Expert Systems With Applications 311 (2026) 131315

Contents lists available at ScienceDirect

Expert Systems With Applications

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

# A Modular Simos-Roy-Figueira framework for tailored weight elicitation in multi-Criteria decision aiding

River Huang [a,*], Miłosz Kadziński [b], José Rui Figueira [c], Salvatore Corrente [d], Eleftherios Siskos [e], Peter Burgherr [a,f]

[a] Laboratory for Energy Systems Analysis, PSI Centers for Nuclear Engineering & Sciences and Energy & Environmental Sciences, Forschungsstrasse 111, Villigen PSI, 5232, Switzerland
[b] Institute of Computing Science, Poznan University of Technology, Piotrowo 2, Poznań, 60–965, Poland
[c] CEGIST, Instituto Superior Técnico, Universidade de Lisboa, , Lisboa, 1049-001, Portugal
[d] Department of Economics and Business, University of Catania, Corso Italia, 55, Catania, 95129, Italy
[e] Decision Support Systems Laboratory, School of Production Engineering and Management, Technical University of Crete, Akrotiri Campus, Chania, 73100, Greece
[f] Department of Mechanical and Process Engineering, Chair of Energy Systems Analysis, ETH Zurich, Clausiusstrasse 33, Zurich, 8092, Switzerland

ARTICLE INFO

*Keywords:*
Multiple criteria decision aiding
SRF Method
Deck-of-Cards method
Consistency restoration
Weight elicitation

ABSTRACT

The Simos-Roy-Figueira (SRF) deck-of-cards procedure is widely recognized as an intuitive method for eliciting criteria weights in outranking Multiple Criteria Decision Aiding (MCDA) methods. However, its numerous extensions, covering imprecise or fuzzy inputs, robustness analysis, hierarchical criteria structures, and diverse communication protocols, have developed independently. Although these variants share the same underlying deck-of-cards mechanism and are, in principle, mutually compatible, the literature offers little guidance on how to combine them coherently or design new SRF configurations for a given decision context. This paper fills that gap by introducing a modular SRF framework that decomposes existing and new methods into a set of interoperable building blocks, each with an explicit linear (or mixed-integer) programming formulation. The main novelty of the framework is threefold: (i) it unifies the main SRF variants as special cases of a common constraint system; (ii) it embeds generic models for consistency checking, inconsistency diagnosis and minimal restoration directly in that system; and (iii) it operationalizes these modules through a guided questionnaire that helps analysts and decision makers select appropriate configurations and derive the corresponding optimization models. The resulting modular SRF framework offers MCDA practitioners, researchers and actual Decision Makers (DMs) a transparent, user-configurable approach that preserves the rigor of outranking methods, while significantly enhancing practical flexibility and user experience. We demonstrate the framework's versatility through two case studies, each providing custom SRF results, enriched with additional considerations, such as uncertainty and robustness analysis.

## 1. Introduction

Criteria weight elicitation is a fundamental and challenging step in MCDA, when building a preference model (Greco et al., 2016). The importance weights assigned to criteria directly influence the ranking or selection of alternative options, yet DMs often find it challenging to articulate precise weight values. Over the years, a variety of weight elicitation techniques have been developed, each with its own cognitive demands and assumptions (Riabacke et al., 2012). A persistent goal in DM is, thereby, to design weight elicitation procedures that are both easy for DMs to understand and use, while robust enough to capture the nuances of preferences and deliver a stable and accurate model.

In the context of outranking preference methods, such as the ELECTRE and PROMETHEE family methods, weight elicitation is of particular significance (Brans & De Smet, 2005; Govindan & Jepsen, 2016). It is well recognized that the choice of weights in outranking models can critically affect the resulting recommendations and consequent decisions (Belton & Stewart, 2002). Researchers have therefore explored specialized techniques to derive criteria weights that DMs can provide intuitively, without needing extensive technical knowledge of the methods and algorithms. Among these, a "deck-of-cards" elicitation procedure proposed by Jean Simos has become especially popular. Initially introduced in 1990 to support ELECTRE-type methods, the Simos approach requires the DM to rank-order criteria cards from least to most

* Corresponding author.
*E-mail addresses:* river.huang@psi.ch (R. Huang), milosz.kadzinski@cs.put.poznan.pl (M. Kadziński), figueira@tecnico.ulisboa.pt (J.R. Figueira), salvatore.corrente@unict.it (S. Corrente), esiskos@tuc.gr (E. Siskos), peter.burgherr@psi.ch (P. Burgherr).

https://doi.org/10.1016/j.eswa.2026.131315
Received 29 October 2025; Received in revised form 10 January 2026; Accepted 20 January 2026
Available online 25 January 2026

important and insert blank cards to reflect perceived gaps between importance levels (Simos, 1990). Criteria, whose cards are clipped together (ex aequo criteria), are considered of equal importance, while adding more blank cards between criteria indicate a larger difference in importance. This simple "card game" is easy to understand and implement, and for this reason it has been extensively used by practitioners in real-world decision problems. The method was later refined by Figueira and Roy to include an additional parameter $z$, i.e. the ratio between the most and least important criterion, ensuring that it can be anchored to a prespecified value, rather than being left entirely relative and uncontrollable (Figueira & Roy, 2002).

In the field of MCDA, this enhanced Deck-of-Cards Method (DCM), also referred to as SRF procedure is considered an effective and user-friendly technique for assessing criteria weights (Siskos & Tsotsolas, 2015). Along with it, an extensive family of SRF variants has emerged. The Robust-SRF formulation tackles the imprecision of the basic model; instead of delivering a single weight vector, it explores the whole convex polyhedron of admissible weights, inferred by the criteria rank-ordering information that is modeled as a set of linear inequality constraints. It then quantifies the degree the final results can change under that uncertainty and provides certain robustness rules and recommendations (Siskos & Tsotsolas, 2015). Some other variants focus on easing the task of elicitation. SRF-II introduces a hypothetical "zero-importance" criterion, so that the DM can define the global ratio $z$ indirectly (Abastante et al., 2022). In Assessment Through Prioritization (WAP) method, the global $z$ is substituted by sub-ratios between adjacent criteria (Tsotsolas et al., 2019). In another variant of SRF, imprecision in inputs, such as blank-card counts or $z$ values is handled through interval, fuzzy, or probabilistic models, yielding methods like imprecise-SRF and Belief-degree SRF formulations that embed ranges or belief distributions in the constraints (Corrente et al., 2017; Zhang & Liao, 2023). Finally, (Corrente et al., 2016; Huang et al., 2025) suggested ways to handle multi-level criteria trees, i.e., hierarchical criteria structure. Altogether, these methodological developments have transformed the classical SRF into a versatile toolkit that can be adapted to a wide variety of decision-making contexts.

These innovations, however, have made the landscape of weight elicitation more fragmented. Each new method variant and extension has been typically developed in isolation, accompanied by its own procedural steps and mathematical formulations, even though all share the common foundation of the original Simos card technique. From a practitioner's perspective, it may be unclear which variant best fits a decision problem at hand, or whether it is possible to combine exclusive features of the different methods. There is, thereby, an evident research and practical gap, as no integrated framework has been developed, unifying the various SRF-based approaches and guiding DM in configuring a tailor-made weight elicitation process. For example, if a DM requires a hierarchical weight model with fuzzy judgments and a guarantee against any single criterion dominating the decision, there is no easy way to accomplish this without customizing the SRF algorithm. The absence of a unified structure also complicates researchers' efforts to compare methods or build upon each other's work, as each extension is presented as an independent technique.

In this paper, we address this gap by proposing a modular SRF framework that consolidates the many extensions of the SRF method into a single, flexible system. The key insight is that the core SRF algorithm can be broken down into a set of modular components or building blocks with different focuses:

- how the criteria are structured,
- in what form is the input of the DM,
- how the solution space is analyzed.

By offering interchangeable modules in each building block, we enable DMs to mix and match different features, stemming from the original methods, to create a customized weight elicitation procedure that best fits their decision context. To achieve that, a short questionnaire is developed and filled by the DM to guide the analyst in configuring the SRF variant that will be applied. The framework is then complemented with the appropriate linear or mixed-integer programming formulations that incorporate all the chosen modules' constraints, yielding either a unique weight vector or a characterization of the weight space. Finally, a procedure has been integrated into the framework to automatically check the consistency of the DM's preference statements. This step is designed to detect and correct any contradictory judgments. If inconsistencies are found, the procedure can identify the minimal adjustments needed in the DM's input to restore feasibility, offering constructive guidance throughout the process.

The proposed modular SRF framework transforms the once static DCM into a configurable toolbox for criteria weighting. It provides a single mathematical grammar that can reproduce the standard SRF, SRF-II, Robust SRF, Hierarchical SRF, and others simply by activating/deactivating specific constraint blocks. It preserves the intuitive appeal of the original Simos procedure, while offering the versatility to handle imprecise information, ensure normative considerations, and produce richer insights into weight uncertainty. Furthermore, by decoupling the structural, procedural, and informational modules, the framework allows for combinations that have never been proposed before. We demonstrate these advantages through two real-world case studies from the literature: a financial portfolio assessment and an evaluation of public hospital quality. In each case, we demonstrate how the modular approach can yield the results of the SRF analysis by incorporating the DMs' additional preferences that were previously unmodeled. These examples highlight that the unified framework not only bridges multiple methods but also provides practical value by yielding more informative and consistent weight assessments.

Our proposal incorporates mathematical programming, thereby contributing to the large body of MCDA methods, for which this technique provides the optimization backbone. In particular, classical goal programming (Charnes & Cooper, 1977) and compromise programming (Zeleny, 1973, 1974) formulate achievement or distance-to-ideal criteria and solve the resulting problems as linear or nonlinear programs, thus producing compromise solutions in the presence of conflicting goals. Preference disaggregation methods such as UTA/UTADIS, infer compatible preference model instances from indirect preference information (e.g., holistic rankings or assignments of reference alternatives) by solving linear programs that minimize various types of errors (Jacquet-Lagrèze & Siskos, 1982). Under incomplete or imprecise preference information, robustness-oriented approaches like preference programming (Salo & Hämäläinen, 2004), Robust Ordinal Regression (Greco et al., 2008), robust weight-interval analysis (Barron, 1992), and worst-case regret models use (mixed-)linear programming to characterize the whole set of parameters compatible with the available preference statements and derive robust results. Weight-elicitation procedures such as the Best-Worst Method (Rezaei, 2015) cast the determination of criteria weights as an optimization problem balancing consistency with the decision maker's judgments, while model-fitting approaches like MACBETH rely on linear programming to construct cardinal value scales from qualitative judgments of differences in attractiveness (Bana e Costa & Vansnick, 1994). In this way, mathematical programming serves both as a solver of explicit multiobjective formulations and as an inferential engine that learns and stress-tests MCDA models from (possibly indirect or imprecise) preference information. In our proposal, it is used in the same spirit, as a unifying optimization language, encoding the new preference learning, robustness analysis, and recommendation models we introduce.

The remainder of this paper is organized as follows. Section 2 reviews the SRF weight elicitation method and its main extensions for flat and hierarchical criteria, presenting each variant in a unified notation and highlighting its key features. This review lays the groundwork for our modular approach by pinpointing the common elements and unique mechanisms of each variant. Section 3 then introduces the modular SRF framework in detail: we define the modular components, present the

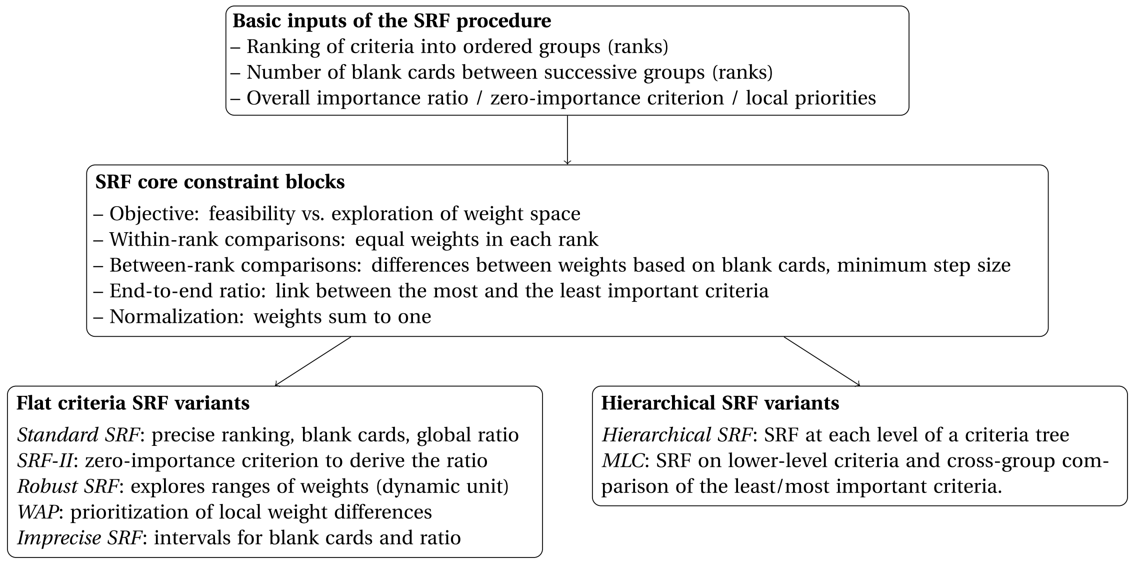


**Fig. 1.** Overview of SRF-based methods reviewed in Section 2: deck-of-cards preference information provided by the DM is encoded through generic SRF constraint blocks, which are instantiated differently in flat and hierarchical SRF variants.

guiding questionnaire for method configuration, and explain how different module choices map onto a final mathematical formulation. We also describe the integrated consistency checking and restoration procedure in this section. Section 4 demonstrates the use of the modular framework on two case studies from the literature, illustrating how the framework can be applied to redesign the weight elicitation process and what benefits this yields in terms of results and insights. Section 5 offers a discussion on the implications of the modular approach, including its flexibility, transparency, and potential limitations, as well as guidance for practitioners. We conclude the paper, summarizing our contributions and suggesting directions for future research.

## 2. Method review and reformulation

This section surveys the main variants of the SRF approach, identifies their key features, and then reformulates each variant in a unified algorithmic notation. Presenting the methods side by side highlights the common core they share with the standard SRF, as well as the mechanisms that set them apart, laying thereby the methodological foundation for the proposed modular SRF framework. The summary of reviewed approaches is provided in Fig. 1.

### 2.1. SRF And its extensions for flat criteria structure

SRF was originally designed for flat criteria structures, where all criteria exist at the same level. Based on SRF, several extensions were developed for the flat criteria structure MCDA problems. Let us use the following notation:

- $G = \{g_1, \dots, g_j, \dots, g_k, \dots, g_n\}$ denotes the set of evaluation criteria, each representing a distinct perspective of the alternatives under assessment.
- $R = \{R_1, \dots, R_s, \dots, R_r, \dots, R_t\}$ is the set of ranking positions (with $t \leqslant n$). The placement of criterion $g_j$ at rank $R_r$ is denoted by $g_j \in R_r$.
- $\mathcal{E} = \{e_1, \dots, e_s, \dots, e_r, \dots, e_{t-1}\}$ represents the set of blank cards inserted between successive ranks; specifically, each $e_s \geqslant 0$ indicates the number of blank cards between ranking positions $R_s$ and $R_{s+1}$. When $e_s = 0$, the weight difference between these ranks is taken as the basic unit; more generally, we define $\bar{e}_s = e_s + 1$ to represent the total number of unit weight intervals between $R_s$ and $R_{s+1}$.
- $W = \{w_1, \dots, w_j, \dots, w_k, \dots, w_n\}$ denotes the set of elicited weights corresponding to the criteria.

