Incremental Learning of Simple Ranking Method Using Reference Profiles Models

Arwa Khannoussi[1], **Alexandru-Liviu Olteanu**[2], **Catherine Dezan**[3], **Jean-Philippe Diguet**[4], **Christophe Labreuche**[5], **Jacques Petit-Frère**[6], **Patrick Meyer**[7]

Abstract. In this article we study several heuristics for selecting pairs of alternatives that a decision maker must compare during an incremental preference elicitation process of the *Simplified Ranking based on Multiple reference Profiles* (SRMP) model. This work is motivated by a real-world context in which a SRMP model has been integrated in the on-board system of an Unmanned Aerial Vehicle (UAV), and where an off-line preference elicitation process is performed with the operator (or decision maker in this case) before the mission.

Our goal is therefore to determine a heuristic, with respect to various problem characteristics, that leads as quickly as possible to a good-enough representation of the decision maker's real preferences, in order to minimize his / her cognitive effort.

Keywords. multi-criteria decision aiding, SRMP model, incremental elicitation process, unmanned aerial vehicles

1 Introduction

When modelling complex decision problems, Multi-Criteria Decision Aiding usually uses three types of problem formulations (see for instance [18]): choice, ranking and sorting. Many aggregation methods have been proposed in the literature to solve these problems (see for instance [6]). In this paper, we focus on a method based on outranking relations, called *Simplified Ranking based on Multiple reference Profiles* (SRMP) [15], which provides a ranking of alternatives by comparing them to a set of reference profiles, instead of directly to each other.

When used in practice, the SRMP model should be tuned so that it accurately reflects the decision maker's (DM) preferences. This is done through a preference elicitation phase, in which an analyst and the DM interact in order to set the values for the model parameters. Very often, an indirect approach is used to set these values, by using holistic (or a posteriori) information (see e.g. [10]). Such an approach, in which the DM is asked to give his / her preferences on pairs of alternatives, has recently been proposed for SRMP, in order to reduce the cognitive effort of the DM [14].

In this paper, we suspect that this effort might still be too large, and that it could be reduced in a progressive, incremental learning process, where the pairs of alternatives to be compared by the DM are selected according to a heuristic, which leads as quickly as possible to a good-enough representation of the DM's real preferences. The work is motivated by a real-world context presented hereafter, in which a drone operator is involved in a preference elicitation process, in order to tune the parameters of an SRMP model that will be embedded in a drone.

The paper is structured in the following way: Section 2 presents the real-world motivation of this work. Then, Section 3 presents a quick overview on preference elicitation, and more specifically incremental learning. In Section 4 we first present the SRMP method, illustrate our discourse with an example taken from our use-case and present the existing non-incremental algorithm used to learn its parameters. In Section 5 we present the heuristics that we have explored so far for an incremental elicitation and in Section 6 we detail two experiments that we have performed in order to determine the heuristic to use in an incremental inference of an SRMP model approach. Finally, in Section 7 we discuss our findings and present our ongoing and future work.

2 Practical motivation

Unmanned Aerial Vehicles (UAVs), also known as drones, are employed in various civilian or military contexts (maritime surveillance, aerial surveying to update maps, tracking of smaller drones, ...). To accomplish a mission, they have to make decisions (e.g. *land*, *loiter*, *return to base*, ...) autonomously by taking into account multiple, potentially conflicting, criteria. Furthermore, UAVs operating with a high level of autonomy require that the operator / DM has a high confidence in the decisions they make.

Consequently, we propose to integrate models of human preferences into the automated control of UAVs. To do so, we first elicit the preferences of the operator in an off-line simulated context, during which the operator is confronted with multiple possible actions that the UAV might face, and for which he is asked to express his / her preferences. The learned model is then integrated into the on-board system of the UAV, having confidence that its on-line decisions will resemble as much as possible the preferences of the operator.

In order to ensure that the drone's decisions are consistent with the operator's preferences, even in situations not seen during the learning phase, and to avoid cognitive fatigue during this learning phase, we propose using an incremental preference elicitation strategy, based

[1] Lab-STICC, UMR 6285, CNRS, Université de Bretagne Occidentale, France & IMT Atlantique, Lab-STICC, Univ. Bretagne Loire, F-29238 Brest, France, arwa.khannoussi@{univ-brest.fr, imt-atlantique.fr}
[2] Lab-STICC, Université Bretagne Sud, Lorient, France, alexandru.olteanu@univ-ubs.fr
[3] Lab-STICC, UMR 6285, CNRS, Université de Bretagne Occidentale, France, catherine.dezan@univ-brest.fr
[4] Lab-STICC, Université Bretagne Sud, Lorient, France, jean-philippe.diguet@univ-ubs.fr
[5] Thales Group, christophe.labreuche@thalesgroup.com
[6] Thales Group, jacques.petit-frere@thalesgroup.com
[7] IMT Atlantique, Lab-STICC, Univ. Bretagne Loire, F-29238 Brest, France, patrick.meyer@imt-atlantique.fr

on heuristic for selecting the pairs of actions needed to compare. Moreover, the set of potential alternatives that can be used in the incremental elicitation process are predefined – as it is classically done in this field. In the context of the UAV application, the alternatives simply correspond to the choices that the UAV has to make (among actions such as "continue the mission", "return to the base", "skip a waypoint", ...) in different scenarios, contexts that are representative of standard or critical situations.

3 Background on preference elicitation

We are all facing the difficulty of making decisions on a day to day basis, whether the goal is to make a choice among a set of alternatives, to sort them or to rank them. In real-word problems, several conflicting evaluation criteria are involved in these decisions, which makes them inherently difficult. Multi-Criteria Decision Aiding (MCDA) [19, 16] aims to help the decision maker (DM) to solve such decision problems by taking into account his / her preferences. In the literature, MCDA methods are classified into three approaches (1) Multi-Attribute Value Theory (MAVT) [11], (2) outranking-based approaches [17] and (3) rule-based models [8].

