Integrating operators' preferences into decisions of unmanned aerial vehicles : multi-layer decision engine and incremental preference elicitation

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

[1] Lab-STICC, UMR 6285, CNRS, Université de Bretagne Occidentale, France
{arwa.khannoussi,catherine.dezan}@univ-brest.fr
[2] IMT Atlantique, Lab-STICC, Univ. Bretagne Loire, F-29238 Brest, France
{arwa.khannoussi,patrick.meyer}@imt-atlantique.fr
[3] Lab-STICC, Université Bretagne Sud, Lorient, France
{alexandru.olteanu, jean-philippe.diguet}@univ-ubs.fr
[4] Thales Group, France
{christophe.labreuche,jacques.petit-frere}@thalesgroup.com
[5] University of the West of England, Bristol, United Kingdom
pritesh.narayan@uwe.ac.uk

Abstract. Due to the nature of autonomous Unmanned Aerial Vehicle~~s~~ (UAV) missions, it is important that the decisions of a UAV stay consistent with the priorities of an operator, while at the same time allowing them to be easily audited and explained. We therefore propose a multi-layer decision engine that follows the logic of an operator and integrates its preferences through a Multi-Criteria Decision Aiding model. We also propose an incremental approach to elicit the operator's preferences, in view of minimizing his/her cognitive fatigue during this task.

Keywords: autonomous UAV · multi-layer decision engine · multi-criteria decision aiding · operator~~'s~~ preferences · traceable decisions.

1 Introduction

Autonomous Unmanned Aerial Vehicles (UAVs) are capable of carrying out various types of missions (military or civilian). Throughout the mission, they are facing multiple choices and have to make many decisions without any human input. This decision making task requires that multiple, potentially conflicting, criteria are taken into account in order to achieve the mission's and the operator's objectives. In addition, in order to increase his/her confidence in the UAV's behavior, its decisions should be consistent with the priorities of the operator.

A lot of research deals with the decisions of autonomous UAVs through the perspective of trajectory calculation taking into account different constraints and objectives, by using optimization techniques [4, 13, 2]. Recently, deep learning techniques have also been used to tackle decision problems of autonomous

UAVs, as, e.g., the path planning problem [6] or the selection of high level directives [24, 15]. In [17] the authors suggest to integrate the operator's perspective into the calculation of trajectories for autonomous UAVs, and propose to use techniques from the field of Multi-Criteria Decision Aiding to model the operator's preferences. They ground their proposal on the hypothesis that an operator will trust the behavior of an autonomous UAV if it makes decisions that are consistent with his/her priorities.

We start from the same observation as [17], but propose to integrate the preferences of the operator even more deeply in the various decision making tasks of autonomous UAVs. We therefore develop in this contribution :

- a multi-layer decision engine for autonomous UAVs, which mimics the logic adopted by operators during a non-autonomous mission,
- the integration of a Multi-Criteria Decision Aiding model, called SRMP [19], into this decision engine, which allows the autonomous UAV to select the appropriate high-level action to be executed during the mission,
- an incremental preference elicitation approach to tune the SRMP decision model according to the preferences of the operator, while minimizing his/her cognitive fatigue during the learning process.

We also demonstrate the interest of our proposal through a simulator, in which we can test the influence of different operator profiles on the UAV's behaviour.

The article is structured in the following way. Section 2 provides a state of the art on UAVs and their decision making processes, as well as existing work on Multi-Criteria Decision Aiding and incremental preference elicitation. Section 3 presents an overview of the proposed multi-layer decision making engine while a detailed description of the considered preference model is given in Section 4, next to our proposal for an incremental preference elicitation procedure. A validation of our proposal, through a UAV simulator that we have developed is then presented in Section 5, before finishing with some concluding remarks and perspectives for future work in Section 6.

2 State of the art

2.1 Decisions in autonomous UAVs

From a general point of view, plenty of work focuses on trajectory calculation while taking into account different constraints and objectives. Blackmore et al. [4] present an approach to calculate the optimal trajectory in the presence of obstacles and uncertain information, while Kabama et al. [13] illustrate and approach to calculate the optimal trajectory for combat UAVs by avoiding radars. Some approaches also address the calculation of the trajectory in a non-convex environment with uncertainties [2]. All these methods construct objective functions that integrate the different quality measures of the solution in order to find the optimal trajectory, however, the weighting of these measures is generally done more or less arbitrarily.

Delmerico et al. [6] propose to use deep learning techniques, and more specifically Convolutional Neural Networks (CNNs), for path planning in the context of collaborative search and rescue missions. Deep learning is also used to present a solution for UAV localization and cross-view localization of images [25]. Concerning the UAV navigation task, some advances have led to the application of CNNs in order to map images to high-level behavior directives (e.g., turn left, turn right, rotate left, and rotate right) [24, 15]. Due to resource limitations, the learned model is executed off-board, on the GPU of an external laptop. The work presented by Tipaldi and Glielmo [27] integrates Markovian Decision Processes (MDP) for spacecraft reconfiguration in order to deal with the uncertainty in the outcome of actions and is applied to autonomous mission planning and execution.

Using a drone with a high level of autonomy to perform a mission requires that the human operator has a high degree of confidence in the capacity of the drone to make the "right" decisions. This observation motivated Narayan et al. [17] to integrate preferences into the calculation of the objective function in order to generate trajectories that more accurately represent the preferences of an operator (and therefore may differ from one operator to another). They use a decision model from the field of Multi-criteria Decision Aiding.