In the standard SRF method, the weight elicitation can be formulated as the following linear feasibility problem. Moreover, the SRF variants can be concluded within a modular framework with several components marked as $[Ox]$s that generalizes the initial algorithm:

$$\left.\begin{array}{ll}
[O1] & \text{no objective function,} \\
[O2] & w_j = w_k,\ g_j, g_k \in R_s,\ s = 1, \dots, t, \\
[O3] & w_j = w_k + C \cdot \bar{e}_s,\ g_j \in R_{s+1},\ g_k \in R_s,\ s = 1, \dots, t-1, \\
[O4] & w_j = z \cdot w_k,\ g_j \in R_t,\ g_k \in R_1, \\
[O5] & \sum_{j=1}^{n} w_j = 1, \\
[O6] & C \geqslant \varepsilon
\end{array}\right\} E(SRF) \tag{1}$$

where $\varepsilon$ is an arbitrarily small positive constant, usually defined by the analyst. In this formulation, the parameter $C$ denotes a unit of weight. It guarantees that the minimum weight difference between any two adjacent criteria is $C$; specifically, inserting $e_s$ blank cards between ranks $R_s$ and $R_{s+1}$ increases the difference to $(e_s + 1) \cdot C$. Combined with the ratio $z$ and the normalization constraint, this Constraint Programming (CP) formulation determines the weights of the evaluation criteria with precise inputs.

#### 2.1.1. Numerical illustration of the standard SRF procedure

To make the standard SRF formulation in Eq. (1) more concrete, consider a decision problem with four criteria $G = \{g_1, g_2, g_3, g_4\}$. Suppose the DM provides the following deck-of-cards information (from least to most important):

- Rank $R_1 = \{g_4\}$ (least important),
- Rank $R_2 = \{g_2, g_3\}$ (intermediate, ex aequo),
- Rank $R_3 = \{g_1\}$ (most important).

Between $R_1$ and $R_2$ the DM inserts no blank cards, i.e., $e_1 = 0$, while between $R_2$ and $R_3$ she inserts two blank cards, i.e., $e_2 = 2$. Hence, the corresponding unit intervals are

$$\bar{e}_1 = e_1 + 1 = 1, \qquad \bar{e}_2 = e_2 + 1 = 3.$$

Finally, the DM states that the most important criterion $g_1$ is five times as important as the least important one $g_4$, so that $z = 5$. Let $W = \{w_1, w_2, w_3, w_4\}$ denote the weights of $g_1, g_2, g_3, g_4$ and let $C$ be the unknown unit of weight. Instantiating Eq. (1) with this input yields the following system:

$$\left.\begin{array}{l}\text{Ex aequo set at } R_2\ ([O2]):\ w_2 = w_3,\\ \text{Gap between } R_1 \text{ and } R_2\ ([O3]):\ w_2 = w_4 + C\cdot\bar{e}_1 = w_4 + C,\\ \text{Gap between } R_2 \text{ and } R_3\ ([O3]):\ w_1 = w_2 + C\cdot\bar{e}_2 = w_2 + 3C,\\ \text{Global ratio ([O4]): } w_1 = z\cdot w_4 = 5w_4,\\ \text{Normalization ([O5]): } w_1 + w_2 + w_3 + w_4 = 1,\\ [O6]\, C \geq \varepsilon > 0.\end{array}\right\} \tag{2}$$

We set a small positive constant $1e-6$ for $\varepsilon$. From the model we obtain $C = 0.1$, and the final weights are

$$w_1 = 0.50,\quad w_2 = 0.20,\quad w_3 = 0.20,\quad w_4 = 0.10.$$

This example illustrates how the qualitative deck-of-cards information (ranking, blank cards, and the $z$ ratio) is converted, via the linear system Eq. (2), into a unique normalized weight vector consistent with the DM's judgments.

### 2.1.2. Modular decomposition of the SRF formulation

Eq. (1) comprises five essential components, along with an additional set of extra constraints:

- **[O1] Objective:** This component establishes the goal of the method. When a unique solution is desired, the objective is fixed at a constant value (e.g., 1), thereby converting the Linear Programming (LP) into a feasibility problem. Alternatively, different objective functions can be adopted to explore various regions of the solution space.
- **[O2] Comparison Rules within a subset of ex aequo criteria:** This element specifies the comparison of weights among criteria within the same rank.
- **[O3] Comparison Rules between successive criteria or subsets of ex aequo criteria:** This component specifies the comparison between criteria (or subsets of ex aequo criteria) of two successive ranks.
- **[O4] Ratio:** This constraint specifies the ratio $z$, indicating how many times more important the last criterion (or subset of ex aequo criteria) is compared to the first in the ranking.
- **[O5] Weight Normalization:** This constraint ensures that the sum of all criteria weights is normalized to 1.
- **[O6] Extra Constraints:** Additional constraints are introduced here. In particular, when a unit of weight is fixed, the constant $C$ is applied to define its minimum permissible value.

In certain decision contexts, the DM may wish to impose additional constraints. For example, to prevent any single criterion from "dictating" the outcome, i.e., having a weight larger than the sum of all others (Roy & Mousseau, 1996). The following dictatorship constraint can be added in that case (Corrente et al., 2017):

$$[OD]\, w_k \leq \sum_{j\neq k} w_j,\quad \forall k = 1,\ldots,n. \tag{3}$$

Alternatively, DM may require that all criteria are assigned at least a minimum weight, $\varepsilon^m$. This can be formulated as:

$$[OM]\, w_k \geq \varepsilon^m,\quad \forall k = 1,\ldots,n, \tag{4}$$

where $\varepsilon^m \in (0,1)$ is a user-specified lower bound on each criterion's weight. It should be noted that if the DM decides to include this information, the constraints required to translate it could render the problem infeasible.

From this perspective, every block in the set, including the optional components, can be regarded as a reusable module. By selectively redesigning, relaxing, or enriching these components, one can create new extensions of SRF, and systematically construct new ones tailored to specific decision contexts. Subsequent sections exploit this modular view to organize known SRF variants and support the design of customized procedures within a unified framework.

### 2.1.3. SRF-II

In the SRF-II method proposed by (Abastante et al., 2022), a "zero-importance criterion" is introduced as an alternative to defining the $z$ value. The DM specifies the number of blank cards, $e_0$, to be placed between this "zero-importance criterion" and the least important criterion. Consequently, the $[O4]$ component can be modified as follows:

$$[O4]\left.\begin{array}{l} z = \dfrac{\sum_{s=1}^{t-1}\bar{e}_s + e_0 + 1}{e_0+1},\\ w_j = z\cdot w_k,\ \forall g_j \in R_t, g_k \in R_1,\end{array}\right\} E(SRF-II) \tag{5}$$

### 2.1.4. Robust SRF

The robust SRF method, proposed by (Siskos & Tsotsolas, 2015), adopts a dynamic unit of weight and explores the weight space. This leads to modifications in constraints [O1], [O3], and [O6]. The modified formulation is presented in Eq. (6):

$$\left.\begin{array}{ll}[O1] & \max\,\&\,\min w_j, j = 1,\ldots,n \text{ subject to,}\\ [O3] & w_j > w_k,\ \forall g_j \in R_t, g_k \in R_s, t > s,\\ [O6] & w_{j+1} - w_j > w_{k+1} - w_k,\ \text{if } e_j > e_k\end{array}\right\} E(Robust-SRF) \tag{6}$$

In this formulation, constraint $[O1]$ is redefined to determine the maximum and minimum weights among all criteria, to calculate the maximum range of permissible weights. $[O3]$ and $[O6]$ are adapted to reflect the absence of a fixed weight unit and to capture the influence of the inserted blank cards. $[O3]$ enforces strict inequalities between the weights of criteria, assigned to different ranking sets, while $[O6]$ specifies the weight differences between successive criteria or subsets of ex aequo criteria.

### 2.1.5. WAP

The WAP method, proposed by (Tsotsolas et al., 2019), adapts the Robust SRF method by employing an alternative approach for defining the $z$ value. Rather than specifying a single $z$ value directly, the DM is asked to assign a value $z_s$ for each pair of successive criteria or subsets of ex aequo criteria, $g_s$ and $g_{s+1}$. The overall $z$ is then computed as:

$$z = z_1 \cdot z_2 \cdot \cdots \cdot z_{t-1}.$$

Moreover, as part of the WAP method, the DM can provide imprecise estimates for each $z_s$, expressing them as intervals $\left[z_s^{\text{low}}, z_s^{\text{upp}}\right]$. A subsequent robustness analysis is then applied to account for this induced imprecision. The modified formulation is presented in Eq. (7):

$$\left.\begin{array}{ll}[O1] & \max\,\&\,\min w_j,\quad j = 1,\ldots,n,\quad \text{subject to,}\\ [O3,O4,O6] & \left.\begin{array}{l} w_j \geqslant z^{\text{low}}\cdot w_k,\\ w_j \leqslant z^{\text{upp}}\cdot w_k,\end{array}\right\}\ \begin{array}{l}\forall g_j \in R_{s+1},\ g_k \in R_s,\\ \forall s = 1,\ldots,t-1,\end{array}\end{array}\right\} E(WAP) \tag{7}$$

In this formulation, constraints [O3], [O4], and [O6] are consolidated into a single constraint that defines the interval for $z_r$.

### 2.1.6. Imprecise SRF

In the imprecise SRF method, proposed by (Corrente et al., 2017), instead of specifying an exact number of blank cards between criteria or a precise $z$ value, possible intervals for these values are defined. These intervals are represented as $\left[e_s^{\text{low}}, e_s^{\text{upp}}\right]$ for blank cards and $\left[z^{\text{low}}, z^{\text{upp}}\right]$ for the $z$ value. Consequently, the $[O3]$ and $[O4]$ components are modified as follows:

$$\left.\begin{array}{ll}[O3] & \left.\begin{array}{l} w_j \geqslant w_k + C\cdot\left(e_s^{\text{low}}+1\right),\\ w_j \leqslant w_k + C\cdot\left(e_s^{\text{upp}}+1\right),\end{array}\right\}\begin{array}{l}\forall g_j \in R_{s+1}, g_k \in R_s,\\ \forall s = 1,\ldots,t-1,\end{array}\\ [O4] & \left.\begin{array}{l} w_j \geqslant z^{\text{low}}\cdot w_k,\\ w_j \leqslant z^{\text{upp}}\cdot w_k,\end{array}\right\}\forall g_j \in R_t, g_k \in R_1,\end{array}\right\} E(Imprecise\,SRF) \tag{8}$$

Here, constraint $[O3]$ is reformulated as two inequalities that specify the lower and upper bounds for the weight difference between successive criteria (or subsets of ex aequo criteria), based on the interval for blank cards. Likewise, constraint $[O4]$ is redefined to enforce that the weight ratio between the successive criteria falls within the interval $\left[z^{\text{low}}, z^{\text{upp}}\right]$. [O1] can use the same objective function from robust SRF to obtain the feasible weight space.

In (Corrente et al., 2021) , the authors extend both the standard SRF method and its imprecise variant by considering an enriched model, in which the DM can provide the number of blank cards $e_{pq}$, not only between consecutive levels but between any pair of levels $R_p$ and $R_q$. Also in this setting, the DM is again not required to give an exact value for the number of blank cards; instead, they may specify an interval $\left[e^{low}_{pq}, e^{upp}_{pq}\right]$ for $e_{pq}$. From a technical point of view, constraints [O3] are replaced by the following:

$$[O3]\quad \left.\begin{array}{l} w_j \geqslant w_k + C\cdot\left(e^{low}_{pq}+1\right),\\ w_j \leqslant w_k + C\cdot\left(e^{upp}_{pq}+1\right),\end{array}\right\}\quad \forall g_j\in R_p, g_k\in R_q, \forall p<q.$$

#### 2.1.7. Belief-degree imprecise SRF

Zhang and Liao (2023) further extends the imprecise SRF method within a Group Decision-Making (GDM) framework by considering the belief degree of DMs. Still, it is possible to apply it in a single DM context. In addition to specifying ranges of the number of blank cards and the $z$ parameter, the DM is required to express their degree of belief regarding these values. Specifically, the DM's belief about the number of blank cards inserted between $R_s$ and $R_{s+1}$, $s=1,2,\ldots,t-1$ is represented as a set of pairs, $\left\{(e_{s_1},\beta_{s_1}),(e_{s_2},\beta_{s_2}),\ldots\right\}$, where each pair $(e_{s_i},\beta_{s_i})$ indicates that the DM assigns a belief degree $\beta_{s_i}$ to the possibility of inserting $e_s=e_{s_i}$ blank cards between $R_s$ and $R_{s+1}$, with the probabilities satisfying $\sum_i \beta_{s_i}=1$. For example, the set $\{(1,0.4),(2,0.6)\}$ means that the DM believes there is a 40% probability that one blank card is inserted and a 60% probability that two blank cards are inserted between the relevant ranking positions. Similarly, a belief distribution, $\{(z_1,\beta_1),\ldots\}$, is specified for the $z$ parameter. Finally, the Stochastic Multi-criteria Acceptability Analysis (SMAA) (Lahdelma et al., 1998) is applied to run the standard SRF algorithm with these stochastic inputs, thereby deriving the corresponding weight information. It is possible to employ the same objective function for [O1] as the robust SRF method, thereby defining the feasible weight space.

#### 2.1.8. Hesitant Fuzzy Linguistic (HFL)-SRF

The HFL-SRF method, proposed by (Wu et al., 2022), employs Hesitant Fuzzy Linguistic Term Set (HFLTS) to assist the DM in expressing hesitation among linguistic terms, thereby capturing the inherent uncertainty in their evaluations (Rodriguez et al., 2011). The HFLTS $\mathcal{L}$ can be uniformly and symmetrically defined as follows for example:

$$\mathcal{L}=\left\{\begin{array}{ll} l_{-2}: \text{not preferred}, & l_{-1}: \text{slightly not preferred},\\ l_0: \text{equally preferred}, & \\ l_1: \text{slightly preferred}, & l_2: \text{preferred}\end{array}\right\}$$

The modified formulation is presented in Eq. (9):

$$\left.\begin{array}{lll} [O1] & \max\ \epsilon \quad \text{subject to,} & \\ [O3] & \left.\begin{array}{l} w_j \geqslant w_k + C\cdot\tau^{\text{low}}_s,\\ w_j \leqslant w_k + C\cdot\tau^{\text{upp}}_s,\end{array}\right\} & \begin{array}{l}\forall g_j\in R_{s+1},\ g_k\in R_s,\\ \forall s=1,\ldots,t-1,\end{array}\\ [O4] & \left.\begin{array}{l} w_j \geqslant \tau^{\text{low}}\cdot w_k,\\ w_j \leqslant \tau^{\text{upp}}\cdot w_k,\end{array}\right\} & \forall g_j\in R_t,\ g_k\in R_1.\end{array}\right\} E(HFL-SRF) \tag{9}$$

In this formulation, $\tau^{\text{low}}_s$ and $\tau^{\text{upp}}_s$ denote the smallest and largest linguistic terms, respectively, used to express the importance difference between ranking positions $R_s$ and $R_{s+1}$-analogous to the role of blank cards. Similarly, $\tau^{\text{low}}$ and $\tau^{\text{upp}}$ represent the smallest and largest linguistic terms for capturing the importance difference between the highest and lowest ranking positions ($R_1$ and $R_t$), serving a role similar to that of the $z$ value. The algorithm seeks to maximize the unit weight $\epsilon$, from which the weight vector is subsequently derived.

### 2.2. SRF Extension for hierarchical criteria structure

Many practical decision problems require a hierarchical organization of criteria, which facilitates the decomposition of complex evaluations into multiple levels of detail (Corrente et al., 2016). To address these challenges, researchers have developed several SRF variants, adapted for hierarchical criteria structures. Although these methods can be applied to complex structures of more than two levels, for simplicity reasons and without loss of generality, we present here the modeling corresponding to two levels. Suppose the higher-level criteria set, denoted as $G=\{g_1,\ldots,g_j,\ldots,g_k,\ldots,g_n\}$. We have:

- $\left\{R_1,\ldots,R_t\right\}$ is the set of ranking positions of criteria in $G$, with $t\leqslant n$.
- $\left\{e_1,\ldots,e_{t-1}\right\}$ represents the number of blank cards inserted between successive ranks, and $\bar{e}_s=e_s+1$, $s=1,\ldots,t-1$.
- $\left\{w_1,\ldots,w_n\right\}$ denotes the set of elicited weights for criteria in $G$.