The DM's preferences are usually transformed into values or constraints on the parameters of the MCDA model at hand (e.g. criteria weights and value functions for MAVT models). They can be given directly by the DM through a direct preference elicitation approach, however, such an approach is usually too difficult in real life, as the DM needs to have a very good understanding of the MCDA model. Therefore, a second approach is to start from a partial knowledge on the output of the method, such as, for example, pair-wise comparisons of alternatives in the ranking context, or assignment examples in the sorting context and then infer the model parameters.

The second approach so-called *indirect* preference elicitation has received much attention from researchers. For example, in the MAVT field, the seminal work of Jacquet-Lagreze et.al. [10] proposes the UTA method to infer additive value functions from a given ranking of a reference set through mathematical programming techniques. Later, such learning algorithms have been used on various different aggregation operators. Next to that, for outranking methods, Simos [20] proposes a simple procedure that allows to determine indirectly the weights of ELECTRE methods by using a set of cards. This method has been improved in [5]. Ngo The et.al. [21] propose a mixed integer linear program in order to infer the profiles of an ELECTRE TRI model with fixed weights and thresholds. Greco et.al. [7] define the ELECTRE$^{\text{GKMS}}$ method, which uses robust ordinal regression to construct a set of outranking models compatible with the provided preference information.

These techniques are by nature not incremental, as they start from the partial knowledge on the desired output of the MCDA method, as expressed by the DM, and determine in one shot a parameter configuration compatible with it. Regarding this progressiveness, in the MAVT context, Durbach [4] and Lahdelma et.al. [12] use an index that quantifies the volume of the polyhedron of the constraints specifying the possible value functions. They try to reduce this volume by adding constraints representing pair-wise comparisons of alternatives, until they converge to the best solution. Holloway et.al. [9] show the importance of the order of the pair-wise comparisons in decreasing the number of questions for reducing the cognitive effort of the DM. Ciomek et.al. [2] present a set of heuristics to minimize the number of elicitation questions and prioritize them in the context of *single choice* decision problems. They conclude that the best performing heuristic depends on the problem settings (e.g. number of criteria and alternatives). In the same context, Benabbou et.al. [1] select a set of pair-wise questions using a minimax regret strategy. This strategy reduces the number of pair-wise questions but the performance guarantee is weakened (with some acceptable bounds to the ideal situation).

To our knowledge, no previous work deals with incremental elicitation for outranking methods, and more specifically with the Simple Ranking Method Using Reference Profiles (SRMP).

4 Ranking with SRMP

4.1 Background and notation

In outranking methods, an "at least as good as" relation is built between pairs of alternatives evaluated on multiple criteria. This binary relation, called "outranking relation" [17] is often denoted by $\succsim$. An alternative a outranks another one, b, i.e. $a \succsim b$, if there are strong enough arguments to declare that a is at least as good as b and if there is no essential reason to refute that statement. Unfortunately, comparing all possible alternatives according to such a relation may result in cycles in the outranking relation, thus making it impossible to create a ranking. It has therefore been proposed by Rolland [15] to use a so-called reference point in the comparison of two alternatives : a is considered as strictly preferred to b if and only if the outranking relation between a and the reference point is stronger than the outranking relation between b and the reference point. Let us now show how this is implemented more formally.

We denote with $\mathcal{A}$ a set of n alternatives and with $M = \{1, \dots, m\}$ the indexes of m criteria. The evaluation of an alternative $a \in \mathcal{A}$ on criterion $j \in M$ is denoted with a_j.

SRMP [15] is an outranking method which uses so-called reference points to rank alternatives. The method is defined by several parameters, which may differ from one DM to another and which need to be identified beforehand, either directly, or through a learning algorithm. These parameters are:

- the reference profiles: $\mathcal{P} = \{p^h, h = 1..k\}$ where $p^h = \{p_1^h, \dots, p_j^h, \dots, p_m^h\}$ corresponds to the evaluations of p^h on all criteria and $p_j^h \succsim p_j^l, \forall h, l \in 1..k, h > l$;
- the lexicographic order of the profiles: σ, which corresponds to a permutation on $1..k$;
- the criteria weights: $w_1, w_2, \dots, w_m$, where $w_j \geq 0$ and $\sum_{j \in M} w_j = 1$

SRMP consists in a three-steps procedure as follows:

1. compute $C(a, p^h) = \{j \in M : a_j \geq p_j^h\}$ with $a \in \mathcal{A}, h = 1..k$, the set of criteria on which alternative a is at least as good as profile p^h.
2. compare all pairs of alternatives $a, b \in \mathcal{A}$ to the reference profiles in order to define the following relations:
   - $a \succsim_{p^h} b \Leftrightarrow \sum_{j \in C(a,p^h)} w_j \geqslant \sum_{j \in C(b,p^h)} w_j$
   - $a \succ_{p^h} b \Leftrightarrow \sum_{j \in C(a_i,p^h)} w_j > \sum_{j \in C(b,p^h)} w_j$
   - $a \sim_{p^h} b \Leftrightarrow \sum_{j \in C(a_i,p^h)} w_j = \sum_{j \in C(b,p^h)} w_j$
3. rank two alternatives $a, b \in \mathcal{A}$ by sequentially considering the relations $\succsim_{p^{\sigma(1)}}, \succsim_{p^{\sigma(2)}}, \dots, \succsim_{p^{\sigma(k)}}$ (according to the lexicographic order σ):

- a is preferred to b iff:

$$a \succ_{p^{\sigma(1)}} b \text{ or}$$
$$a \sim_{p^{\sigma(1)}} b \text{ and } a \succ_{p^{\sigma(2)}} b \text{ or}$$
$$\ldots$$
$$a \sim_{p^{\sigma(1)}} b \text{ and } \ldots \text{ and } a \sim_{p^{\sigma(k-1)}} b \text{ and } a \succ_{p^{\sigma(k)}} b$$

- a is indifferent to b iff:

$$a \sim_{p^{\sigma(1)}} b \text{ and } \ldots \text{ and } a \sim_{p^{\sigma(k)}} b$$

4.2 Illustrative example

For this example, we return to the practical motivation of this article, which is integrating autonomous decisions into drones based on the preferences of an operator / human decision maker. Classically, in a mission, such an UAV has to pass through a certain number of so-called way-points. For simplicity reasons, we will suppose here that at each way-point, the drone has to make a decision, which is to select between multiple actions or alternatives, like, e.g., continue the mission, land, interrupt the mission and return to base, . . . Consequently, we will suppose here that the decision model that is embedded in the drone is an SRMP model, which, for this illustration, is tuned according to the operator's preferences and which has been learned beforehand.

