparison approach. Indeed, the result of a pairwise comparison based DSS requires a significant number of judgments and other calculations [42], which entails a greater cost than a corresponding holistic-based DSS. Thus, a cost-benefit evaluation of DSS represents a promising avenue for future research, provided that the probabilistic convolutions of the evaluated systems are obtained.

Funding

This work was supported by the Spanish Ministry of Economy and Competitiveness, [ECO2017-86305-C4-3-R], the Diputación General de Aragón (DGA) and the European Social Fund [CREVALOR], the CUD (UZCUD2017-SOC -04), by the Spanish State Research Agency under Project PID2020-113338RB-I00, by C054/2024_1 Catedras-Universidad de Zaragoza, and by the Fundación Seminario de Investigación por la Paz de Zaragoza (SIP Foundation) under Project[2024].

CRediT authorship contribution statement

Carlos Sáenz-Royo: Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Software, Resources, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization. **Francisco Chiclana:** Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Software, Resources, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Carlos Saenz Royo reports financial support, administrative support, article publishing charges, equipment, drugs, or supplies, travel, and writing assistance were provided by Spanish Ministerio de Economía y Competitividad. Carlos Saenz Royo reports financial support, administrative support, article publishing charges, equipment, drugs, or supplies, travel, and writing assistance were provided by Diputación General de Aragón (DGA) and the European Social Fund. Carlos Saenz Royo reports financial support, administrative support, article publishing charges, equipment, drugs, or supplies, travel, and writing assistance were provided by Spanish State Research Agency. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

To Professor Antonio Elipe for the enriching discussions held, and for his selfless help.

Data availability

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

References

[1] R.L. Keeney, Value-Focused Thinking: A Path to Creative Decisionmaking, Harvard Univ. Press, Cambridge, Mass, 1992.
[2] J.M. Moreno-Jiménez, J. Aguarón-Joven, M.T. Escobar-Urmeneta, A. Turón-Lanuza, Multicriteria procedural rationality on SISDEMA, Eur. J. Oper. Res. 119 (1999) 388–403, https://doi.org/10.1016/S0377-2217(99)00141-1.
[3] E. Turban, Decision Support and Business Intelligence Systems, Pearson Education India, 2011.
[4] M. Cinelli, The art of supporting decision-making, Exchanges: The Interdiscip. Res. J. 4 (2017) 298–308.
[5] H.A. Simon, The New Science of Management Decision, Harper & Brothers, New York, 1960, https://doi.org/10.1037/13978-000.
[6] F. Liu, M.-Y. Qiu, W.-G. Zhang, An uncertainty-induced axiomatic foundation of the analytic hierarchy process and its implication, Expert. Syst. Appl. 183 (2021) 115427, https://doi.org/10.1016/j.eswa.2021.115427.
[7] C. Sáenz-Royo, F. Chiclana, E. Herrera-Viedma, Intentional bounded rationality methodology to assess the quality of decision-making approaches with latent alternative performances, Inf. Fusion 89 (2023) 254–266, https://doi.org/10.1016/j.inffus.2022.08.019.
[8] C. Sáenz-Royo, F. Chiclana, E. Herrera-Viedma, Steering committee management. Expertise, diversity, and decision-making structures, Inf. Fusion (2023) 101888.
[9] S.A. Zarghami, O. Zwikael, Measuring project resilience – learning from the past to enhance decision making in the face of disruption, Decis. Support. Syst. 160 (2022) 113831, https://doi.org/10.1016/j.dss.2022.113831.
[10] A. Avetyan, G. Martirosyan, S. Baghdasaryan, General principles and potential organizational gains of customer relationship management system implementation for municipal government structures in Armenia, Acciones e Investigaciones Sociales (2025). https://doi.org/10.26754/ojs_ais/accionesinvestigsoc.20254610444.
[11] T. Comes, M. Hiete, N. Wijngaards, F. Schultmann, Decision maps: a framework for multi-criteria decision support under severe uncertainty, Decis. Support. Syst. 52 (2011) 108–118.
[12] I.J. Pérez, F.J. Cabrerizo, S. Alonso, Y.C. Dong, F. Chiclana, E. Herrera-Viedma, On dynamic consensus processes in group decision making problems, Inf. Sci. 459 (2018) 20–35, https://doi.org/10.1016/j.ins.2018.05.017.

