Annals of Operations Research (2021) 306:173–207
https://doi.org/10.1007/s10479-020-03833-z

S.I.: REGRESSION METHODS BASED ON OR TECHNIQUES

Enriched preference modeling and robustness analysis for the ELECTRE Tri-B method

Miłosz Kadziński[1] · **Magdalena Martyn**[1]

Accepted: 13 October 2020 / Published online: 27 October 2020
© The Author(s) 2020

Abstract
We consider multiple criteria sorting problems with preference-ordered classes delimited by a set of boundary profiles. While significantly extending the ELECTRE Tri-B method, we present an integrated framework for modeling indirect preference information and conducting robustness analysis. We allow the Decision Maker (DM) to provide the following three types of holistic judgments: assignment examples, assignment-based pairwise comparisons, and desired class cardinalities. A diversity of recommendation that can be obtained given the plurality of outranking-based sorting models compatible with the DM's preferences is quantified by means of six types of results. These include possible assignments, class acceptability indices, necessary assignment-based preference relation, assignment-based outranking indices, extreme class cardinalities, and class cardinality indices. We discuss the impact of preference information on the derived outcomes, the interrelations between the exact results computed with mathematical programming and stochastic indices estimated with the Monte Carlo simulations, and new measures for quantifying the robustness of results. The practical usefulness of the approach is illustrated on data from the Financial Times concerning MBA programs.

Keywords Multiple criteria decision aiding · Multiple criteria sorting · Outranking relation · Preference modeling · Robustness analysis · Monte Carlo simulation

1 Introduction

Multiple criteria sorting problems involve an assignment of alternatives to pre-defined and ordered classes in the presence of numerous pertinent viewpoints. The definition of decision classes is related to how the alternatives placed in a given class should be processed or treated. For example, in portfolio decision analysis, the stocks can be assigned to three categories: attractive, to be studied further, and non-attractive (Hurson and Zopounidis 1995). When

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s10479-020-03833-z) contains supplementary material, which is available to authorized users.

✉ Miłosz Kadziński
milosz.kadzinski@cs.put.poznan.pl

1 Institute of Computing Science, Poznań University of Technology, Piotrowo 2, 60-965 Poznań, Poland

evaluating software producing technology companies, the latter ones can be sorted into predefined efficiency classes (Bilich and da Silva 2008). Furthermore, the assessment of natural gas pipelines involves assigning different accident scenarios into risk categories (Brito et al. 2010). In nanotechnology, one can predict a precaution level while handling nanomaterials in certain conditions (Kadziński et al. 2020). As far as medical diagnosis is concerned, patients can be judged in terms of the risk levels of developing different illnesses (e.g., diabetic retinopathy Saleh et al. 2018). In the context of a comprehensive assessment of insulating materials, we can consider preference-ordered sustainability classes (Kadziński et al. 2018).

The field of Multiple Criteria Decision Aiding (MCDA) offers a plethora of methods that support Decision Makers (DMs) in providing recommendations regarding sorting problems. These approaches differ in terms of the underlying assumptions, contexts of application, required preference information, incorporated models, and applied sorting rules. Nonetheless, all of them expect the DM to interact in the specification of a consistent family of criteria, a set of alternatives and their performances, as well as a set of parameters that represent his/her preferences.

One of the most prevailing sorting methods is ELECTRE Tri-B (Yu 1992), which incorporates an outranking-based preference model (Figueira et al. 2005) and defines decision classes by means of limiting profiles. Such profiles are composed of performances that can be seen as natural boundaries between a pair of successive classes. The intuitiveness of ELECTRE Tri-B was found appealing in numerous real-world case studies in such various application areas as climate change (Diakoulaki and Hontou 2003), economy and finance (Dimitras et al. 1995), land-use suitability assessment (Joerin et al. 2001), skills accreditation system (Siskos et al. 2007), water resources management (Raju et al. 2000), or zoning risk analysis (Merad et al. 2004).

The basic variant of ELECTRE Tri-B has also been advanced in numerous ways with the aim of extending its applicability. In particular, it was extended to account for both a hierarchical structure of criteria and interaction effects (Corrente et al. 2016). Moreover, it was generalized to admit both a set of limiting class profiles for better characterization of the boundaries between categories (Fernandez et al. 2017) and interval data (Fernandez et al. 2019). Also, the requirement of defining class boundaries was alleviated by admitting the use of characteristic (i.e., the most typical) profiles (Almeida-Dias et al. 2010) or defining the class boundaries implicitly by means of example assignment examples (Köksalan et al. 2009). Nevertheless, the greatest number of methodological advancements of ELECTRE Tri-B was related to the elicitation of parameters compatible with the DM's preferences. Since the DMs may have some difficulties in specifying directly a coherent set of such parameters (involving criteria weights, credibility level, comparison thresholds, and limiting profiles), a preference disaggregation approach was envisaged to deal with the DM's holistic preference information and indirect inference of parameters.

The pioneering preference disaggregation approach for ELECTRE Tri-B was provided in Mousseau and Słowiński (1998). The authors proposed an interactive non-linear optimization model to infer all parameters of the sorting method from the desired assignments for a subset of reference alternatives relatively well-known to the DM. Then, the linear counterpart of this model was introduced to derive only the criteria weights and credibility level (Mousseau et al. 2001). On the contrary, Mixed Integer Linear Programming was used in Ngo The and Mousseau (2002) to learn the category limits. Furthermore, the definition of the credibility of an outranking relation was revised in Dias and Mousseau (2006) for allowing easier indirect inference of veto-related parameters. A group decision preference disaggregation setting was considered in Damart et al. (2007) with the aim of preserving consistency of sorting examples at both individual and collective levels. Moreover, when considering

assignment examples of multiple DMs, (Cailloux et al. 2012) discussed how to infer collective class boundaries. While all aforementioned works considered the pessimistic sorting rule of ELECTRE Tri-B, (Zheng et al. 2014) studied the efficiency of inference procedures in the context of an optimistic procedure. Finally, the algorithms for dealing with inconsistent holistic preference judgments were proposed in Mousseau et al. (2003) and Mousseau et al. (2006).

Apart from inferring a set of compatible parameters from the DM's assignment examples, an inherent part of the preference disaggregation paradigm consists in using such parameters through a sorting model to derive well-justified decisions. Traditionally, the analysis of the DM's holistic decisions resulted in constructing a single set of compatible parameter values, whose employment led to the univocal class assignments of decision alternatives. However, when using indirect and/or imprecise preference information, there usually exist infinitely many compatible sets of parameters. Although their application on the set of reference alternatives leads to reproducing the DM's desired assignments, the sorting recommendation derived for the non-reference alternatives which were not judged by the DM may differ from one set of parameters to another.

In this perspective, (Dias et al. 2002) combined a preference disaggregation approach with robustness analysis providing for each alternative a set of classes confirmed by at least one compatible set of parameters. In this case, the variables of proposed mathematical programming models were the criteria weights and credibility level. Furthermore, while admitting imprecision in the specification of admissible parameter values related to the formulation of sorting model (including the boundary profiles), (Tervonen et al. 2007) and (Tervonen et al. 2009) adapted Stochastic Multicriteria Acceptability Analysis (SMAA) to ELECTRE Tri-B. Specifically, the SMAA-TRI method incorporated the simulation techniques to estimate a proportion of feasible sets of parameters that assign each alternative to a given class.

This paper advances ELECTRE Tri-B by proposing an enriched framework for preference modeling and robustness analysis. When it comes to preference learning, we account for three types of indirect and imprecise preference information. They refer to a small part of a desired final recommendation or requirements to be imposed on the derived assignments. Specifically, apart from the assignment examples which have been traditionally incorporated into preference disaggregation approaches for ELECTRE Tri-B (see Mousseau and Słowiński 1998; Dias et al. 2002), we consider assignment-based pairwise comparisons (Kadziński et al. 2015) and desired class cardinalities (Kadziński and Słowiński 2013). The former refer to the relative class assignments desired for pairs of alternatives without specifying any concrete classes (e.g., alternative a should be assigned to a class better than alternative b), whereas the latter specify the bounds on the number of alternatives that can be assigned to a particular class (e.g., at least 5 alternatives should be assigned to the most preferred class, and at most half of the alternatives can be sorted into the worst class). While pairwise comparisons are strictly related to the DM's preferences, the constraints on the number of alternatives that can be assigned to each class are rather implied by a particular decision aiding context. For all types of preference information, we discuss the mathematical programming models that contribute to the definition of all compatible sets of parameters. Accounting for such diverse forms of preference is beneficial in terms of offering greater flexibility to the DM, reducing the space of feasible parameter sets, and making the proposed framework suitable for various sorting contexts.

As far as robustness analysis is concerned, we incorporate a number of robust results, which are derived from the exploitation of a set of parameter sets compatible with the DM's preferences. The diversity of these outcomes comes from using different tools for assessing the robustness of the same part of the recommendation and from considering various per-

spectives on the stability of recommendation for a given problem. Specifically, we jointly use linear programming—for constructing the necessary (certain), possible, and extreme outcomes in an exact manner—and the Monte Carlo simulation—for estimating the probability of results, which are possible, though not certain. Indeed, some recent case studies indicate that a joint consideration of the robust and stochastic results can be even more beneficial for real-world decision making, as these results offer complementary perspectives on the stability of recommendation (see Dias et al. 2018; Kadziński et al. 2018).

The constructed results differ in terms of an adopted perspective while representing multidimensional and interrelated conclusions (Kadziński and Ciomek 2016). For the individual alternatives, we consider possible assignments and class acceptability indices, for pairs of alternatives—necessary assignment-based preference relation (Kadziński and Słowiński 2013) and assignment-based outranking indices, whereas for decision classes—extreme class cardinalities and class cardinality indices (Kadziński et al. 2016). Note that the so-far existing robustness analysis methods for ELECTRE Tri-B focussed on the stability of class assignments. Accounting for various types of results, including assignment-based preference relation and class cardinalities, is beneficial for understanding one's own preferences, informing decision making, enhancing the DM's learning on the problem, stimulating his/her reaction to interactively enrich preference information, and arriving at a more credible recommendation.

When considering different types of preference information, the space of compatible sets of parameters is non-convex. Thus, to derive the exact robust results, we propose some Mixed-Integer Linear Programming models, whereas to estimate the acceptability indices, we combine Hit-And-Run (HAR) (Tervonen et al. 2013) with the rejection sampling technique. The additional contribution of the paper comes from discussing an impact of the provided preference information on the truth or falsity of robust results as well as the interrelations between the exact and stochastic outcomes obtained with mathematical programming and the Monte Carlo simulations. The practical usefulness of the approach is illustrated on data from the Financial Times concerning MBA programs.

Overall, the novelty of the paper derives from the following developments:

- introducing original models based on mathematical programming for defining a set of outranking models compatible with three types of indirect and imprecise preference information concerning alternatives, pairwise comparisons, and decision classes; in particular, we propose two alternative models for representing assignment-based pairwise comparisons, which have not been used before in the context of ELECTRE Tri-B;
- introducing novel procedures based on mathematical programming and the Monte Carlo simulation for exploiting a set of compatible outranking models and deriving six types of complementary results; the exact and stochastic outcomes concerning assignment-based preference relations and class cardinalities are considered for the first time in view of ELECTRE Tri-B;
- implementing an original idea for combining Hit-And-Run and the rejection sampling techniques for efficient exploitation of a non-convex space of compatible outranking models;
- comprehensive discussion of the theoretical properties of the results and interrelations between inputs and outputs which enhance interactive elicitation of preferences;
- providing the richest framework for preference modeling and robustness analysis, which can be generalized to other sorting methods;
- introducing novel measures for quantifying the robustness of sorting results.

The remainder of the paper is organized as follows. In Sect. 2, we remind the basic variant of ELECTRE Tri-B. Section 3 introduces an integrated framework for preference modeling and robustness analysis with ELECTRE Tri-B. In Sect. 4, we discuss the properties of the model outcomes, which are relevant for real-world applications. Section 5 demonstrates the method's applicability on a realistic data set. The last section concludes and provides avenues for future research.

2 Reminder on the ELECTRE Tri-B method

We use the following notation throughout the paper:

- $A = \{a_1, a_2, \ldots, a_i, \ldots, a_n\}$—a finite set of n alternatives;
- $A^R = \{a_1^*, a_2^* \ldots\}$—a set of reference alternatives on which the DM accepts to express preferences; $A^R \subseteq A$;
- $C_1, C_2, \ldots, C_p$—p pre-defined preference ordered classes, where C_{h+1} is preferred to C_h, $h = 1, \ldots, p-1$ ($H = \{1, .., p\}$);
- $B = \{b_0, \ldots, b_h, \ldots, b_p\}$—a set of limiting class profiles such that b_{h-1} and b_h are, respectively, the lower and upper boundaries of class C_h, $h = 1, \ldots, p$;
- $G = \{g_1, \ldots, g_j, \ldots, g_m\}$—a set of m evaluation criteria, $g_j : a \rightarrow \mathbb{R}$ for all $j \in J = \{1, 2, \ldots, m\}$; performances of $a_i \in A$ or $b_h \in B$ on g_j are denoted by, respectively, $g_j(a_i)$ and $g_j(b_h)$; if different is not stated explicitly, we assume that all criteria are of gain type (i.e., greater performances are more preferred);
- $q_j(a_i)$, $p_j(a_i)$, $v_j(a_i)$ for $j = 1, 2, \ldots, m$ and $a_i \in A$—indifference, preference, and veto thresholds for criterion g_j and alternative a_i such that $v_j(a_i) > p_j(a) \geq q_j(a) \geq 0$;
- w_j for $j = 1, 2, \ldots, m$—weight of criterion g_j, expressing its relative power in set G; the weights are non-negative $w_j \geq 0$, for $j = 1, \ldots, m$, and normalized to sum up to one, i.e., $\sum_{j=1}^{m} w_j = 1$;
- $\lambda \in [0.5, 1]$—a credibility level indicating a minimal value of outranking credibility $\sigma(a, b_h)$ implying the truth of a crisp outranking relation S of alternative a_i over profile b_h ($a_i S b_h$) (since $\lambda \geq 0.5$, the support of a weighted majority of all criteria is always required to instantiate $a_i S b_h$).