Moreover, for each higher-level criterion $g_j\in G$, we define the following sets:

- $g_j=\left\{g_{j1},\ldots,g_{jj'},\ldots,g_{jk'},\ldots,g_{jn_j}\right\}$ is the set of lower-level criteria under $g_j$. Note that the number of lower-level criteria, $n_j$, may vary across different higher-level criteria.
- $\left\{R_{j1},\ldots,R_{js'},\ldots,R_{jr'},\ldots,R_{jt_j}\right\}$ is the set of ranking positions (with $t_j\leqslant n_j$).
- $E_j=\left\{e_{j1},\ldots,e_{js'},\ldots,e_{jr'},\ldots,e_{jt_j-1}\right\}$ represents the set of blank cards inserted between successive ranks, $\bar{e}_{js'}=e_{js'}+1$.
- $W_j=\left\{w_{j1},\ldots,w_{jj'},\ldots,w_{jk'},\ldots,w_{jn_j}\right\}$ is the set of elicited weights corresponding to the lower-level criteria.

#### 2.2.1. Hierarchical SRF

The hierarchical SRF method proposed by Corrente et al. (2016) extends the SRF framework to hierarchically structured criteria by applying the deck-of-cards approach at every level. In a two-level structure, the DM performs SRF on the higher-level set $G$, as well as on each lower-level set $g_j$. The formulation is as follows:

$$\left.\begin{array}{ll} [O^H1] & \text{no objective function},\\ \left.\begin{array}{ll}[O^H2] & w_{jj'}=w_{jk'},\quad g_{jj'},g_{jk'}\in R_{js'},\ s'=1,\ldots,t_j,\\ [O^H3] & w_{jj'}=w_{jk'}+C_j\cdot\bar{e}_{js'},\quad g_{jj'}\in R_{js'+1},\\ & \qquad g_{jk'}\in R_{js'},\ s'=1,\ldots,t_j-1,\\ [O^H4] & w_{jj'}=z_j\cdot w_{jk'},\quad g_{jj'}\in R_{jt_j},\ g_{jk'}\in R_{j1},\\ [O^H5] & C_j\geqslant\varepsilon\end{array}\right\}\forall j=1,\ldots,n, & \\ [O^H6] & w_j=w_k,\quad g_j,g_k\in R_s,\ s=1,\ldots,t,\\ [O^H7] & w_j=w_k+C\cdot\bar{e}_s,\quad g_j\in R_{s+1},\ g_k\in R_s,\ s=1,\ldots,t-1,\\ [O^H8] & w_j=z\cdot w_k,\quad g_j\in R_t,\ g_k\in R_1,\\ [O^H9] & C\geqslant\varepsilon\\ [O^H10] & \displaystyle\sum_{j'=1}^{n_j} w_{jj'}=w_j,\quad \forall j=1,\ldots,n\\ [O^H11] & \displaystyle\sum_{j=1}^{n} w_j=1,\end{array}\right\} E(Hierarchical-SRF) \tag{10}$$

In this formulation, constraints $[O^H2]-[O^H4]$ capture the DM's input for the lower-level criteria, while constraints $[O^H6]-[O^H8]$ reflect the information provided for the higher-level criteria. Constraint $[O^H10]$ ensures that the sum of the weights of the lower-level criteria under each higher-level criterion equals the weight of that higher-level criterion, and $[O^H11]$ normalizes the final weights so that their total sum is

1. Note that the $z$ values for the lower and higher levels may differ, and the weight units $C$ and $C_j$, which are ensured to be positive by $[O^H5]$ and $[O^H9]$, respectively, can also be distinct.

#### 2.2.2. Multi-Level Cards (MLC) Method

The MLC method proposed by Huang et al. (2025) employs a different weight elicitation approach, compared to the Hierarchical SRF. Instead of performing elicitation on multiple levels, MLC applies SRF directly to the lower-level criteria. Then, by selecting the sets of most and least important lower-level criteria from each ranking, SRF is applied again to compare these criteria across different higher-level categories. This approach is particularly useful when the higher-level criteria (eg. social, environmental, economic criteria) are inherently vague and difficult to compare directly.

Formally, we can use the notation introduced in the section above for the higher-level criteria $g_j$, $g = 1, \dots, n$. The information that differentiates MLC from the hierarchical SRF is modeled, as follows:

- The set composed of the least and most important lower-level criteria is: $\left\{R_{11}, R_{1t_1}, R_{21}, R_{2t_2}, \dots, R_{n1}, R_{nt_n}\right\}$.
- $\left\{R_1, \dots, R_t\right\}$ is the set of criteria ranks with $t \leqslant 2 \cdot n$.
- $\left\{e_1, \dots, e_{t-1}\right\}$ is the set of blank cards inserted between successive ranks, and $\bar{e}_s = e_s + 1$, $s = 1, \dots, t-1$.

The resulting formulation is given by:

$$\left.\begin{array}{ll}
[O^H1] & \text{no objective function,} \\
\left.\begin{array}{ll}
[O^H2] & w_{jj'} = w_{jk'}, \quad g_{jj'}, g_{jk'} \in R_{js'}, s' = 1, \dots, t_j, \\
[O^H3] & w_{jj'} = w_{jk'} + C_j \cdot \bar{e}_{js'}, \quad g_{jj'} \in R_{js'+1}, \\
 & \quad g_{jk'} \in R_{js'}, s' = 1, \dots, t_j - 1, \\
[O^H4] & w_{jj'} = z_j \cdot w_{jk'}, \quad g_{jj'} \in R_{jt_j}, g_{jk'} \in R_{j1}, \\
[O^H5] & C_j \geqslant \varepsilon
\end{array}\right\} & \forall j = 1, \dots, n, \\
\left.\begin{array}{ll}
[O^H6] & w_{jj'} = w_{kk'}, \quad g_{jj'}, g_{kk'} \in R_s, s = 1, \dots, t, \\
[O^H7] & w_{jj'} = w_{kk'} + C \cdot \bar{e}_s, \quad g_{jj'} \in R_{s+1}, \\
 & \quad g_{kk'} \in R_s, s = 1, \dots, t-1, \\
[O^H8] & w_{jj'} = z \cdot w_{kk'}, \quad g_{jj'} \in R_t, g_{kk'} \in R_1,
\end{array}\right\} & \\
[O^H9] \quad C \geqslant \varepsilon & \\
[O^H11] \quad \sum_{j=1}^{n}\sum_{j'=1}^{n_j} w_{jj'} = 1, &
\end{array}\right\} E(MLC) \tag{11}$$

In this approach, constraints $[O^H1]$ – $[O^H5]$ remain unchanged with respect to (11), meaning that the lower-level elicitation method is identical to that used in the Hierarchical SRF. Constraints $[O^H6]$ – $[O^H8]$ compare the most and least important lower-level criteria across different higher-level criteria. Since the higher-level criteria are not directly compared, $[O^H10]$ is omitted, and $[O^H11]$ normalizes all lower-level criteria so that their sum equals 1.

### 2.3. Consistency checking and restoration

Corrente et al. (Corrente et al., 2016) noted that inconsistencies can arise when accommodating hierarchically structured criteria in the SRF method. The same issue applies in the modular SRF algorithm, either in a flat criteria structure or a hierarchical criteria structure. These inconsistencies may result from conflicting rank-ordering preferences or incompatible $z$ values given by the DM. Inconsistencies in this context can result in infeasible models, preventing the proper determination of criteria weights and undermining the decision-making process. The techniques proposed by Mousseau et al. (Mousseau et al., 2003) for identifying conflicting constraints, along with [placeholder]'s extension for identifying potential resolution operations, are particularly important to implement.

These consistency checking and restoration techniques can be applied to the modular SRF model. To this end, three generalized models have been developed: the consistency checking model $\left(E^C\right)$, the inconsistency identification model $\left(E^I\right)$, and the consistency restoration model $\left(E^R\right)$. The consistency checking model $E^C$ should be applied immediately after the decision maker provides the deck-of-cards information, to verify that the input is consistent. This is achieved by modifying constraint $[O1]$ in the flat-criteria-structure SRFs and constraint $[O^H1]$ in the hierarchical-criteria-structure SRF variants:

$$[O1] \quad \varepsilon^* = \max \varepsilon \} \; E^C$$

Here, $E^C$ determines the maximum possible unit weight $\varepsilon^*$. If $\varepsilon^* > 0$, the model is deemed feasible, indicating that the constraints derived from the DM's input are consistent. In this case, the standard SRF can be applied by selecting a unit weight $\varepsilon$ from the interval $(0, \varepsilon^*]$. Conversely, if $\varepsilon^* \leqslant 0$, the model is infeasible, and the inconsistency identification model $\left(E^I\right)$ should be employed. $E^I$ identifies the conflicting information provided by the DM, which may necessitate revising different constraints depending on the specific SRF variant. As a general illustration, we present an example that applies the inconsistency identification model $E^I$ to the standard SRF formulation:

$$\left.\begin{array}{ll}
[O1] & \min \left[\sum_{s=1}^{t-1}(y_s^+ + y_s^-) + y_z^+ + y_z^-\right], \\
[O2] & w_j = w_k, \quad g_j, g_k \in R_s, \; s = 1, \dots, t, \\
[O^+3] & w_j - w_k - C \cdot \bar{e}_s \geqslant -M \cdot y_s^+, \quad g_j \in R_{s+1}, \\
 & \quad g_k \in R_s, \; s = 1, \dots, t-1, \\
[O^-3] & w_j - w_k - C \cdot \bar{e}_s \leqslant M \cdot y_s^-, \quad g_j \in R_{s+1}, \\
 & \quad g_k \in R_s, \; s = 1, \dots, t-1, \\
[O^+4] & w_j - z \cdot w_k \geqslant -M \cdot y_z^+, \quad g_j \in R_t, \; g_k \in R_1, \\
[O^-4] & w_j - z \cdot w_k \leqslant M \cdot y_z^-, \quad g_j \in R_t, \; g_k \in R_1, \\
[O5] & \sum_{j=1}^{n} w_j = 1, \\
[O6] & C \geqslant \varepsilon \\
[O7] & y_s^+, y_s^-, y_z^+, y_z^- \in \{0, 1\}, \quad s = 1, \dots, t-1, \\
[O8] & y_s^+ + y_s^- \leqslant 1, \quad y_z^+ + y_z^- \leqslant 1
\end{array}\right\} E^I(SRF) \tag{12}$$

Here, the original constraints $[O3]$, and $[O4]$ are replaced by $[O^+3]$, $[O^-3]$, $[O^+4]$, and $[O^-4]$, which incorporate binary variables $y_s^+$, $y_s^-$, $y_z^+$, and $y_z^-$ to detect instances of High Estimation Inconsistency Error (HEIE) and Low Estimation Inconsistency Error (LEIE) in the deck-of-cards information. Constraints $[O7]$ and $[O8]$ ensure that both HEIE and LEIE cannot co-occur for a given pair of criteria or $z$ value. Specifically, when $y_s^+ = 1$, the gap between criteria in $R_s$ and criteria in $R_{s+1}$ is too large (i.e., too many blank cards have been inserted), and when $y_z^+ = 1$, the $z$ value is excessively high; these conditions are referred to as HEIE. Conversely, $y_s^- = 1$ or $y_z^- = 1$ indicate that the gap or $z$ value is too small, corresponding to LEIE. $[O6]$ defines an extremely small value for the unit weight. The objective in $[O1]$ identifies the minimum set of inconsistencies in the decision maker's input. When the optimal value of $[O1]$ is zero, the model is consistent.

This model identifies the minimal set of constraints that, once removed, make the model feasible. However, since more than one of these subsets could exist, to identify all of them, an iterative procedure can be developed. This involves solving the model iteratively and progressively adding new constraints to limit previously identified subsets, allowing the discovery of new infeasible subsets. The pseudocode for this iterative Mixed-Integer Linear Programming (MILP) model for standard SRF is outlined below:

Algorithm 1 systematically identifies all subsets of constraints that, when removed, make the model $E(SRF)$ feasible. Once an inconsistent subset is identified, it is recorded in a designated set, and a new constraint is introduced to prevent the same subset from being flagged again. In the extreme case, this procedure continues until all pieces of preference information provided by the DM have been removed. The stopping condition can be easily modified to consider only those subsets of constraints with minimal cardinality.

Based on Algorithm 1, we identify the subsets of constraints that lead to model infeasibility, indicating inconsistencies in the DM's input

**Algorithm 1** Iterative $E^I$.

**Ensure:** Minimal sets of indices of constraints $\mathcal{X}$ that once removed restore the model feasibility.

1: Initialize an empty set $\mathcal{X}$ to remove minimal constraint sets.
2: Construct and solve model $E^I(SRF)$:
3: **if** $[O1] = 0$ **then**
4: Exit: no constraints need to be relaxed. Go to step 17:
5: **else**
6: $i = 1$;
7: **repeat**
8: Denote by $\mathbf{y}^i = \left[y_1^{i,+}, y_1^{i,-}, y_2^{i,+}, y_2^{i,-}, \ldots, y_t^{i,+}, y_t^{i,-}, y_z^{i,+}, y_z^{i,-}\right]$ the vector of binary variables found solving $E^I(SRF)$.
9: Identify the set of indices of violated constraints $x^i = x_{blank}^{i,+} \cup x_{blank}^{i,-} \cup x_{ratio}^{i,+} \cup x_{ratio}^{i,-}$ where
10:

$$x_{blank}^{i,+} = \left\{s : y_s^{i,+} = 1\right\}, \qquad x_{blank}^{i,-} = \left\{s : y_s^{i,-} = 1\right\}, \qquad x_{ratio}^{i,+} = \begin{cases} z & \text{if} \quad y_z^{i,+} = 1, \\ \emptyset & \text{if} \quad y_z^{i,+} = 0 \end{cases} \qquad x_{ratio}^{i,-} = \begin{cases} z & \text{if} \quad y_z^{i,-} = 1, \\ \emptyset & \text{if} \quad y_z^{i,-} = 0 \end{cases}$$

11: Add the set of indices $x^i$ to $\mathcal{X}$:

$$\mathcal{X} = \mathcal{X} \cup x^i$$

12: Introduce a new constraint to avoid finding the same set of indices of constraints causing the infeasibility in the next iteration:

$$\left[New^i\right] \qquad \sum_{s \in x_{blank}^{i,+}} y_s^{i,+} + \sum_{s \in x_{blank}^{i,-}} y_s^{i,-} + \sum_{z \in x_{ratio}^{i,+}} y_z^{i,+} + \sum_{z \in x_{ratio}^{i,-}} y_z^{i,-} \leqslant \left|x_{blank}^{i,+}\right| + \left|x_{blank}^{i,-}\right| + \left|x_{ratio}^{i,+}\right| + \left|x_{ratio}^{i,-}\right| - 1$$

13: Solve the MILP problem $E^I(SRF) = E^I(SRF) \cup \left\{New^i\right\}$
14: $i = i + 1$
15: **until** No further feasible solutions exist.
16: **end if**
17: Return all minimal sets of relaxed constraints $\mathcal{X}$, which represent the subsets of constraints that need to be relaxed for feasibility.

regarding the relationships between criteria. To address this, $E^R$ is applied to identify minimal modifications to the DM's input to restore feasibility.

$$\left.\begin{array}{ll} [O1] & \min \left[\sum_{s=1}^{t-1} (\sigma_s^+ + \sigma_s^-) + \sigma_z^+ + \sigma_z^-\right], \\ [O^R3] & w_j = w_k + C \cdot (\bar{e}_s - \sigma_s^+ + \sigma_s^-), \quad g_j \in R_{s+1}, \\ & \quad g_k \in R_s,\ s \in x^i, \\ [O^R4] & w_j = (z - \sigma_z^+ + \sigma_z^-) \cdot w_k, \quad g_j \in R_t,\ g_k \in R_1,\ z \in x^i \\ [O7] & \sigma_s^+, \sigma_s^-, \sigma_z^+, \sigma_z^- \in \mathbf{N}, \quad s = 1, \ldots, t-1, \\ [O8] & \bar{e}_s - \sigma_s^+ + \sigma_s^- \geqslant 1, \quad z - \sigma_s^+ + \sigma_s^- \geqslant 2 \end{array}\right\} E^R(SRF) \tag{13}$$

Here, $\sigma_s^+$, $\sigma_s^-$, $\sigma_z^+$, and $\sigma_z^-$ are introduced to determine the appropriate number of blank cards and the corresponding $z$ value needed for feasibility. The objective function in $[O1]$ identifies the minimum modifications required to restore consistency in the model. In a simple case, one might assume that the effect of one blank card is equivalent to a unit change in the $z$ value; however, it is straightforward to assign distinct weights to these two components to reflect the relative priority of minimizing changes in $z$ versus adjustments in the number of blank cards.