[13] Y. Ji, Y. Li, C. Wijekoon, Robust two-stage minimum asymmetric cost consensus models under uncertainty circumstances, Inf. Sci. 663 (2024) 120279.
[14] S. Qu, Y. Zhou, Y. Ji, Z. Dai, Z. Wang, Robust maximum expert consensus modeling with dynamic feedback mechanism under uncertain environments, J. Ind. Manag. Optim. (2024), 0–0.
[15] H. Zhang, I. Palomares, Y. Dong, W. Wang, Managing non-cooperative behaviors in consensus-based multiple attribute group decision making: an approach based on social network analysis, Knowl. Based. Syst. 162 (2018) 29–45, https://doi.org/10.1016/j.knosys.2018.06.008.
[16] E.S. Berner, T.J. La Lande, Overview of clinical decision support systems. Clinical Decision Support Systems: Theory and Practice, 2016, pp. 1–17.
[17] S. Pant, A. Kumar, M. Ram, Y. Klochkov, H.K. Sharma, Consistency indices in analytic hierarchy process: a review, Mathematics 10 (2022) 1206.
[18] P.Z. Lappas, A.N. Yannacopoulos, A machine learning approach combining expert knowledge with genetic algorithms in feature selection for credit risk assessment, Appl. Soft. Comput. 107 (2021) 107391.
[19] A. Saltelli, S. Tarantola, F. Campolongo, M. Ratto, Sensitivity Analysis in Practice: A Guide to Assessing Scientific Models, Wiley Online Library, 2004.
[20] S. Opricovic, G.-H. Tzeng, Extended VIKOR method in comparison with outranking methods, Eur. J. Oper. Res. 178 (2007) 514–529, https://doi.org/10.1016/j.ejor.2006.01.020.
[21] J. Wallenius, J.S. Dyer, P.C. Fishburn, R.E. Steuer, S. Zionts, K. Deb, Multiple criteria decision making, multiattribute utility theory: recent accomplishments and what lies ahead, Manag. Sci. 54 (2008) 1336–1349, https://doi.org/10.1287/mnsc.1070.0838.
[22] E. Triantaphyllou, Multi-Criteria Decision Making Methods: A Comparative Study, Springer US, Boston, MA, 2000. http://www.springerlink.com/content/978-1-4757-3157-6 (accessed July 5, 2020).
[23] V. Belton, A comparison of the analytic hierarchy process and a simple multi-attribute value function, Eur. J. Oper. Res. 26 (1986) 7–21, https://doi.org/10.1016/0377-2217(86)90155-4.
[24] S.H. Zanakis, A. Solomon, N. Wishart, S. Dublish, Multi-attribute decision making: a simulation comparison of select methods, Eur. J. Oper. Res. 107 (1998) 507–529, https://doi.org/10.1016/S0377-2217(97)00147-1.
[25] H.A. Simon, Administrative Behavior, Free Press, New York, 1947.
[26] H.A. Simon, The architecture of complexity, in: G.J. Klir (Ed.), Facets of Systems Science, Springer US, Boston, MA, 1991, pp. 457–476, https://doi.org/10.1007/978-1-4899-0718-9_31.
[27] R.L. Keeney, R.S. Gregory, Selecting attributes to measure the achievement of objectives, Oper. Res. 53 (2005) 1–11.
[28] M. Cinelli, M. Kadziński, G. Miebs, M. Gonzalez, R. Słowiński, Recommending multiple criteria decision analysis methods with a new taxonomy-based decision support system, Eur. J. Oper. Res. 302 (2022) 633–651.
[29] R. Gutiérrez, V. Rampérez, H. Paggi, J.A. Lara, J. Soriano, On the use of information fusion techniques to improve information quality: taxonomy, opportunities and challenges, Inf. Fusion 78 (2022) 102–137.
[30] B. Roy, Decision-aid and decision-making, Eur. J. Oper. Res. 45 (1990) 324–331.
[31] B. Roy, R. Słowiński, Questions guiding the choice of a multicriteria decision aiding method, EURO J. Decis. Process. 1 (2013) 69–97.
[32] O.F. Morera, D.V. Budescu, Random error reduction in analytic hierarchies: a comparison of holistic and decompositional decision strategies, J. Behav. Decis. Making 14 (2001) 223–242, https://doi.org/10.1002/bdm.374.
[33] C. Shang, R. Zhang, X. Zhu, An adaptive consensus model in large-scale group decision making with noncooperative and compromising behaviors, Appl. Soft. Comput. 149 (2023) 110944.