Construction of outranking relation ELECTRE Tri-B supports the assignment of each alternative $a_i \in A$ to one of decision classes in C. For this purpose, the method constructs an outranking relation S, which verifies whether a_i is at least as good as boundary profile b_h, $h = 0, \ldots, p$. The truth of $a_i S b_h$ is captured by means of the concordance and discordance tests. The former involves computation of a comprehensive concordance index $C(a_i, b_h)$, which quantifies the strength of a coalition of criteria confirming that a_i is at least as good as b_h, i.e.:

$$C(a_i, b_h) = \sum_{j=1}^{m} w_j \varphi_j(a_i, b_h), \tag{1}$$

where $\varphi_j(a_i, b_h)$ is the marginal concordance index indicating a degree to which criterion g_j supports the hypothesis $a_i S b_h$:

$$\varphi_j(a_i, b_h) = \begin{cases} 0, & \text{if } g_j(b_h) - g_j(a_i) \geq p_j(a_i), \\ 1, & \text{if } g_j(b_h) - g_j(a_i) \leq q_j(a_i), \\ \frac{p_j(a_i) - (g_j(b_h) - g_j(a_i))}{p_j(a_i) - q_j(a_i)}, & \text{if } q_j(a_i) < g_j(b_h) - g_j(a_i) < p_j(a_i). \end{cases} \tag{2}$$

In particular, if a_i is not worse than b_h by more than indifference threshold $q_j(a_i)$ on all criteria, then $C(a_i, b_h) = 1$, which means that the hypothesis $a_i Sb_h$ is fully supported. On the contrary, if a_i is worse than b_h by at least preference threshold $p_j(a_i)$ for $j = 1, \ldots, m$, then $C(a_i, b_h) = 0$, indicating an absolute lack of support for validating $a_i Sb_h$.

The discordance index measures the strength of opposition against the truth of outranking on each criterion g_j, $j = 1, \ldots, m$, in the following way:

$$d_j(a_i, b_h) = \begin{cases} 0, & \text{if } g_j(b_h) - g_j(a_i) \leq p_j(a_i), \\ 1, & \text{if } g_j(b_h) - g_j(a_i) \geq v_j(a_i), \\ \frac{(g_j(b_h) - g_j(a_i)) - p_j(a_i)]}{[v_j(a_i) - p_j(a_i)}, & \text{if } p_j(a_i) < g_j(b_h) - g_j(a_i) < v_j(a_i). \end{cases} \tag{3}$$

Specifically, when a_i is worse than b_h on g_j by at least veto threshold $v_j(a_i)$, criterion g_j strongly opposes to $a_i Sb_h$, hence invalidating this relation.

The credibility of outranking aggregates the concordance and discordance indices to measure an overall support given to $a_i Sb_h$ by all criteria:

$$\sigma(a_i, b_h) = C(a_i, b_h)[1 - max_j d_j(a_i, b_h)]. \tag{4}$$

We apply a revised credibility index proposed in Dias and Mousseau (2006) so that to account for the veto-related effects within the preference disaggregation procedure. Note that when using $\sigma(a_i, b_h)$ originally coupled with ELECTRE Tri-B, the inference models would become non-linear.

The truth of $a_i Sb_h$ is confirmed iff $\sigma(a_i, b_h)$ is not lesser than credibility level λ, i.e.:

$$a_i Sb_h \Leftrightarrow \sigma(a_i, b_h) \geq \lambda. \tag{5}$$

Otherwise, i.e., if $\sigma(a_i, b_h) < \lambda$, a_i cannot be judged comprehensively at least as good as b_h ($a_i S^c b_h$).

Exploitation of outranking relation The assignment procedures of ELECTRE Tri-B exploit the outranking relation for all pairs of $(a_i, b_h) \in A \times B$. Specifically, a pair of disjoint sorting rules can be applied. The optimistic one assigns a_i to class C_h in case b_{h+1} is the least preferred profile such that $a_i S^c b_{h+1}$ and $b_{h+1} Sa_i$. In this case, b_{h+1} is strictly preferred to a_i, thus preventing its assignment to a class better than C_h. The pessimistic rule assigns a_i to class C_h in case b_h is the most preferred profile such that $a_i Sb_h$. Then, a_i is at least as good as b_h, which hampers its assignment to a class worse than C_h. Throughout the paper, we will use the pessimistic rule.

3 Integrated framework for preference modeling and robustness analysis with ELECTRE Tri-B

In this section, we extend the ELECTRE Tri-B method by providing the mathematical programming models that allow the DM to provide three types of indirect and imprecise preference information. We also discuss the algorithms for deriving six types of results, which contribute to the multi-dimensional robustness analysis.

3.1 Preference modeling

ELECTRE Tri-B employs an outranking-based sorting model. We expect the DM to provide the indifference q_j, preference p_j, and veto v_j thresholds for each criterion. These thresholds

can be defined as constants or affine functions. The latter ones make the threshold values dependent on the performance of particular alternatives. Moreover, the DM is asked to specify directly the boundary profiles $B = \{b_1 \ldots, b_h, \ldots, b_{p-1}\}$, which serve as frontiers between classes. The extreme profiles b_0 and b_p are composed of, respectively, the worst and the best performances on all criteria. Overall, we assume the variables in our model are criteria weights w_j, $j = 1, \ldots, m$, and credibility level λ. Hence, the basic set of constraints can be formulated as follows:

$$\left.\begin{array}{l} [B_1]\ w_j \geq 0,\ \ j = 1, \ldots, m, \\ [B_2]\ \sum_{j=1}^{m} w_j = 1, \\ [B_3]\ 0.5 \leq \lambda \leq 1. \end{array}\right\} E^{BASE} \tag{6}$$

The admissible values of weights and credibility level are delimited by the indirect and imprecise preference information. The use of such statements reduces the cognitive effort on the part of DM. In what follows, we present the mathematical programming models that allow to incorporate the following three types of preferences: assignment examples, assignment-based pairwise comparisons, and desired class cardinalities. These models involve binary variables, which are subsequently used to minimize the number of preference information pieces that need to be removed to ensure consistency between the DM's preference information and an assumed outranking-based sorting model of ELECTRE Tri-B.

Assignment examples Assignment example consists of a reference alternative $a_i^* \in A^R \subseteq A$ and its desired assignment $a_i^* \rightarrow [C_{L^{DM}(a_i^*)}, C_{U^{DM}(a_i^*)}]$, with $C_{L^{DM}(a_i^*)}$ and $C_{U^{DM}(a_i^*)}$ being, respectively, the least and the most preferred classes admissible for a_i^*. The example statements of this type are as follows: "a_1^* needs to be assigned to at least class C_3 and at most class C_5" or "a_2^* should be assigned to C_1". The following constraint set allows to model $a_i^* \rightarrow [C_{L^{DM}(a_i^*)}, C_{U^{DM}(a_i^*)}]$ in terms of the parameters incorporated by the pessimistic rule of ELECTRE Tri-B:

$$\left.\begin{array}{l} \text{for all } a_i^* \in A^R: \\ [AE_1]\ \sigma(a_i^*, b_{L^{DM}-1}) + Mu(a_i^*) \geq \lambda, \\ [AE_2]\ \sigma(a_i^*, b_{U^{DM}}) + \varepsilon \leq \lambda + Mu(a_i^*), \\ [AE_3]\ u(a_i^*) \in \{0, 1\}, \end{array}\right\} E^{ASS-EX} \tag{7}$$

where M and ε are, respectively, arbitrarily large and small positive constants (this interpretation is valid also for the constraints sets that follow). In case $u(a_i^*) = 0$, constraint $[AE_1]$ stands for assigning a_i^* to class at least $C_{L^{DM}(a_i^*)}$, whereas constraint $[AE_2]$ implies that a^* will be assigned to class at most $C_{U^{DM}(a_i^*)}$. In case $u(a_i^*) = 1$, the set of constraints is always satisfied, whichever the values of all remaining parameters.

Assignment-based pairwise comparisons Assignment-based pairwise comparison involves specification of a desired difference between classes to which a pair of reference alternatives $(a_i^*, a_j^*) \in P^R \subseteq A^R \times A^R$ should be assigned. We distinguish the following two sub-types of such comparisons:

– Alternative a_i^* is better than alternative a_j^* by at least $k \geq 0$ classes ($a_i^* \succ^{\rightarrow}_{\geq k, DM} a_j^*$), e.g., "$a_1^*$ is better than a_3^* by at least 2 classes". The following constraint set translates $a_i^* \succ^{\rightarrow}_{\geq k, DM} a_j^*$ to the parameters of an assumed sorting model:

$$\left.\begin{array}{l}\text{for all } a_i^*, a_j^* \in A^R : a_i^* \succ^{\rightarrow}_{\geq k, DM} a_j^* : \\ \quad \text{for all } h = 1, \ldots, p-k : \\ \quad [PL_1]\ \sigma(a_i^*, b_{h-1+k}) \geq \lambda - M(1 - u(a_i^*, a_j^*, \geq k, h)), \\ \quad [PL_2]\ \sigma(a_j^*, b_h) + \varepsilon \leq \lambda + M(1 - u(a_i^*, a_j^*, \geq k, h)), \\ \quad [PL_3]\ \sum_{h=1}^{p-k} u(a_i^*, a_j^*, \geq k, h) \geq 1 - Mu(a_i^*, a_j^*), \\ \quad [PL_4]\ u(a_i^*, a_j^*, \geq k, h) \in \{0, 1\}, h = 1, \ldots, p-k, \\ [PL_5]\ u(a_i^*, a_j^*) \in \{0, 1\}. \end{array}\right\} E^{PCL} \tag{8}$$

In case $u(a_i^*, a_j^*) = 1$, all constraints are satisfied, being eliminated. Otherwise, if $u(a_i^*, a_j^*, \geq k, h) = 1$, constraint $[PL_1]$ guarantees that a_i^* is assigned to class at least C_{h+k}, whereas constraint $[PL_2]$ ensures that a_j^* is placed in class at most C_h. Constraint $[PL_3]$ imposes that at least one variable $u(a_i^*, a_j^*, \geq k, h)$, $h = 1, \ldots, p-k$, is equal to one, implying the occurrence of the above mentioned scenario for some h.

– Alternative a_i^* is better than alternative a_j^* by at most $l \geq 0$ classes, ($a_i^* \succ^{\rightarrow}_{\leq l, DM} a_j^*$), e.g., "$a_1^*$ is by at most 1 class better than a_3^*". The following constraint set translates $a_i^* \succ^{\rightarrow}_{\leq l, DM} a_j^*$ in terms of the parameters of an assumed sorting model:

$$\left.\begin{array}{l}\text{for all } a_i^*, a_j^* \in A^R : a_i^* \succ^{\rightarrow}_{\leq l, DM} a_j^* : \\ \quad \text{for all } h = 1, \ldots, p-l : \\ \quad [PU_1]\ \sigma(a_i^*, b_{h+l}) + \epsilon \leq \lambda + M(1 - u(a_i^*, a_j^*, \leq l, h)), \\ \quad [PU_2]\ \sigma(a_j^*, b_h) \geq \lambda - M(1 - u(a_i^*, a_j^*, \leq l, h)), \\ \quad [PU_3]\ \sum_{h=1}^{p-l} u(a_i^*, a_j^*, \leq l, h) \geq 1 - Mu(a_i^*, a_j^*), \\ \quad [PU_4]\ u(a_i^*, a_j^*, \leq l, h) \in \{0, 1\}, h = 1, \ldots, p-l, \\ [PU_5]\ u(a_i^*, a_j^*) \in \{0, 1\}. \end{array}\right\} E^{PCU} \tag{9}$$

In case $u(a_i^*, a_j^*) = 1$, all constraints are satisfied, irrespective of the remaining parameter values. Otherwise, if $u(a_i^*, a_j^*, \leq l, h) = 1$, constraint $[PU_1]$ guarantees that a_i^* is assigned to class at most C_{h+l}, whereas constraint $[PU_2]$ ensures that a_j^* is placed in class at least C_h. Constraint $[PU_3]$ imposes that at least one variable $u(a_i^*, a_j^*, \leq l, h)$, $h = 1, \ldots, p-k$, is equal to one.

In case $a_i^* \succ^{\rightarrow}_{\geq 0, DM} a_j^*$ and $a_i^* \succ^{\rightarrow}_{\leq 0, DM} a_j^*$, alternatives a_i^* and a_j^* should be assigned to the same class.

Desired class cardinalities In classical sorting problems, each alternative is comprehensively judged in terms of its intrinsic value through its comparison with some norms or references (in the case of ELECTRE Tri-B, these are defined by a set of boundary profiles separating the classes) (see Almeida-Dias 2010 and Dias et al. 2003). Under such a scenario, sorting incorporates an absolute judgment of each alternative to be assigned. As a result, depending on the performances of alternatives on particular criteria, some classes may remain empty, whereas other classes may accommodate the vast majority of alternatives. However, the context of some decision problems may suggest that the number of alternatives assigned to some or all classes should be bounded. The example problems involving this type of constraints have been discussed in, e.g., (Almeida-Dias 2010; Dias et al. 2003; Kadziński and Słowiński 2013; b). They concern credit risk assessment, assignment of bonus packages, the evaluation of the performance of retails or R&D projects, accreditation of qualification and skills, establishing national priorities in the energy sector, or student admission programs. Such requirements may be linked to both problem's nature and the DM's preferences. A variety of these examples confirms the need for methods that tolerate the lower and/or upper bounds on the category sizes.

In this regard, desired class cardinalities indicate extreme numbers of alternatives $C^{min}_{h,DM}$ and $C^{max}_{h,DM}$ which can be assigned to a given class C_h. The example preference information pieces of this kind are as follows: "at least 3 and at most 7 alternatives can be assigned to C_1" or "class C_2 needs to accommodate at least 5 alternatives". The respective mathematical programming model can be formulated as follows:

$$\left.\begin{array}{l}\text{for all } a_i \in A: \\ [CC_0]\ \sum_{h=1}^{p} u(a_i,h) = 1, \\ \text{for all } h \in H \text{ with specified desired class cardinalities}: \\ [CC_1]\ \sum_{a_i \in A} u(a_i,h) \geq C^{min}_{h,DM} - Mu(C_h), \\ [CC_2]\ \sum_{a_i \in A} u(a_i,h) \leq C^{max}_{h,DM} + Mu(C_h), \\ [CC_3]\ u(a_i,h) \in \{0,1\},\ \text{for all } a_i \in A, \\ [CC_4]\ u(C_h) \in \{0,1\},\end{array}\right\} E^{CARD} \tag{10}$$

where $u(a_i, h)$ is equal to one in case $a_i \in A$ is assigned to class C_h (similarly as imposed in constraint set E^{ASS-EX}). In case $u(C_h) = 1$, all constraints are satisfied, being eliminated. Otherwise, constraints $[CC_1]$ and $[CC_2]$ ensure that there are, respectively, at least C^{min}_h and at most C^{max}_h alternatives assigned to C_h. Obviously, for a given outranking model each alternative can be assigned to a single class only (see constraint $[CC_0]$).