2.2 Multi-criteria Decision Aiding and (incremental) preference elicitation

In this article, we start from the same observation as [17], but propose to integrate the operator's preferences into higher level decisions rather than the calculation of trajectories, as, e.g., the choice between decision actions that the drone has to perform, as landing, returning to the base, aborting the mission, skipping a waypoint, etc.

Multi-Criteria Decision Aiding (MCDA) [22, 20] is the study of decision problems, methods and tools which may be used in order to assist a decision maker (DM) in reaching a decision when faced with a set of so-called alternatives (or decision actions), described via multiple, often conflicting, criteria. Various methodologies and preference models have been proposed to support DMs facing a multi-criteria decision problem. *Outranking methods* [21] compare any two alternatives, based on the preferences of the DM on the set of criteria, using a majority rule. Alternately, methods based on *multi-attribute value theory* (MAVT) [14] aim to construct a numerical representation of the DM's preference on the set of alternatives. In our autonomous UAV context, the DM is the operator, whose preference model is integrated into the drone and is thus guiding its decisions.

The preference parameters of MCDA decision models can be given directly by the DM through a direct preference elicitation approach. However, such an approach is usually too difficult to implement in practice, as the DM needs to have a very good understanding of the MCDA model. Therefore, a second approach is to start from 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. This second approach so-called *indirect* preference elicitation has received much attention from researchers, as for example in the seminal works of Jacquet-Lagreze et.al. [12] in the MAVT context and of Ngo The et.al. [26] in the outranking context. These techniques generally determine in one shot a parameters configuration compatible with the input provided by the DM, and are therefore *not incremental* by nature.

Incremental learning focuses on learning the parameters of a decision model in a streaming setting. Incremental learning algorithms receive learning data sequentially, one by one or chunk by chunk, and use this data with the previously learned model to produce a new, better one, that encapsulates information held by the data seen so far. Regarding this progressiveness, in the MAVT context, Durbach [8] and Lahdelma et.al. [16] 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. [11] 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. [5] 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. [3] 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).

An incremental learning of the parameters of MCDA models should reduce the cognitive effort of the DM, as he/she is facing only to a limited number of questions. As we will show in this article, the decision model that we are integrating into autonomous UAVs is learned incrementally before the mission, and it is important that the operator is not overly stressed during this phase.

3 Onboard multi-layer decision engine

The starting point of our proposal is the hypothesis that an operator will trust the behavior of an autonomous UAV if it makes decisions, which are consistent with his/her priorities. Furthermore, we need a model which can be easily explained and whose outcomes (decisions) are easily interpretable, so that the operator can validate the decision engine implemented in the UAV.

Consequently we focus on the logic adopted by the operator during a mission, in order to define the model of the autonomous decision engine. We first suppose that, in a non-autonomous context, the human operator does not make decisions continuously during the flight, but that the decision making act is triggered by events (e.g. the appearance of an obstacle, a breakdown, a change in the weather conditions . . .). Second, still in a non-autonomous context, we also suppose that when operators have to deal with a complex decision, triggered by an event, they tend to decompose it into a sequence of sub-decisions. Consequently, in case of such an event, the operator will take into consideration possible trajectories (i.e. which is a sub-decision) while choosing a high level action (e.g. land, continue the mission, skip a waypoint, . . .).

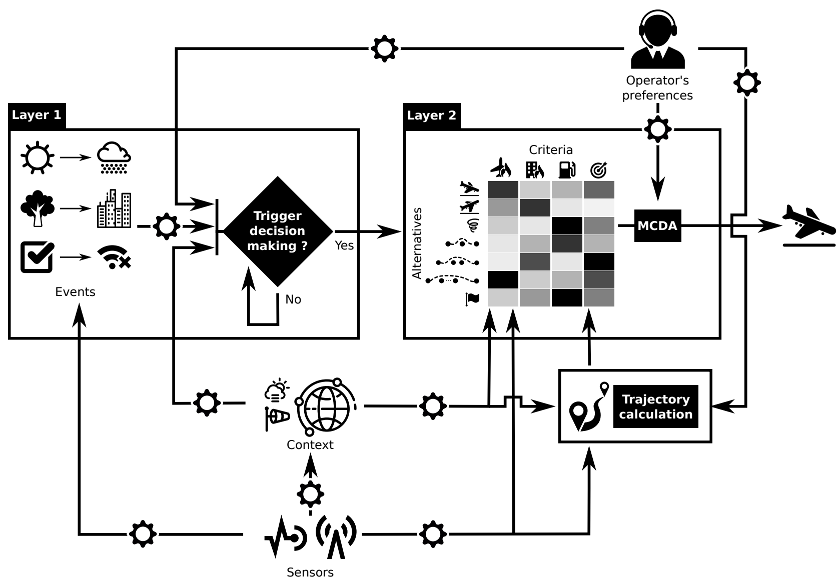


Fig. 1. Proposed Multi-layer decision engine

Following this reasoning, we propose to decompose the decision-making process of the autonomous UAV during the mission into two layers (Figure 1).