[34] L.L. Thurstone, A law of comparative judgment, Psychol. Rev. 34 (1927) 273–286, https://doi.org/10.1037/h0070288.
[35] W.W. Wittmann, H.-M. Süß, Investigating the paths between working memory, intelligence, knowledge, and complex problem-solving performances via Brunswik symmetry., (1999).
[36] M. Yokoyama, The nature of the affective judgment in the method of paired comparisons, Am. J. Psychol. 32 (1921) 357, https://doi.org/10.2307/1414000.
[37] H.V. Ravinder, Random error in holistic evaluations and additive decompositions of multiattribute utility — an empirical comparison, J. Behav. Decis. Making 5 (1992) 155–167, https://doi.org/10.1002/bdm.3960050302.
[38] H.V. Ravinder, D.N. Kleinmuntz, J.S. Dyer, The Reliability of subjective probabilities obtained through decomposition, Manag. Sci. 34 (1988) 186–199, https://doi.org/10.1287/mnsc.34.2.186.
[39] T.L. Saaty, The Analytic Hierarchy Process: Planning, Priority Setting, Resource Allocation, McGraw-Hill International Book Co, New York ; London, 1980.
[40] J.R. Figueira, S. Greco, B. Roy, R. Słowiński, ELECTRE methods: main features and recent developments. Handbook of Multicriteria Analysis, 2010, pp. 51–89.
[41] J.-P. Brans, P. Vincke, B. Mareschal, How to select and how to rank projects: the PROMETHEE method, Eur. J. Oper. Res. 24 (1986) 228–238.
[42] C. Sáenz-Royo, F. Chiclana, E. Herrera-Viedma, Ordering vs. AHP. Does the intensity used in the decision support techniques compensate? Expert. Syst. Appl. 238 (2024) 121922 https://doi.org/10.1016/j.eswa.2023.121922.
[43] S. Highhouse, Stubborn reliance on intuition and subjectivity in employee selection, Ind. Organ. Psychol. 1 (2008) 333–342.
[44] G. Gigerenzer, H. Brighton, Homo heuristicus: why biased minds make better inferences, Top. Cogn. Sci. 1 (2009) 107–143.
[45] J. Rezaei, Best-worst multi-criteria decision-making method, Omega (Westport) 53 (2015) 49–57, https://doi.org/10.1016/j.omega.2014.11.009.
[46] I.N. Durbach, T.J. Stewart, Using expected values to simplify decision making under uncertainty, Omega (Westport) 37 (2009) 312–330.
[47] Y.-M. Wang, T.M. Elhag, An approach to avoiding rank reversal in AHP, Decis. Support. Syst. 42 (2006) 1474–1480.
[48] O.I. Larichev, Cognitive validity in design of decision-aiding techniques, J. Multi-Criteria Decis. Anal. 1 (1992) 127–138.
[49] C. Sáenz-Royo, V. Salas-Fumás, Á. Lozano-Rojo, Authority and consensus in group decision making with fallible individuals, Decis. Support. Syst. 153 (2022) 113670, https://doi.org/10.1016/j.dss.2021.113670.
[50] T.J. Stewart, F.B. Losa, Towards reconciling outranking and value measurement practice, Eur. J. Oper. Res. 145 (2003) 645–659.
[51] I.N. Durbach, T.J. Stewart, Modeling uncertainty in multi-criteria decision analysis, Eur. J. Oper. Res. 223 (2012) 1–14.
[52] J. Butler, J. Jia, J. Dyer, Simulation techniques for the sensitivity analysis of multi-criteria decision models, Eur. J. Oper. Res. 103 (1997) 531–546.
[53] J.C. Butler, J.S. Dyer, J. Jia, K. Tomak, Enabling e-transactions with multi-attribute preference models, Eur. J. Oper. Res. 186 (2008) 748–765.
[54] J. Davies, G. Katsirelos, N. Narodytska, T. Walsh, L. Xia, Complexity of and algorithms for the manipulation of Borda, Nanson's and Baldwin's voting rules, Artif. Intell. 217 (2014) 20–42.
[55] H.P. Young, Condorcet's theory of voting, Am. Political Sci. Rev. 82 (1988) 1231–1244.
[56] D.G. Saari, Basic Geometry of Voting, Springer Science & Business Media, 2012.
[57] W.V. Gehrlein, D. Lepelley, Voting Paradoxes and Group Coherence: The Condorcet Efficiency of Voting Rules, Springer Science & Business Media, 2010.
[58] H. Nurmi, Comparing Voting Systems, Springer Science & Business Media, 2012.
[59] N. Tideman, Collective Decisions and Voting: The Potential for Public Choice, Routledge, 2017.