By running the model over all subsets in $\mathcal{X}$, various solutions for restoring consistency can be identified. The DM can then select a preferred solution, adjust the input accordingly, and generate the criteria weights. Conversely, if no solution is found by modifying the inconsistent subsets, it indicates that the preferences provided by the DM are excessively inconsistent. In such cases, it would be more beneficial to restart the elicitation process rather than attempting to "patch" the constructed model.

## 3. Proposition of a modular SRF framework

Existing SRF variants have expanded the original deck-of-cards idea in valuable ways, yet several important gaps remain. First, there is no unified formalism that groups these methods according to their common mechanisms and clarifies which components they actually share; most variants are presented in isolation, rather than as instances of a broader architecture. Second, the literature offers little guidance to analysts and DM on which SRF variant to choose for a given decision context, or on how to combine complementary features from different variants into a coherent procedure; in practice, DM must improvise whenever their requirements do not match exactly the assumptions of a single published method. Third, although inconsistency analysis and robustness assessment have been studied for particular SRF extensions, they are not provided in a systematic, method-agnostic way that can be applied across the full family of SRF models. Overall, the field lacks a structured, theoretically grounded framework that both unifies existing SRF variants and supports the principled design of tailor-made SRF configurations, equipped with explicit inconsistency handling and robustness analysis from the outset.

Therefore, we proposed a four-module framework that addresses these gaps by recasting SRF and its extensions within a single, compositional grammar of constraints, which is structured in subsection 3.1. The framework is complemented with a questionnaire, presented in subsection 3.2, that supports and guides analysts and DM in selecting the components they need.

### 3.1. Overview and rationale

The modular SRF framework rebuilds the classical SRF procedure as a set of interoperable modules that can be assembled to fit a specific decision context and provide targeted solutions. The idea is to separate what is being structured from how information is elicited and how results can be analyzed. Concretely, the following four modules drive the construction of a consolidated SRF model:

1. Structural module: defines the organization of criteria (flat or hierarchical);
2. Procedural module: selects the deck-of-cards mechanism (standard SRF, SRF-II with a "zero-importance" criterion, WAP-style successive ratios, etc.);

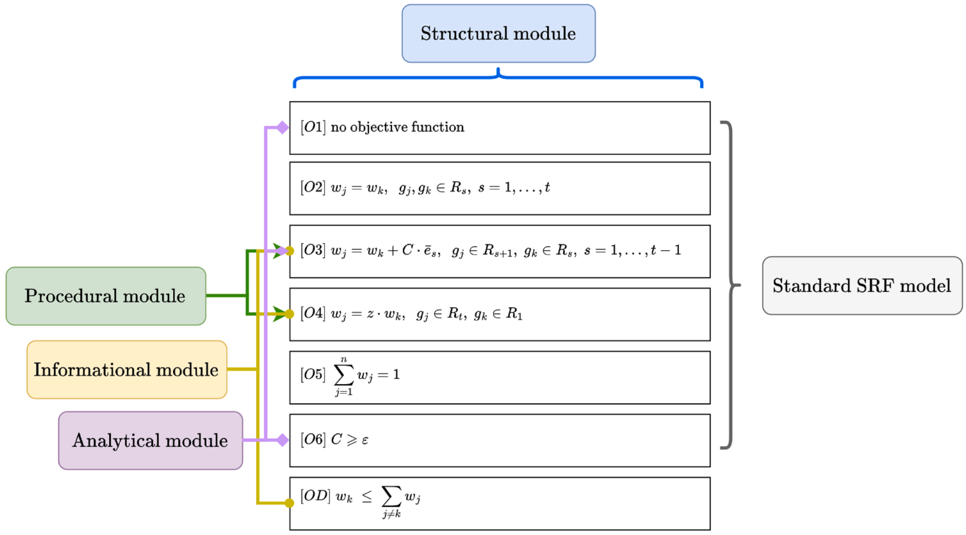


**Fig. 2.** Architecture of the modular SRF framework. SRF is decomposed into four plug-in modules: Structural, Procedural, Informational, and Analytical.

3. Informational module: Encodes the DM's inputs (precise or imprecise gaps/ratios; interval, fuzzy/linguistic, or stochastic formats; and optional normative constraints, such as no-dictatorship and minimum weight);
4. Analytical module: Determines the output form and related analysis that needs to be performed.

These choices compile into a single linear (or mixed) model by plugging the appropriate constraints, after which optional consistency checking/restoration and robustness analysis are applied.

Fig. 2 illustrates the architecture and information flow between the four distinct modules. This pathway starts from the basic SRF (Eq. (1)) and guides the analyst to (i) define the criteria structure, (ii) choose the most suitable deck-of-cards mechanism, (iii) specify the format and fidelity of the required preference information, and (iv) select the appropriate post-optimality/robustness analyses. The chosen components are combined into a single model, enabling consistency checking, restoration, and robustness analysis.

Within the modular SRF framework, robustness analysis focuses on assessing and enhancing the stability of the model under construction, in line with the propositions and concerns of Roy (Roy, 2010). Specialized techniques and robustness indicators can be applied to measure, interpret, and visualize the instability or variability of the criteria weights. These indicators provide guidance on whether additional preference information should be requested from the DM to improve the robustness of the model.

One such indicator is Average Stability Index (ASI), which synthesizes the dispersion of the sampled weight vectors into a single measure:

$$\mathrm{ASI} = 1 - \frac{1}{n} \sum_{i=1}^{n} \frac{\sqrt{h \sum_{x=1}^{h} \left(w_i^x\right)^2 - \left(\sum_{x=1}^{h} w_i^x\right)^2}}{\frac{h}{n}\sqrt{n-1}}, \tag{14}$$

where $n$ is the total number of criteria, $h$ is the number of sampled weight vectors, and $w_i^x$ denotes the weight of criterion $i$ in sample $x$. The index takes values in $[0, 1]$, with values close to 1 indicating high stability (low weight variability) and values close to 0 indicating instability.

Another useful robustness indicator is Average Range of the Preference Parameters (ARP) (Siskos et al., 2022), which captures the average spread of the weights:

$$\mathrm{ARP} = \frac{1}{n} \sum_{i=1}^{n} \left(w_i^{\max} - w_i^{\min}\right), \tag{15}$$

where $w_i^{\max}$ and $w_i^{\min}$ denote, respectively, the maximum and minimum feasible values of each criterion $i$ that can be calculated using the results of the min-max procedure of $[O1]$ as part of the Robust-SRF method as shown in Eq. (6). ARP index takes values in $[0, 1]$, with values close to 0 indicating high robustness (Roy, 2010).

However, it is important to note that the proposed SRF framework does not assess the robustness of the model's results, as in the paradigm of Robust Ordinal Regression (Kadziński, 2022), because such analyses are dependent on the problematic of the decision and the actual MCDA method that is applied. Practitioners interested in implementing robustness analysis can refer to Kadziński and Tervonen (2013) for additive value models, to Greco et al. (2010) for classification problems, to Greco et al. (2011) for outranking methods, and to Siskos et al. (2025) for the Choquet integral models.

### 3.2. A detailed questionnaire to personalize the modular SRF

The four modules of the proposed framework, as well as the alternative SRF solutions and analyses, can be selected by the users via the tailor-made questionnaire that we designed. The questionnaire describes the different required components and asks for targeted information to help DMs understand and eventually select the structural, procedural, informational, and analytical modules they need. Inspired by the MCDA-MSS questionnaire structure (Cinelli et al., 2020), we categorize these design choices into four key dimensions. They transform the SRF method from a fixed elicitation tool into a customizable framework, offering flexibility in both modeling assumptions and data requirements. The questionnaire can be administered in a paper or online format during the elicitation process, or integrated into dedicated software to guide, record, and validate the DM's responses, thereby streamlining the workflow. Additionally, when necessary, the questionnaire offers concise descriptions and references to the different SRF variants, providing guidance to users unfamiliar with their specifics. The conceptual form of the questionnaire, split into the four different types of questions, is as follows:

- **Structural questions** concern the organization of the evaluation criteria. DMs begin by choosing between a flat or a hierarchical structure, and in the latter case, indicate whether SRF should be applied independently at each level or aggregated in a bottom-up way.

(a) **Which structure best describes your evaluation criteria?**
  i. Flat: all criteria reside on a single level.
  ii. Hierarchical: criteria are organized into multiple levels.
(b) **If your criteria are hierarchical, how should SRF be applied?**
  i. Level-wise: apply SRF independently at each level.
  ii. Bottom-up aggregation: first, apply SRF to each set of lower-level criteria; then, apply SRF to the most/least important of those to derive higher-level weights.

At Q1, the DM chooses between a flat or a hierarchical criteria structure, thereby selecting the base SRF variant for all subsequent modules.

- If flat is chosen (Q1 (a)), the procedure proceeds with $E(SRF)$ as its foundation (Figueira & Roy, 2002).
- If hierarchical is chosen (Q1 (b)), Q2 then asks whether to
  * apply SRF independently at each level (Q2 (a)), yielding $E(Hierarchical - SRF)$ (Corrente et al., 2016), or
  * aggregate bottom-up via the MLC procedure (Q2 (b)), yielding $E(MLC)$ (Huang et al., 2025).

- **Procedural question** lets the DMs select among three variants of the deck-of-cards elicitation mechanism.

(a) **Which deck-of-cards mechanism would you like to use?**
  i. Standard SRF deck (blank cards + $z$ value) (Figueira & Roy, 2002).
  ii. Zero-criterion variant: introduce a "zero-importance criterion" as a null benchmark, so no $z$ value needs to be defined (Abastante et al., 2022).
  iii. Direct-ratio variant: ask for the $z$ value between successive criteria (or subsets of ex aequo criteria) ranks directly, so that no blank cards are used (Tsotsolas et al., 2019).

At Q3, the DM selects one of three deck-of-cards mechanisms, which in turn determines which comparison constraints to include in the SRF model (and analogously in the Hierarchical-SRF or MLC variants):

- Standard SRF deck (Q3 (a)): use blank cards plus a global $z$ value, retaining both $[O3]$ and $[O4]$ (or $[O^H3]$, $[O^H4]$, $[O^H7]$, $[O^H8]$ in $E(Hierarchical - SRF)$ or $E(MLC)$ (Corrente et al., 2016; Figueira & Roy, 2002; Huang et al., 2025)[1]).
- Zero-criterion variant (Q3 (b)): introduce a "zero-importance criterion" as a null benchmark, eliminating the need for $z$ and replacing $[O4]$ with the modified constraint from $E(SRF - II)$ (Abastante et al., 2022).
- Direct-ratio variant (Q3 (c)): elicit $z$ directly between successive criteria (or subsets of ex aequo criteria) ranks (no blank cards), thereby substituting both $[O3]$ and $[O4]$ with the corresponding constraints in $E(WAP)$ (Tsotsolas et al., 2019).

- **Informational questions** concern how the DMs' judgments are captured and represented. Here, the questionnaire explores whether inputs are precise or imprecise, and in the latter case, the DMs are asked to select a preferred format. Also, two normative design options are introduced, giving the possibility to avoid criteria dominance and/or criteria receiving negligible weighting levels.

(a) **What type of input will you provide for the distances between successive criteria ranks (i.e., blank cards, or $z$ if Q3 (c) is chosen)?**
  i. Precise (exact number of cards).
  ii. Imprecise (uncertain information).
(b) **If your distance inputs are imprecise, which format will you use?**
  i. Interval (bounds only).
  ii. Fuzzy numbers (e.g., HFLTS).
(c) **What type of input will you provide for the $z$ ratio?**
  i. Precise (exact value).
  ii. Imprecise (uncertain information).
(d) **If your $z$ ratio is imprecise, which format will you use?**
  i. Interval (bounds only).
  ii. Fuzzy numbers (e.g., HFLTS).
(e) **For any imprecise inputs, would you rather specify a probability distribution?**
  i. No
  ii. Yes
(f) **Would you like to enforce a "no-dictatorship" constraint among the criteria (no criterion receives weight above 50%)?**
  i. No
  ii. Yes
(g) **Would you like to impose a minimum allowable weight for all criteria?**
  i. No
  ii. Yes

These questions determine how the DM's input is interpreted and which SRF variant or constraints are activated.

- At Q4, the DM specifies whether importance gaps between successive criteria (or subsets of ex aequo criteria) ranks (or $z$ when Q3(c)) are precise or imprecise:
  * Precise (Q4 (a)): keep the distance constraint $[O3]$ decided by previous questions.
  * Imprecise (Q4 (b)): proceed to Q5 to choose a representation.
- At Q5, the DM chooses how to model imprecise distances:
  * Interval (Q5 (a)): further revise $[O3]$ based on the interval-based constraint of $E(Imprecise - SRF)$ $[O3]$ (Corrente et al., 2017).
  * Fuzzy (Q5 (b)): further revise $[O3]$ based on the Fuzzy-logic, for example the HFL constraints of $E(HFL - SRF)$ (Wu et al., 2022).
- At Q6, the DM indicates if the $z$ ratio is precise or imprecise:
  * Precise (Q6 (a)): keep the ratio constraint $[O4]$ decided by previous questions.
  * Imprecise (Q6 (b)): proceed to Q7 for representation.
- At Q7, the DM selects a format for imprecise $z$:
  * Interval (Q7 (a)): further revise $[O4]$ based on the interval-based constraint of $E(Imprecise - SRF)$ $[O4]$ (Corrente et al., 2017).
  * Fuzzy (Q7 (b)): further revise $[O4]$ based on the Fuzzy-logic, for example the HFL constraints of $E(HFL - SRF)$ (Wu et al., 2022).
- At Q8, the DM decides whether to endow imprecise inputs with probability distributions:
  * (Q8 (a)) retain the interval or fuzzy constraints chosen above.
  * (Q8 (b)) DM needs define probability of the inputs.
- At Q9, the DM chooses whether to enforce a no-dictatorship rule:
  * (Q9 (a)) results in no additional constraint.
  * (Q9 (b)) add the dictatorship constraint $[OD]$.
- At Q10, the DM decides whether to impose a minimum weight on each criterion:
  * (Q10 (a)) results in no additional constraint.
  * (Q10 (b)) add the minimum-weight constraint $[OM]$.

- **Analytical questions** determine how the results of the SRF procedure will be expressed and interpreted. The DMs first choose between using a single, deterministic weight vector and conducting an output variability analysis. If the latter is selected, further choices involve whether the unit weight should remain fixed or be allowed to vary dynamically, and what method should be used to conduct the analysis: exact or probabilistic.

(a) **What type of weight output do you prefer?**
  i. Single (univocal) weight vector.
  ii. Variability analysis (range or distribution).

[1] To avoid repetition, only the flat-structure modifications are detailed in all questions after; hierarchical variants substitute the corresponding $[O^H]$ constraints.

(b) **Would you like to assume a fixed unit weight $C$ or allow it to vary dynamically?**
  i. Fixed unit weight.
  ii. Dynamic unit weight.

(c) **If you choose variability analysis, how should it be performed?**
  i. Extreme scenarios: identify best- and worst-case weight vectors.
  ii. Sampling: approximate the full weight distribution via Monte Carlo or other sampling methods.

These questions determine whether and how to perform variability analysis on the elicited weights:

- At Q11, the DM chooses between a single univocal weight vector or a variability analysis as the output:
  - Single vector (Q11 (a)): retain the original objective $[O1]$.
  - Variability analysis (Q11 (b)): proceed to Q12.
- At Q12, the DM decides whether the unit weight $C$ is fixed or dynamic:
  - Fixed (Q12 (a)): no further change needed.
  - Dynamic (Q12 (b)): further revise $[O3]$, $[O6]$ based on the non-fixed unit distance in from $E(Robust-SRF)$ $[O3]$, $[O6]$ (Siskos & Tsotsolas, 2015).
- At Q13, the DM selects the method for variability analysis:
  - Extreme scenarios (Q13 (a)): use $E(Robust-SRF)[O1]$ to compute best- and worst-case weight vectors (Siskos & Tsotsolas, 2015).
  - Sampling (Q13 (b)): integrate Stochastic Multicriteria Acceptability Analysis (SMAA) to generate a sample-based weight distribution (Siskos & Tsotsolas, 2015).