Layer 1 consists in the constant monitoring of the progress of the mission and all the information that might impact its success. Based on the occurrence of certain events, the second layer may be triggered. These events could be related to the UAV's environment (e.g. a change of the flight zone, the appearance of an obstacle, the detection of heavy rainfall, a mechanical breakdown, . . .) or to risk levels (e.g. exceeding a certain threshold of risk towards the drone, the mission, or the environment, . . .). While we do not tackle this topic in our current proposal, a series of rules or even a preference model, that is previously tuned to the perspective of an operator, may be integrated into the drone.

Layer 2 consists in determining which high-level action (e.g. *takeoff*, *continue*, *skip* one or more waypoints, *return to base*, *loiter*, *land*, . . .) is the best answer to the risks generated by the event from the first layer. The evaluation of these actions is supplied by the context [1] but also by the trajectories that the drone will take. The context, which is taken into account within the second layer, is a set of elements that describes the UAV's environment. It can include information about the mission, its objective (e.g. to monitor a target, protect a convoy, ...) or its current state. Other information related to the drone are also included in this context which are given by the UAV's sensors. The sensors' outputs can be used directly (e.g. GPS coordinates, altitude) or they can be processed before, while other information regarding its surroundings can also be used (flight zone map, invisibility zone, weather conditions). A trajectory calculation module is also included, and can, for example, be implemented through the work of [17], where an additive MAVT method is used to compute the best trajectory by respecting multiple criteria and is based on the operator's preferences.

In our case, we have retained four such elements, or criteria, but they could be more diverse, depending on the mission:

- energy consumption [9], corresponding to the amount of energy left after executing the selected action,
- risk to the drone [23], i.e. the risk associated with flying over different areas such as forests, sea, military zones,
- risk to the environment, such as people, buildings, in the case of a crash,
- mission progress, which is a weighted percentage of the achieved sub-objectives.

Each operator may view these elements, i.e. the consequences of the possible decision actions, differently. As a result, a model of the operator's preferences has to be constructed prior to the mission. We propose to rank the different decision actions with respect to their evaluations on the multiple criteria and the preferences of the operator by integrating a multi-criteria decision model.

4 The multi-criteria decision aiding model

As already mentioned, the DM is the operator, whose preferences must be modeled before including them in the decision engine of the autonomous drone. In this section, we focus more specifically on layer 2, and show how an MCDA model can help the drone to make high level decisions, by considering the operator's preferences. To guarantee a certain level of trust in the UAV's decision making process, the preference model and its consequences should be presented to the operator in order to be validated beforehand. It is therefore of high importance that this model is easy to explain to a non-expert of MCDA and that the decision recommendations (the recommended UAV actions which will ultimately influence the UAV's behavior) can be easily justified. We chose to implement a method from the outranking paradigm of MCDA methods, called SRMP (Simple Ranking Method using Reference Profiles). This choice is motivated by the following reasons:

- both qualitative and quantitative criteria can be easily integrated;
- the output always corresponds to a pre-order of the alternatives (thus a transitive relation);
- the output can be easily explained through a series of rules that can be audited by the operator.

The last point is of particular interest due to the critical nature of the decisions that an autonomous drone must make during operation.

4.1 Simple Ranking Method using Reference Profiles (SRMP)

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" [21] 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 [19] 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.

The SRMP method is defined by several preference parameters which need to be identified beforehand. These parameters are:

- the reference profiles: $\mathcal{P} = \{p^h, h = 1, \dots, k\}$ where $p^h = \{p^h_1, \dots, p^h_j, \dots, p^h_m\}$ corresponds to the evaluations of p^h on all criteria and $p^h_j \succsim_j p^l_j, \forall h, l \in \{1, \dots, k\}, h > l$, and $\succsim_j$ representing the preferential pre-order on the values of criterion j;
- the lexicographic order of the profiles: σ, which corresponds to a permutation on $1, \dots, k$;
- the criteria weights: $w_1, w_2, \dots, w_m$, where $w_j \geq 0$ and $\sum\limits_{j \in M} w_j = 1$

SRMP consists in a three-steps procedure as follows:

1. compute $C(a, p^h) = \{j \in M : a_j \succsim_j p^h_j\}$ with $a \in \mathcal{A}$, $h = 1, \dots, k$, 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:
   - $a \succ_{p^h} b \Leftrightarrow \sum\limits_{j \in C(a_i, p^h)} w_j > \sum\limits_{j \in C(b, p^h)} w_j$
   - $a \sim_{p^h} b \Leftrightarrow \sum\limits_{j \in C(a_i, p^h)} w_j = \sum\limits_{j \in C(b, p^h)} w_j$
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:

$$
\begin{aligned}
&(a \succ_{p^{\sigma(1)}} b) \text{ or} \\
&(a \sim_{p^{\sigma(1)}} b \text{ and } a \succ_{p^{\sigma(2)}} b) \text{ or} \\
&\dots \\
&(a \sim_{p^{\sigma(1)}} b \text{ and } \dots \text{ and } a \sim_{p^{\sigma(k-1)}} b \text{ and } a \succ_{p^{\sigma(k)}} b)
\end{aligned}
$$

   - a is indifferent to b iff: $\quad a \sim_{p^{\sigma(1)}} b$ and $\dots$ and $a \sim_{p^{\sigma(k)}} b$

4.2 Illustrative example

Let us show on a small example how the UAV could use this SRMP model to make decisions. Imagine that the UAV has to select among 3 high level actions x, y and z, like for example "land", "loiter" and "skip a waypoint", once the second layer of our decision engine has been triggered.

Table 1. Evaluations of the decision actions and SRMP parameters.