[60] M. Risse, Why the count de Borda cannot beat the Marquis de Condorcet, Soc. Choice Welfare 25 (2005) 95–113.
[61] H.A. Simon, A study of decision-making processes in administrative organizations: a study of decision-making processes in administrative organisations. Administrative Behavior, 4th Edition, Simon & Schuster, New York, 1997. Edición: 4 Rev ed.
[62] R.D. Luce, Semiorders and a theory of utility discrimination, Econometrica 24 (1956) 178, https://doi.org/10.2307/1905751.
[63] J.T. Alander, On optimal population size of genetic algorithms, in: CompEuro 1992 Proceedings Computer Systems and Software Engineering, IEEE, 1992, pp. 65–70.
[64] W. Li, X. Zhou, C. Yang, Y. Fan, Z. Wang, Y. Liu, Multi-objective optimization algorithm based on characteristics fusion of dynamic social networks for community discovery, Inf. Fusion 79 (2022) 110–123.
[65] H. Mühlenbein, D. Schlierkamp-Voosen, Predictive models for the breeder genetic algorithm I. Continuous parameter optimization, Evol. Comput. 1 (1993) 25–49.
[66] B. Robson, S. Boray, Studies in the extensively automatic construction of large odds-based inference networks from structured data. Examples from medical, bioinformatics, and health insurance claims data, Comput. Biol. Med. 95 (2018) 147–166.
[67] E. Levrat, E. Thomas, B. Iung, Odds-based decision-making tool for opportunistic production-maintenance synchronization, Int. J. Prod. Res. 46 (2008) 5263–5287.
[68] A. Hussain, Y.S. Muhammad, Trade-off between exploration and exploitation with genetic algorithm using a novel selection operator, Complex. Intell. Systems. 6 (2020) 1–14.
[69] J.G. Cally, D. Stuart-Fox, L. Holman, Meta-analytic evidence that sexual selection improves population fitness, Nat. Commun. 10 (2019) 2017.
[70] K.E. Train, Discrete Choice Methods with Simulation, Cambridge University Press, 2009.
[71] D. McFadden, Economic choices, Am. Econ. Rev. 91 (2001) 351–378.
[72] I. Erev, T.S. Wallsten, D.V. Budescu, Simultaneous over- and underconfidence: the role of error in judgment processes, Psychol. Rev. 101 (1994) 519–527, https://doi.org/10.1037/0033-295X.101.3.519.
[73] I. Erev, A.E. Roth, Predicting how people play games: reinforcement learning in experimental games with unique, mixed strategy equilibria, Am. Econ. Rev. 88 (1998) 848–881.
[74] D.V. Budescu, T.S. Wallsten, Subjective estimation of precise and vague uncertainties., (1987).
[75] D.V. Budescu, I. Erev, T.S. Wallsten, On the importance of random error in the study of probability judgment. Part I: new theoretical developments, J. Behav. Decis. Mak. 10 (1997) 157–171.
[76] C. Sáenz-Royo, F. Chiclana, E. Herrera-Viedma, Functional representation of the intentional bounded rationality of decision-makers: a laboratory to study the decisions a priori, Mathematics 10 (2022) 739, https://doi.org/10.3390/math10050739.
[77] F.A. Csaszar, Organizational structure as a determinant of performance: evidence from mutual funds, Strat. Mgmt. J. 33 (2012) 611–632, https://doi.org/10.1002/smj.1969.
[78] R. Garud, P. Nayyar, Z. Shapira, Technological choices and the inevitability of errors, in: R. Garud, P.R. Nayyar, Z.B. Shapira (Eds.), Technological Innovation, 1st ed., Cambridge University Press, 1997, pp. 20–40, https://doi.org/10.1017/CBO9780511896613.005.
[79] B. Jiang, Y. Zhao, J. Dong, J. Hu, Analysis of the influence of trust in opposing opinions: an inclusiveness-degree based signed Deffuant–Weisbush model, Inf. Fusion 104 (2024) 102173.
[80] V. Salas-Fumás, C. Sáenz-Royo, Á. Lozano-Rojo, Organisational structure and performance of consensus decisions through mutual influences: a computer simulation approach, Decis. Support. Syst. 86 (2016) 61–72, https://doi.org/10.1016/j.dss.2016.03.008.