Note that the use of desired class cardinalities makes sense when the applied sorting model parameters are not defined precisely. Otherwise, i.e., if ELECTRE Tri-B is parameterized with precise parameter values, the method would assign each alternative to some class in a deterministic way. Then, the class cardinalities would simply correspond to the number of alternatives accommodated in each class. Whether they would meet the pre-defined constraints cannot be controlled by the method. On the contrary, when the model parameters are defined through indirect or imprecise preference information (such as assignment examples or assignment-based pairwise comparisons), the inclusion of the requirements on the category size has an impact on which parameter sets can be used for deriving a sorting recommendation and which should be neglected, because of not being compatible with the DM's requirements of all types.

As a result, in constrained sorting problems, an assignment does not depend solely on the intrinsic value of a given alternative, but also on the assignment of remaining alternatives (Almeida-Dias 2010). Thus, a partial dependence in the considered set of alternatives occurs, incorporating relative comparisons, typical for choice and ranking, into sorting. Indeed, even if the assignment of some alternative to a given class would be possible without cardinality constraints, their incorporation may exclude such possibility in favor of the other possible assignments. This is, however, consistent with the specificity of thus defined sorting problems.

Once the desired class cardinalities are specified and the binary variables $u(a_i, h)$ are incorporated into the model, the assignment-based preference relations can be modeled in a different, more transparent way. Specifically, $a^*_i \succ^{\rightarrow}_{\geq k,DM} a^*_j$ can be translated into the parameters of an assumed sorting model in the following way:

$$\left.\begin{array}{l}\text{for all } a^*_i, a^*_j \in A^R : a^*_i \succ^{\rightarrow}_{\geq k,DM} a^*_j: \\ [PL_1] \sum_{h=1}^{p} h \cdot u(a^*_i,h) \geq \sum_{h=1}^{p} h \cdot u(a^*_j,h) + k - Mu(a^*_i,a^*_j); \\ [PL_2]\ u(a^*_i,a^*_j) \in \{0,1\}.\end{array}\right\} E^{PCL} \tag{11}$$

In turn, $a_i^* \succ^{\rightarrow}_{\leq l,DM} a_j^*$ can be modeled as follows:

$$\left.\begin{array}{l}\text{for all } a_i^*, a_j^* \in A^R : a_i^* \succ^{\rightarrow}_{\leq l,DM} a_j^* : \\ [PU_1] \sum_{h=1}^{p} h \cdot u(a_i^*, h) \leq \sum_{h=1}^{p} h \cdot u(a_j^*, h) + l + Mu(a_i^*, a_j^*); \\ [PU_2]\, u(a_i^*, a_j^*) \in \{0, 1\}.\end{array}\right\} E^{PCU} \tag{12}$$

3.2 Consistency verification

A set of parameter sets compatible with the DM's preference information is denoted by $E(A^R) = E^{BASE} \cup E^{ASS-EX} \cup E^{PCL} \cup E^{PCU} \cup E^{CARD}$. To ensure that it is non-empty, the following optimization problem needs to be solved:

$$Minimize \sum_{a_i^* \in A^R} u(a_i^*) + \sum_{(a_i^*, a_j^*) \in P^R} u(a_i^*, a_j^*) + \sum_{h \in H} u(C_h), \quad \text{subject to} \quad E(A^R).$$

It minimizes the number of preference information pieces (i.e., assignment examples for $a_i^* \in A^R$, assignment-based pairwise comparisons for $(a_i^*, a_j^*) \in P^R$, and desired class cardinalities for $C_h, h \in H$) that need to be removed to restore consistency. The instantiation of each statement is controlled by the respective binary variable, i.e., $u(a_i^*)$ for assignment examples, $u(a_i^*, a_j^*)$ for assignment-based pairwise comparisons, and $u(C_h)$ for desired class cardinalities. In case all preference information pieces can be reproduced together, the optimal value of the objective function would be equal to zero. Otherwise, the binary variables equal to one would indicate the DM's statements that need to be removed to reinstate compatibility with an assumed outranking-based sorting model.

3.3 Robustness analysis

In case a set of indirect and imprecise preference information pieces is not contradictory, infinitely many sorting models can reproduce it. Their exploitation can be conducted in two ways. On the one hand, a representative set of parameters can be selected using some arbitrary procedure. The application of thus constructed sorting model on the set of alternatives implies univocal recommendation. On the other hand, all compatible sets of parameters can be exploited to conduct robustness analysis. In the proposed approach, we implement the latter idea, while extending it to provide multiple complementary results quantifying the stability of sorting recommendation given the plurality of outranking-based models compatible with the DM's incomplete preferences. Since we treat criteria weights and credibility levels as variables, all compatible sets of parameters defined by $E(A^R)$ are denoted by $\mathcal{S}(w, \lambda)$.

3.3.1 Robustness analysis with mathematical programming

In this section, we focus on deriving exact robust results by means of mathematical programming. These outcomes include:

- Possible assignment $C_P(a_i)$ indicating a set of classes to which alternative $a_i \in A$ is assigned by at least one compatible set of parameters ($a_i \rightarrow^P C_h$), i.e. $C_P(a_i) = \{C_h : \exists(w, \lambda) \in \mathcal{S}(w, \lambda)$ for which $a_i \rightarrow C_h\}$;
- Necessary assignment-based preference relation $\succeq^{\rightarrow,N}$ which holds for pairs of alternatives $(a_i, a_j) \in A \times A$, such that a_i is assigned to a class at least as good as a_j

for all compatible sets of parameters, i.e., $a_i \succeq^{\rightarrow,N} a_j \Leftrightarrow \forall(w,\lambda) \in \mathcal{S}(w,\lambda)\ a_i \rightarrow C_h$ and $a_j \rightarrow C_k$ where $h \geq k$;

– Extreme class cardinalities indicating the minimal C_h^{min} and maximal C_h^{max} numbers of alternatives assigned to class C_h for some compatible set of parameters, i.e., $C_h^{min} = min_{(w,\lambda)\in\mathcal{S}(w,\lambda)}|C_h|$ and $C_h^{max} = max_{(w,\lambda)\in\mathcal{S}(w,\lambda)}|C_h|$, where $|C_h|$ denotes the size of class C_h.

Possible assignments For $a_i \in A$, $C_P(a_i)$ is composed of classes $C_h, h = 1, \ldots, p$, such that $E(a_i \rightarrow^P C_h)$ given below is feasible and $\varepsilon^* = \max \varepsilon$, subject to $E(a_i \rightarrow^P C_h)$ is greater than 0.

$$\left.\begin{array}{l} [PA_1]\ \sigma(a_i, b_{h-1}) \geq \lambda,\ \text{if } h > 1, \\ [PA_2]\ \sigma(a_i, b_h) + \varepsilon \leq \lambda,\ \text{if } h < p, \\ E(A^R). \end{array}\right\} E(a_i \rightarrow^P C_h) \tag{13}$$

Constraints $[PA_1]$ and $[PA_2]$ are responsible for assigning alternative a_i to class C_h by ensuring that a_i outranks the lower profile and does not outrank the upper profile of C_h. Let us denote the indices of the least and the most preferred classes to which a_i can be possibly assigned by, respectively, $L(a_i)$ and $R(a_i)$.

Necessary assignment-based preference relation The purpose of the assignment-based preference relation is to enable the pairwise comparisons of alternatives in a manner compatible with the sorting method. This is not possible when referring to the assignment- or cardinality-oriented results. The analysis for pairs may be particularly useful when the DM is interested in the recommendation obtained for some particular alternatives or when there are numerous alternatives possibly assigned to the same class range.

By comparing the alternatives in terms of the robust assignment-based relations, one can arrive at conclusions such as:

– "all compatible outranking models assign a to a class at least as good as b" (in case of analyzing necessary assignment-based preference relation), or
– "for 50% of compatible outranking models a is assigned to a class better than b, for 20% of feasible models the order of classes is inverse, and for the remaining 30%—a and b are placed in the same class" (in case of analyzing Assignment-based Outranking Indices).

These conclusions are compatible with the sorting problem, and cannot be obtained through analyzing the stability of an outranking or preference relation considered in the context of ranking problems.

When it comes to the necessary assignment-based preference relation $\succeq^{\rightarrow,N}$, for $a_i, a_j \in A$, $a_i \succeq^{\rightarrow,N} a_j$ holds if $\forall h \in H$, $E_{h,C}(a_i \succeq^{\rightarrow,N} a_j)$ given below is not feasible or, in case of feasibility, $\varepsilon^* = \max \varepsilon$ subject to $E_{h,C}(a_i \succeq^{\rightarrow,N} a_j)$ is not greater than 0.

$$\left.\begin{array}{l} [PN_1]\ \sigma(a_j, b_h) \geq \lambda, \\ [PN_2]\ \sigma(a_i, b_h) + \varepsilon \leq \lambda, \\ E(A^R) \end{array}\right\} E_{h,C}(a_i \succeq^{\rightarrow,N} a_j) \tag{14}$$

Constraints $[PN_1]$ and $[PN_2]$ ensure that a_j is assigned to class at least C_{h+1}, whereas a_i is assigned to class at most C_h. Hence, to verify if a_i is assigned to a class at least as good as a_j for all compatible sets of parameters, we check if a_j can be assigned to class better than a_i for at least one feasible parameter set.

Extreme class cardinalities The minimal C_h^{min} and maximal C_h^{max} cardinalities of class C_h correspond to the optimal solutions of the following Mixed-Integer Linear Programming

(MILP) problems:

$$Minimize/Maximize: \sum_{a_i \in A} u(a_i, h), \text{ subject to } E(A^R), \tag{15}$$

where $u(a_i, h)$ is equal to one in case the conditions justifying an assignment of a_i to C_h are satisfied. Hence, by minimizing (maximizing) the sum of $u(a_i, h)$ for all $a_i \in A$, we identify an outranking model in the set of all compatible models defined by $E(A^R)$ that minimizes (maximizes) the number of alternatives for which the conditions supporting an assignment to C_h are met at the same time.

The possible assignments, necessary assignment-based preference relation, and extreme class cardinalities offer complementary perspectives on the robustness of results. None of them can be fully captured based on the analysis of the remaining ones. This justifies their joint employment. For example, when alternative a_i is possibly assigned to classes between C_3 and C_5 and alternative a_j is possibly assigned to classes between C_2 and C_4, a_i is not guaranteed to be always assigned to a class at least as good as a_j, so it may only be verified by means of the necessary assignment-based preference relation. Furthermore, there may exist numerous alternatives that are possibly assigned to class C_h, but this may happen for different compatible sets of parameters. Thus, only the analysis of extreme class cardinalities captures how many alternatives can be placed in the same class for some feasible set of parameters.

3.3.2 Robustness analysis with the Monte Carlo simulations

The exploitation of a set of compatible sorting models using mathematical programming techniques leads to the exact robust results. They reveal what happens for some, all, and extreme compatible sets of parameters. Nonetheless, the possible assignment often includes multiple classes, the necessary relation leaves many pairs of alternatives, whereas the range of observable class cardinalities can be wide. In this perspective, it would be useful to compute the results which quantify the probability of attaining different outcomes given the multiplicity of compatible sorting models. Specifically, we consider the following results, which are defined as multi-dimensional integrals over the space of uniformly distributed weights and credibility levels compatible with the DM's preference information:

– **Class Acceptability Index** $CAI(a_i, h)$ indicating the proportion of compatible sets of parameters assigning alternative $a_i \in A$ to class C_h, i.e.

$$CAI(a_i, h) = \int_{(w,\lambda) \in \mathcal{S}(w,\lambda)} m(w, \lambda, i, h) d(w, \lambda), \tag{16}$$

where $m(w, \lambda, i, h)$ is the class membership function:

$$m(w, \lambda, i, h) = \begin{cases} 1, & \text{if } a_i \to C_h \text{ for } (w, \lambda), \\ 0, & \text{otherwise.} \end{cases}$$

– **Assignment-Based Outranking Index** $AOI(a_i, a_j)$ indicating the share of compatible sets of parameters for which alternative $a_i \in A$ is assigned to a class at least as good as alternative $a_j \in A$, i.e.:

$$AOI(a_i, a_j) = \int_{(w,\lambda) \in \mathcal{S}(w,\lambda)} r(w, \lambda, i, j) d(w, \lambda), \tag{17}$$

where $r(w, \lambda, i, j)$ is the assignment-based outranking relation confirmation function:

$$r(w, \lambda, i, j) = \begin{cases} 1, & \text{if } a_i \to C_h \text{ and } a_j \to C_k \text{ and } h \geq k \text{ for } (w, \lambda), \\ 0, & \text{otherwise.} \end{cases}$$

– **Class Cardinality Index** $CCI(h, r)$ indicating the share of compatible sets of parameters for which the size of class C_h (i.e., the number of alternatives assigned to it) is equal to r, i.e.

$$CCI(h, r) = \int_{(w,\lambda)\in\mathcal{S}(w,\lambda)} c(w, \lambda, h, r) d(w, \lambda), \tag{18}$$

where $c(w, \lambda, h, r)$ is the class cardinality function:

$$c(w, \lambda, h, r) = \begin{cases} 1, & \text{if } |C_h| = r \text{ for } (w, \lambda), \\ 0, & \text{otherwise.} \end{cases}$$

Since $\mathcal{S}(w, \lambda)$ is multi-dimensional and non-convex, it is not possible to compute the stochastic acceptability indices accurately. Computing the exact value of these indices would require us to (1) calculate the volume of the full space exactly, (2) for each alternative, pair, or class to partition the space in regions that grant a particular result, and (3) to compute the volumes of these regions. Exact computation of the volume of a polytope is extremely hard from the computational viewpoint (Kadziński and Tervonen 2013). In turn, to estimate CAIs, AOIs, and CCIs, we use the Monte Carlo simulation to sample from the space of uniformly distributed criteria weights and credibility levels. Since there do not exist efficient algorithms for sampling the compatible outranking models from the non-convex space, we combine the HAR algorithm (Tervonen et al. 2013) with a naive rejection technique. Specifically, we first sample from a convex space of compatible parameter sets defined by basic normalization and non-negativity constraints as well as the DM's assignment examples, i.e., $E^{BASE} \cup E^{ASS-EX}$. Note that these constraints introduce the linear cuts in the space of feasible weights and credibility levels, which makes HAR suitable for efficiently sampling from such a space (Lovász 1999). However, the constraints introduced by the assignment-based pairwise comparisons and desired class cardinalities involve the binary variables, thus making the space of compatible parameters sets non-convex. In this regard, to find the samples which are compatible with all types of DM's preference information, we reject these samples returned by HAR that do not align with $E^{PCL} \cup E^{PCU} \cup E^{CARD}$. The latter is justifiable because a rejection sampling is feasible in low dimensionality problems, and acceptable error limits for the stochastic indices can always be achieved, given sufficient computation time, with a certain number of Monte Carlo iterations (Kadziński and Tervonen 2013).