	R	E	M		R	E	M
x	low	80%	20%	p^1	low	70%	80%
y	high	55%	90%	p^2	high	35%	40%
z	medium	20%	50%				
		σ	$\{1,2\}$				

With each of these actions, a trajectory is associated, which has been calculated beforehand by the trajectory calculation module. The three actions are evaluated on three criteria, the risk (R), the energy consumption (E) and the mission progress (M), and the result is presented in Table 1. The preference parameters of the SRMP model, which model the preferences of an operator, are also given in this table. They have been learned from a prior preference elicitation process such as the one we will present in Section 4.3.

The two reference profiles allow to define three intervals on the performances on each criterion: better than p^2; between p^1 and p^2; worse than p^1. This allows to identify intervals of performances as illustrated in Figure 2, such that:

- "good" performances are above p^2,
- "intermediate" performances are between p^1 and p^2 on each criterion,
- "insufficient" performances are below p^1 on each criterion.

Let us now follow the steps presented earlier to rank the three alternatives x, y and z. First we compute $C(a,p^h)=\{j\in M : a_j \geqslant p^h_j\}$, $\forall a\in\mathcal{A}=\{x,y,z\}$, $h\in\{1,2\}$, $M=\{R,E,M\}$ and then compare each alternative to the others by using the profiles p^h, and finally rank the alternatives by considering the lexicographic order $\sigma=\{1,2\}$

p^1 :

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

p_2 :

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

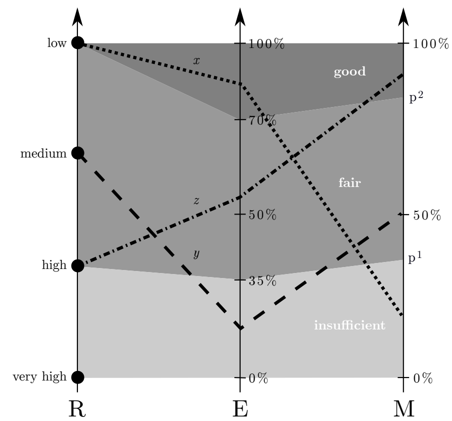


Fig. 2. SRMP example

The final ranking is thus $y \succ x \succ z$, hence y is globally the best alternative, followed by x and then z. This result can be explained to the operator in the following way: y is better than all other alternative because it does not have any "insufficient" evaluations, while x and z have one "insufficient" evaluation on criterion M, respectively E; x is better than z because it has "good" evaluations on criteria R and E while z does not have any "good" evaluations. The drone will thus implement decision y and its corresponding trajectory.

4.3 Incremental preference elicitation for SRMP models

Olteanu et al. propose in [18] to learn the preference parameters of SRMP models from a set of pairwise comparisons of alternatives given by a DM in one iteration. They thus formulate SRMP preference elicitation as a mixed integer linear program (MIP), and show that to obtain an expressive preference model, this learning algorithm requires quite a few pairwise comparisons of alternatives as inputs.

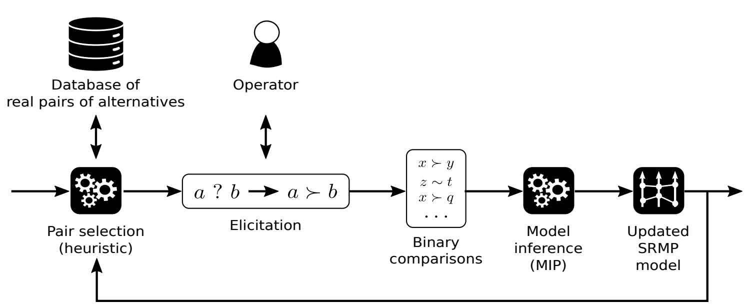


Fig. 3. Incremental learning process

In order to reduce the cognitive effort of the operator during the preference elicitation process, we propose an *incremental learning process* for SRMP models presented in Figure 3, which should reduce the number of pairwise comparisons of alternatives that the operator has to evaluate. This process is performed before the mission, and its output (the preference parameters of the SRMP model) is then integrated into the second layer of our decision engine to configure the SRMP algorithm presented in Section 4.1. The input of the learning process is a database $\mathcal{D}$ of pairs of alternatives / decision actions (typically, two actions among which the autonomous UAV would have to choose during a mission). At each iteration a heuristic selects a pair of alternatives (a, b) from $\mathcal{D}$ and the operator 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 MIP which infers the new SRMP model parameters. This procedure is repeated, as depicted by the continuous arrow until a "good enough" preference model is obtained.

The selection heuristic that we propose in this article, and which we name $\mathcal{H}_{\text{mp}}$, works as follows. The first iteration is a random selection of a pair of alternatives from $\mathcal{D}$. Then, at each iteration i, we use the preference model M_{i-1} generated in the previous iteration to select the next pair of alternatives. The idea is to select a pair which, in the current model M_{i-1} uses the highest possible number of profiles in its comparison (ideally a pair considered as indifferent by M_{i-1}). By confronting the operator with such a pair, we hope that his/her answer will generate a new constraint for the MIP which will reduce the size of the search space, and thus the possible values that the preference parameters may take.