[81] C. Shang, R. Zhang, X. Zhu, Y. Liu, An adaptive consensus method based on feedback mechanism and social interaction in social network group decision making, Inf. Sci. 625 (2023) 430–456.

[82] C. Sáenz-Royo, Á. Lozano-Rojo, Authoritarianism versus participation in innovation decisions, Technovation 124 (2023) 102741, https://doi.org/10.1016/j.technovation.2023.102741.

[83] C. Shang, R. Zhang, X. Zhu, Comparison and design of organizational decision mechanisms, Decis. Support. Syst. 178 (2024) 114141.

[84] C. Sáenz-Royo, J. Fleta-Asín, Evaluating blockchain as a participatory organisational system: looking for transaction efficiency, Int. Trans. Oper. Res. (2023).

[85] R.L. Keeney, H. Raiffa, Decisions with Multiple Objectives: Preferences and Value Tradeoffs, John Wiley & Sons Inc, New York, 1976.

[86] R.W. Saaty, The analytic hierarchy process—what it is and how it is used, Math. Model. 9 (1987) 161–176, https://doi.org/10.1016/0270-0255(87)90473-8.

[87] B. Feng, X. Hu, I.J. Orji, Multi-tier supply chain sustainability in the pulp and paper industry: a framework and evaluation methodology, Int. J. Prod. Res. 61 (2023) 4657–4683.

[88] A. Majumdar, S.K. Sinha, M. Shaw, K. Mathiyazhagan, Analysing the vulnerability of green clothing supply chains in South and Southeast Asia using fuzzy analytic hierarchy process, Int. J. Prod. Res. 59 (2021) 752–771.

[89] Z. Li, H. Guo, A.V. Barenji, W.M. Wang, Y. Guan, G.Q. Huang, A sustainable production capability evaluation mechanism based on blockchain, LSTM, analytic hierarchy process for supply chain network, Int. J. Prod. Res. 58 (2020) 7399–7419.

[90] C. Bai, H. Zhou, J. Sarkis, Evaluating industry 4.0 technology and sustainable development goals–a social perspective, Int. J. Prod. Res. (2023) 1–21.

[91] C. Sáenz-Royo, F. Chiclana, The value of expert judgments in decision support systems, Appl. Soft. Comput. (2025) 112806.

[92] S. Chinchanachokchai, P. Thontirawong, P. Chinchanachokchai, A tale of two recommender systems: The moderating role of consumer expertise on artificial intelligence based product recommendations, J. Retail. Consum. Serv. 61 (2021) 102528, https://doi.org/10.1016/j.jretconser.2021.102528.

[93] J. Wu, L. Dai, F. Chiclana, H. Fujita, E. Herrera-Viedma, A minimum adjustment cost feedback mechanism based consensus model for group decision making under social network with distributed linguistic trust, Inf. Fusion 41 (2018) 232–242, https://doi.org/10.1016/j.inffus.2017.09.012.

[94] H. Sellak, B. Ouhbi, B. Frikh, B. Ikken, Expertise-based consensus building for MCGDM with hesitant fuzzy linguistic information, Inf. Fusion 50 (2019) 54–70, https://doi.org/10.1016/j.inffus.2018.10.003.