The questionnaire introduced above is visualized in Fig. 3 as a decision-flowchart, which guides the DM step-by-step through the modular SRF framework.

The DM does not need to answer all questions, and by default the answer to any question is (a). Therefore, if no questions are answered, the standard SRF is produced. By tracing a path from the top of the chart down to various components, the DM knows which mix of structural, procedural, informational, and analytical modules to plug into the final algorithm formulation. In this way, as shown in the detailed flowchart of Fig. 3, the questionnaire does not only organize the many SRF variants but also provides a straightforward recommendation for building new, tailor-made SRF algorithms.

In what follows, we discuss how to instantiate the modular mathematical programming model based on the vector $q$ of answers to Q1–Q13. The complete constraint system induced by $q$ can be written as:

$$E^{\text{mod}}(q) = E^{\text{elic}}(q) \cup E^{\text{norm}}(q) \cup E^{\text{anal}}(q),$$

where $E^{\text{elic}}(q)$ collects elicitation constraints (structural, procedural, informational), $E^{\text{norm}}(q)$ contains optional normative safeguards, and $E^{\text{anal}}(q)$ contains the constraints and objectives used for robustness or variability analysis. The structural part selects a base model according to Q1–Q2,

$$E^{\text{struct}}(q) = \begin{cases} E(\text{SRF}), & \text{if Q1(a)}, \\ E(\text{Hierarchical-SRF}), & \text{if Q1(b), Q2(a)}, \\ E(\text{MLC}), & \text{if Q1(b), Q2(b)}. \end{cases}$$

It is then locally modified by the procedural and informational choices (Q3–Q8) through replacements of the comparison blocks $[O3]$ and $[O4]$. In particular, $E(\text{SRF-II})$ and $E(\text{WAP})$ provide alternative versions of these blocks, while $E(\text{Imprecise-SRF})[O3]$, $E(\text{Imprecise-SRF})[O4]$, $E(\text{HFL-SRF})[O3]$ and $E(\text{HFL-SRF})[O4]$ replace the corresponding parts of the active base model when interval or HFL inputs are selected. Formally,

$$E^{\text{elic}}(q) = E^{\text{struct}}(q) \cup E^{\text{info}}(q),$$

$$E^{\text{info}}(q) \subseteq \{E(\text{SRF-II}), E(\text{WAP}), E(\text{Imprecise-SRF})[O3], E(\text{Imprecise-SRF})[O4], E(\text{HFL-SRF})[O3], E(\text{HFL-SRF})[O4]\},$$

with the subset determined by Q3–Q8 and summarized in Fig. 3. Normative options Q9–Q10 add the constraints $[OD]$ and/or $[OM]$:

$$E^{\text{norm}}(q) \in \{\emptyset, \{[OD]\}, \{[OM]\}, \{[OD],[OM]\}\}.$$

In turn, analytical options Q11–Q13 determine whether and how the robust constraints of $E(\text{Robust-SRF})$ (in particular blocks $[O1]$, $[O3]$ and $[O6]$) and a sampling algorithms are used:

$$E^{\text{anal}}(q) \subseteq \{E(\text{Robust-SRF})[O1], E(\text{Robust-SRF})[O3], E(\text{Robust-SRF})[O6]\},$$

with $E^{\text{anal}}(q) = \emptyset$ when a single representative weight vector is computed.

The consistency-analysis models of Subsection 2.3 are then obtained by instantiating the generic templates $E^C(\cdot)$, $E^I(\cdot)$ and $E^R(\cdot)$ on the same structural choice as $E^{\text{struct}}(q)$, yielding, for example, $E^C(\text{SRF})$, $E^I(\text{SRF})$, $E^R(\text{SRF})$ or $E^C(\text{MLC})$, $E^I(\text{MLC})$, $E^R(\text{MLC})$. All these mathematical programming models reuse the constraint blocks selected in $E^{\text{mod}}(q)$.

Note that all constraint sets used in this paper – namely $E(\text{SRF})$, $E(\text{SRF-II})$, $E(\text{WAP})$, $E(\text{Imprecise-SRF})$, $E(\text{HFL-SRF})$, $E(\text{Hierarchical-SRF})$, $E(\text{MLC})$, $E(\text{Robust-SRF})$, as well as the consistency-analysis models $E^C(\text{SRF})$, $E^I(\text{SRF})$, $E^R(\text{SRF})$ and their hierarchical/MLC counterparts introduced in Subsection 2.3, are defined on the same variable space. The modular design only turns constraint blocks on or off (or replaces one block by another) and never duplicates variables. Hence, the selected blocks can be safely assembled into a unified LP/MILP formulation. However, it is necessary to note that when combining SRF model with imprecision, i.e., $E(\text{Imprecise-SRF})$ or $E(\text{HFL-SRF})$ with $E(\text{Robust-SRF})$, it introduces a nontrivial modeling issue: $E(\text{Imprecise-SRF})$, $E(\text{HFL-SRF})$ represents the number of blank cards as interval-valued inputs, whereas $E(\text{Robust-SRF})$ does not enforce a fixed unit gap and instead requires only ordinal relations between successive weight differences. When the number of blank cards may vary across ranks, the implied ordering of gap sizes becomes conditional. To preserve the intended semantics, auxiliary variables (typically binary variables in a MILP reformulation) are required to linearize these conditional relations and ensure logical consistency across all admissible realizations of the interval inputs. This implementation is also presented in Subsection 4.1 of the following case study.

## 4. Framework demonstration with case studies

To illustrate the flexibility and added value of the modular SRF framework, we revisit two published case studies drawn from different application domains that initially employed the standard SRFs. Moreover, the two studies we selected are recent, and they report on the weight elicitation procedures they followed in full detail, including the complete deck-of-cards settings and the resulting weight vectors, making them ideal for replication. For each study, we redesign the weighting procedure by assembling an SRF variant from our module library. By comparing the new weight vectors with those obtained in the original analyses, and reporting on the decision support experience from the perspective of the DM, we can highlight the practical advantages and potential trade-offs introduced by the modular approach.

### 4.1. Case study 1: Financial equity-portfolio construction

This study proposes a unified framework that couples MCDA with time-series forecasting to improve equity-portfolio construction (Xidonas et al., 2025). The authors show that forward-looking forecasts can be transformed into multicriteria scores, whose informational content is distinct from both the raw market data and the forecasts themselves.

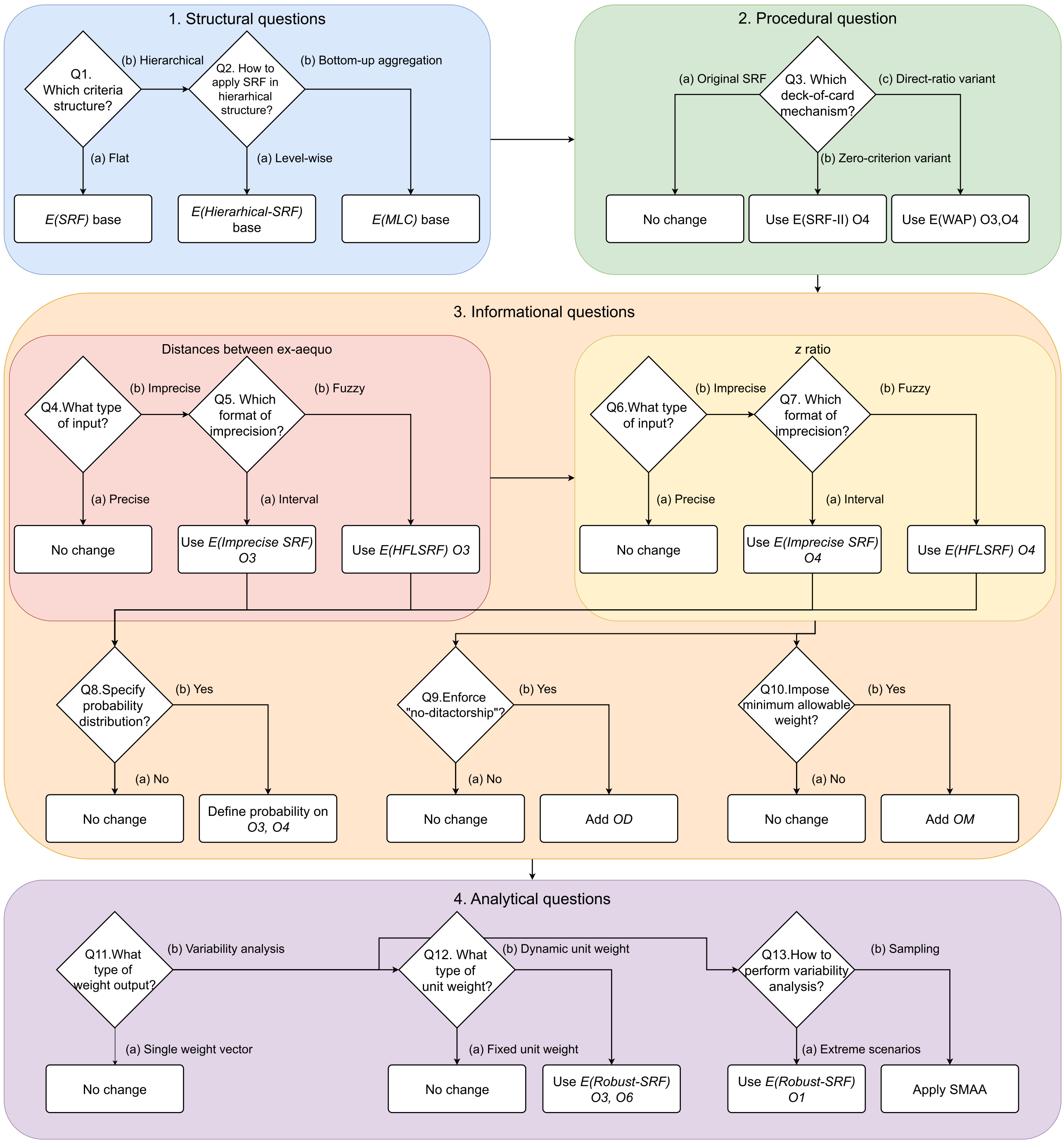


**Fig. 3.** Modular SRF workflow.

Using 22 years of monthly data for the eleven S&P 500 sector ETFs, they demonstrate that portfolios built on the integrated scores outperform those based on traditional metrics alone. The study evaluates the portfolio using six criteria: $g_1$ (Wealth), $g_2$ (Volatility), $g_3$ (Maximum Drawdown), $g_4$ (Mean-Squared Error of the forecasts, MSE), $g_5$ (Mean Absolute Error of the forecasts, MAE), and $g_6$ (Sign-Success Ratio of the forecasts, SSR). Criteria weights were elicited using the standard SRF method, and the expert's rank order of the six criteria is depicted in Fig. 4.

In our case, during the re-assessment of the preference model using the Modular-SRF framework, theDM filled the questionnaire, and the responses are shown in Table 1. A flat criteria structure is retained, and the DM prefers to use a "zero-importance criterion" as a benchmark instead of specifying $z$ value directly (Q3 (b)). The DM is, however, uncertain about both the number of blank cards: distances are treated as imprecise intervals (Q4 (b)-Q5 (a)), and the $z$ value related questions are not applicable. Two additional constraints are requested: a no-dictatorship rule (Q9 (b)) and a minimum weight for every criterion (Q10 (b)). For validation of the stability of the results, the DM requests a variability analysis (Q11 (b)). In the analysis, the weight unit $C$ is kept free/unconstrained (Q12 (b)), and samples are taken from the feasible region (Q13 (b)).

The configuration chosen in this study can be summarized as follows. The DM selects a flat structure (Q1(a)), so the base model is $E$(SRF); chooses the zero-criterion mechanism (Q3(b)), which replaces block [$O4$] by its SRF-II counterpart in $E$(SRF-II); declares imprecise distances

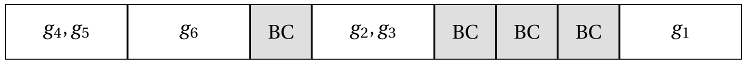


**Fig. 4.** Deck-of-cards representation of the elicited preference structure for case study 1. Criteria appear from left (least important) to right (most important). Grey blocks are blank cards. Ex aequo criteria are grouped within the same block.

**Table 1**
Decision maker's answers to the Modular SRF questionnaire for case study 1.

| Question | Q1 | Q2 | Q3 | Q4 | Q5 | Q6 | Q7 | Q8 | Q9 | Q10 | Q11 | Q12 | Q13 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Answer** | (a) | N/A | (b) | (b) | (a) | N/A | N/A | (a) | (b) | (b) | (b) | (b) | (b) |

(Q4(b)-Q5(a)), which replace block $[O3]$ by $E$(Imprecise-SRF)$[O3]$; and activates both normative options (Q9(b), Q10(b)), which add anti-dictatorship and minimum-weight requirements $[OD]$ and $[OM]$. The elicitation and normative constraints, therefore, read

$$E^{\text{mod}}_{\text{CS1}} = E(\text{SRF}) \cup E(\text{SRF-II})[O4] \cup E(\text{Imprecise-SRF})[O3] \cup E(\text{Robust-SRF})[O6] \cup \{[OD],[OM]\}.$$

Eventually, the modular SRF model, constructed for the DM, is presented in Eq. (16), which is an instantiated LP formulation of $E^{\text{mod}}_{\text{CS1}}$. The complete formulation, which encodes the full set of elicited preference information and introduces the auxiliary variables needed to linearize the conditional constraints, is reported in Appendix A.1.

$$\left.\begin{array}{ll}
[O1] & \text{no objective function,} \\
[O2] & w_j = w_k, \quad g_j, g_k \in R_s,\ s = 1,\dots,t, \\
[O3] & w_j > w_k, \quad \forall g_j \in R_t, g_k \in R_s, t > s, \quad g_j \in R_{s+1}, \\
 & \qquad g_k \in R_s,\ s = 1,\dots,t-1, \\
[O4] & \left.\begin{array}{l} z \le \dfrac{\sum_{s=1}^{t-1} \overline{e}_s + e_0^{\text{low}} + 1}{e_0^{\text{low}} + 1}, \\ z \ge \dfrac{\sum_{s=1}^{t-1} \overline{e}_s + e_0^{\text{upp}} + 1}{e_0^{\text{upp}} + 1}, \\ w_j = z \cdot w_k, \quad \forall g_j \in R_t, g_k \in R_1, \end{array}\right\} \\
[O5] & \sum_{j=1}^{n} w_j = 1, \\
[O6] & w_{j+1} - w_j > w_{k+1} - w_k, \quad \text{if } e_j > e_k \\
[OD] & w_k \le \sum_{j \neq k} w_j, \quad k = 1,\dots,n, \\
[OM] & w_k \ge \epsilon^m, \quad k = 1,\dots,n.
\end{array}\right\} \qquad (16)$$

Feasibility is checked via $E^C$(SRF). Since no inconsistencies are found, the admissible weight region defined by $E^{\text{mod}}_{\text{CS1}}$ is explored by hit-and-run sampling in the spirit of $E$(Robust-SRF). In particular, the DM specifies:

- $e_0 \in [0,1]$ (between "zero-importance criterion" and $g_4, g_5$);
- $e_{(4,5),6} = 0$ (between $g_4, g_5$ and $g_6$);
- $e_{6,(2,3)} \in [0,1]$ (between $g_6$ and $g_2, g_3$);
- $e_{(2,3),1} \in [2,3]$ (between $g_2, g_3$ and $g_1$).