4 Properties of the model outcomes

In this section, we present the properties of robust sorting results. We first discuss the impact of preference information on the outcomes. Then, we present how these results evolve with an incremental specification of preference information. We exhibit the interrelations between the exact results derived with mathematical programming and the stochastic acceptability indices estimated with the Monte Carlo simulations. All these properties are of high practical relevance. However, since their proofs are obvious, we omit them. Finally, we discuss some measures for quantifying comprehensive robustness of results.

Relations between provided preference information and model outcomes The approach proposed in this paper has been designed so that to ensure a correspondence between the

preference information and the robust results. The relations between the provided input and the derived outputs exist on three levels concerning individual alternatives, pairs of alternatives, and decision classes. They can be summarized as follows:

- for each reference alternative $a_i^* \in A^R$ its possible assignment is a subset of the desired assignment, i.e. $C_P(a_i^*) \subseteq [C_{L_{DM}(a_i^*)}, C_{U_{DM}(a_i^*)}]$; that is, the set of classes confirmed by at least one compatible parameter set can be narrower than the range specified by the DM, but never wider;
- for each pair of reference alternatives $a_i^*, a_j^* \in A^R$ such that $a_i^* \succ^{\rightarrow}_{\geq k,DM} a_j^*$ with $k \geq 0$, the necessary assignment-based preference relation holds, i.e. $a_i^* \succ^{\rightarrow}_{\geq k,DM} a_j^*$, $k \geq 0 \Rightarrow a_i^* \succeq^{\rightarrow,N} a_j^*$;
- for each class C_h with specified desired class cardinalities, the extreme class cardinalities are within the bounds provided by the DM, i.e., $C_{h,DM}^{min} \leq C_h^{min}$ and $C_h^{max} \leq C_{h,DM}^{max}$.

Evolution of robust results with incremental specification of preference information The analysis of the robust result stimulates the DM to enrich his/her preference information. For example, the analysis of $C_P(a_i)$ and CAIs may lead to new or more precise assignment examples, $\succeq^{\rightarrow,N}$ and AOIs may enhance the DM to provide additional assignment-based pairwise comparisons, whereas C_h^{min}, C_h^{max}, and CCIs may stimulate more constrained desired class cardinalities. With each new additional piece of preference information, the set of compatible parameter sets becomes smaller. Thus, when considering the outcomes obtained in two subsequent iterations t and $t+1$ with preference information in $(t+1)$th iteration enriching the DM's requirements in tth iteration, the following properties concerning the evolution of results hold:

- for each alternative $a_i \in A$ its possible assignment can become narrower, but not wider, i.e., $C_P^{t+1}(a_i) \subseteq C_P^t(a_i)$;
- the assignment-based necessary preference relation is enriched, i.e., $\succeq_t^{\rightarrow,N} \subseteq \succeq_{t+1}^{\rightarrow,N}$;
- the difference between extreme class cardinalities can become lesser, but not greater, i.e., $C_{h,t}^{min} \leq C_{h,t+1}^{min}$ and $C_{h,t+1}^{max} \leq C_{h,t}^{max}$.

Interrelations between the exact outcomes and estimates of stochastic acceptability indices The results computed with mathematical programming are exact. Specifically, the necessary outcomes are confirmed by all compatible parameter sets, the possible results—by at least one feasible set of parameters, whereas the extreme ones indicate the recommendation obtained in the most and the least advantageous cases. On the contrary, the results derived from the Monte Carlo simulation are only estimates of the exact acceptability indices. Thus, although they can be estimated up to an arbitrarily selected accuracy threshold by analyzing a required number of samples, they would never be exact. In what follows, we list the interrelations between the necessary, possible, and extreme outcomes and the estimates of stochastic acceptability indices. Let us start with the results concerning the class assignments of individual alternatives:

- $h \notin C_P(a_i) \Rightarrow CAI(a_i, h) = 0$ and $C_P(a_i)$ is a singleton $\Rightarrow CAI(a_i, h) = 1$;
- $h \in [C_P(a_i)] \Rightarrow CAI(a_i, h) \in [0, 1]$ and $\sum_{h \in C_P(a_i)} CAI(a_i, h) = 1$;
- $CAI(a_i, h) > 0 \Rightarrow h \in C_P(a_i)$.

When it comes to the results observed for pairs of alternatives, the following properties hold:

- $a_i \succsim^{\rightarrow,N} a_j \Rightarrow AOI(a_i, a_j) = 1$;
- $AOI(a_i, a_j) < 1 \Rightarrow \neg(a_i \succsim^{\rightarrow,N} a_j)$.

As far as the class cardinalities are concerned, the following interrelations can be considered:

- $r < C_h^{min}$ or $r > C_h^{max} \Rightarrow CCI(h, r) = 0$ and $C_h^{min} = C_h^{max} = r \Rightarrow CCI(h, r) = 1$;
- $C_h^{min} \leq r \leq C_h^{max} \Rightarrow CCI(h, r) \in [0, 1]$ and $\sum_{r=C_h^{min},\ldots,C_h^{max}} CCI(h, r) = 1$;
- $CCI(h, r) > 0 \Rightarrow C_h^{min} \leq r \leq C_h^{max}$.

The above properties imply that the DM's preference information has an impact also on the stochastic acceptability indices. As a result, the latter ones evolve with an incremental specification of preferences. However, it is not possible to formulate some general conclusions on how a particular acceptability index would be modified in case the respective exact result would be either possible, though not necessary, or it would be contained within the extreme, non-precise bounds.

One needs to bear in mind that CAIs, AOIs, and CCIs are not sufficient to derive the exact results. Since we consider the estimates of acceptability indices, $CAI(a_i, h) = 0$ and $CCI(h, r) = 0$ do not exclude $h \in C_P(a_i)$ or $C_h^{min} \leq r \leq C_h^{max}$. Moreover, $CAI(a_i, h) = 1$ and $AOI(a_i, a_j) = 1$ do not imply $C_P(a_i)$ is a singleton or $a_i \succsim^{\rightarrow,N} a_j$, respectively.

Robustness measures for multiple criteria sorting Let us introduce a few measures quantifying comprehensive robustness of results in view of six types of outcomes obtained from the analysis of $\mathcal{S}(w, \lambda)$:

- an average difference between indices of the most and the least preferred possible classes, built on the possible assignments $C_P(a_i)$ of all alternatives $a_i \in A$:

$$f_{CPW}(\mathcal{S}(w, \lambda)) = \frac{1}{n} \sum_{i=1}^{n} [R(a_i) - L(a_i) + 1]; \tag{19}$$

- a mean class assignment entropy, built on the class acceptability indices for all alternatives $a_i \in A$ and all classes C_h, $h = 1, \ldots, p$:

$$f_{CAI}(\mathcal{S}(w, \lambda)) = \frac{1}{n} \sum_{i=1}^{n} - \sum_{h=1}^{p} CAI(a_i, h) log_2 CAI(a_i, h); \tag{20}$$

- the share of ordered pairs of alternatives for which the order of classes is not fixed, which means that the necessary assignment-based preference relation is not true, but the possible relation holds (which is approximated with $APOI(a_i, a_j) > 0$):

$$f_{NAPOI}(\mathcal{S}(w, \lambda)) = \frac{\sum_{(a_i,a_j) \in A \times A,\ i \neq j} POS(a_i, a_j)}{n \cdot (n-1)}, \tag{21}$$

  where:

$$POS(a_i, a_j) = \begin{cases} 1, & \text{if } not(a_i \succeq^{\rightarrow,N} a_j) \text{ and } APOI(a_i, a_j) > 0, \\ 0, & \text{otherwise;} \end{cases} \tag{22}$$

- a mean assignment-based outranking entropy, built on the assignment-based outranking indices for all pairs of alternatives $(a_i, a_j) \in A \times A,\ i \neq j$:

$$f_{AOI}(\mathcal{S}(w, \lambda)) = \sum_{(a_i,a_j) \in A \times A,\ i \neq j} \frac{-APOI(a_i, a_j) \cdot log_2 APOI(a_i, a_j)}{n \cdot (n-1)}; \tag{23}$$

Table 1 A set of decision alternatives: indices and names of MBA programs

a_i	Program name	a_i	Program name
a_1	London Business School	a_{16}	Insead
a_2	MIT: Sloan	a_{17}	New York Un.: Stern
a_3	Yale School of Management	a_{18}	Northwestern Un.: Kellogg
a_4	Un. of North Carolina: Kenan-Flagler	a_{19}	Rotterdam School of Management
a_5	Un. of Toronto: Rotman	a_{20}	Stanford Un. GSB
a_6	UCLA: Anderson	a_{21}	UC Berkeley: Haas
a_7	Un. of Oxford: Said	a_{22}	Un. of Chicago GSB
a_8	Un. of Maryland: Smith	a_{23}	Un. of Iowa: Tippie
a_9	Carnegie Mellon: Tepper	a_{24}	Un. of Michigan: Ross
a_{10}	Columbia Business School	a_{25}	Un. of Pennsylvania: Wharton
a_{11}	Cornell University: Johnson	a_{26}	Un. of Rochester: Simon
a_{12}	Duke University: Fuqua	a_{27}	Un. of Virginia: Darden
a_{13}	Esade Business School	a_{28}	Un. of Western Ontario: Ivey
a_{14}	Harvard Business School	a_{29}	Vanderbilt University: Owen
a_{15}	Iese Business School	a_{30}	York University: Schulich

– an average difference between the greatest and the least class cardinalities for all classes $C_h, h = 1, \ldots, p$:

$$f_{CCW}\left(\mathcal{S}(w, \lambda)\right) = \frac{1}{p}\sum_{h=1}^{p}[C_h^{max} - C_h^{min} + 1]; \tag{24}$$

– a mean class cardinality entropy, built on the class cardinality indices for all classes C_h, $h = 1, \ldots, p$, and all possible cardinalities $r = 1, \ldots, n$:

$$f_{CCI}\left(\mathcal{S}(w, \lambda)\right) = \frac{1}{p}\sum_{h=1}^{p} - \sum_{r=1}^{n} CCI(h, r) log_2 CCI(h, r). \tag{25}$$

The above measures can be analyzed by the DM, who incrementally provides preference information. All measures have been designed in such a way that lesser values indicate greater robustness of results. In this perspective, the DM can judge, after each iteration, whether the interactive process should be continued, in case the robustness is satisfying, or stopped, otherwise.

5 Illustrative study: multiple criteria evaluation of MBA programs

In this section, we present an illustrative study to demonstrate the applicability and practical usefulness of the proposed framework for preference modeling and robustness analysis with ELECTRE Tri-B. The actual data concerns evaluation of 30 MBA programs (Köksalan et al. 2009) (see Table 1). The considered gain-type criteria are defined on a 0–100 scale, and involve alumni career progress (g_1), diversity (g_2) and idea generation (g_3) (the performances are provided in Table 2).

Table 2 Performances of the MBA programs on three criteria

a_i	g_1	g_2	g_3	a_i	g_1	g_2	g_3
a_1	68.78	62.03	59.87	a_{16}	56.59	70.74	65.45
a_2	60.01	24.1	85.81	a_{17}	68.24	26.74	80.4
a_3	79.01	25.98	51.84	a_{18}	68.68	24.52	72.43
a_4	67.8	22.45	62.4	a_{19}	51.2	52.91	42.98
a_5	59.34	36.68	60.79	a_{20}	76.77	28.52	81.8
a_6	55.46	21.52	74.54	a_{21}	61.26	31.04	73.69
a_7	57.04	43.49	47.12	a_{22}	75.64	23.57	77.73
a_8	55.33	25.28	64.13	a_{23}	58.25	26.25	46.81
a_9	54.02	18.69	71.93	a_{24}	63.23	28.36	69.61
a_{10}	83.44	32.94	71.63	a_{25}	77.11	31.51	91.59
a_{11}	60.54	30.31	55.99	a_{26}	60.25	29.6	49.25
a_{12}	64.05	27.25	64.68	a_{27}	76.49	19.05	37.68
a_{13}	54.61	60.15	14.45	a_{28}	50.04	40.55	53.55
a_{14}	77.84	29.05	93.91	a_{29}	64.67	26.33	35.14
a_{15}	61.99	54.61	40.04	a_{30}	57.03	50.73	47.55

Table 3 Performances of the boundary class profiles and indifference q_j, preference p_j, and veto v_j thresholds for all criteria

g_j	Boundary profiles						Thresholds		
	b_0	b_1	b_2	b_3	b_4	b_5	q_j	p_j	v_j
g_1	0	56	58	60	65	100	3	5	10
g_2	0	24	30	33	37	100	3	5	9
g_3	0	50	55	60	63	100	3	5	9

We aim at assigning the MBA programs to five quality classes C_1, C_2, C_3, C_4, C_5, with C_1 and C_5 indicating the least and the most preferred categories, respectively. The performances of boundary profiles b_0, b_1 …, b_5 as well as the indifference, preference, and veto thresholds are provided in Table 3.

We will illustrate the use of the proposed methodological framework in its most general form, i.e., by referring to all types of indirect and imprecise preference information and robust outcomes. However, there is no obligation to use all kinds of inputs and outputs in case some of them are not found appealing by the DM (e.g., in case of assignment-based pairwise comparisons) or they are not imposed by the problem's nature or DM's preferences (e.g., in case of desired class cardinalities). The versatile robustness analysis can be conducted for a set of outranking models compatible with any preference information.