We start by selecting a pair (a, b) that is indifferent using M_{i-1}. This means that for (a, b) all the k profiles have been tested in the SRMP procedure, and model M_{i-1} was not able to say whether a is preferred to b or b is preferred to a. If there is no such indifferent pair, a pair that uses k profiles in M_{i-1} to be discriminated will be selected. If no such pair exists, we search for a pair using $k-1$ profiles, and so on, until reaching the case of one profile. More formally, the $\mathcal{H}_{\text{mp}}$ selection heuristic selects a pair (a, b) from $\mathcal{D}$ such that:

$$
\begin{aligned}
&(a \sim_{p^{\sigma(1)}} b \text{ and } \dots \text{ and } a \sim_{p^{\sigma(k)}} b) \text{ or}\\
&(a \sim_{p^{\sigma(1)}} b \text{ and } \dots \text{ and } a \sim_{p^{\sigma(k-1)}} b \text{ and } a \succ_{p^{\sigma(k)}} b) \text{ or}\\
&\dots\\
&(a \sim_{p^{\sigma(1)}} b \text{ and } a \succ_{p^{\sigma(2)}} b) \text{ or}\\
&(a \succ_{p^{\sigma(1)}} b).
\end{aligned}
$$

In order to validate empirically that the $\mathcal{H}_{\text{mp}}$ selection heuristic allows to find a good preference model with a limited number of pairwise comparisons, we perform some experiments, and compare it to a random selection of the pairs from the database (we call this heuristic $\mathcal{H}_{\text{rnd}}$).

These experiments follow the incremental learning process presented in Figure 3 with an additional test phase to evaluate the quality of the obtained SRMP model. A database $\mathcal{D}$ of 100 pairs of alternatives is used as input for the proposed heuristic. $\mathcal{H}_{\text{mp}}$ selects a pair of alternatives from $\mathcal{D}$ at each iteration. The operator of Figure 3 is replaced for our experiments with a randomly generated SRMP model M_{OP}. It is used to compare pairs of alternatives, which in turn generate a new constraints for model M_i.

To test the quality of a model, we generated a test database D_{test} composed of 5000 alternatives. These alternatives are ranked both by the original SRMP model M_{OP} and the current one M_i. The quality of M_i is then evaluated through *Kendall*'s rank correlation measure τ between these two rankings. 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 inverted then $\tau = -1$.

We execute this process for 100 different artificial databases $\mathcal{D}$, composed each of 100 pairs of alternatives, for different problem sizes ($m = 3, 5, 7$). We also fix the number of profiles to 2. This generates 3 problem configurations which we call (2P 3C), (2P 5C) and (2P 7C) (for k profiles and m criteria).

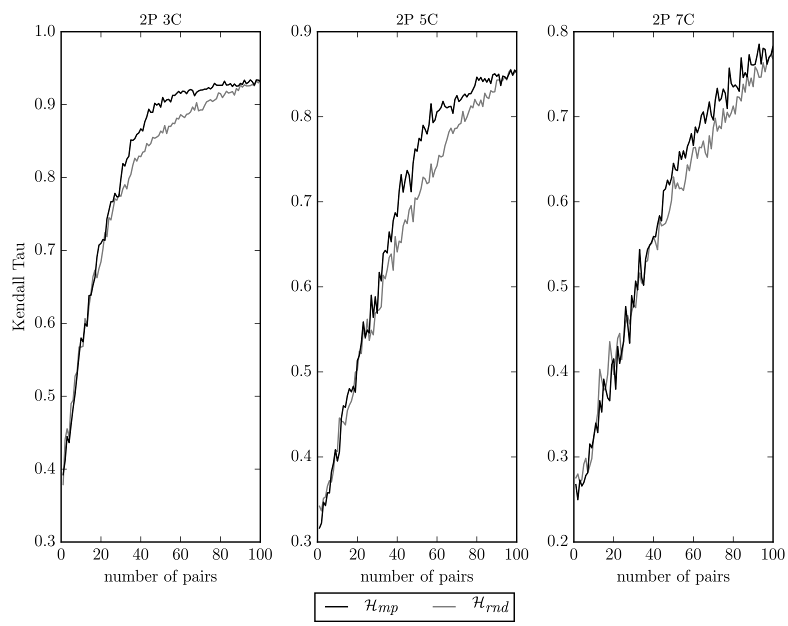


Fig. 4. Average Kendall tau for 2P 3C, 2P 5C and 2P 7C [6]

[6] At the time of writing the tests for the larger sizes of problems have not finished yet. Consequently, in the figure, the number of tests for 2P 3C equals 100, for 2P 5C it equals 87 and for 2P 7C it equals 53.

Each of the plots of Figure 4 depicts the average value of the *Kendall tau* across the 100 different artificial databases as a function of the number of pairwise comparisons submitted to M_{OP}. For example, for problems containing 3 criteria, after asking the M_{OP} to compare 40 pairs of alternatives, selected with the $\mathcal{H}_{\text{mp}}$ heuristic, we can obtain on average a preference model which ranks the test data quite similarly to the way M_{OP} ($\tau \sim 85\%$).

What we can observe here is that the Kendall τ increases when adding new preferences of the operator (i.e. pairs of alternatives). The plots show that the $\mathcal{H}_{\text{mp}}$ heuristic dominates the $\mathcal{H}_{\text{rnd}}$ one, which is confirmed by *Kolmogorov-Smirnov* statistical tests allowing to compare two samples [7]. We also notice that for the first few iterations both curves ($\mathcal{H}_{\text{mp}}$ and $\mathcal{H}_{\text{rnd}}$) behave similarly (for the different problem sizes), which is due to the small number of learning pairs involved, and which tend to produce not very expressive SRMP models. Then, both curves separate clearly in favour of $\mathcal{H}_{\text{rnd}}$. For the last few iterations, the curves become again similar, which is explained by the fact that the set of learning pairs is almost the same, independently of the selection heuristic ($\mathcal{D}$ is finite and fixed to 100 pairs of alternatives).