Consider that the drone has arrived in a way-point and that it faces three possible actions, x, y and z, which are evaluated on three criteria: risk (R), energy consumption (E) and mission progress (M). The evaluation of these alternatives on the criteria is presented in Table 1, as well as the preference parameters, learned beforehand from the operator.

Table 1. EVALUATIONS OF THE ALTERNATIVES AND SRMP PARAMETERS

	R	E	M
x	low	80%	20%
y	high	55%	90%
z	medium	20%	50%
p^1	low	70%	80%
p^2	high	35%	40%
w	1/3	1/3	1/3
σ		$(1, 2)$	

The two reference profiles allow to define three segments on the performances on each criterion: better than p^2; between p^1 and p^2; worse than p^1. In other terms, the reference profiles allow to identify an ordered encoding for each criterion defined by three ordered intervals of performances (A, B, and C) as illustrated in Fig. 1, such that:

A performances above p^2 on each criterion are denoted A (which can be interpreted as "good").
B performances between p^1 and p^2 on each criterion are denoted B (which can be interpreted as "intermediate" or "fair").
C performances below p^1 on each criterion are denoted C (interpreted as "insufficient").

The evaluations of the alternatives through this encoding is presented in Table 2.

We are going to follow the steps presented before to rank the three alternatives x, y and z. First we compute $C(a, p^h) = \{j \in M : a_j \geqslant p_j^h\}$, $\forall a \in \mathcal{A} = \{x, y, z\}$, $h = \{1, 2\}$, $M = \{R, E, M\}$. Then we compare the alternatives one to each other by using the profiles p^h, before ranking them with respect to the lexicographic order $\sigma = \{1, 2\}$.

Table 2. ENCODING OF THE EVALUATIONS OF THE ALTERNATIVES

	R	E	M
x	A	A	C
y	B	B	A
z	B	C	B

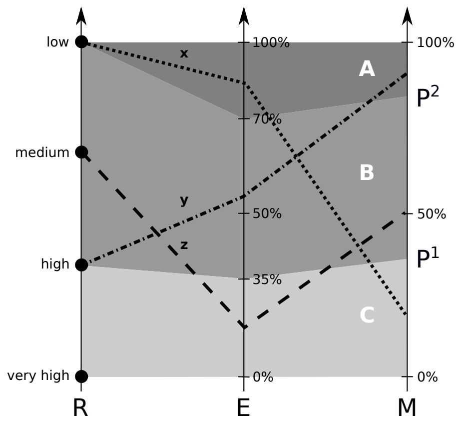

Figure 1. SRMP example

p^1 :

$$\left.\begin{array}{l}\sum_{j\in C(x,p^1)} w_j = 1/3 + 1/3 + 0 = 2/3\\ \sum_{j\in C(y,p^1)} w_j = 1/3 + 1/3 + 1/3 = 1\\ \sum_{j\in C(z,p^1)} w_j = 1/3 + 0 + 1/3 = 2/3\end{array}\right\} \Rightarrow \begin{array}{l} y \succ_{p^1} x\\ y \succ_{p^1} z\\ x \sim_{p^1} z\end{array}$$

p_2 :

$$\left.\begin{array}{l}\sum_{j\in C(x,p^2)} w_j = 1/3 + 1/3 + 0 = 2/3\\ \sum_{j\in C(z,p^2)} w_j = 0 + 0 + 0 = 0\end{array}\right\} \Rightarrow x \succ_{p^2} z$$

The final ranking is thus $y \succ x \succ z$, hence y is globally the best alternative, followed by x and then z. In other words, the drone will choose action y.

This small examples underlines two features of the decision model that are important for this application. First of all, the evaluations of the possible actions of the drone in each way-point might be of very different types (an ordinal risk level vs. a quantitative evaluation of remaining energy). This pleads in favour of an outranking model, where these evaluations do not have to be brought to the same scale. Second, if we want the operator to trust the drone's decisions, any decision it makes must be easily explainable. In our case, we can tell the operator that : action y is better than all other alternatives because it does not have any "bad" evaluations (C), while x and z have one "bad" evaluation on criterion M, respectively E. Also, action x is better than action z because it has "good" evaluations (A) on criteria R and E while z does not have any "good" evaluations.

4.3 Non-incremental elicitation of SRMP models

In [14] the authors propose to elicit the preference parameters of SRMP models from pairwise comparisons of alternatives given by the DM. They formulate SRMP preference elicitation as a mixed integer linear optimization problem (MIP). Furthermore, they propose Algorithm 1 which iteratively adjusts the number of reference points. It has as inputs (1) a set of pair-wise comparisons of alternatives $\mathcal{L} = \mathcal{P}r \cup \mathcal{I}n$, where $\mathcal{P}r$ (resp. $\mathcal{I}n$) is the set of pairs of alternatives where the DM strictly prefers the first alternative over the second one (resp. that the DM considers as indifferent), and (2) the evaluation of each alternative $a_j, j \in M, a \in A^*$. The solution of the algorithm is a set of SRMP parameters, together with the number of reference profiles k.

Algorithm 1 Procedure to elicit an SRMP model.