Also, the proposed framework for robustness analysis may support the elicitation of indirect preference information already in the first iteration. In particular, it would be possible to determine the results when only the boundary profiles and comparison thresholds are provided. Subsequently specified holistic judgments (e.g., assignment examples) or requirements (e.g., desired class cardinalities) should be consistent with the outcomes derived from such an analysis (e.g., the possible assignments or extreme class cardinalities, respectively). In this way, they can be linked to a prior specification of an outranking model through profiles and thresholds.

Table 4 Desired class cardinalities in the first iteration

C_h	C_1	C_2	C_3	C_4	C_5
$C_{h,DM}^{min}$	3	5	4	2	3
$C_{h,DM}^{max}$	7	10	8	10	7

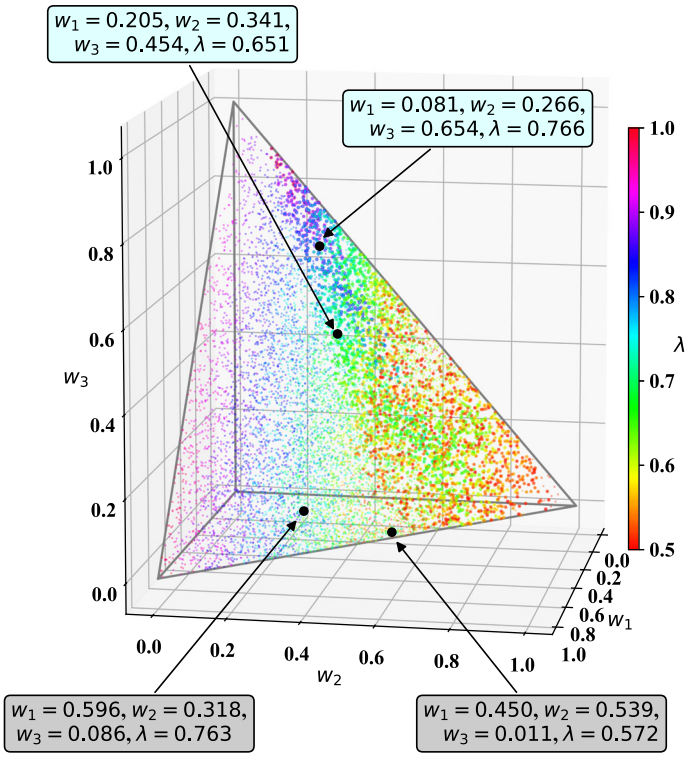

Fig. 1 Compatible sets of parameters derived with the Monte Carlo simulation in the first iteration (small dots are compatible with the assignment examples, whereas large dots are additionally consistent with assignment-based pairwise comparisons and desired class cardinalities)

5.1 Robustness analysis in the first iteration

We start with different statements representative for the three allowed types of preference information. In particular, we consider five assignment examples, one for each class: $a_9 \rightarrow C_1$, $a_{23} \rightarrow C_2$, $a_3 \rightarrow C_3$, $a_{21} \rightarrow C_4$, and $a_1 \rightarrow C_5$. Moreover, we account for three assignment-based pairwise comparisons: a_3 is better than a_2 by at most one class, a_5 is better than a_{17} by at least two classes, and a_{27} is worse than alternative a_{23} by at most two classes. Finally, we tolerate the extreme class cardinalities provided in Table 4. They impose that each class accommodates at least a few alternatives and prevent too many MBA programs to be classified in the same way. The provided preference information is consistent with an assumed sorting model. Thus, the set of compatible weights and credibility levels is non-empty and can be exploited in the spirit of robustness analysis.

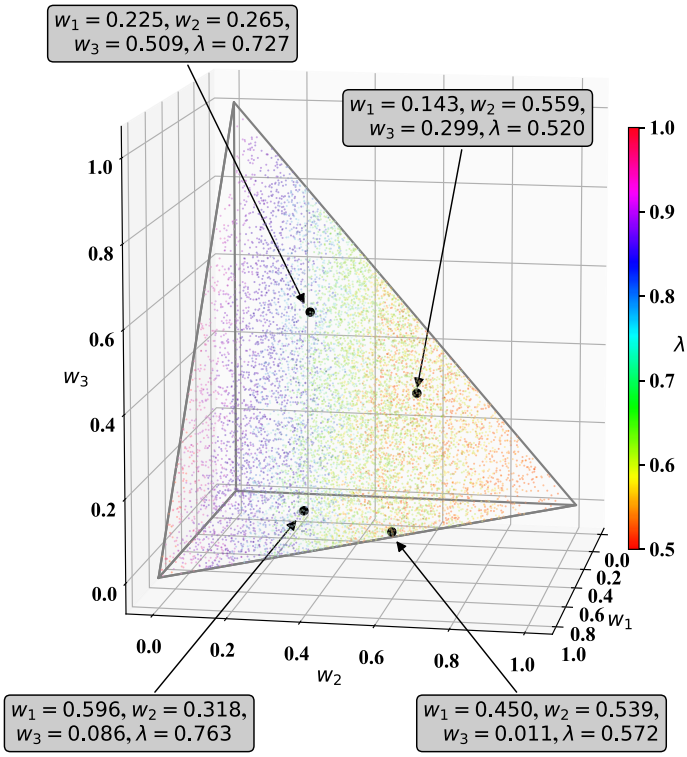

Fig. 2 Sets of parameters derived with the Monte Carlo simulation compatible with the assignment examples in the first iteration

The possible, necessary, and extreme results are computed with mathematical programming techniques, whereas the values of acceptability indices are estimated with the Monte Carlo simulation. In Fig. 1, we present the samples considered in the first iteration. Each dot corresponds to a compatible set of parameters composed of three weights (w_1, w_2, w_3) corresponding to the coordinates according to the three axes and credibility level λ, whose value is represented with a color on a pre-defined scale. For clarity of presentation, some randomly selected dots are described with precise parameter values. All dots, including small and large ones, are counterparts of 10,000 sets of parameters compatible with the assumptions of the outranking-based model as well as assignment examples. They were obtained with HAR. The analysis of their distribution in the space of all feasible parameters indicates that the space of compatible weights and credibility levels is not constrained much by the assignment examples (see dots in Fig. 2). The large dots in Fig. 1 represent the sets of parameters compatible with all pieces of preference information, including assignment-based pairwise comparisons and desired class cardinalities (see Fig. 3). Hence, the small dots correspond to the samples that were neglected due to the application of a rejection sampling technique, whereas the large ones are used to estimate the stochastic results. As indicated by Fig. 1, assignment-based pairwise comparisons and desired class cardinalities significantly reduced the space of compatible weights and credibility levels.

In Table 5, we provide the possible assignments for all alternatives. They are precise for 24 alternatives. For example, 3 alternatives (a_1, a_5, a_{10}) are univocally assigned to the best

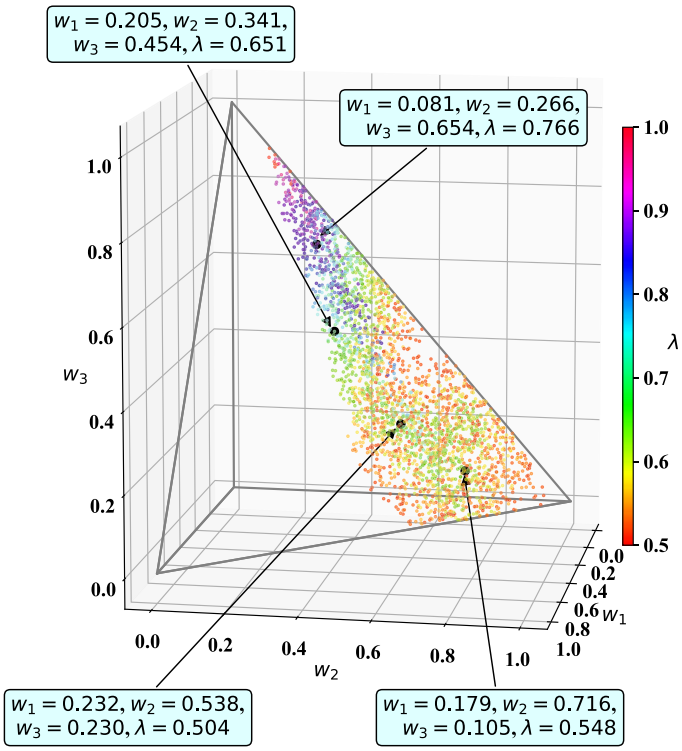

Fig. 3 Sets of parameters derived with the Monte Carlo simulation compatible with all types of preference information provided in the first iteration

Table 5 Possible assignments for all alternatives in the first iteration

a_i	$C_P(a_i)$	a_i	$C_P(a_i)$	a_i	$C_P(a_i)$
a_1	C_5	a_{11}	C_3, C_4	a_{21}	C_4
a_2	C_2	a_{12}	C_3	a_{22}	C_2
a_3	C_3	a_{13}	C_1	a_{23}	C_2
a_4	C_2	a_{14}	C_4	a_{24}	C_3, C_4
a_5	C_5	a_{15}	C_1	a_{25}	C_4
a_6	C_2	a_{16}	C_4	a_{26}	C_2, C_3
a_7	C_2	a_{17}	C_3	a_{27}	C_1
a_8	C_2, C_3	a_{18}	C_2	a_{28}	C_1, C_2
a_9	C_1	a_{19}	C_1	a_{29}	C_1
a_{10}	C_5	a_{20}	C_3, C_4	a_{30}	C_2

class (C_5), whereas 6 alternatives ($a_9, a_{13}, a_{15}, a_{19}, a_{27}, a_{29}$) are assigned to the worst class (C_1) with all compatible sets of parameters. All DM's assignment examples are reproduced in the possible assignments (e.g., since a_{23} was assigned by the DM to class C_2, we have $C_P(a_{23}) = \{C_2\}$). For the remaining 6 alternatives (a_8, a_{12}, a_{20}, a_{24}, a_{26}, and a_{28}), the possible assignments are imprecise. Specifically, these alternatives are sorted into one of two consecutive classes depending on the selected compatible set of parameters.

When it comes to CAIs, for the alternatives with precise possible assignments, they are equal to one for the sole admissible class and zero for all remaining classes. The analysis of CAIs is the most informative for the alternatives with imprecise possible assignments. Indeed, for some of them, the probability of being assigned to either of two admissible classes is similar. For example, the shares of compatible sets of parameters assigning a_{24} to C_3 and C_4 are equal to, respectively, 51.09% and 48.91%. On the contrary, for some other alternatives, the vast majority of feasible weights and credibility levels assign them to a single class. In particular, for a_{11}, the probability of being assigned to C_4 is eight times greater than the respective probability for C_3.

The other perspective on the robustness of sorting recommendation concerns the relative comparison of assignments for pairs of alternatives. A Hasse diagram indicating a partial pre-order imposed on the set of alternatives by the necessary assignment-based preference relation is presented in Fig. 4. It respects the DM's assignment-based pairwise comparisons. For example, since a_{23} and a_5 were required to be assigned to a class better than, respectively, a_{27} and a_{17}, $\succsim^{\rightarrow,N}$ is true for these pairs.

To support understanding of the interrelations between $\succsim^{\rightarrow,N}$ and $C_P(a_i)$, the nodes representing alternatives with the same possible assignments are marked with the same color. In this regard, some relations follow the possible assignments directly. For example:

- alternatives assigned to C_5 with all compatible sets of parameters (a_1, a_5, a_{10}) are preferred over all remaining alternatives, whereas alternatives possibly assigned only to C_1 $(a_9, a_{13}, a_{15}, a_{19}, a_{27}, a_{29})$ are worse than all other alternatives;
- alternatives assigned to C_3 (a_3, a_{12}, a_{17}) are preferred over those possibly assigned to either C_3 or C_2 (a_8 and a_{26});
- alternatives precisely assigned to the same class are indifferent in terms of the necessary assignment-based preference relation, and hence they are grouped in the same node in Fig. 4.

However, some relations cannot be derived solely from the analysis of possible assignments. For example, when considering alternatives possibly assigned to class C_3 or C_4, a_{11} is incomparable with a_{20}. This means that for some feasible sets of parameters, a_{11} is assigned to a class better than a_{20}, whereas for some other compatible weights and credibility levels—the order of classes for this pair is inverse. Furthermore, although the possible assignments for a_{20} and a_{24} are the same, a_{20} is always is assigned to a class at least as good as a_{24}, whereas the opposite is not true. This suggests that for some feasible sorting models, a_{20} is assigned to a class strictly better than a_{24}.

The analysis of the necessary assignment-based preference relation can be enriched with consideration of Assignment-based Outranking Indices. We report them in the form of a heatmap, which translates the index values between 0 and 1 on a pre-defined color scale (see Fig. 5). As indicated by the associated color scale, the darker is a specific cell in a matrix, the greater is the proportion of compatible sets of parameters for which an alternative from the row is assigned to a class at least as good as an alternative from the column.

For pairs of alternatives related by the necessary assignment-based preference relations, AOIs are equal to one. In particular, $AOI(a_1, a_5) = 1$ and $AOI(a_5, a_1) = 1$. Furthermore, for pairs with disjoint sets of possible assignments, $APOI$ is equal to zero in case one of them is always assigned to a class strictly worse than the other (e.g., $AOI(a_{12}, a_{14}) = 0$). Let us emphasize that even if an alternative is not necessarily preferred to some other alternative, it can be assigned to a class at least as good for some feasible sets of parameters. In particular, a_3 is assigned to a class not worse than a_{11}, a_{20}, and a_{24} by, respectively, 11%, 41%, and 51% of compatible weights and credibility levels. Finally, the analysis of AOIs is the most

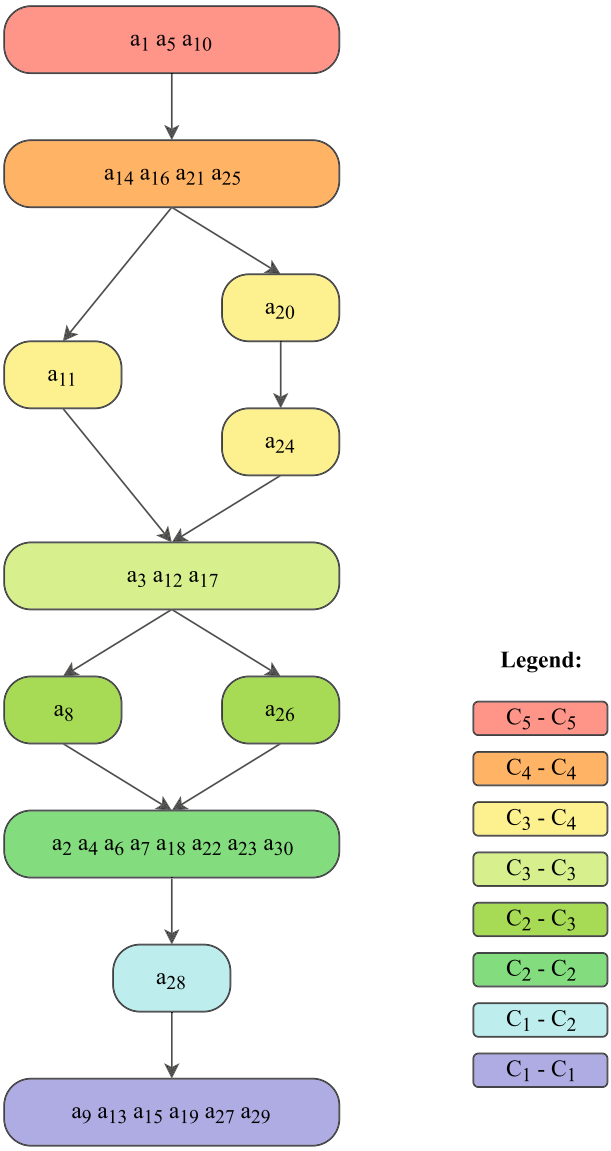

Fig. 4 A Hasse diagram representing the necessary assignment-based preference relation in the first iteration

informative for pairs, which are incomparable in terms of $\succsim^{\rightarrow,N}$. For example, the proportion of compatible sets of parameters for which a_{11} is assigned to a class not worse than a_{20} is greater than the share of feasible models for which the inverse relation holds. The exact values of Assignment-based Outranking Indices are provided for all pairs in the e-Appendix (supplementary material available online).