The standard deviations associated with the average values depicted in these figures are small (on average $\sim$ 0.075 for 2P 3C for both $\mathcal{H}_{\text{mp}}$ and $\mathcal{H}_{\text{rnd}}$, $\sim$ 0.095 for 2P 5C for both $\mathcal{H}_{\text{mp}}$ and $\mathcal{H}_{\text{rnd}}$, and $\sim$ 0.109 for 2P 7C for both $\mathcal{H}_{\text{mp}}$ and $\mathcal{H}_{\text{rnd}}$) and they also decrease with the addition of more learning pairs. They have not been included in these illustrations for these reasons.

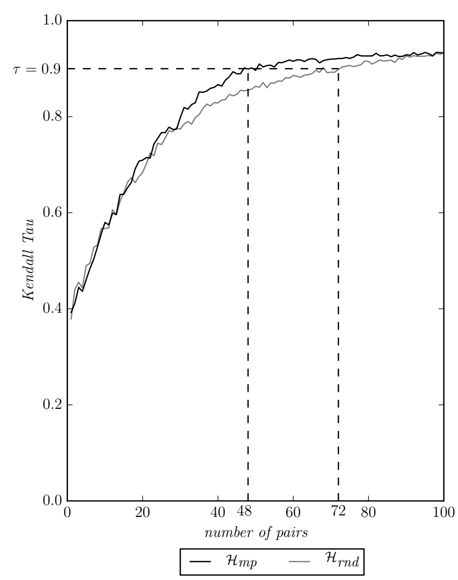


Fig. 5. Average Kendall Tau for 2P 3C

This result can be used in a practical case and gives an answer to the research question which is how many learning pairs / iterations are needed to achieve a *"good"* SRMP model with an objective to reduce the cognitive effort of the operator. An SRMP model is considered as a *"good"* one by the operator if this model can reach a given *Kendall tau* value. Once this valus is given we use the average *Kendall tau* curves to find the number of pairs required to reach τ.

For example, Figure 5 depicts the average *Kendall tau* for *2P 3C* where the operator fixed $\tau = 0.9$ we can see that we need about 48 learning pairs by using the $\mathcal{H}_{\mathrm{mp}}$ heuristic while about 72 for the $\mathcal{H}_{\mathrm{rnd}}$ heuristic.

5 Experimental validation of the decision model

In order to validate our proposal and show the importance of integrating the preferences of the operator into the automated decisions of a UAV, we developed an UAV simulator. It simulates the flight of an UAV which contains the previously presented decision engine. The graphical user interface (GUI) of the simulator is presented in Figure 6.

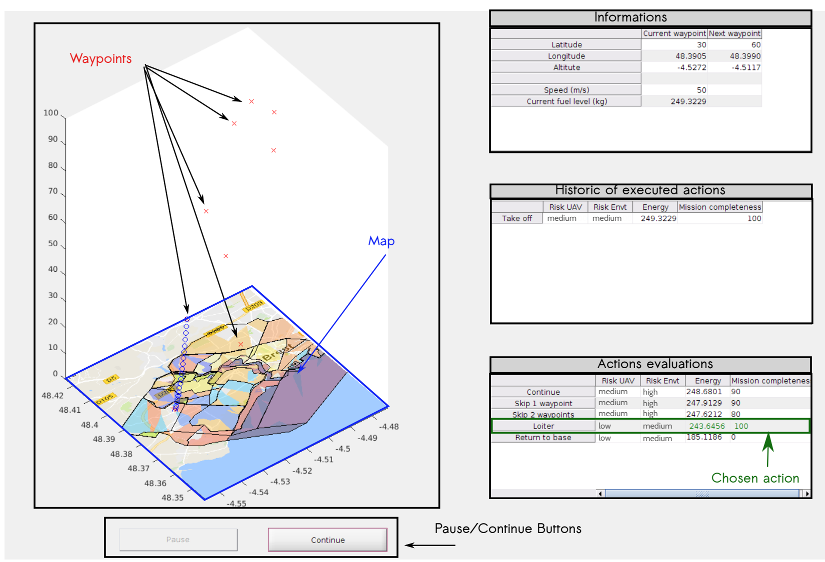


Fig. 6. Graphical User Interface for the autonomous UAV simulator

The modeled autonomous UAV is a Watchkeeper Unmanned Aircraft System from Thales [10], represented by a point which is submitted to physical constraints. The simulated UAV is able to navigate through a set of waypoints and execute different high-level actions (e.g. take off, loiter, ...). The simulator is also able to evaluate these actions on different criteria presented in Section 3.

The GUI presented in Figure 6 is composed of four parts. The left window plots the details of the mission, such as the different waypoints (as red crosses), the mission map as well as two maps of the risk associated with the UAV and the environment respectively. They allow the evaluation of the risk of a trajectory using a weighted average of the risk of the different zones overflown by the drone. The top right box provides information about the current waypoint and the next ones, as well as information on the current speed and the amount of energy left. The middle right box presents a history of the executed actions. Finally, the bottom right box shows the evaluations of all the possible high-level actions for the current waypoint. The SRMP model is executed in the background in order to decide which action will be chosen next (highlighted in green).