[95] X. You, F. Hou, F. Chiclana, A reputation-based trust evaluation model in group decision-making framework, Inf. Fusion 103 (2024) 102082.

[96] E. Herrera-Viedma, F. Herrera, F. Chiclana, M. Luque, Some issues on consistency of fuzzy preference relations, Eur. J. Oper. Res. 154 (2004) 98–109, https://doi.org/10.1016/S0377-2217(02)00725-7.

[97] E. Triantaphyllou, Multi-criteria decision making methods, in: E. Triantaphyllou (Ed.), Multi-Criteria Decision Making Methods: A Comparative Study, Springer US, Boston, MA, 2000, pp. 5–21, https://doi.org/10.1007/978-1-4757-3157-6_2.

[98] A. Ishizaka, D. Balkenborg, T. Kaplan, Influence of aggregation and measurement scale on ranking a compromise alternative in AHP, J. Oper. Res. Soc. 62 (2011) 700–710, https://doi.org/10.1057/jors.2010.23.

[99] L.G. Vargas, Reciprocal matrices with random coefficients, Math. Model. 3 (1982) 69–81. https://doi.org/10.1016/0270-0255(82)90013-6.

[100] B. Han, C.L. Liu, W.J. Zhang, A method to measure the resilience of algorithm for operation management, IFAC-PapersOnLine 49 (2016) 1442–1447.

[101] M. Cai, Y. Lin, B. Han, C. Liu, W. Zhang, On a Simple and efficient approach to probability distribution function aggregation, IEEE Trans. Syst. Man Cybern.: Syst. 47 (2017) 2444–2453, https://doi.org/10.1109/TSMC.2016.2531647.

[102] A.T. de Almeida, E.A. Frej, L.R.P. Roselli, Combining holistic and decomposition paradigms in preference modeling with the flexibility of FITradeoff, Cent. Eur. J. Oper. Res. 29 (2021) 7–47.

[103] N. De Condorcet, Essai sur l'application de l'analyse à la probabilité des décisions rendues à la pluralité des voix, Cambridge University Press, Cambridge, 1785, https://doi.org/10.1017/CBO9781139923972.

[104] J.-C. de Borda, Mémoire sur les élections au scrutin, Histoire de l'Académie Royale des Sciences 12 (1781).

[105] F.H. Barron, B.E. Barrett, The efficacy of SMARTER—simple multi-attribute rating technique extended to ranking, Acta Psychol. (Amst.) 93 (1996) 23–36.

[106] T.L. Saaty, A scaling method for priorities in hierarchical structures, J. Math. Psychol. 15 (1977) 234–281, https://doi.org/10.1016/0022-2496(77)90033-5.

[107] K. Ciomek, M. Kadziński, T. Tervonen, Heuristics for selecting pair-wise elicitation questions in multiple criteria choice problems, Eur. J. Oper. Res. 262 (2017) 693–707.

[108] J. Tu, Z. Wu, Analytic hierarchy process rank reversals: causes and solutions, Ann. Oper. Res. (2023) 1–25.

[109] Z. Wu, J. Tu, Managing transitivity and consistency of preferences in AHP group decision making based on minimum modifications, Inf. Fusion 67 (2021) 125–135.

[110] W. Waegeman, B. De Baets, L. Boullart, Kernel-based learning methods for preference aggregation, 4OR 7 (2009) 169–189.

[111] V. Banken, Q. Ilmer, I. Seeber, S. Haeussler, A method for smart idea allocation in crowd-based idea selection, Decis. Support. Syst. 124 (2019) 113072.

[112] A. Chernev, U. Böckenholt, J. Goodman, Choice overload: a conceptual review and meta-analysis, J. Consum. Psychol. 25 (2015) 333–358.

[113] A. Sharma, S.K. Nair, Category ratio: a search for an optimal solution to reduce choice overload, J. Consum. Behav. 22 (2023) 880–890, https://doi.org/10.1002/cb.2180.

[114] U. Yuksel, N.T. Thai, M.S.W. Lee, Boycott them! No, boycott this! Do choice overload and small-agent rationalization inhibit the signing of anti-consumption petitions? Psychol. Mark. 37 (2020) 340–354, https://doi.org/10.1002/mar.21293.