The last perspective concerns class cardinalities. In Table 6, we present the extreme numbers of alternatives that can be assigned to each class by some compatible sets of parameters. For all classes, they are more precise than the cardinalities desired by the DM. For example,

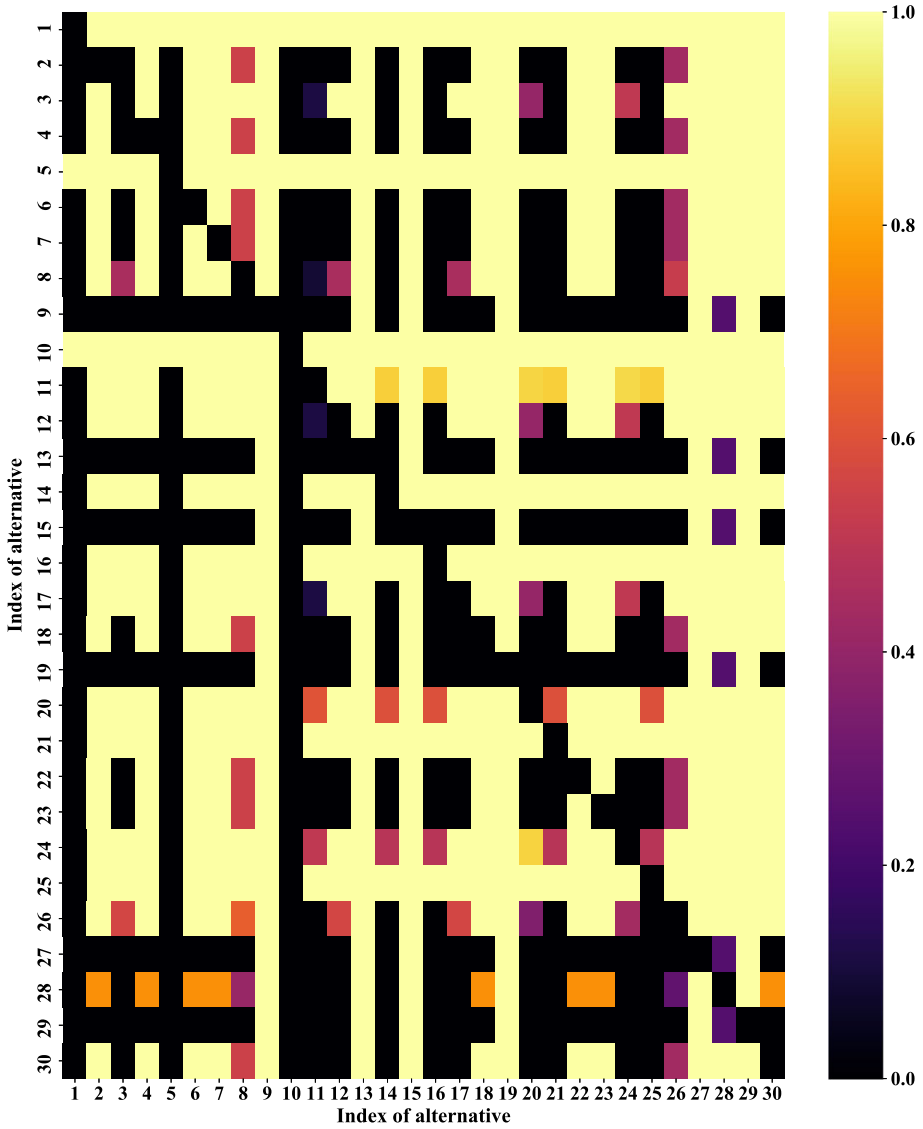

Fig. 5 A heatmap of the Assignment-based Outranking Indices in the first iteration (for the numerical values, see e-Appendix)

Table 6 Extreme class cardinalities in the first iteration

	C_1	C_2	C_3	C_4	C_5
C_h^{min}	6	8	4	4	3
C_h^{max}	7	10	6	7	3

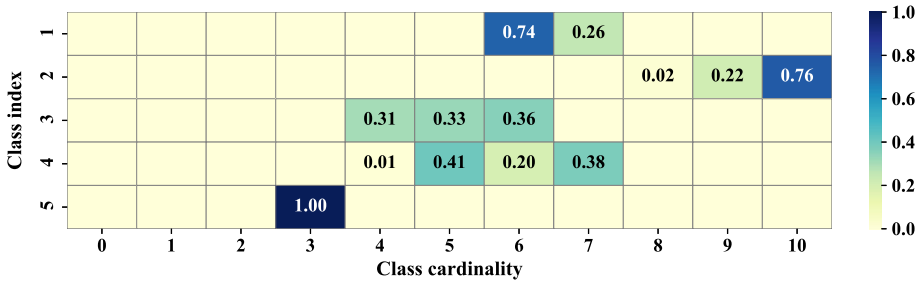

Fig. 6 A heatmap of Class Cardinalities Indices in the first iteration

the minimal number of alternatives that were allowed to be assigned to C_5 was 5, whereas the least observable cardinality for this class is 8. In the same spirit, the maximal number of alternatives that could be placed in C_3 was 8, whereas the greatest size of C_3 that is observed for some feasible set of parameters is 6.

Furthermore, for all feasible sets of weights and credibility levels, only 3 alternatives are assigned to C_5. This is understandable as they were the only alternatives that could possibly be assigned to the most preferred class. For the remaining classes, the minimal and maximal sizes differ. For class C_1 these extreme numbers ($C_h^{min} = 6$ and $C_h^{max} = 7$) agree with the analysis of possible assignments as they were 7 alternatives which could be possibly assigned to C_1, but only six of them were assigned to this class for all compatible sets of parameters. However, for some other classes, these extreme cardinalities cannot be derived from the analysis of other results. For example, the minimal and maximal numbers of alternatives assigned to C_3 are equal to, respectively, 4 and 6. However, the numbers of alternatives that are always or sometimes assigned to C_3 are equal to 3 and 8, respectively. These sizes are not observed for any feasible sets of parameters.

The analysis of extreme class cardinalities can be supported with consideration of class cardinality indices. Table 7 represents a heatmap of CCIs. The cells without a number correspond to the cardinalities, which are not attained for any compatible set of parameters (i.e., $CCI(h, r) = 0$). Clearly, for C_5—the size of 3 is confirmed by all considered samples ($CCI(5, 3) = 1$). For the remaining classes, the analysis of CCIs offers rich insights on how many alternatives they accommodate for different compatible sets of parameters. In particular, the vast majority of feasible sets of weight and credibility levels indicate that there are 6 and 10 alternatives assigned to, respectively, C_1 and C_2 (i.e., $CCI(1, 6) = 0.74$ and $CCI(2, 10) = 0.76$). Furthermore, for probabilities of 4, 5, and 6 alternatives being assigned to C_3 are very similar. Finally, for C_4—the size of 4 is the least probable, being confirmed only by 1% of compatible sets of parameters.

5.2 Robustness analysis in the second iteration

Let us assume that the analysis of robust results in the first iteration stimulated the DM to provide additional preference information of different types:

- alternative a_{11} should be assigned to class C_3 implied by the analysis of its imprecise possible assignment and Class Acceptability Indices;
- alternative a_{24} should be assigned to a class better than a_{26} based on the analysis of Assignment-based Outranking Indices;

Table 7 Class Acceptability Indices in the first iteration

	C_1	C_2	C_3	C_4	C_5
a_1	0	0	0	0	1
a_2	0	1	0	0	0
a_3	0	0	1	0	0
a_4	0	1	0	0	0
a_5	0	0	0	0	1
a_6	0	1	0	0	0
a_7	0	1	0	0	0
a_8	0	0.5444	0.4556	0	0
a_9	1	0	0	0	0
a_{10}	0	0	0	0	1
a_{11}	0	0	0.1088	0.8912	0
a_{12}	0	0	1	0	0
a_{13}	1	0	0	0	0
a_{14}	0	0	0	1	0
a_{15}	1	0	0	0	0
a_{16}	0	0	0	1	0
a_{17}	0	0	1	0	0
a_{18}	0	1	0	0	0
a_{19}	1	0	0	0	0
a_{20}	0	0	0.4081	0.5919	0
a_{21}	0	0	0	1	0
a_{22}	0	1	0	0	0
a_{23}	0	1	0	0	0
a_{24}	0	0	0.5109	0.4891	0
a_{25}	0	0	0	1	0
a_{26}	0	0.4479	0.5521	0	0
a_{27}	1	0	0	0	0
a_{28}	0.2602	0.7398	0	0	0
a_{29}	1	0	0	0	0
a_{30}	0	1	0	0	0

– at least 5 alternatives can be assigned to C_3 and at most 5 should be placed in C_4 motivated by the analysis of extreme class acceptabilities and Class Cardinality Indices.

The compatible sets of parameters derived with the Monte Carlo simulation for thus enriched preference information are presented in Fig. 7. The sets of parameters compatible solely with assignment examples (see Fig. 8) as well as with the remaining types of preference information (see Fig. 9) have been vastly constrained when compared with the first iteration. This is understandable in view of how the preferences were formulated in the second iteration, i.e., they excluded some results previously deemed as possible.

The outcomes that were confirmed by all compatible sets of parameters in the previous iteration are also validated in the next iteration (e.g., a univocal possible assignment of a_{14} to C_3; a necessary assignment-based preference of a_2 over a_{28}, or a precise cardinality of class

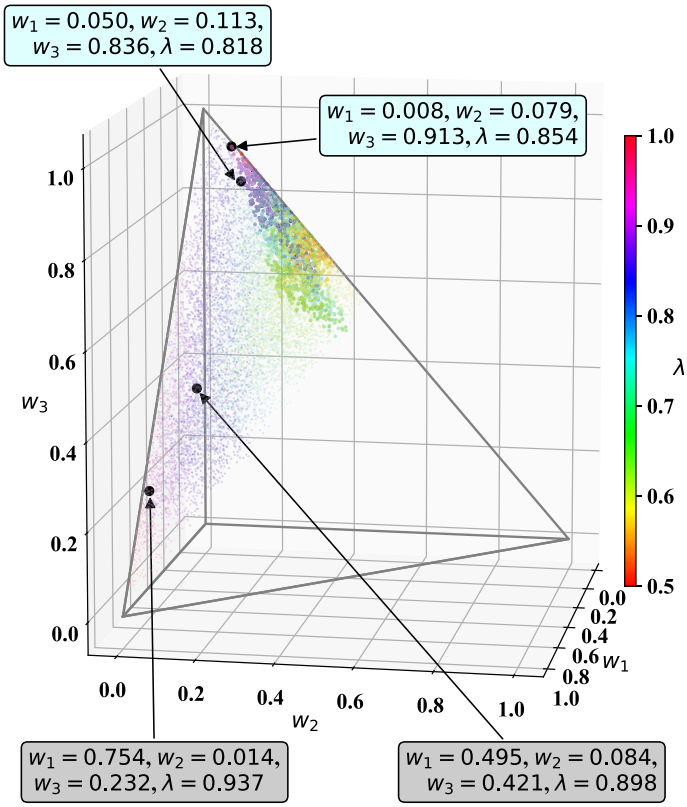

Fig. 7 Compatible sets of parameters derived with the Monte Carlo simulation in the second iteration (small dots are compatible with the assignment examples, whereas large dots are additionally consistent with assignment-based pairwise comparisons and desired class cardinalities)

C_5). Thus, in what follows, we focus only on the evolution of results that were not validated for all feasible sets of weights and credibility levels in the first iteration.

When it comes to the possible assignments (see Table 8), they have become precise for another five alternatives. For example, a_8 and a_{26} are assigned to C_2 for all compatible sets of parameters, whereas previously, they could also be placed in C_3. In fact, there is only one alternative (a_{20}) for which the possible assignment is still imprecise. The analysis of CAIs indicated that 53.17% compatible models assign it to C_3, whereas 46.83% sort a_{20} to C_4. Thus, class C_3 is slightly more probable for a_{20} given indeterminacy of the DM's preference model due to the incompleteness of his/her preference information, whereas in the previous iteration—class C_4 was indicated by almost 60% of feasible sorting models.

The necessary assignment-based preference relation is presented in the form of a Hasse diagram in Fig. 10. The subsets of alternatives which are assigned to the same class by all feasible sets of parameters have now become more numerous. For example, 10 alternatives are indifferent in terms of $\succsim^{\rightarrow,N}$ due to being precisely assigned to C_2. Moreover, there is no single pair of alternatives that would be incomparable in terms of the necessary relation. For example, a_{20} is now assigned to a class at least as good as a_{11} for all compatible weights and credibility levels.