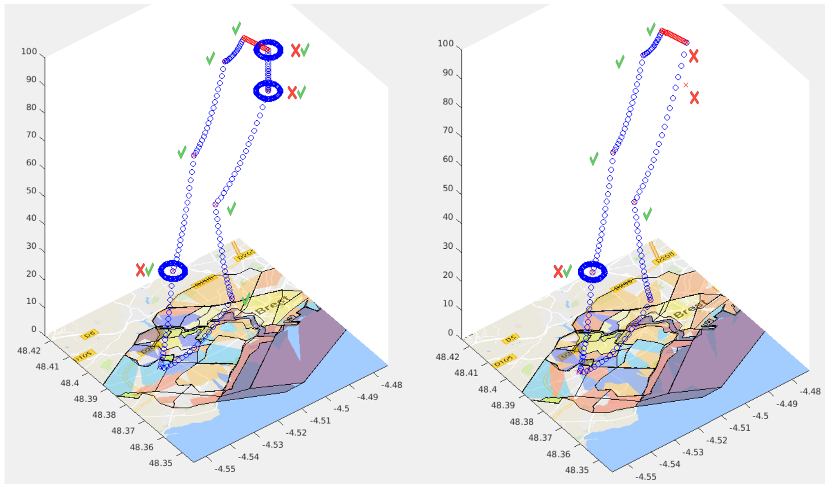
To illustrate our work, we provide here an example, where the UAV has to accomplish a mission consisting of flying through a set of nine waypoints and taking photos at each of them. We suppose that for waypoints 1, 5 and 6, these photos are missed, which requires the UAV to loiter for a second shot in order to complete the mission at 100%. We execute the mission according to two different operator profiles, represented by two different sets of preference parameters.

	operator 1				operator 2			
	R_{UAV}	R_{Env}	E	M	R_{UAV}	R_{Env}	E	M
p^1	high	v.high	30%	30%	high	v.high	30%	30%
p^2	low	medium	60%	99%	low	medium	60%	70%
w	0.1	0.1	0.1	0.7	0.25	0.25	0.25	0.25
σ	$\{2,1\}$				$\{2,1\}$			

Table 2. Preference parameters for the two operators.

The first operator is mainly focusing on completing the mission, placing as secondary objectives the risk and the fuel consumption. The incremental preferences elicitation phase (Section 4.3) leads to the SRMP parameters presented in the left half of Table 2. The second second operator gives a more uniform importance to the mission completeness, risks and fuel consumption objectives. The preference parameters representing this operator are summarized in the right half of Table 2, and have again been obtained using our incremental elicitation process.

As expected the UAV configured with first operator's preferences accomplishes the mission with success, as shown in Figure 7 on the left, by flying through all the waypoints and loitering at waypoints 1, 5 and 6 in order to take another round of photos, without taking into account the risk linked to the underlying zones. The right side of Figure 7 shows the execution of the mission with respect to the preferences of the second operator. We can observe that the UAV, even if only one photo was taken at waypoints 5 and 6, did not loiter



Operator 1 Operator 2

Fig. 7. Mission simulation with the preferences of operators 1 and 2

(because it considered it too risky to fly again over the same zone), and even skipped waypoint 6 (for the same reason).

6 Conclusion and perspectives

In this work, we propose a new approach for integrating an operator's perspective within the decision engine of autonomous UAVs, through a multi-layer decision engine, a traceable MCDA technique and an incremental process which minimizes the cognitive effort of the operator during the preference elicitation. Dividing the decision process of the autonomous UAV into several layers allows us to integrate the perspective of the operator in different elements of the autonomous decision making process, and thus provides an autonomy of the UAV guided by the preferences of a human operator.

Depending on the characteristics of the decision problem (as for example the number of considered criteria), the resolution of the MIP which is used iteratively in the elicitation process can take some time. This could limit its use in practice in an incremental elicitation process, which motivates us to study in a next step approximate algorithms (meta-heuristics) for the determination of the parameters of the SRMP model. Next to that, we wish to confront the incremental elicitation process with real operators, to validate our approach from a practical point of view, before integrating our decision engine into a real UAV.

References

1. Ajami, A., Balmat, J.F., Gauthier, J.P., Maillot, T.: Path planning and ground control station simulator for uav. In: 2013 IEEE Aerospace Conference. pp. 1–13