Input:
The sets of pairwise comparisons $\mathcal{L}$;
The evaluations of alternatives $g_j(a), j \in M, a \in A^*$;
Output:
k the number of reference profiles;
A set of reference profiles $p^1, p^2, \ldots, p^k$;
A lexicographic order on the reference profiles;
Criteria weights $w_j, j \in M$;

```
k ← 1
Problem solved ← false
while not Problem solved do
   for all lexicographic order on reference profiles σ do
      if all preference statements are restored by a S-RMP model
      using σ  then
         Problem solved = true
         Break
   k ← k + 1
```

The set of linear constraints that define a mixed integer linear program (MIP) used to infer an SRMP model is presented in Table 3. The program includes parameters that are related to the set of alternatives, the criteria, the number of profiles and the binary comparisons (preference / indifference) of alternatives provided by the decision maker (1). Some variables are used for SRMP parameters (weights (3) and profiles (4)) and other binary variables for the modelling of the constraints (5-7). The linear constraints listed in this program serve different purposes. The first ones, (8) and (9), ensure that the sum of the weights is equal to 1 and that each weight is greater then 0. The following three constraints ensure that the profiles evaluations are in a relation of dominance and that they are bounded within the [0,1] interval. Constraints 13 and 14 model are used to model $\delta_j(a, p^h)$ as a binary indicator for a being at least as good as profile p^h, while the following three constraints model $\omega_j(a, p^h)$ as the equivalent of $\delta_j(a, p^h) \cdot w_j$, for all $a \in \mathcal{A}^*$, $h = 1..k$ and $j \in M$. Constraints 21 to 23 are used to model the input of the DM in the form of a preference between $(a, b) \in \mathcal{P}r$, i.e. $a \succ b$. Finally, the last constraint is used to model the input of the DM in the form of an indifference between $(a, b) \in \mathcal{I}n$, i.e. $a \sim b$.

5 Heuristics for an incremental learning of SRMP models

In order to reduce the cognitive effort of the DM during the preference elicitation process, we propose here to study how an incremental learning of the SRMP model parameters can be performed. In practice this means that the DM participates in an iterative process, where at each iteration he / she expresses his / her preferences by answering a pair-wise comparison question: do you strictly prefer alternative a to alternative b, b to a, or are you indifferent between a and b? His / her answer is then added as a supplementary constraint in the mixed integer linear program which infers the model parameters. The number of such pair-wise questions consequently needs to be minimized and the pairs well selected.

For now, our work focuses on finding the heuristic of selecting the pairs of alternatives to present to the DM. We assume that the number of profiles and their lexicographic order are fixed in advance. We are thus using the Algorithm 1 presented in Section 4.3 with a given k. As mentioned in Section 2, we consider that an existing database contains all the feasible alternatives for our problem and so our task is to find a strategy of which pairs to select among them.

The proposed heuristics can be divided into two categories:

1. the ones related to the space of the alternatives, where the pair-wise questions are selected without using the model generated in the previous iteration,
2. the ones related to the space of the solutions (and thus the parameters) by considering the previous preferences of the DM in the choice of the next pair.

The alternatives are defined by their evaluations on different criteria, which can be wither qualitative, quantitative or both in real cases. However, to facilitate our experimental work here, we focus on the case where all evaluations inside the $[0, 1]$ numerical scale.

5.1 Random heuristic

This first heuristic that we wish to study simply selects sequentially and randomly a pair of alternatives in the set $\mathcal{L}$ containing the possible pairs.

$$(a, b)_{\text{random}} = \text{random}\{(a, b) \in \mathcal{L}\}$$

5.2 Heuristics using the distance between alternatives

In this category of heuristics, the DM is asked to compare either “similar” alternatives (having thus close evaluations on the criteria), or “dissimilar” ones (having thus very contrasted evaluations on the criteria).

In our context, these two heuristics can use an *L1 norm* $d_{(a,b)}$, defined for two alternatives a and b by

$$d_{(a,b)} = \frac{\sum_{j=1}^{m} |a_j - b_j|}{m}.$$

The selected pair is the one with smallest distance for the “similar” alternatives $(a, b)_{\text{similar}}$ and the largest distance for those how are “dissimilar” $(a, b)_{\text{dissimilar}}$.

$$(a, b)_{\text{similar}} = \operatorname*{argmin}_{(x,y)\in\mathcal{L}} d_{(x,y)}$$

$$(a, b)_{\text{dissimilar}} = \operatorname*{argmax}_{(x,y)\in\mathcal{L}} d_{(x,y)}$$

5.3 Heuristics using the distance between alternatives and the profile(s) of the current SRMP model

This heuristic is built in two steps. First we select a fixed number of pair-wise questions (e.g. 10% of the total number of pairs) randomly

Parameters:		
$A^*, \mathcal{L}, \mathcal{P}r, \mathcal{I}n, M, k, \sigma, \gamma$		(1)
$a_j \in [0,1]$	$\forall (a,b) \in \mathcal{L}, \forall j \in M$	(2)
Variables:		
$w_j \in [0,1]$	$\forall j \in M$	(3)
p_j^h	$\forall h \in 1..k, \forall j \in M$	(4)
$\delta_j(a, p^h) \in \{0,1\}$	$\forall a \in A^*, \forall h \in 1..k, \forall j \in M$	(5)
$\omega_j(a, p^h) \in [0,1]$	$\forall a \in A^*, \forall h \in 1..k, \forall j \in M$	(6)
$s(a, b, p^h) \in \{0,1\}$	$\forall (a,b) \in \mathcal{P}r, \forall h \in 0..k$	(7)
Constraints:		
$s.t. \sum_{j \in M} w_j = 1$		(8)
$w_j \geqslant \gamma$	$\forall j \in M$	(9)
$p_j^1 \geqslant 0$	$\forall j \in M$	(10)
$1 \geqslant p_j^k$	$\forall j \in M$	(11)
$p_j^{h+1} \geqslant p_j^h$	$\forall h \in 1..k-1, \forall j \in M$	(12)
$a_j - p_j^h \geqslant \delta_j(a, p^h) - 1$	$\forall a \in A^*, \forall h \in 1..k, \forall j \in M$	(13)
$\delta_j(a, p^h) \geqslant a_j - p_j^h + \gamma$	$\forall a \in A^*, \forall h \in 1..k, \forall j \in M$	(14)
$w_j \geqslant \omega_j(a, p^h)$	$\forall a \in A^*, \forall h \in 1..k, \forall j \in M$	(15)
$\omega_j(a, p^h) \geqslant 0$	$\forall a \in A^*, \forall h \in 1..k, \forall j \in M$	(16)
$\delta_j(a, p^h) \geqslant \omega_j(a, p^h)$	$\forall a \in A^*, \forall h \in 1..k, \forall j \in M$	(17)
$\omega_j(a, p^h) \geqslant \delta_j(a, p^h) + w_j - 1$	$\forall a \in A^*, \forall h \in 1..k, \forall j \in M$	(18)
$s(a, b, p^0) = 1$	$\forall (a,b) \in \mathcal{P}r$	(19)
$s(a, b, p^k) = 0$	$\forall (a,b) \in \mathcal{P}r$	(20)
$\sum_{j \in M} \delta_j(a, p^{\sigma(h)}) \geqslant \sum_{j \in M} \delta_j(b, p^{\sigma(h)}) + \gamma - s(a, b, p^{\sigma(h)}) \cdot (1 + \gamma) - (1 - s(a, b, p^{\sigma(h-1)}))$	$\forall (a,b) \in \mathcal{P}r, \forall h \in 1..k$	(21)
$\sum_{j \in M} \delta_j(a, p^{\sigma(h)}) \geqslant \sum_{j \in M} \omega_j(b, p^{\sigma(h)}) - (1 - s(a, b, p^{\sigma(h)})) - (1 - s(a, b, p^{\sigma(h-1)}))$	$\forall (a,b) \in \mathcal{P}r, \forall h \in 1..k$	(22)
$\sum_{j \in M} \omega_j(b, p^{\sigma(h)}) + (1 - s(a, b, p^{\sigma(h)})) + (1 - s(a, b, p^{\sigma(h-1)})) \geqslant \sum_{j \in M} \omega_j(a, p^{\sigma(h)})$	$\forall (a,b) \in \mathcal{P}r, \forall h \in 1..k$	(23)
$\sum_{j \in M} \omega_j(a, p^{\sigma(h)}) = \sum_{j \in M} \omega_j(b, p^{\sigma(h)})$	$\forall (a,b) \in \mathcal{I}n, \forall h \in 1..k$	(24)