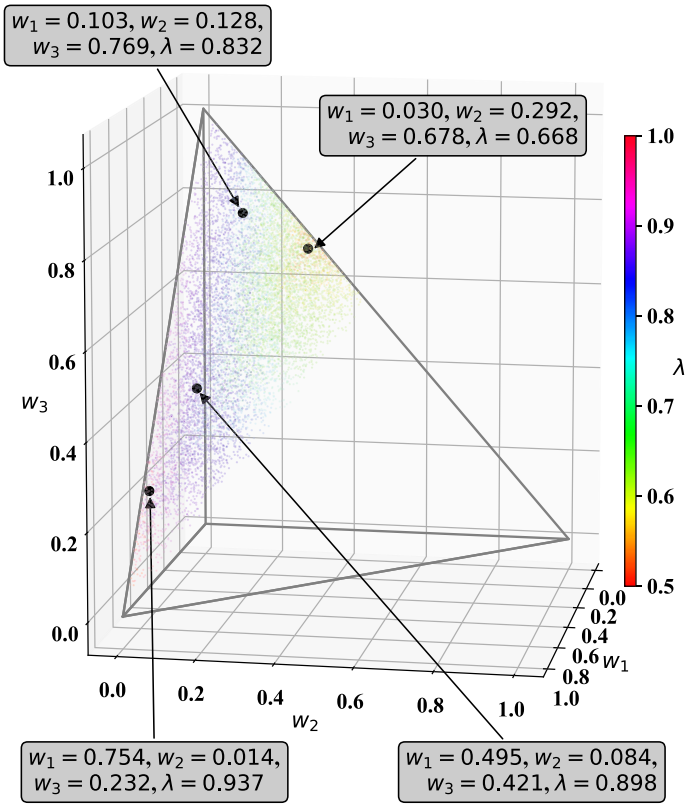

Fig. 8 Sets of parameters derived with the Monte Carlo simulation compatible with the assignment examples in the second iteration

When it comes to the Assignment-based Outranking Indices, the only values which are neither 0 nor 1 can be observed for the comparison of a_{20} with other alternatives. Specifically, it is assigned to the same class (i.e., C_4) as a_{14}, a_{16}, a_{21}, and a_{25} for less than half of feasible sorting models, whereas a_3, a_{11}, a_{12}, a_{17}, and a_{24} are assigned to the same class (i.e., C_3) for more than 50% of compatible parameter sets (Fig. 11).

The extreme class cardinalities indicate that the numbers of alternatives assigned to each class are now more precise for all classes. On the one hand, there are always 7, 10, and 3 alternatives which are assigned to, respectively, C_1, C_2, and C_5. On the other hand, C_3 accommodates from 5 to 6 alternatives, whereas the cardinality of C_4 is between 4 and 5. The respective CCIs are given in Fig. 12. They confirm that it is slightly more probable that 4 and 6 alternatives are assigned to, respectively, C_4 and C_3, which is consistent with the previously discussed results related to the assignment of a_{20}.

The DM should analyze the obtained class assignments and cardinalities as well as assignment-based preference relation after each iteration. (S)he should judge whether the suggested recommendation is convincing and decisive enough, and the correspondence between the output of the model and the preferences (s)he has at the moment is satisfying. This decision can be supported by the analysis of robustness measures (see Table 9). For example, the average difference between indices of the most and the least preferred possible classes has decreased from 1.2 to 1.033 (see f_{CPW}). In fact, the possible assignment is now imprecise

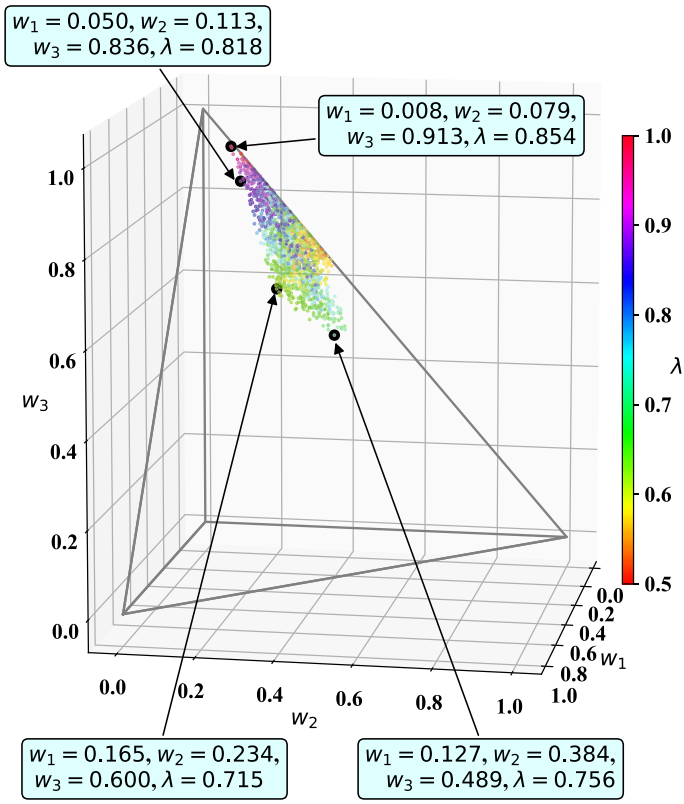

Fig. 9 Sets of parameters derived with the Monte Carlo simulation compatible with all types of preference information provided in the second iteration

Table 8 Possible assignments for all alternatives in the second iteration

a_i	$C_P(a_i)$	a_i	$C_P(a_i)$	a_i	$C_P(a_i)$
a_1	C_5	a_{11}	C_3	a_{21}	C_4
a_2	C_2	a_{12}	C_3	a_{22}	C_2
a_3	C_3	a_{13}	C_1	a_{23}	C_2
a_4	C_2	a_{14}	C_4	a_{24}	C_3
a_5	C_5	a_{15}	C_1	a_{25}	C_4
a_6	C_2	a_{16}	C_4	a_{26}	C_2
a_7	C_2	a_{17}	C_3	a_{27}	C_1
a_8	C_2	a_{18}	C_2	a_{28}	C_1
a_9	C_1	a_{19}	C_1	a_{29}	C_1
a_{10}	C_5	a_{20}	C_3, C_4	a_{30}	C_2

only for 1 out of 30 alternatives as compared to 6 alternatives in the first iteration. In the same spirit, the share of ordered pairs of alternatives f_{NAPOI} for which the order of classes is not fixed has decreased from 0.0793 to 0.0103 (i.e., from 69 to 9 out of 870 ordered pairs), whereas an average difference f_{CCW} between the greatest and the least class cardinalities has dropped from 2.6 to 1.4. We can also observe significant improvement of the entropy-based

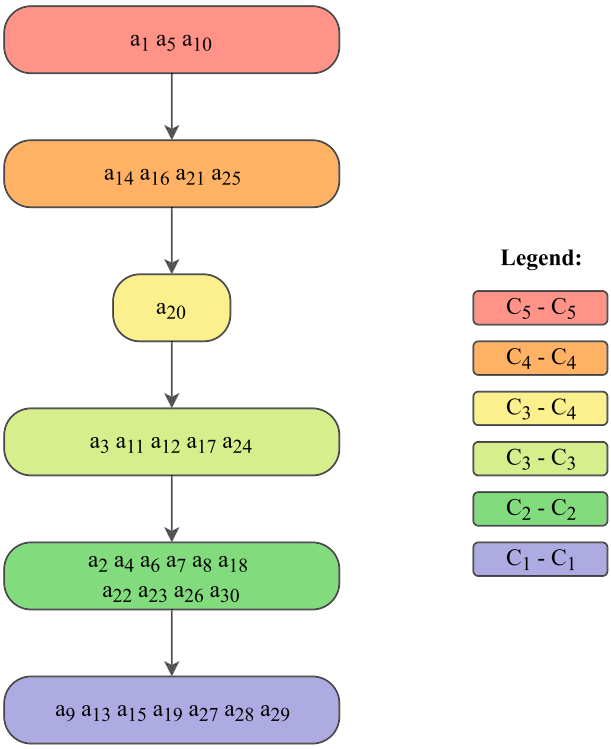

Fig. 10 A Hasse diagram representing the necessary assignment-based preference relation in the second iteration

measures exploiting the stochastic indices. For example, the entropy of class assignment indices $f_{CAI} = 0.0332$ indicates an extremely low variability of results, i.e., over five times lesser than in the first iteration.

If the DM judges the robustness of results as satisfying, the interactive process stops. We terminate the illustration after the second iteration because the recommendation is already very robust. The only uncertainty is related to the assignment of a_{20}. Once the DM would assign it to either C_3 or C_4, all compatible models would already recommend the same sorting decision.

In case the robustness of sorting outcomes would not be found sufficient by the DM, (s)he should pursue the exchange of preference information. If (s)he changed her/his mind or discovered that the expressed judgments were inconsistent with some previous judgments that (s)he considers more important, (s)he may backtrack to one of the earlier iterations and continue from this point. In this way, the process of preference construction is either continued or restarted (Corrente et al. 2013).

6 Summary and directions of future research

In this paper, we presented a novel integrated framework for preference modeling and robustness analysis with the ELECTRE Tri-B method. On the one hand, we discussed mathematical

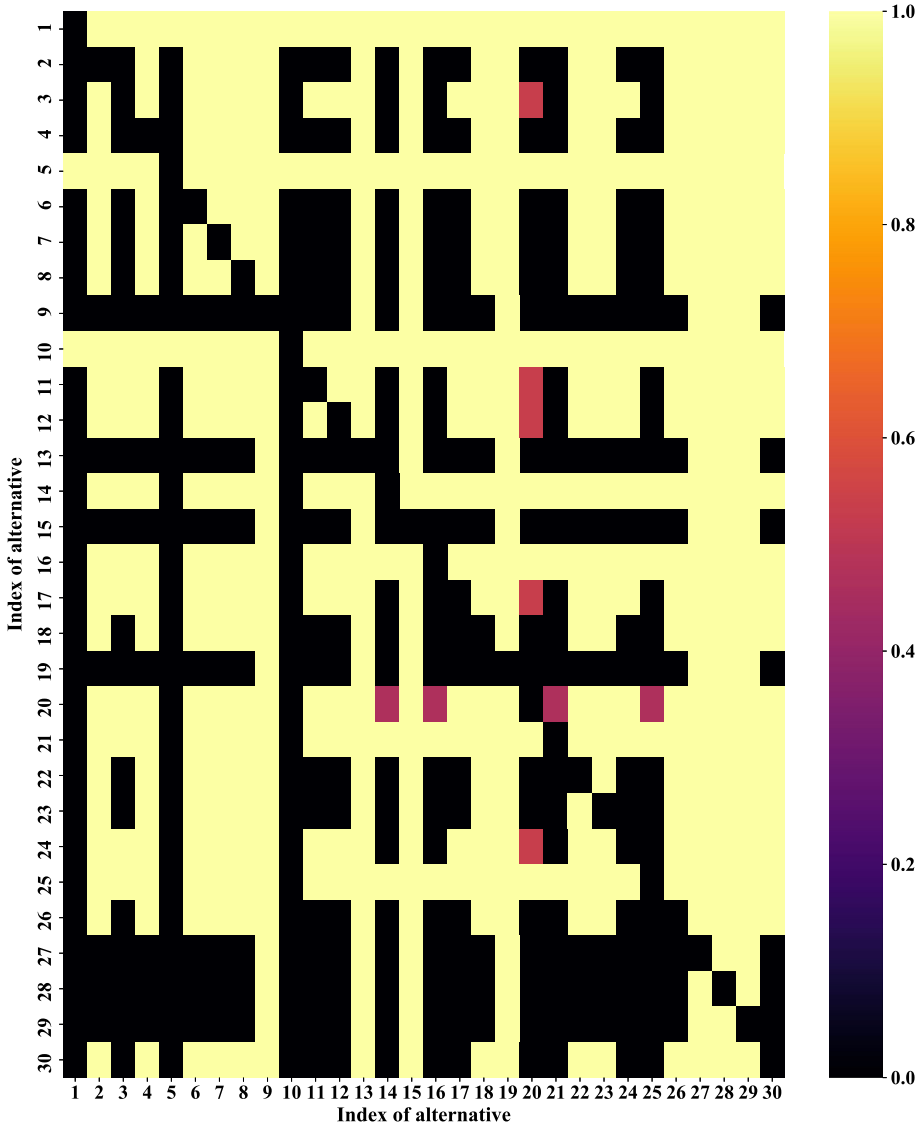

Fig. 11 A heatmap of the Assignment-based Outranking Indices in the second iteration (for the numerical values, see e-Appendix)

programming models for incorporating three types of indirect and imprecise preferences. These include assignment examples (e.g., "alternative a_1 should be assigned to C_2" or "alternative a_2 should be placed in either C_3 or C_4"), assignment-based pairwise comparisons (e.g., "alternative a_3 should be assigned to a class better than alternative a_4" or "alternatives a_5 and a_6 need to be assigned to the same class"), and desired class cardinalities (e.g., "the number of alternatives assigned to class C_1 should be between 3 and 6"). On the other hand, we provided the algorithms for computing multiple robust results. The exact outcomes, including the possible assignments, necessary assignment-based preference relation, and extreme

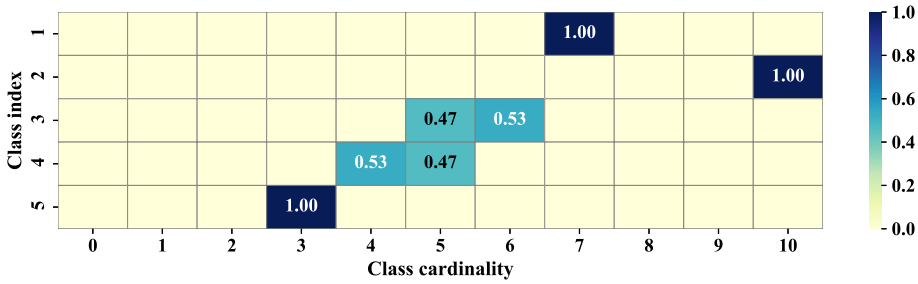

Fig. 12 A heatmap of Class Cardinalities Indices in the second iteration

Table 9 Robustness measures in the first and second iterations

Iteration	f_{CPW}	f_{CAI}	f_{NAPOI}	f_{AOI}	f_{CCW}	f_{CCI}
First	1.200	0.1762	0.0793	0.0340	2.6	0.9784
Second	1.033	0.0332	0.0103	0.0051	1.4	0.3989

class cardinalities, can be derived with mathematical programming techniques. Furthermore, the stochastic acceptability indices quantifying the probability of a particular assignment, the truth of assignment-based relation, or some specific number of alternatives assigned to a given class are estimated with the Monte Carlo simulation. For this purpose, we used an original combination of the Hit-And-Run algorithm and the rejection sampling technique. Overall, we provide the richest MCDA framework for robustness analysis. Indeed, (Kadziński and Ciomek 2016) and (Kadziński et al. 2015) considered only exact robust outcomes in the context of outranking- and value-based sorting procedures. Let us emphasize that when applying the method to real-world decision problems, the DMs can use only preference information that they find relevant for a particular context and refer to the results that they consider useful for supporting decision making.