With the additional lower bound of 0.05, the hit-and-run algorithm draws 10,000 weighting samples and produces the weight distribution in Fig. 5 and their mean values, listed in Table 2. In the figure, each violin illustrates the empirical distribution of the sampled feasible weights for criteria $g_1$-$g_6$. The thick horizontal line inside every violin corresponds to the mean values of the weights. The results clearly indicate that Criterion $g_1$ is the most important, with most of its probability mass being between 0.30 and 0.40. Criteria $g_2$ and $g_3$ exhibit nearly symmetric distributions centered around 20%. It is important to note that the sampled weights are derived from a new set of elicitation inputs and, therefore, are not directly comparable to the point estimates from the original study. The purpose of Table 2 and Fig. 5 is to show the successful integration of the no-dictatorship and minimum-weight preference statements into the construction of the modular SRF model and how such input can influence the resulting weights. In addition, the results highlight the inherent imprecision of the constructed model, which stems from the SRF elicitation process and is further amplified by the use of interval-valued blank cards. If these levels of imprecision are deemed unacceptable by the analyst and the DM, additional preference information can be requested from the DM, as detailed in Section 3.

**Table 2**
Case study 1 – original point estimates versus mean weights obtained from 10 000 hit-and-run samples of the modular SRF algorithm.

| Criterion | Original Weight (%) | Average Weight from Samples (%) |
|---|---|---|
| $g_1$ | 40.00 | 35.07 |
| $g_2$ | 20.00 | 19.56 |
| $g_3$ | 20.00 | 19.56 |
| $g_4$ | 5.00 | 6.70 |
| $g_5$ | 5.00 | 6.70 |
| $g_6$ | 10.00 | 12.40 |

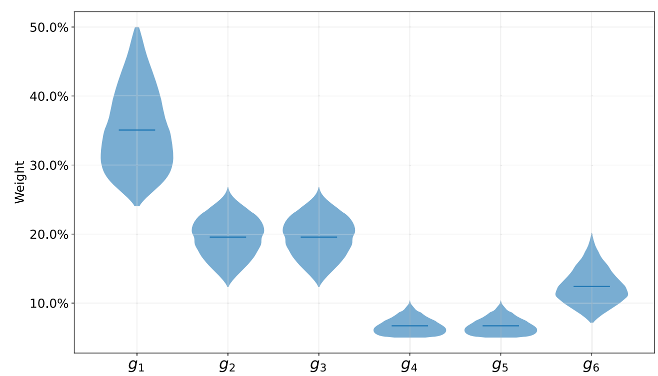


**Fig. 5.** Case study 1 – violin plots of the feasible weight distributions generated by hit-and-run sampling; the solid line inside each violin marks the sample mean reported in Table 2.

### 4.2. Case study 2: Public hospital quality assessment

Building upon recent reforms in Portugal's National Health Service, the second case study developed a hierarchical MCDA framework to benchmark the quality of 25 public hospitals over the 2019-2022 period (Ricardo et al., 2024). Four top-level dimensions are retained: $g_1$ (Timeliness), $g_2$ (Effectiveness), $g_3$ (Safety), and $g_4$ (Efficiency). Each criterion is decomposed into elementary performance indicators, i.e., subcriteria, drawn from the ACSS national benchmarking database. The full set of criteria and subcriteria is shown in Table 3. Criterion weights were elicited through the Hierarchical-SRF procedure, according to the expert's rank order of the criteria, which is depicted in Fig. 6.

**Table 3**
Hierarchy of criteria and associated performance indicators.

| Criterion | Subcriterion | Indicator (elementary criterion) |
|---|---|---|
| $g_1$ Timeliness | $g_{11}$ | % of hip-fracture surgeries performed within the first 48 h |
| | $g_{12}$ | % of first consultations performed within the recommended time |
| | $g_{13}$ | % of LIC registrations completed within the TMRG |
| | $g_{14}$ | Average time before surgery |
| $g_2$ Effectiveness | $g_{21}$ | % of outpatient-clinic surgeries within total scheduled surgeries (DRG) |
| | $g_{22}$ | % of deliveries by Caesarean section |
| | $g_{23}$ | Standard patients per physician (ETC) |
| | $g_{24}$ | Standard patients per nurse (ETC) |
| $g_3$ Safety | $g_{31}$ | % of readmissions within 30 days (different calendar years) |
| | $g_{32}$ | % of hospitalizations with a delay exceeding 30 days |
| | $g_{33}$ | Pressure-ulcer rate |
| | $g_{34}$ | Rate of central-venous-catheter bloodstream infections |
| | $g_{35}$ | Post-operative pulmonary embolism or deep vein thrombosis per 100 000 |
| | $g_{36}$ | Post-operative sepsis per 100 000 |
| | $g_{37}$ | % of instrumented vaginal deliveries with 3rd/4th-degree lacerations |
| $g_4$ Efficiency | $g_{41}$ | Annual inpatient occupancy rate |
| | $g_{42}$ | Operating costs per standard patient |

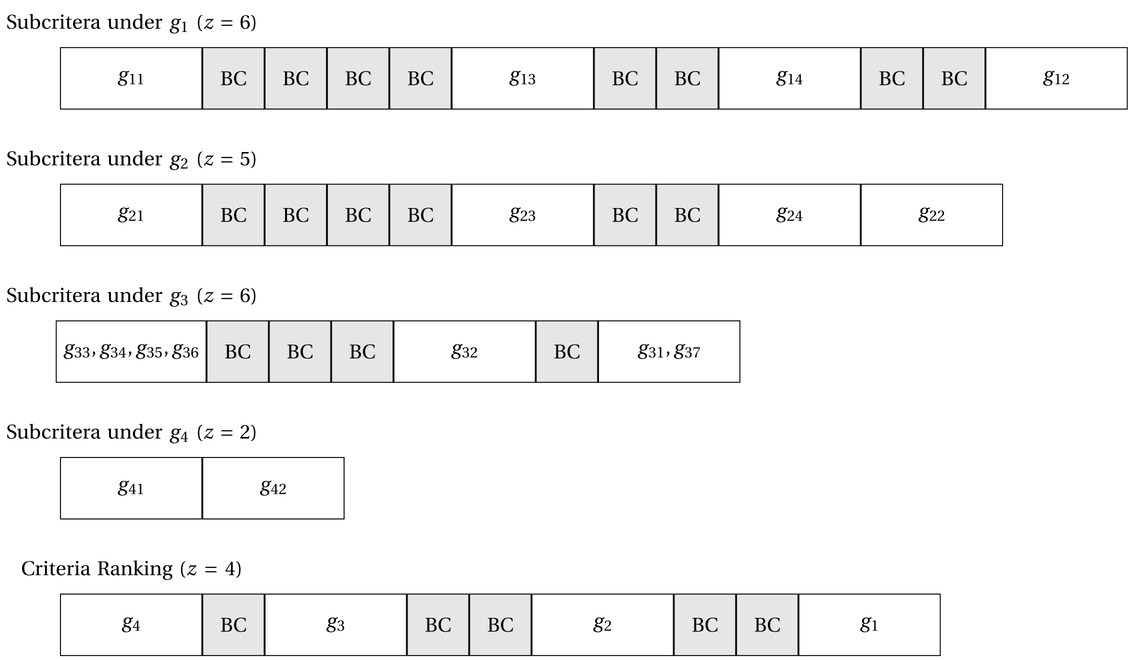


**Fig. 6.** Deck-of-cards representation of the elicited preference structure for case study 2. Criteria appear from left (least important) to right (most important). Grey blocks are blank cards. Ex aequo criteria are grouped within the same block.

**Table 4**
Decision maker's answers to the Modular SRF questionnaire for case study 2.

| Question | Q1 | Q2 | Q3 | Q4 | Q5 | Q6 | Q7 | Q8 | Q9 | Q10 | Q11 | Q12 | Q13 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Answer** | (b) | (b) | (a) | (a) | N/A | (a) | N/A | N/A | (b) | (b) | (b) | (b) | (a) |

In a new elicitation round, the DM completed the Modular-SRF questionnaire and the answers are summarized in Table 4. A hierarchical criteria structure is retained (Q1 (b)), and DM prefers a bottom-up aggregation of weights (Q2 (b)). The standard deck-of-cards procedure is maintained for all levels. Two extra constraints are imposed: a *no-dictatorship* rule (Q9 (b)) and a minimum weight for every criterion (Q10 (b)). For the exploitation phase, the DM requests a variability analysis (Q11 (b)) with a dynamic unit weight $C$ (Q12 (b)) and prefers to perform some sampling in the weight space rather than merely calculate the extreme values of the weights (Q13 (b)), implying the use of a SMAA-type technique.

Since the DM opts for a hierarchical structure (Q1(b)) with bottom-up aggregation (Q2(b)), the base model is $E$(MLC). Then, we use the standard SRF deck (Q3(a)) with precise distance and ratio inputs (Q4(a), Q6(a)), so neither $E$(Imprecise-SRF) nor $E$(HFL-SRF) is activated. The selection of both normative constraints (Q9(b), Q10(b)) adds anti-dictatorship $[OD]$ and minimum-weight $[OM]$ constraints. The elicitation and normative part of the model is therefore

$$E^{\text{mod}}_{\text{CS2}} = E(\text{MLC}) \cup E(\text{Robust-SRF})[O^H 3] \cup E(\text{Robust-SRF})[O^H 5] \cup E(\text{Robust-SRF})[O^H 7] \cup E(\text{Robust-SRF})[O^H 9] \cup \{[OD], [OM]\}.$$

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| $g_{41}$ | $g_{33}, g_{34}, g_{35}, g_{36}, g_{42}$ | $g_{21}$ | $g_{11}$ | $g_{31}, g_{37}$ | BC | $g_{22}$ | $g_{12}$ |

**Fig. 7.** Deck-of-cards representation of inter-group comparison of case study 2 with MLC procedure. Criteria appear from left (least important) to right (most important). Grey blocks are blank cards. Ex aequo criteria are grouped within the same block.

Eq. (17) provides a concrete hierarchical LP formulation of $E^{\text{mod}}_{\text{CS2}}$, while the detailed preferences of the DM can be found in Eq. (A.2) in the Appendix.

$$\left.\begin{array}{ll}
[O^H1] & \text{no objective function,} \\
\left.\begin{array}{ll}
[O^H2] & w_{jj'} = w_{jk'},\ \ g_{jj'}, g_{jk'} \in R_{js'}, s' = 1,\dots,t_j, \\
[O^H3] & w_{jj'} > w_{jk'} \cdot \bar{e}_{js'},\ \ g_j \in R_{s+1}, g_k \in R_s, s = 1,\dots,t-1, \\
[O^H4] & w_{jj'} = z_j \cdot w_{jk'},\ \ g_{jj'} \in R_{jt_j}, g_{jk'} \in R_{j1}, \\
[O^H5] & w_{jj'+1} - w_{jj'} > w_{jk'+1} - w_{jk'},\ \ \text{if } e_{jj'} > e_{jk'}
\end{array}\right\} & \forall j = 1,\dots,n, \\
\left.\begin{array}{ll}
[O^H6] & w_{jj'} = w_{kk'},\ \ g_{jj'}, g_{kk'} \in R_s, s = 1,\dots,t, \\
[O^H7] & w_{jj'} > w_{kk'},\ \ g_{jj'} \in R_{s+1}, g_{kk'} \in R_s, s = 1,\dots,t-1, \\
[O^H8] & w_{jj'} = z \cdot w_{kk'},\ \ g_{jj'} \in R_t, g_{kk'} \in R_1 \\
[O^H9] & w_{jj'+1} - w_{jj'} > w_{kk'+1} - w_{kk'},\ \ \text{if } e_{j'} > e_{k'}
\end{array}\right\} & \begin{array}{l} g_{jj'} \in \{R_{j1}, R_{jt_j}\}, \\ g_{kk'} \in \{R_{k1}, R_{kt_k}\}, \end{array} \\
[O^H10] & \sum_{j=1}^{n} \sum_{j'=1}^{n_j} w_{jj'} = 1, \\
[OD] & w_{kk'} \le \sum_{\substack{j=1 \\ (j,j') \ne (k,k')}}^{n} \sum_{j'=1}^{n_j} w_{jj'},\ \ k = 1,\dots,n, k' = 1,\dots,n_k. \\
[OM] & w_{kk'} \ge \epsilon^m.
\end{array}\right\} \tag{17}$$

The consistency analysis of Subsection 2.3 is instantiated as $E^C$(MLC), $E^I$(MLC), and $E^R$(MLC), solved in sequence, and the revised feasible region is then sampled using the robustness machinery of $E$(Robust-SRF).

In this elicitation round, DM no longer specifies a complete deck-of-cards ordering for every elementary criterion. Instead, following the MLC protocol, the DM selects the least- and most-important subcriteria inside each main criterion and then compares these "extreme" subcriteria across dimensions, yielding the inter-group hierarchy depicted in Fig. 7. The value of $z$, indicating a ratio of weights associated with the global most important criterion $g_{12}$ and the global least important criterion $g_{41}$, is 12. This two-step procedure captures subtler inter-group relationships with a process that does not increase the DM's cognitive load. To comply with the $[OM]$ constraint, every criterion is subject to a uniform lower bound of 0.01.

The initial set of preference statements provided by the DM proved inconsistent after applying the consistency-checking model. By running the inconsistency identification model $E^I$ (see Subsection 2.3) with the structure of Eq. (17), we identified 46 minimal sets of constraints whose joint activation renders the model infeasible. Table 5 reports the partial such sets, expressed in terms of the binary inconsistency indicators introduced in $E^I$.

In this notation, a superscript " + " (e.g. $y^+_{\text{Intra},g_{24},g_{22}}$) flags a HEIE, meaning that the corresponding gap or scaling constraint is too demanding (for instance, too many blank cards or an excessively large ratio), whereas a superscript " - " (e.g. $y^-_{\text{Intra},g_{11},g_{13}}$) indicates a LEIE, i.e., a gap or ratio that is too small to be compatible with the remaining statements. The first subscript label (Intra, Inter, Scale, GlobalScale) identifies whether the affected constraint concerns intra-group gaps, inter-group comparisons, local $z_j$-ratios or the global scaling $z$, while the remaining subscripts specify the criteria involved.

For example, Set 1 isolates a single intra-group gap between $g_{24}$ and $g_{22}$ that must be relaxed, while Sets 3, 4 and 5 reveal pairs of mutually incompatible intra-group gaps inside the Timeliness, Effectiveness and Safety dimensions. Sets 6-10 involve combinations of intra-, inter-group and (global) scaling constraints, showing that the joint specification of local $z$-ratios, global $z$ and cross-group comparisons is overly tight.

However, not all inconsistent subsets can be repaired under our assumptions. Consider, for instance, Set 1, where $y^+_{\text{Intra},g_{24},g_{22}} = 1$ indicates that the gap between $g_{24}$ and $g_{22}$ is too large. If blank cards had been inserted between these two criteria, the restoration model could simply recommend removing one or more of them. In our case, however, no blank card separates $g_{24}$ and $g_{22}$. Reducing the gap would therefore force these criteria to become an ex aequo set, which would contradict the ranking information provided by the DM. In the proposed framework, we explicitly allow modifications to the number of blank cards and to the $z$ values, since these inputs are typically more imprecise. By contrast, we treat the rank order itself as hard ordinal information that cannot be altered by the repair procedure.

Algorithm 1 is then applied to identify all constraint sets that can be adjusted to restore feasibility without changing the ranking information. Among the admissible repair patterns, two of the simplest are summarized by the adjustment variables $\sigma^+$ and $\sigma^-$ as follows:

$$\begin{cases} \sigma^-_{\text{Inter},g_{31},g_{22}} = 1,\ \sigma^-_{\text{Inter},g_{37},g_{22}} = 1,\ \sigma^-_{\text{Scale},g_{12},g_{11}} = 2,\ \sigma^+_{\text{GlobalScale},g_{12},g_{41}} = 4, \\ \sigma^+_{\text{Inter},g_{22},g_{12}} = 1,\ \sigma^+_{\text{GlobalScale},g_{12},g_{41}} = 13, \end{cases} \tag{18}$$

which correspond to the following two minimal repair actions:

1. increase the global ratio $z$ from 12 to 16, decrease the intra-group ratio $z_1$ (within $g_1$) from 4 to 2, and remove one blank card from the inter-group comparison between $\{g_{31}, g_{37}\}$ and $g_{22}$;
2. increase the global $z$ to 25 and insert one additional blank card between $g_{22}$ and $g_{12}$.