Table 3. Mixed integer linear program for inferring the parameters of an S-RMP model given a lexicographic order σ;

from $\mathcal{L}$ to construct a first SRMP model M_0. Second, we select iteratively the following pair using the heuristic detailed bellow.

The concept of using a distance is also maintained here but used in a different context. Since we are using the current model M_i, i.e., the model generated in a previous iteration, we will use its parameters to select the next pair of alternatives. To reflect this, the distance between alternative $a \in A^*$ and a profile $p \in \{p, p, ..., p^k\}$ is a weighted distance defined as follows :

$$d_{(a,p)} = \frac{\sum_{j=1}^{m} w_j |a_j - p_j|}{m},$$

where w corresponds to criteria weights from model M_i.

Since we are dealing with a pair of alternatives of $\mathcal{L}$, we consider here the average distance $d_{(a,b),p}$ between $d_{(a,p)}$ and $d_{(b,p)}$:

$$d_{((a,b),p)} = \frac{(d_{(a,p)} + d_{(b,p)})}{2}$$

Two strategies can be mentioned here. The first one, where we select the closest pair $(a, b)_{\text{close}}$ to the profile and the second one, where the farthest pair is chosen $(a, b)_{\text{far}}$.

$$(a, b)_{\text{close}} = \underset{(x,y) \in \mathcal{L}}{\operatorname{argmin}}\, d_{((x,y),p)}$$

$$(a, b)_{\text{far}} = \underset{(x,y) \in \mathcal{L}}{\operatorname{argmax}}\, d_{((x,y),p)}$$

Once this selection is done for a given profile, the MIP is run, and in the next iteration, the next profile of the lexicographic order is chosen (until we reach the last one, after which we restart from the beginning of the lexicographic order).

5.4 Volume heuristic

The DM's answer to a pairwise comparison question is unpredictable. In this heuristic, in order to select the next pairwise comparison, we use an indicator which estimates the volume of the search space for the preferential parameters (i.e. the polyhedron of the constraints of the MIP). The answer to the pairwise comparison question "is a preferred to b or b preferred to a?" generates a supplementary constraint in the MIP. In most of the cases, this question splits the search space in two parts (one for the answer $a \succ b$, one for the answer $b \succ a$). If we estimate the volume of each of the two parts, the best pairwise comparison to choose from the set $\mathcal{L}$, without any further knowledge on the possible answer, is the one which splits the search space into two equal parts.

As the volume of this search space is too hard to compute, we approximate it, as in Leistedt [13] by the polyhedron of solutions based on the minimum and maximum values that the variables can take.

So for each possible pair of $\mathcal{L}$, we calculate these volumes, and determine in the end the one which splits the original space into two, mostly equal, parts.

This heuristic is obviously computationally very costly, as it requires to calculate the minimal and maximal values each variable can take for the current version of the MIP.

6 Experiments

In order to limit the number of computational experiments needed to validate the proposed approach and study the different strategies of choosing the pairs of alternatives to compare, we initially carry out an experiment seeking to study the relevance of having fewer or more reference profiles in an SRMP model. Given these limits, we then proceed to testing the aforementioned strategies.

6.1 Expressiveness of the SRMP model

In order to determine the expressiveness of the S-RMP models, we begin by generating a series of 100 SRMP models with $m \in \{3, 5, 7\}$ criteria. We additionally pick a large number of reference profiles, $k = 10$, in order to simulate the creation of preference models of high complexity. Using these models, we generate and compare a total of $5,000$ alternatives with randomly generated criteria evaluations. We illustrate in Table 4 the distribution of the number of profiles required for constructing the relations between these alternatives. We remind here that, each pair of alternatives is compared to the reference profiles one by one, stopping at the first profile in the lexicographic order which places the two alternatives in a relation of preference.