(March 2013)
2. Arantes, M.S., Arantes, J.S., Toledo, C.F.M., Williams, B.C.: A hybrid multi-population genetic algorithm for UAV path planning. In: Proceedings of the Genetic and Evolutionary Computation Conference 2016. pp. 853–860. GECCO '16, ACM, New York, NY, USA (2016)
3. Benabbou, N., Perny, P., Viappiani, P.: Incremental elicitation of choquet capacities for multicriteria choice, ranking and sorting problems. Artificial Intelligence **246**, 152 – 180 (2017). https://doi.org/https://doi.org/10.1016/j.artint.2017.02.001, http://www.sciencedirect.com/science/article/pii/S0004370217300152
4. Blackmore, L., Ono, M., C. Williams, B.: Chance-constrained optimal path planning with obstacles. IEEE Transactions on Robotics **27**(6), 1080–1094 (Dec 2011)
5. Ciomek, K., Kadziński, M., Tervonen, T.: Heuristics for selecting pair-wise elicitation questions in multiple criteria choice problems. European Journal of Operational Research **262**(2), 693 – 707 (2017). https://doi.org/https://doi.org/10.1016/j.ejor.2017.04.021, http://www.sciencedirect.com/science/article/pii/S0377221717303545
6. Delmerico, J., Mueggler, E., Nitsch, J., Scaramuzza, D.: Active autonomous aerial exploration for ground robot path planning. IEEE Robotics and Automation Letters **2**(2), 664–671 (April 2017)
7. Dodge, Y.: Kolmogorov–Smirnov Test, pp. 283–287. Springer New York, New York, NY (2008)
8. Durbach, I.: The use of the SMAA acceptability index in descriptive decision analysis. European Journal of Operational Research **196**(3), 1229 – 1237 (2009). https://doi.org/https://doi.org/10.1016/j.ejor.2008.05.021, http://www.sciencedirect.com/science/article/pii/S0377221708004189
9. Franco, C.D., Buttazzo, G.: Energy-aware coverage path planning of uavs. In: 2015 IEEE International Conference on Autonomous Robot Systems and Competitions. pp. 111–117 (April 2015)
10. Group, T.: Watchkeeper. Tech. rep., Thales Group (2010), https://www.thalesgroup.com/en/global/activities/defence/unmanned-aerial-vehicles-systems/tactical-uav
11. Holloway, H., White III, C.C.: Question selection for multi-attribute decision-aiding. European Journal of Operational Research **148**(3), 525 – 533 (2003). https://doi.org/https://doi.org/10.1016/S0377-2217(02)00436-8, http://www.sciencedirect.com/science/article/pii/S0377221702004368
12. Jacquet-Lagreze, E., Siskos, J.: Assessing a set of additive utility functions for multicriteria decision-making, the uta method. European Journal of Operational Research **10**(2), 151 – 164 (1982). https://doi.org/https://doi.org/10.1016/0377-2217(82)90155-2, http://www.sciencedirect.com/science/article/pii/0377221782901552
13. Kabamba, P., Meerkov, S., Zeitz, F.: Optimal path planning for unmanned combat aerial vehicles to defeat radar tracking. Journal of Guidance, Control, and Dynamics **29**(2), 279–288 (3 2006)
14. Keeney, R., Raiffa, H.: Decisions with Multiple Objectives: Preferences and Value Trade-Offs. Wiley series in probability and mathematical statistics. Applied probability and statistics, Cambridge University Press (1993), https://books.google.fr/books?id=GPE6ZAqGrnoC
15. Kim, D., Chen, T.: Deep neural network for real-time autonomous indoor navigation. CoRR **abs/1511.04668** (2015), http://arxiv.org/abs/1511.04668

16. Lahdelma, R., Hokkanen, J., Salminen, P.: SMAA - stochastic multiobjective acceptability analysis. European Journal of Operational Research **106**(1), 137 – 143 (1998). https://doi.org/https://doi.org/10.1016/S0377-2217(97)00163-X, http://www.sciencedirect.com/science/article/pii/S037722179700163X
17. Narayan, P., Meyer, P., Campbell, D.: Embedding human expert cognition into autonomous uas trajectory planning. IEEE Transactions on Cybernetics **43**(2), 530–543 (April 2013)
18. Olteanu, A.L., Belahcène, K., Mousseau, V., Ouerdane, W., Rolland, A., Zheng, J.: Preference Elicitation for a Ranking Method based on Multiple Reference Profiles (Aug 2018), https://hal.archives-ouvertes.fr/hal-01862334, working paper or preprint
19. Rolland, A.: Reference-based preferences aggregation procedures in multi-criteria decision making. European Journal of Operational Research **225**(3), 479–486 (2013). https://doi.org/10.1016/j.ejor.2012.10.013, https://doi.org/10.1016/j.ejor.2012.10.013
20. Roy, B.: Méthodologie Multicritère d'Aide la Décision. Economica (01 1985)
21. Roy, B.: The outranking approach and the foundations of electre methods. Theory and Decision **31**(1), 49–73 (Jul 1991), https://doi.org/10.1007/BF00134132
22. Roy, B.: Multicriteria Methodology for Decision Aiding. Kluwer Academic, Dordrecht (1996)
23. Ruz, J.J., Arevalo, O., de la Cruz, J.M., Pajares, G.: Using milp for uavs trajectory optimization under radar detection risk. In: 2006 IEEE Conference on Emerging Technologies and Factory Automation. pp. 957–960 (Sept 2006)
24. Sadeghi, F., Levine, S.: Real single-image flight without a single real image. CoRR **abs/1611.04201** (2016), http://arxiv.org/abs/1611.04201
25. Taisho, T., Enfu, L., Kanji, T., Naotoshi, S.: Mining visual experience for fast cross-view uav localization. In: 2015 IEEE/SICE International Symposium on System Integration (SII). pp. 375–380 (Dec 2015)
26. The, A., Mousseau, V.: Using assignment examples to infer category limits for the ELECTRE TRI method. Journal of Multi-Criteria Decision Analysis **11**(1), 29–43 (2002)
27. Tipaldi, M., Glielmo, L.: A survey on model-based mission planning and execution for autonomous spacecraft. IEEE Systems Journal **PP**(99), 1–13 (2017)