The DM preferred option 1, reasoning that it left the inner structure of the other dimensions untouched. With this single adjustment, the model became feasible and we were thereby able to explore the resulting weighting space and obtain the minimum, maximum, and mean weight vectors. The results are juxtaposed with the original point estimates in Table 6.

According to the new results, it is evident that most criteria now display pronounced weight intervals. In particular, the Safety indicators $g_{31}$ and $g_{37}$ become markedly heavier, whereas the Efficiency indicators $g_{41}$ and $g_{42}$ contract toward the imposed 5% floor, an outcome that is fully coherent with the DM's revised judgments. Every feasible weight vector fulfills both the no-dictatorship requirement and the 5% mini-

**Table 5**
Partial minimal inconsistent constraint sets identified by $E^I$ for case study 2.

| Set | Inconsistent binary variables |
|---|---|
| 1 | $\{y^+_{\text{Intra},g_{24},g_{22}}\}$ |
| 2 | $\{y^+_{\text{Intra},g_{14},g_{12}}\}$ |
| 3 | $\{y^+_{\text{Intra},g_{13},g_{14}}, y^-_{\text{Intra},g_{11},g_{13}}\}$ |
| 4 | $\{y^+_{\text{Intra},g_{23},g_{24}}, y^-_{\text{Intra},g_{21},g_{23}}\}$ |
| 5 | $\{y^+_{\text{Intra},g_{32},g_{31}}, y^+_{\text{Intra},g_{32},g_{37}}\}$ |
| 6 | $\{y^+_{\text{Inter},g_{22},g_{12}}, y^-_{\text{GlobalScale},g_{12},g_{41}}\}$ |
| 7 | $\{y^-_{\text{Scale},g_{42},g_{41}}, y^-_{\text{Inter},g_{41},g_{42}}, y^+_{\text{Inter},g_{42},g_{21}}\}$ |
| 8 | $\{y^-_{\text{Inter},g_{31},g_{22}}, y^-_{\text{Inter},g_{11},g_{37}}, y^+_{\text{Scale},g_{12},g_{11}}, y^-_{\text{Inter},g_{11},g_{31}}\}$ |
| 9 | $\{y^-_{\text{Inter},g_{21},g_{11}}, y^-_{\text{Inter},g_{31},g_{22}}, y^+_{\text{Scale},g_{12},g_{11}}, y^-_{\text{Inter},g_{37},g_{22}}\}$ |
| 10 | $\{y^-_{\text{Inter},g_{37},g_{22}}, y^-_{\text{Inter},g_{31},g_{22}}, y^+_{\text{GlobalScale},g_{12},g_{41}}, y^-_{\text{Scale},g_{22},g_{21}}\}$ |
| ... | ... |

**Table 6**
Case study 2 – comparison between the original point estimates and the feasible weight range obtained with the modular SRF after minimal consistency restoration.

| | Hierarchical SRF | Modular SRF | | |
|---|---|---|---|---|
| Criterion | Weight (%) | Min (%) | Mean (%) | Max (%) |
| $g_{11}$ | 2.79 | 2.81 | 3.14 | 3.48 |
| $g_{12}$ | 16.72 | 11.25 | 12.54 | 13.92 |
| $g_{13}$ | 9.12 | 6.33 | 9.38 | 11.78 |
| $g_{14}$ | 12.92 | 7.99 | 10.06 | 12.82 |
| $g_{21}$ | 2.17 | 1.97 | 2.21 | 2.54 |
| $g_{22}$ | 10.84 | 9.83 | 11.04 | 12.72 |
| $g_{23}$ | 6.99 | 5.27 | 7.46 | 10.35 |
| $g_{24}$ | 9.88 | 7.99 | 9.55 | 11.76 |
| $g_{31}$ | 5.37 | 8.44 | 9.41 | 10.44 |
| $g_{32}$ | 3.87 | 5.30 | 7.18 | 9.55 |
| $g_{33}$ | 0.89 | 1.41 | 1.57 | 1.74 |
| $g_{34}$ | 0.89 | 1.41 | 1.57 | 1.74 |
| $g_{35}$ | 0.89 | 1.41 | 1.57 | 1.74 |
| $g_{36}$ | 0.89 | 1.41 | 1.57 | 1.74 |
| $g_{37}$ | 5.37 | 8.44 | 9.41 | 10.44 |
| $g_{41}$ | 3.46 | 0.70 | 0.78 | 0.87 |
| $g_{42}$ | 6.93 | 1.41 | 1.57 | 1.74 |

mum, while the adjusted ratios $z = 16$ (global) and $z_1 = 2$ (within $g_1$) re-establish consistency throughout the hierarchy.

For the assessment of the robustness of the constructed model, we calculate the ASI and ARP indicators, as shown in Eq. (14) and Eq. (15). For the present case study, we obtain $ASI = 0.973$ and $ARP = .021$, which signifies that the feasible weight region exhibits low instability. This fact allows us to confidently apply the predefined MCDA method, using the mean weights, and obtain a ranking of the alternatives under evaluation.

As discussed in Section 3, while the weighting results exhibit a high degree of stability, it remains essential for practitioners to conduct a robustness analysis on the results before utilizing them for decision-making. Additionally, it is important to emphasize that these weighting results are derived from a new set of preference inputs; thus, they should be interpreted as complementary rather than directly comparable to the single-point weights reported in the original study.

### *4.3. Adaptability to complex decision environments*

The two case studies presented in Section 4 demonstrated how the modular SRF framework can be successfully applied to redesign existing weight elicitation procedures, yielding tailored results that respect specific normative constraints and criteria structures. However, decision-making environments often present even greater complexity, involving ambiguous preference statements, mixed-fidelity data, or intricate hierarchical dependencies that go beyond standard applications. To further illustrate the adaptability and innovation of the proposed framework, we delineate five distinct decision scenarios characterized by challenging working conditions. In each case, we specify how the modules are assembled to provide a solution where a static, "one-size-fits-all" approach would likely fail.

**Scenario 1: hierarchy, level-wise, imprecise intensities.** Consider a decision problem structured as a hierarchy with top-level groups such as *Sustainability*, *Service Quality*, and *Cost*. *Sustainability* is decomposed into *Emissions*, *Energy Use*, and *Waste*, while *Service Quality* includes *Reliability*, *Responsiveness*, and *Accessibility*. The DM can rank criteria within each group (e.g., *Emissions* ≻ *Energy Use* ≻ *Waste*) but can only express intensities in an imprecise, comparative manner. For instance, statements may take the form: "*Emissions* is slightly more important than *Energy Use*, and this difference is not larger than the gap between *Energy Use* and *Waste*." To address this, the tailored SRF variant combines: (a) a hierarchical structural module with level-wise elicitation; (b) a standard deck-of-cards procedural module; and (c) an informational module designed for imprecise distances (e.g., interval blank-card counts). Since aggregating multiple local interval statements increases the risk of infeasibility, the *inconsistency detection module* is applied immediately after elicitation. If needed, the *restoration module* proposes minimal preference adjustments to recover feasibility. The final output is a feasible set of weights, summarized by ranges and, where beneficial, a probability distribution of admissible solutions.

**Scenario 2: hierarchy, bottom-up, precise local statements.** In this scenario, the hierarchy is again central, containing groups such as *Technical Merit* (*Performance*, *Maintainability*, *Interoperability*) and *Socioeconomic Impact* (*Local Employment*, *Supplier Diversity*). Here, the DM prefers a bottom-up approach. Within each group, they provide a clear ordering and precise card placements, calibrating the scale solely via the SRF ratio between the most and least important criterion in that specific set (e.g., "the most important is 5 times the least important"). This necessitates a hierarchical structural module with *bottom-up aggregation*. It is paired with the standard SRF protocol applied locally at each node and an informational module set for precise distances. The expected output is a single, deterministic weight vector derived by propagating the local weights up through the hierarchy.

**Scenario 3: flat structure, normative requirements, robust SRF.** The decision problem is flat, comprising *Accident Risk*, *Compliance*, *Delivery Time*, and *Cost*, but is constrained by strict normative requirements defined by the DM, such as "no criterion should be negligible" and "no single criterion should dominate." The tailored SRF variant employs the standard elicitation procedure for distances but augments the model with informational modules encoding a minimum-weight floor and anti-dominance constraints. Because the DM requires assurance that recommendations hold true under all weights satisfying these legalistic constraints, the analysis relies on the *robust SRF* module. Rather than committing to a single weight vector, the method returns the min–max admissible weights consistent with both the elicited preferences and the normative constraints.

**Scenario 4: hierarchy, confidence-qualified intensities.** A hierarchy is used to evaluate complex dimensions, such as *Patient-Centredness* (*Communication*, *Waiting Time*) and *Clinical Effectiveness* (*Outcomes*, *Safety*). While the DM provides rankings and intensity comparisons, these statements vary in reliability. For example, "*Safety* is more important than *Outcomes* (High Confidence)," versus "*Equipment* is slightly more important than *Staffing* (Low Confidence)." The customized SRF variant integrates a hierarchical structural module with an extended elicitation procedure that records confidence levels alongside card placements. The informational module translates these confidence levels into admissible bounds (narrow for high confidence, wide for low) or soft restrictions. The result is a distributional summary of weights that reflects the certainty of the input, showing tighter concentration around strongly supported preferences.

**Scenario 5: multi-level structure, Monte Carlo simulations.** This scenario involves a deep, multi-level structure (three or more levels), such as an *Infrastructure Appraisal* branching into *Economic*, *Environmental*, and *Social* impacts, which are further decomposed into specific indicators like *Biodiversity* or *Displacement*. The DM accepts imprecise intensity statements (interval blank cards) at various nodes and calibrates the scale only via ratios at selected key nodes. Given the depth of the tree, the goal is not a single vector but an assessment of global stability. The tailored workflow employs the *Monte Carlo simulation* analytical module. By repeatedly sampling admissible parameterizations within the declared bounds and propagating them through the multi-level structure, the framework generates global weight distributions. This enables a comprehensive robustness report that highlights how local uncertainties propagate to downstream scores.

## 5. Discussion, conclusions, and future work

This paper introduces a novel, modular SRF framework that retains the strengths of existing SRF variants while offering enhanced flexibility and adaptability for real-world MCDA applications. By decomposing the standard SRF and its extensions into interchangeable components, we empower analysts and DMs to customize the weight elicitation process to align precisely with their decision context. This modularity is particularly valuable for practitioners seeking personalized solutions, ensuring that the method adapts to specific problem characteristic and organizational preferences without sacrificing its robust decision-making principles.

From an innovation standpoint, the proposed modular framework is the first SRF-based toolkit that simultaneously unifies existing variants in a single constraint system, embeds generic consistency-diagnosis and minimal-restoration procedures, and supports robustness-aware, customized weight-elicitation schemes through a practical questionnaire-guided workflow. Particular theoretical novelty stems from three contributions. First, we show that the main SRF variants are recoverable as special cases of a common constraint system, and that customized procedures constructed by mixing modules remain within a tractable linear (or mixed-integer) class. Second, consistency checking ($E^C$), inconsistency identification ($E^I$), and minimal restoration ($E^R$) are embedded in the same formal layer as elicitation, providing provable diagnostics and least-change repair without leaving the SRF scheme. Third, the framework embeds several complementary robustness analyses to guarantee the stability of the results from customized SRFs.

A key benefit of decomposing the SRF into distinct modules is the enhanced understanding it provides. The modular design clarifies the role and function of each part of the framework, making the overall method more transparent. DMs can examine each component, such as uncertainty modeling, preference aggregation, or solution optimization in isolation, which facilitates more transparent communication and reasoning about the weight elicitation process. This decomposition not only aids in comprehending how the SRF works but also simplifies the task of tailoring the method. Users can identify which parts of the framework need adjustment for a particular situation and then adjust or replace the corresponding module without overhauling the entire approach. In this way, the modular structure serves both as a pedagogical instrument and a practical toolkit that offers user-friendliness and full control over the results, even when advanced weight calculation techniques are employed.

Moreover, the modular framework supports a high degree of model consistency. Each SRF module has well-defined inputs and outputs, and integration of modules is governed by a coherent set of design rules. This means that when components are assembled into a full weight elicitation model, they fit together in a logically consistent manner. The guided selection process, implemented via the questionnaire, reinforces this consistency by recommending only compatible components. By following the questionnaire's guidance, DM avoids combining elements that may be theoretically or practically incompatible, preserving, thereby, the internal coherence and validity of the constructed model. In summary, the modular SRF framework unites the theoretical rigor of robust decision analysis with the practical need for personalized solutions.

Looking ahead, the development of software tools to support the implementation of the modular SRFis an important next step. Such platforms would embed the modular components and the questionnaire, providing an interactive environment for users to configure, validate, and deploy tailored decision models efficiently. By providing a user-friendly interface, software support can significantly lower the barrier to adoption and help ensure a sound, efficient and automated application of the framework. Future research can also investigate recent advances in MCDA, such as modeling interactions among criteria, in the form of synergies or redundancies. In this way, the current SRF modules can be extended to incorporate capacity-based aggregators, such as the Choquet integral (as in (Siskos et al., 2025)). Additionally, expanding the framework to GDM scenarios (Kadziński et al., 2018), through consensus or probabilistic pooling methods (see for instance (Shanian et al., 2008), presents promising avenues, with software tools facilitating inputs from multiple DM and ensuring consistency across diverse perspectives.

## CRediT authorship contribution statement

**River Huang:** Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Data curation, Writing – original draft, Writing – review & editing, Visualization; **Miłosz Kadziński:** Conceptualization, Methodology, Validation, Formal analysis, Investigation, Writing – original draft, Writing – review & editing; **José Rui Figueira:** Conceptualization, Methodology, Validation, Writing – original draft, Writing – review & editing; **Salvatore Corrente:** Conceptualization, Methodology, Validation, Investigation, Writing – review & editing; **Eleftherios Siskos:** Conceptualization, Methodology, Validation, Formal analysis, Writing – review & editing; **Peter Burgherr:** Conceptualization, Writing – review & editing, Funding acquisition.

## Data availability

All data are available from the corresponding author upon reasonable request.

## 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.

## Acknowledgments

The work of River Huang, Eleftherios Siskos and Peter Burgherr, published in this article, was carried out with the support of the Swiss Federal Office of Energy SFOE as part of the SWEET project SURE. The authors bear sole responsibility for the conclusions and results presented in this publication. Miłosz Kadzi ´nski acknowledges support from the PolishMinistry of Education and Science (grant no. 0311/SBAD). The work by José Rui Figueira is financed by Portuguese funds through the FCT - Foundation for Science and Technology, I.P., under project UID/97/2025 (CEGIST). The work of Salvatore Corrente has been funded by the European Union - NextGenerationEU, Mission 4, Component 2, in the framework of the GRINS - Growing Resilient, INclusive and Sustainable project (GRINS PE00000018 - CUP E63C22002120006). The views and opinions expressed are solely those of the authors and do not necessarily reflect those of the European Union, nor can the European Union be held responsible for them.

## Appendix A. Exact models used in the case studies

This appendix reports the exact (non-generic) optimization programs solved for the two case studies. Strict inequalities are implemented as $\geq$ with a small tolerance $\delta = 10^{-6}$. All weights are non-negative and normalized to one; the no-dictatorship ($[OD]$) and minimum-weight ($[OM]$) safeguards are included where indicated.