The usefulness of multiple accounted inputs and outputs was discussed theoretically based on several properties and demonstrated on a realistic example concerning the evaluation of MBA programs. Firstly, we revealed the impact of provided preference relation on the results. It is enhanced by the correspondence between different perspectives concerning individual alternatives, pairs of alternatives, and decision class. Secondly, we showed how the results evolve with an incremental specification of preference information. In this perspective, diverse robust outcomes play a crucial role in stimulating the DM to supply additional preference information pieces, whereas the contraction of the set of compatible parameter sets implies that the results become more and more robust. Thirdly, we focussed on why it is useful to analyze the outcomes concerning assignments, assignment-based relation, and class sizes. Although they are interrelated, none of them can be fully derived from the analysis of the remaining ones. Fourthly, we discussed the benefits of using exact robust outcomes and stochastic acceptability indices. Specifically, we demonstrated how these two types of outcomes complement each other with exact results indicating which recommendation is certain, possible, or extreme and stochastic indices filling a gap, in a probabilistic way, between what is possible and certain. Finally, we proposed some measures for quantifying the comprehensive robustness of sorting results. Their analysis may stimulate the decision on terminating an interaction with the method or pursuing an incremental specification of preference information.

The idea of processing diverse human-readable preference information and providing sophisticated, robust results at the method's output could be transferred to other MCDA approaches. The most appealing extension refers to the ELECTRE Tri-C method (Almeida-Dias et al. 2010), which – unlike ELECTRE Tri-B—requires the DM to provide characteristic rather than boundary class profiles and employs two conjoint rules to work out an assignment for each alternative.

In what follows, we discuss the limitations of our approach and the related directions for future research. First, in our sorting model, we considered only criteria weights and credibility levels as variables whose values need to be inferred from indirect and imprecise preference information. However, an outranking-based sorting model requires specification of other parameters such as boundary profiles or indifference, preference, and veto thresholds. As demonstrated by a rich literature on preference disaggregation in the context of ELECTRE Tri-B, it is feasible to learn these parameters from the DM's holistic judgments. However, it is not possible to infer them all at once with linear programming techniques. Thus, it would require the development of dedicated non-linear optimization algorithms.

Second, our method has been designed for an interactive specification of relatively small subsets of indirect and imprecise statements. However, when larger sets of preferences are involved, inconsistencies occur more frequently. Moreover, the number of binary variables involved in specifying compatible sorting models is proportional to the number of alternatives and classes. When considering the insufficient efficiency of existing solvers, it would be neither possible to identify inconsistencies nor to model the preferences using Mixed-Integer Linear Programming in case the number of binary variables would go into thousands. Hence we envisage an extension of the method for dealing with large sets of preference information pieces and large decision problems by developing some heuristic preference learning approaches.

Third, we left the DM freedom to select the alternatives, pairs of alternatives, or classes for which (s)he specified some desired outcomes or constraints that should be respected by the recommended decision. As demonstrated by the recent research on active learning (see, e.g., Ciomek et al. 2017), it is possible to design some algorithms to indicate the objects that should be judged by the DM so that to increase both information gain from the DM's answer and robustness of the results. These approaches should incorporate the measures for quantifying the robustness of sorting outcomes proposed in this paper.

Fourth, when discussing the results of a case study, we visualized the compatible sets of parameters. This is very useful for understanding how the DM's indirect and imprecise statements constrain the space of an outranking-based sorting model's parameters. In the considered study, the number of model variables was four. Thus, we used three-dimensional coordinate systems to illustrate the weight values and color for reflecting the values of a credibility level. When the fifth variable would be involved, it would be possible to show its values using points' sizes. However, with an even greater number of model parameters, it would be already challenging to ensure the figure's readability.

Finally, we demonstrated the method's applicability on a study concerning the evaluation of MBA programs. We simulated all types of preference information pieces and analyzed all types of robust results. In this way, we could illustrate the capabilities of the proposed approach. The ultimate verification of its practical usefulness will involve applying it in case studies with real-world DMs.

Acknowledgements M. Kadziński acknowledges financial support from the Polish National Science Center under the SONATA BIS project (Grant No. DEC-2019/34/E/HS4/00045). The authors thank the two anonymous reviewers whose remarks allowed us to improve the previous version of the manuscript.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

References

Almeida-Dias, J. (2010). Multiple criteria decision aiding for sorting problems: concepts, methodologies, and applications. Doctoral thesis, Université Paris IX Dauphine, Paris, France.

Almeida-Dias, J., Figueira, J., & Roy, B. (2010). Electre Tri-C: A multiple criteria sorting method based on characteristic reference actions. *European Journal of Operational Research*, *204*(3), 565–580.

Bilich, F., & da Silva, R. (2008). Valuation and optimization of the impact of intellectual capital on organizational performance. *Portuguese Journal of Management Studies*, *13*(3), 341–359.

Brito, A. J., de Almeida, A. T., & Mota, C. M. M. (2010). A multicriteria model for risk sorting of natural gas pipelines based on ELECTRE TRI integrating Utility Theory. *European Journal of Operational Research*, *200*(3), 812–821.

Cailloux, O., Meyer, P., & Mousseau, V. (2012). Eliciting ELECTRE TRI category limits for a group of decision makers. *European Journal of Operational Research*, *223*(1), 133–140.

Ciomek, K., Kadziński, M., & Tervonen, T. (2017). Heuristics for selecting pair-wise elicitation questions in multiple criteria choice problems. *European Journal of Operational Research*, *262*(2), 693–707.

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

Corrente, S., Kadziński, M., Greco, S., & Słowiński, R. (2013). Robust ordinal regression in preference learning and ranking. *Machine Learning*, *93*, 381–422.

Damart, S., Dias, L., & Mousseau, V. (2007). Supporting groups in sorting decisions: Methodology and use of a multi-criteria aggregation/disaggregation DSS. *Decision Support Systems*, *43*(4), 1464–1475.

Diakoulaki, D., & Hontou, V. (2003). A multicriteria approach to burden sharing among industrial branches for combating climate change. *Clean Technologies and Environmental Policy*, *5*(1), 35–40.

Dias, L. C., Henggeler Antunes, C., Dantas, G., de Castro, N., & Zamboni, L. (2018). A multi-criteria approach to sort and rank policies based on Delphi qualitative assessments and ELECTRE TRI: The case of smart grids in Brazil. *Omega*, *76*, 100–111.

Dias, L. C., & Mousseau, V. (2006). Inferring Electre's veto-related parameters from outranking examples. *European Journal of Operational Research*, *170*(1), 172–91.

Dias, L. C., Mousseau, V., Figueira, J., & Clímaco, J. (2002). An aggregation/disaggregation approach to obtain robust conclusions with Electre Tri. *European Journal of Operational Research*, *138*(2), 332–348.

Dias, L. C., Mousseau, V., & Figueira, J. (2003). On the notion of category size in multiple criteria sorting models. DIMACS technical report 2003–02.

Dimitras, A., Zopounidis, C., & Hurson, C. (1995). A multicriteria decision aid method for the assessment of business failure risk. *Foundations of Control Engineering*, *20*(2), 99–112.

Doumpos, M., & Zopounidis, C. (2011). A multicriteria outranking modeling approach for credit rating. *Decision Sciences*, *42*(3), 721–742.

Fernandez, E., Figueira, J., & Navarro, J. (2019). An interval extension of the outranking approach and its application to multiple-criteria ordinal classification. *Omega*, *84*, 189–198.

Fernandez, E., Figueira, J., Navarro, J., & Roy, B. (2017). ELECTRE TRI-nB: A new multiple criteria ordinal classification method. *European Journal of Operational Research*, *261*(1), 214–24.

Figueira, J., Mousseau, V., & Roy, B. (2005). ELECTRE methods. In *Multiple criteria decision analysis: State of the art surveys* (pp. 133–153). New York: Springer.

Hurson, C., & Zopounidis, C. (1995). On the use of multicriteria decision aid methods to portfolio selection. *Journal of Euro-Asian Management*, *1*(2), 69–94.

Joerin, F., Theriault, M., & Musy, A. (2001). Using GIS and outranking multicriteria analysis for land-use suitability assessment. *International Journal of Geographical Information Science*, *15*(2), 153–174.

Kadziński, M., Cinelli, M., Ciomek, K., Coles, S., Nadagouda, M., Varma, R., et al. (2018). Co-constructive development of a green chemistry-based model for the assessment of nanoparticles synthesis. *European Journal of Operational Research*, *264*(2), 472–490.

Kadziński, M., & Ciomek, K. (2016). Integrated framework for preference modeling and robustness analysis for outranking-based multiple criteria sorting with ELECTRE and PROMETHEE. *Information Sciences*, *241*(3), 830–841.

Kadziński, M., Ciomek, K., & Słowiński, R. (2015). Modeling assignment-based pairwise comparisons within integrated framework for value-driven multiple criteria sorting. *European Journal of Operational Research*, *352–353*, 167–187.

Kadziński, M., Greco, S., & Słowiński, R. (2016). Robustness analysis for decision under uncertainty with rule-based preference model. *Information Sciences*, *328*, 321–349.

Kadziński, M., Martyn, K., Cinelli, M., Słowiński, R., Corrente, S., & Greco, S. (2020). Preference disaggregation for multiple criteria sorting with partial monotonicity constraints: Application to exposure management of nanomaterials. *International Journal of Approximate Reasoning*, *117*, 60–80.

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 & Negotiation*, *27*, 33–59.

Kadziński, M., & Tervonen, T. (2013). Stochastic ordinal regression for multiple criteria sorting problems. *Decision Support Systems*, *55*(1), 55–66.

Kadziński, M., & Słowiński, R. (2013). DIS-CARD: a new method of multiple criteria sorting to classes with desired cardinality. *Journal of Global Optimization*, *56*(3), 1143–66.

Köksalan, M., Mousseau, V., Özpeynirci, Ö., & Özpeynirci, S. B. (2009). A new outranking-based approach for assigning alternatives to ordered classes. *Naval Research Logistics*, *56*(1), 74–85.

Lovász, L. (1999). Hit-and-run mixes fast. *Mathematical Programming*, *86*(3), 443–461.

Mailly, D., Abi-Zeid, I., & Pepin, S. (2014). A Multi-Criteria Classification Approach for Identifying Favourable Climates for Tourism. *Journal of Multi-Criteria Decision Analysis*, *21*(1–2), 65–75.

Merad, M., Verdel, T., Roy, B., & Kouniali, S. (2004). Use of multi-criteria decisionaids for risk zoning and management of large area subjected to mining-induced hazards. *Tunnelling and Underground Space Technology*, *19*, 165–178.

Mousseau, V., Dias, L. C., & Figueira, J. (2006). Dealing with inconsistent judgments in multiple criteria sorting models. *4OR*, *4*(3), 145–158.

Mousseau, V., Dias, L. C., Figueira, J., Gomes, C., & Clímaco, J. (2003). Resolving inconsistencies among constraints on the parameters of an MCDA model. *European Journal of Operational Research*, *147*(1), 72–93.

Mousseau, V., Figueira, J., & Naux, J. P. (2001). Using assignment examples to infer weights for ELECTRE TRI method: Some experimental results. *European Journal of Operational Research*, *130*(2), 263–275.

Mousseau, V., & Słowiński, R. (1998). Inferring an ELECTRE TRI model from assignment examples. *Journal of Global Optimization*, *12*(2), 157–174.

Ngo The, A., & Mousseau, V. (2002). Using assignment examples to infer category limits for the ELECTRE TRI method. *Journal of Multi-Criteria Decision Analysis*, *11*(1), 29–43.

Raju, K., Duckstein, L., & Arondel, C. (2000). Multicriterion analysis for sustainable water resources planning: A case study in Spain. *Water Resources Management*, *14*(6), 435–456.

Saleh, E., Błaszczyński, J., Moreno, A., Valls, A., Romero-Aroca, P., de la Riva-Fernandez, S., et al. (2018). Learning ensemble classifiers for diabetic retinopathy assessment. *Artificial Intelligence in Medicine*, *85*, 50–63.

Siskos, Y., Grigoroudis, E., Krassadaki, E., & Matsatsinis, N. (2007). A multicriteria accreditation system for information technology skills and qualifications. *European Journal of Operational Research*, *182*(2), 867–885.

Tervonen, T., Lahdelma, R., Almeida-Dias, J., Figueira, J., & Salminen, P. (2007). SMAA-TRI: A parameter stability analysis method for Electre Tri. In I. Linkov, G. Kiker, & R. Wenning (Eds.), *Environmental security in harbors and coastal areas: Management using comparative risk assessment and multi-criteria decision analysis, nato security through science series c: environmental security* (pp. 217–231). New York: Springer.

Tervonen, T., Figueira, J., Lahdelma, R., Almeida-Dias, J., & Salminen, P. (2009). A stochastic method for robustness analysis in sorting problems. *European Journal of Operational Research*, *192*(1), 236–242.

Tervonen, T., van Valkenhoef, G., Basturk, N., & Postmus, D. (2013). Hit-And-Run enables efficient weight generation for simulation-based multiple criteria decision analysis. *European Journal of Operational Research*, *224*(3), 552–559.

Yu, W. (1992). ELECTRE TRI: Aspects méthodologiques et manuel d'utilisation. Document du LAMSADE 74, Université-Paris-Dauphine

Zheng J., Cailloux O., & Mousseau V. (2011) Constrained multicriteria sorting method applied to portfolio selection. In R.I. Brafman, F.S. Roberts, A. Tsoukias (Eds.), Algorithmic decision theory. ADT 2011. Lecture notes in computer science (Vol. 6992). Berlin: Springer.

Zheng, J., Metchebon Takougang, S., Mousseau, V., & Pirlot, M. (2014). Learning criteria weights of an optimistic Electre Tri sorting rule. *Computers & Operations Research*, *49*, 28–40.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.