Table 4. Distribution of pairwise relations with respect to the number of profiles needed to construct them (in percentages).

k	m = 3		5		7	
1	67.15	(1.84)	82.21	(1.88)	89.21	(1.37)
2	19.39	(1.65)	13.97	(1.72)	8.93	(1.19)
3	6.60	(0.72)	2.86	(0.42)	1.61	(0.32)
4	3.09	(0.46)	0.55	(0.05)	0.19	(0.05)
5	1.29	(0.18)	0.22	(0.03)	0.04	(0.01)
6	0.89	(0.12)	0.09	(0.01)	0.01	(0.00)
7	0.48	(0.05)	0.04	(0.01)	0.00	(0.00)
8	0.30	(0.04)	0.03	(0.00)	0.00	(0.00)
9	0.20	(0.02)	0.01	(0.00)	0.00	(0.00)
10	0.62	(0.03)	0.02	(0.00)	0.00	(0.00)

We observe that, for problem instances with both few and many criteria, the first profile is able to discriminate between a large percentage of alternatives. Furthermore, as we consider problems with a larger number of criteria, this percentage also grows, up to almost 90% for problems with 7 criteria. In all cases, each subsequent profile is able to differentiate between fewer and fewer alternatives. We may therefore conclude that, for problems of this size, inferring models with more than 3 profiles does not offer any significant improvement with respect to the expressiveness of the model. For problems with more than 5 criteria we may even only need to consider SRMP models with at most 2 profiles.

In order to account for any bias caused by dominated alternatives, we further provide in Table 5 the percentage of pair-wise relations between dominated alternatives with respect to the number of profiles needed in order to discriminate them.

Table 5. Distribution of dominated pairs of alternatives with respect to the number of profiles needed to discriminate them (in percentages).

k	m = 3		5		7	
1	16.74	(0.49)	5.23	(0.11)	1.41	(0.02)
2	4.73	(0.38)	0.74	(0.10)	0.14	(0.02)
3	1.83	(0.22)	0.20	(0.03)	0.01	(0.00)
4	0.68	(0.09)	0.06	(0.01)	0.00	(0.00)
5	0.36	(0.04)	0.02	(0.00)	0.00	(0.00)
6	0.19	(0.02)	0.01	(0.00)	0.00	(0.00)
7	0.14	(0.02)	0.00	(0.00)	0.00	(0.00)
8	0.08	(0.01)	0.00	(0.00)	0.00	(0.00)
9	0.05	(0.00)	0.00	(0.00)	0.00	(0.00)
10	0.15	(0.01)	0.00	(0.00)	0.00	(0.00)

Unsurprisingly, for problems with fewer criteria we have a higher total percentage of alternatives that are dominated. We observe, however, that even in the case of dominated alternatives, we may require more than one reference profile in order to construct the preference relation between them, as denoted by the non-zero values below the first line in Table 5.

6.2 Incremental inference of an SRMP model

6.2.1 Experimental setting

We continue by presenting some experimental results related to the proposed heuristics using artificially generated datasets with different numbers of alternatives (n), criteria (m) and profiles (k). To do this, we implemented a test platform which integrates the different steps of the test process, starting from the data generation to the evaluation of heuristics.

The DM is simulated by a random SRMP model M_{DM} with a fixed number of reference profiles that are evaluated on the same criteria as the alternatives. The artificial dataset is decomposed into two subsets, the first one for the elicitation phase (the set of different pairs of alternatives $\mathcal{L}$) and the second for the test phase, composed of 100 alternatives, which will be used to make sure that the generated model is close to the original one.

The next step uses the same set of pair-wise comparisons $\mathcal{L}$ as input of the different heuristics. At each iteration the heuristic selects a pair of alternatives and assigns them a preference relation using M_{DM}. This then adds a new constraint in the MILP, which we use to construct a new SRMP model M_i. The test data is used to compare the original model M_{DM} and the current one M_i by ranking the 100 alternatives using both models, and then comparing these rankings using *Kendall*'s rank correlation measure τ. τ measures the correlation of two rankings, and varies between 1 and -1. If both rankings are identical then $\tau = 1$, while if they are completely different then $\tau = -1$.

This experimental protocol is executed using SRMP models of different sizes (number of criteria and the number of profiles) and for each across 100 different problem instances.

We have so far fixed the number of possible pair-wise comparisons to 45 (i.e. $|\mathcal{L}| = 45$) and the maximum number of criteria to 7 since a DM may find it difficult to compare alternatives evaluated on more criteria in practice (a limit of 6 criteria was determined in [3]). The analysis on the expressiveness of the SRMP models presented in 6.1 shows that we don't need more than 3 reference profiles for accurately modelling the DM's preferences.

6.2.2 *Results*

In Figure 2, we present the average of the Kendall τ and its standard deviation on the test data using 100 different independent executions of the proposed SRMP inference protocol, using problem instances of 1 profile and 3 criteria, and the "Similar" heuristic from Section 5.2.

What we can observe here is that on average the Kendall τ increases when adding new preferences of the DM (i.e. new pairs selected from $\mathcal{L}$) and that at the same time the variance decreases, which means that we have a higher precision towards the end of the experiments.

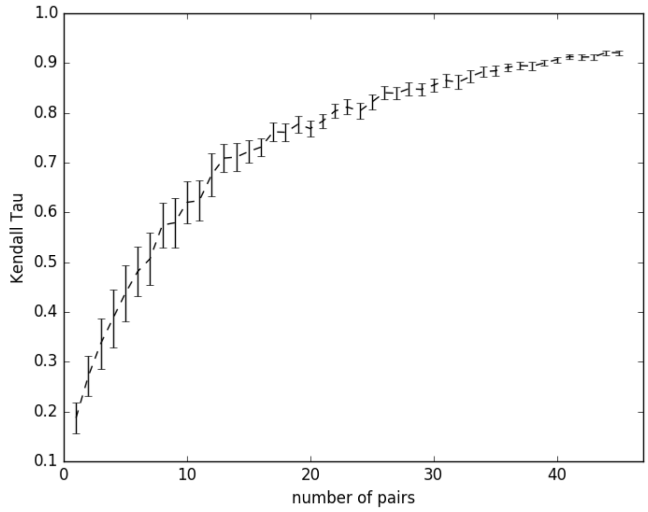

Figure 2. Average and standard deviation of the Kendall τ for the "Similar" heuristic, for $m = 3$ and $k = 1$

The average of the Kendall τ for the different heuristics presented in Section 5 with the same configuration as the previous experiment, 1 profile and 3 criteria, is presented in Figure 3. First we observe that the curves of the various heuristics "converge" to approximately the same average value when the number of pairs becomes large (above 25 pairs). This is explained by the fact that the selected set of learning pairs depends less on the heuristic, as more than half of the available pairs of the dataset is considered. Therefore our goal here is to determine a heuristic which performs better than the other ones, when the number of pairs is small. Let us also note that for the "Volume" heuristic, the data of Figure 3 is only based on 50 executions (due to limited time), which explains the more erratic tendency of this result. While it is difficult to determine a heuristic which clearly dominates the other ones, the "Volume" heuristic looks promising, as it starts with a high Kendall τ and seems to perform better than the other strategies during the first iterations of the learning protocol.