### *A.1. Case study 1: Flat structure, SRF-II with interval inputs and sampling*

We consider weights $w_1, \dots, w_6$ for criteria $(g_1, \dots, g_6)$. From least to most important, the ranks are

$$R_1 = \{g_4, g_5\}, \quad R_2 = \{g_6\}, \quad R_3 = \{g_2, g_3\}, \quad R_4 = \{g_1\}.$$

The number of blank cards (BC) between successive ranks is interval-valued: $e_{12} = 0$ , $e_{23} \in [0, 1]$ (between $R_1$ and $R_2$) and $e_{34} \in [2, 3]$ (between $R_2$ and $R_3$), and a "zero importance criterion" distance $e_0 \in [0, 1]$ (between the zero benchmark and $R_1$). These form the global ratio:

$$z = \frac{\overline{e}_{12} + \overline{e}_{23} + \overline{e}_{34} + e_0 + 1}{e_0 + 1} \in [3.5,\ 8].$$

The exact feasibility model is:

$$
\left.\begin{array}{ll}
\textit{Equalities within ex aequo subsets:} & w_4 = w_5, \quad w_2 = w_3. \\
\textit{Monotonicity across ranks:} & w_6 \geq w_4 + \delta, \quad w_2 \geq w_6 + \delta, \\
 & w_1 \geq w_2 + \delta. \\
\textit{Largest gap from BC information:} & (w_1 - w_2) \;\geq\; (w_2 - w_6) + \delta, \\
 & (w_1 - w_2) \;\geq\; (w_6 - w_4) + \delta. \\
\textit{Conditional gap between } R_2 \text{ and } R_3: & \Delta_{26} := w_2 - w_6, \quad \Delta_{64} := w_6 - w_4, \\
 & \Delta_{26} - \Delta_{64} \;\geq\; -M\,y, \\
 & \Delta_{26} - \Delta_{64} \;\leq\; M\,y, \\
 & \Delta_{26} - \Delta_{64} \;\geq\; \delta - M\,(1-y), \\
 & y \in \{0,1\}. \\
\textit{SRF–II global ratio (from } e_0 \in [0,1]\textit{):} & w_1 \;\geq\; 3.5\,w_4, \qquad w_1 \;\leq\; 8\,w_4 \\
 & \quad (\text{with } w_4 = w_5). \\
\textit{Normalization:} & \displaystyle\sum_{j=1}^{6} w_j = 1. \\
\textit{No–dictatorship:} & w_k \;\leq\; \displaystyle\sum_{j\neq k} w_j, \qquad k = 1,\dots,6. \\
\textit{Minimum weight:} & w_k \;\geq\; 0.05, \qquad k = 1,\dots,6. \\
\textit{Nonnegativity:} & w_k \;\geq\; 0, \qquad k = 1,\dots,6.
\end{array}\right\} \tag{A.1}
$$

where $w_j$ denotes the weight of criterion $g_j$ for $j = 1,\dots,6$, $\delta > 0$ is a small separation constant, and $M > 0$ is a sufficiently large constant used in the big-$M$ linearization. The equalities $w_4 = w_5$ and $w_2 = w_3$ encode the two ex aequo subsets in the ranking. The monotonicity constraints ensure that weights are non-decreasing along the rank order, with at least a margin $\delta$ between successive ranks. The auxiliary variables $\Delta_{26} = w_2 - w_6$ and $\Delta_{64} = w_6 - w_4$ represent, respectively, the gap between the second and third ranks, and the gap between the third and fourth ranks. The binary variable $y \in 0, 1$ encodes the interval-valued blank-card information between $R_2$ and $R_3$, and linearizes the conditional relation between the gaps: it activates either the case where the gap between $R_2$ and $R_3$ equals the gap between $R_3$ and $R_4$ ($e_{23} = 0$) or the case where it is strictly larger ($e_{23} = 1$).

With this model, feasibility is first verified (consistency check). For exploitation, Model Eq. (A.1) is embedded in a hit-and-run sampler to explore the admissible polytope (10 000 draws), producing the distributions and sample means reported in the paper.

### A.2. Case study 2: Hierarchical structure, MLC with minimal repair

We work directly at the elementary (indicator) level with weights $w_{jj'}$ under four top-level dimensions $g_1$ (Timeliness), $g_2$ (Effectiveness), $g_3$ (Safety), $g_4$ (Efficiency), and impose $\sum_j \sum_{j'} w_{jj'} = 1$. Within each branch, the rank orders and BC counts are:

$$
\begin{aligned}
g_1 &: \; g_{11} \xrightarrow{4\,\mathrm{BC}} g_{13} \xrightarrow{2\,\mathrm{BC}} g_{14} \xrightarrow{2\,\mathrm{BC}} g_{12}, \\
g_2 &: \; g_{21} \xrightarrow{4\,\mathrm{BC}} g_{23} \xrightarrow{2\,\mathrm{BC}} g_{24} \xrightarrow{0\,\mathrm{BC}} g_{22}, \\
g_3 &: \; \{g_{33}, g_{34}, g_{35}, g_{36}\} \xrightarrow{3\,\mathrm{BC}} g_{32} \xrightarrow{1\,\mathrm{BC}} \{g_{31}, g_{37}\}, \\
g_4 &: \; g_{41} \xrightarrow{0\,\mathrm{BC}} g_{42}.
\end{aligned}
$$

Using the MLC scheme, only the least and most important indicators from each branch are compared across branches. After minimal consistency restoration, the inter-group chain (least→most) and BC counts become:

$$
g_{41} \xrightarrow{1\,\mathrm{BC}} \{g_{31}, g_{37}\} \xrightarrow{2\,\mathrm{BC}} g_{21} \xrightarrow{2\,\mathrm{BC}} g_{12},
$$

the global top-to-bottom ratio is fixed at $z = 12$ (so $w_{12} = 12\,w_{41}$). The exact feasibility model is given below:

$$
\left.\begin{array}{l}
(\textit{Timeliness } g_1) \\
w_{13} \geq w_{11} + \delta, w_{14} \geq w_{13} + \delta, w_{12} \geq w_{14} + \delta, \\
(w_{13} - w_{11}) \geq (w_{14} - w_{13}) + \delta, \\
(w_{14} - w_{13}) \geq (w_{12} - w_{14}) + \delta, \\
w_{12} = 6\,w_{11}; \\
(\textit{Effectiveness } g_2) \\
w_{23} \geq w_{21} + \delta, w_{24} \geq w_{23} + \delta, w_{22} \geq w_{24} + \delta, \\
(w_{23} - w_{21}) \geq (w_{24} - w_{23}) + \delta, \\
(w_{24} - w_{23}) \geq (w_{22} - w_{24}) + \delta, \\
w_{22} = 5\,w_{21}; \\
(\textit{Safety } g_3) \\
w_{33} = w_{34} = w_{35} = w_{36}, w_{31} = w_{37} \\
w_{32} \geq w_{33} + \delta, w_{31} \geq w_{32} + \delta, \\
(w_{32} - w_{33}) \geq (w_{31} - w_{32}) + \delta, \\
w_{31} = 6\,w_{33}; \\
(\textit{Efficiency } g_4) \\
w_{42} \geq w_{41} + \delta, \\
w_{42} = 2w_{41}; \\
(\textit{Inter-group MLC, extremes}) \\
w_{33} = w_{34} = w_{35} = w_{36} = w_{42}, w_{31} = w_{37} \\
w_{33} \geq w_{41} + \delta, \\
w_{21} \geq w_{33} + \delta, \\
w_{11} \geq w_{21} + \delta, \\
w_{31} \geq w_{11} + \delta, \\
w_{22} \geq w_{31} + \delta, \\
w_{12} \geq w_{22} + \delta, \\
(w_{22} - w_{31}) \geq (w_{33} - w_{41}) + \delta, \\
(w_{22} - w_{31}) \geq (w_{21} - w_{33}) + \delta, \\
(w_{22} - w_{31}) \geq (w_{11} - w_{21}) + \delta, \\
(w_{22} - w_{31}) \geq (w_{31} - w_{11}) + \delta, \\
(w_{22} - w_{31}) \geq (w_{12} - w_{22}) + \delta, \\
w_{12} = 12w_{41}; \\
(\textit{Normalization and safeguards}) \\
\sum_{j\in\{1,2,3,4\}} \sum_{j'} w_{jj'} = 1, \\
w_{kk'} \leq \sum_{(j,j')\neq(k,k')} w_{jj'} \forall (k, k'), \\
w_{jj'} \geq 0.01 \forall (j, j').
\end{array}\right\} \tag{A.2}
$$

The model is first checked for feasibility. After consistency restoration, we then explored min-max optimization to yield the ranges and means reported in the paper.

## References


Abastante, F., Corrente, S., Greco, S., Lami, I. M., & Mecca, B. (2022). The introduction of the SRF-II method to compare hypothesis of adaptive reuse for an iconic historical building. *Operational Research, 22*(3), 2397–2436.

Bana e Costa, C. A., & Vansnick, J. C. (1994). MACBETH – An interactive path towards the construction of cardinal value functions. *International Transactions in Operational Research, 1*(4), 489–500.

Barron, F. H. (1992). Selecting a best multiattribute alternative with partial information about attribute weights. *Acta Psychologica, 80*(1–3), 91–103.

Belton, V., & Stewart, T. J. (2002). Outranking methods. In *Multiple criteria decision analysis: An integrated approach* (pp. 233–259). Springer.

Brans, J.-P., & De Smet, Y. (2005). PROMETHEE Methods. In *Multiple criteria decision analysis: State of the art surveys* (pp. 187–219). Springer.

Charnes, A., & Cooper, W. W. (1977). Goal programming and multiple objective optimizations. *European Journal of Operational Research, 1*(1), 39–54.

Cinelli, M., Kadziński, M., Gonzalez, M., & Słowiński, R. (2020). How to support the application of multiple criteria decision analysis? let us start with a comprehensive taxonomy. *Omega, 96*, 102261.

Corrente, S., Figueira, J. R., & Greco, S. (2021). Pairwise comparison tables within the deck of cards method in multiple criteria decision aiding. *European Journal of Operational Research, 291*(2), 738–756.

Corrente, S., Figueira, J. R., Greco, S., & Słowiński, R. (2017). A robust ranking method extending ELECTRE III to hierarchy of interacting criteria, imprecise weights and stochastic analysis. *Omega, 73*, 1–17.

Corrente, S., Greco, S., & Słowiński, R. (2016). Multiple criteria hierarchy process for ELECTRE tri methods. *European Journal of Operational Research, 252*(1), 191–203.

Figueira, J., & Roy, B. (2002). Determining the weights of criteria in the ELECTRE type methods with a revised Simos' procedure. *European Journal of Operational Research*, *139*(2), 317–326.
Govindan, K., & Jepsen, M. B. (2016). ELECTRE: A comprehensive literature review on methodologies and applications. *European Journal of Operational Research*, *250*(1), 1–29.
Greco, S., Ehrgott, M., & Figueira, J. R. (Eds.) (2016). Multiple Criteria Decision Analysis: State of the Art Surveys (vol. 233). International Series in Operations Research & Management Science. New York, NY: Springer.
Greco, S., Kadziński, M., Mousseau, V., & Słowiński, R. (2011). ELECTRE$^{GKMS}$: Robust ordinal regression for outranking methods. *European Journal of Operational Research*, *214*(1), 118–135.
Greco, S., Mousseau, V., & Słowiński, R. (2008). Ordinal regression revisited: Multiple criteria ranking using a set of additive value functions. *European Journal of Operational Research*, *191*(2), 415–435.
Greco, S., Mousseau, V., & Słowiński, R. (2010). Multiple criteria sorting with a set of additive value functions. *European Journal of Operational Research*, *207*(3), 1455–1470.
Huang, R., Corrente, S., Siskos, E., & Burgherr, P. (2025). The multi-level cards method: Weights elicitation, consistency checking and restoration for hierarchical criteria. *Omega*, (p. 103438).
Jacquet-Lagrèze, E., & Siskos, J. (1982). Assessing a set of additive utility functions for multicriteria decision-Making: The UTA method. *European Journal of Operational Research*, *10*(2), 151–164.
Kadziński, M. (2022). Robust ordinal regression for multiple criteria decision aiding. In S. Greco, V. Mousseau, J. Stefanowski, & C. Zopounidis (Eds.), *Intelligent Decision Support Systems : Combining Operations Research and Artificial Intelligence - Essays in Honor of Roman Słowiński* , pp. 185–205). Cham: Springer International Publishing.
Kadziński, M., Rocchi, L., Miebs, G., Grohmann, D., Menconi, M. E., & Paolotti, L. (2018). Multiple criteria assessment of insulating materials with a group decision framework incorporating outranking preference model and characteristic class profiles. *Group Decision and Negotiation*, *27*(1), 33–59.
Kadziński, M., & Tervonen, T. (2013). Robust multi-criteria ranking with additive value models and holistic pair-wise preference statements. *European Journal of Operational Research*, *228*(1), 169–180.
Lahdelma, R., Hokkanen, J., & Salminen, P. (1998). SMAA - Stochastic multiobjective acceptability analysis. *European Journal of Operational Research*, *106*(1), 137–143.
Mousseau, V., Figueira, J., Dias, L., da Silva, C. G., & Climaco, J. (2003). Resolving inconsistencies among constraints on the parameters of an MCDA model. *European Journal of Operational Research*, *147*(1), 72–93.
Rezaei, J. (2015). Best-worst multi-criteria decision-making method. *Omega*, *53*, 49–57.
Riabacke, M., Danielson, M., & Ekenberg, L. (2012). State-of-the-art prescriptive criteria weight elicitation. *Advances in Decision Sciences*.
Ricardo, A. V., Costa, A. S., & Figueira, J. R. (2024). Enhancing hospital benchmarking: A multidimensional approach to quality assessment in portugal. *Socio-Economic Planning Sciences*, *93*, 101899.
Rodriguez, R. M., Martinez, L., & Herrera, F. (2011). Hesitant fuzzy linguistic term sets for decision making. *IEEE Transactions on Fuzzy Systems*, *20*(1), 109–119.
Roy, B. (2010). Robustness in operational research and decision aiding: A multi-faceted issue. *European Journal of Operational Research*, *200*(3), 629–638.
Roy, B., & Mousseau, V. (1996). A theoretical framework for analysing the notion of relative importance of criteria. *Journal of Multi-Criteria Decision Analysis*, *5*(2), 145–159.
Salo, A., & Hämäläinen, R. P. (2004). Preference Programming. Technical Report Systems Analysis Laboratory, Helsinki University of Technology Helsinki, Finland. Working paper.
Shanian, A., Milani, A. S., Carson, C., & Abeyaratne, R. C. (2008). A new application of ELECTRE III and revised simos' procedure for group material selection under weighting uncertainty. *Knowledge-Based Systems*, *21*(7), 709–720.
Simos, J. (1990). Evaluer l'impact sur l'environnement: Une approche originale par l'analyse multicritère et la négociation (vol. 1015). Presses polytechniques et universitaires romandes Lausanne.
Siskos, E., Desbordes, A., Burgherr, P., & McKenna, R. (2025). Integrated assessment of a robust choquet integral preference model for efficient multicriteria decision support. *European Journal of Operational Research*.
Siskos, E., Kourousias, G., & Siskos, Y. (2022). Bipolar multicriteria aggregation-disaggregation robustness approach: Theory and application on european e-government benchmarking. In *Multiple criteria decision making: Techniques, analysis and applications* (pp. 121–152). Springer.
Siskos, E., & Tsotsolas, N. (2015). Elicitation of criteria importance weights through the simos method: A robustness concern. *European Journal of Operational Research*, *246*(2), 543–553.
Tsotsolas, N., Spyridakos, A., Siskos, E., & Salmon, I. (2019). Criteria weights assessment through prioritizations (WAP) using linear programming techniques and visualizations. *Operational Research*, *19*, 135–150.
Wu, H., Ren, P., & Xu, Z. (2022). Promoting the physician–patient consensus with a hesitant fuzzy linguistic consensus method based on betweenness relation. *Applied Soft Computing*, *124*, 108979.
Xidonas, P., Thomakos, D., & Samitas, A. (2025). On the integration of multiple criteria decision aiding and forecasting: Does it create value in portfolio selection? *European Journal of Operational Research*, *321*(2), 516–528.
Zeleny, M. (1973). Compromise programming. In J. L. Cochrane, & M. Zeleny (Eds.), *Multiple criteria decision making* (pp. 262–301). Columbia, SC: University of South Carolina Press.
Zeleny, M. (1974). A concept of compromise solutions and the method of the displaced ideal. *Computers & Operations Research*, *1*(3–4), 479–496.
Zhang, Z., & Liao, H. (2023). An evidential reasoning-based stochastic multi-attribute acceptability analysis method for uncertain and heterogeneous multi-attribute reverse auction. *Journal of the Operational Research Society*, *74*(1), 239–257.