To make Figure 3 clearer we do not show the variances around the average values. However, they have the same behaviour as for Figure 2, which gives us an indication of the convergence of each heuristic for the same type of data.

In Figure 4 we have represented, for various numbers of criteria and profiles the performances of the various heuristics. Note here, that, due to computational time constraints, the "Volume" heuristic is only illustrated for problems with 3 criteria and 1 profile. For all the presented scenarios, we observe that the various represented heuristics have very similar behaviours.

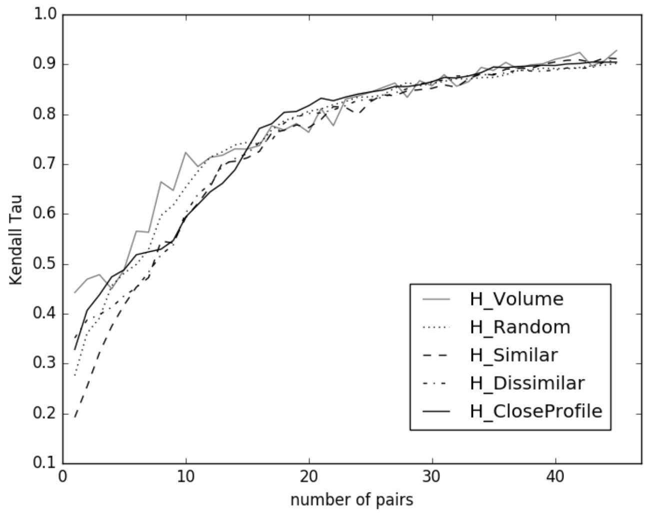

Figure 3. Average of the Kendall τ of all heuristics, for $m = 3$ and $k = 1$

We also observe that the slope of the curves of Kendall τ decreases if we increase the number of profiles (i.e. as seen in Figure 4). This means that models with 1 profile can reach a higher Kendall τ value with fewer pair-wise comparisons, as opposed to models with 2 profiles or more. If we focus on the curves with the same number of criteria but with different numbers of profiles (Figure 4), we notice that if the number of profiles increases the maximum value of Kendall τ decreases. This means that we need more than just 45 pair-wise comparisons for models with more than 1 profile in order to reach a Kendall τ value of around 0.9. We can also see in the same Figure 4 that the average of the starting values of the Kendall τ decreases when the number of criteria increases and the number of profiles stays the same.

7 Discussion and future work

The results presented in the previous section show that the differences in behaviour of the various heuristics that we defined in Section 5 is not influenced by the nature of data (e.g. number of profiles and criteria). We have almost the same curve for all the tested heuristics for the same type of data as shown in Figure 4. However, the volume heuristic seems to be promising, especially during the early steps of the incremental learning process.

One reason that may explain these results is that the studied heuristics do not reduce enough the space of possible solutions, which gives the MIP solver the possibility to choose a solution which is completely different from the one from the previous iteration.

Another observation that we can make is that with the current configuration of the experiment, the test data contains a lot of indifferent alternatives when ranked by the generated models. This might also have an impact on our results.

Currently the volume heuristic seems the most promising. However, due to complexity issues, this heuristic is very costly, and our results, at the time of writing, are only partial. The calculations are ongoing.

Future work will focus on trying to improve the complexity of the "volume" heuristic and its computation time. We also consider to refine the benchmark generation by reducing the amount of indif-

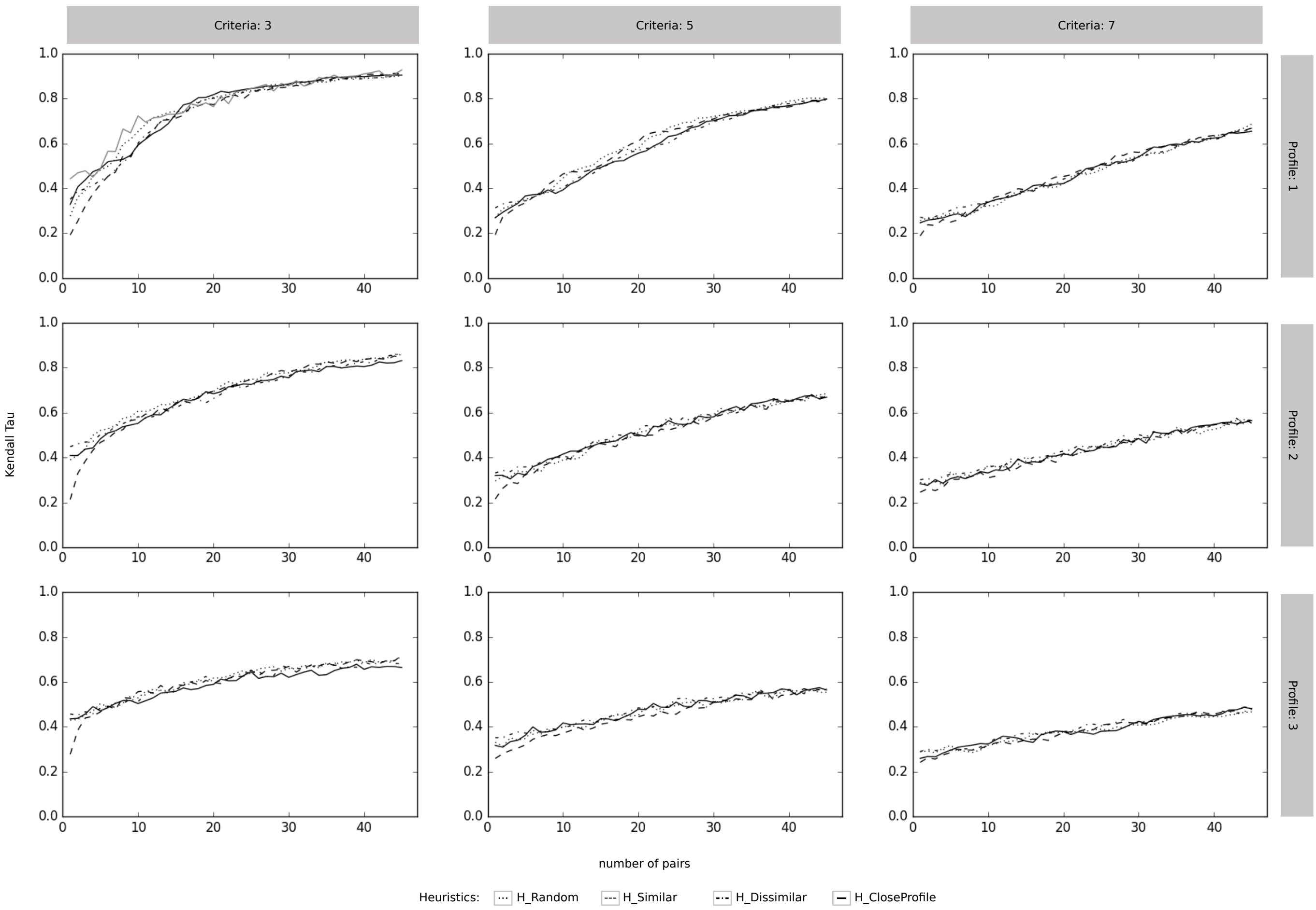

Figure 4. Average of Kendall τ of all heuristics with different numbers of criteria and profiles

ferent alternative pairs. Another aspect which can be considered, is to consider the topic of generating the pairs of alternatives to present to the DM on the spot, as opposed to selecting pairs from an existing database, as is imposed currently by our real-world application context.

Acknowledgement

This work was supported by LATERAL (Lab-STICC Thales Research Alliance).

REFERENCES

[1] N. Benabbou, P. Perny, and P. Viappiani, 'Incremental elicitation of choquet capacities for multicriteria choice, ranking and sorting problems', *Artificial Intelligence*, **246**, 152 – 180, (2017).

[2] K. Ciomek, M. Kadziński, and T. Tervonen, 'Heuristics for selecting pair-wise elicitation questions in multiple criteria choice problems', *European Journal of Operational Research*, **262**(2), 693 – 707, (2017).

[3] J.R. DeShazo and G. Fermo, 'Designing choice sets for stated preference methods: The effects of complexity on choice consistency', *Journal of Environmental Economics and Management*, **44**(1), 123 – 143, (2002).

[4] I.N. Durbach, 'The use of the SMAA acceptability index in descriptive decision analysis', *European Journal of Operational Research*, **196**(3), 1229 – 1237, (2009).

[5] J. Figueira and B. Roy, 'Determining the weights of criteria in the electre type methods with a revised simos' procedure', *European Journal of Operational Research*, **139**(2), 317–326, (2002).

[6] S. Greco, *Multiple Criteria Decision Analysis: State of the Art Surveys*, International Series in Operations Research & Management Science, Springer New York, 2006.

[7] S. Greco, M. Kadziński, V. Mousseau, and R. Słowiński, 'ELECTREGKMS: Robust ordinal regression for outranking methods', *European Journal of Operational Research*, **214**(1), 118 – 135, (2011).

[8] S. Greco, B. Matarazzo, and R. Słowiński, 'Dominance-based rough set approach to decision under uncertainty and time preference', *Annals of Operations Research*, **176**(1), 41–75, (2010).

[9] H.A. Holloway and C. C. White III, 'Question selection for multiattribute decision-aiding', *European Journal of Operational Research*, **148**(3), 525 – 533, (2003).

[10] E. Jacquet-Lagreze and J. Siskos, 'Assessing a set of additive utility functions for multicriteria decision-making, the uta method', *European Journal of Operational Research*, **10**(2), 151 – 164, (1982).

[11] R.L. Keeney and H. Raiffa, *Decisions with Multiple Objectives: Preferences and Value Trade-Offs*, Wiley series in probability and mathematical statistics. Applied probability and statistics, Cambridge University Press, 1993.

[12] R. Lahdelma, J. Hokkanen, and P. Salminen, 'SMAA - stochastic multiobjective acceptability analysis', *European Journal of Operational Research*, **106**(1), 137 – 143, (1998).

[13] B. Leistedt, *Selection of optimal learning sets for preference modeling and decision making*, Master's thesis, Faculty of Engineering, University of Mons, Belgium, 2011.

[14] A-L. Olteanu, K. Belahcene, V. Mousseau, W. Ouerdane, A. Rolland, and J. Zheng, 'Preference Elicitation for a Ranking Method based on Multiple Reference Profiles'. working paper or preprint, August 2018.

[15] A. Rolland, 'Reference-based preferences aggregation procedures in multi-criteria decision making', *European Journal of Operational Research*, **225**(3), 479–486, (2013).

[16] B. Roy, 'Méthodologie multicritère d'aide la décision', (01 1985).

[17] B. Roy, 'The outranking approach and the foundations of electre methods', *Theory and Decision*, **31**(1), 49–73, (Jul 1991).

[18] B. Roy, *Multicriteria Methodology for Decision Aiding*, Nonconvex Optimization and Its Applications, Springer US, 1996.

[19] B. Roy, *Multicriteria Methodology for Decision Aiding*, Kluwer Academic, Dordrecht, 1996.

[20] J. Simos, 'Evaluer l'impact sur l'environnement: Une approche originale par l'analyse multicritère et la négociation', in *Evaluer l'impact sur l'environnement: une approche originale par l'analyse multicritère et la négociation*, Presses polytechniques et universitaires romandes, (1990).

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