*Article*

# Playful Probes for Design Interaction with Machine Learning: A Tool for Aircraft Condition-Based Maintenance Planning and Visualisation

Jorge Ribeiro *, Pedro Andrade, Manuel Carvalho, Catarina Silva, Bernardete Ribeiro and Licínio Roque

CISUC—Centre Informatics and Systems, Informatics Engineering Department, University of Coimbra, 3004-531 Coimbra, Portugal; pmca@student.dei.uc.pt (P.A.); mcc@student.dei.uc.pt (M.C.); catarina@dei.uc.pt (C.S.); bribeiro@dei.uc.pt (B.R.); lir@dei.uc.pt (L.R.)
* Correspondence: jamr@dei.uc.pt

**Abstract:** Aircraft maintenance is a complex domain where designing new systems that include Machine Learning (ML) algorithms can become a challenge. In the context of designing a tool for Condition-Based Maintenance (CBM) in aircraft maintenance planning, this case study addresses (1) the use of Playful Probing approach to obtain insights that allow understanding of how to design for interaction with ML algorithms, (2) the integration of a Reinforcement Learning (RL) agent for Human–AI collaboration in maintenance planning and (3) the visualisation of CBM indicators. Using a design science research approach, we designed a Playful Probe protocol and materials, and evaluated results by running a participatory design workshop. Our main contribution is to show how to elicit ideas for integration of maintenance planning practices with ML estimation tools and the RL agent. Through a participatory design workshop with participants' observation, in which they played with CBM artefacts, Playful Probes favour the elicitation of user interaction requirements with the RL planning agent to aid the planner to obtain a reliable maintenance plan and turn possible to understand how to represent CBM indicators and visualise them through a trajectory prediction.

**Keywords:** design; remaining useful life; visualisation; machine learning; reinforcement learning; condition based maintenance; aircraft maintenance planning

**MSC:** 68U35; 91A12; 68T20

## 1. Introduction

The Aircraft Maintenance (AM) domain poses new challenges for the design of decision support systems such as Condition-Based Maintenance (CBM). Human and Machine Learning (ML) confluence can give rise to new decision support systems that allow the increase in the aircraft's flight time and the cost reduction promised by CBM [1]. This technique exploits ML-based components and systems failure forecasts to perform maintenance only when necessary instead of using a fixed interval approach, increasing aircraft availability and safety while reducing costs.

Introducing ML algorithms in critical and highly regulated operational contexts resists experimentation and raises additional challenges for design approaches to human ownership and control over new ML algorithms. If, on one hand, interacting with ML tools requires an approach that recognises and empowers the user to design new practises, on the other hand, it is necessary to design the technology for a set of practises that are still nonexistent.

The user of the maintenance planning tool, hereinafter referred to as planner, has the great responsibility of constantly having a reliable plan. As such, it is natural that he/she may distrust the operation of a new ML agent. Not only is it important to provide the planner with an interaction with the ML planning algorithm that meets their expectations,

but also ensure that the agent itself responds effectively to the planning challenge by generating a good plan. Aircraft maintenance has a significant influence on the total operating cost of an airline. Therefore, it is imperative to improve the current planning practices and look towards maintenance optimisation. An efficient planning algorithm aims to minimise the fleet ground time and, consequently, increase its availability and enable revenue growth for the airline.

CBM is made possible by using ML components to produce Remaining Useful Life (RUL) estimates for aircraft system components. Still, in addition to producing the estimates, it is necessary to present them to the maintenance planner so that this information empowers the human role in the decision-making process.

To gain insight and better understand how CBM planning could be introduced in a critical industry such as aviation, we designed a CBM Playful Probe. As proposed by Bodker and Kyng [2], we should focus on participation that matters. Therefore, such a probe aims to empower participants to play with alternative scenarios, develop and express their own understandings on the situation in question, to identify the necessary changes for the visualisation of the RUL as the integration of a Reinforcement Learning (RL) agent for collaboration in maintenance planning.

The main research question of this work is: can the use of Playful Probes enable insights that allow the design researchers to understand how to design for interaction with ML algorithms? Through the use of Playful Probes it was possible (1) to obtain a list of design outcomes that allow us to understand how to design for interaction with ML algorithms in a CBM context. Through these design outcomes, it was possible to design (2) the integration of a Reinforcement Learning (RL) agent for human–AI collaboration in maintenance planning and (3) the visualisation of CBM indicators that will be included in a future runnable version of the CBM planning tool.

As a future work, we want to create a runnable Playful Probe to build on the knowledge of domain experts and practitioners, empowering them to speculate on how CBM planning can work for them. Moreover, we wanted to tame the ML object [3] by putting automation of estimates and plan generation to their service. However, first, we need to understand how Playful Probes can better be designed, taking the form of a CBM maintenance simulation game. This paper reports on a Design Research process that runs a Participatory Design Workshop (PDW) for evaluating a proposed Playful Probe design in the form of a "virtual paper prototype".

The next section refers to some background concepts related to information visualisation, explainability and RL. Next are presented works related to Playful Probes. In the Design Case section we present our case of study. The Method section presents the materials and methodology of the playful used in the PDW. In the Results we present content coding and the PDW conversation analysis. In the Design Outcomes section we present the main outcomes designs resulting from the use of Payful Probes as well as the result of these design outcomes in the Integration of the RL agent for human–AI collaboration in maintenance planning and the visualisation of CBM indicators. Finally we present the Discussion and Results.

## 2. Background

To enable planners to understand and interact with the generated maintenance plans, they need to be able to trust and understand the information presented, how it came to be and how likely or accurate the prognostics/estimates are. We aim to do this by relying on Information Visualisation (InfoVis). Based on the definitions and concepts explored in Aigner et al. [4] and Munzner [5] InfoVis is the area that studies the use of visual metaphors and artefacts to more efficiently convey information, making it more accessible and understandable. One of the main subjects of research in InfoVis is precisely how to deal with big quantities of data, specially time-based data.

Keeping in mind that the airline industry is not prone to radical changes, it is important to consider the classical ways of representing time and time-based data, such as the

timeline, the line and area plots [4,5]. To connect the classical InfoVis techniques with this modern problem we relied on the "What? Why? How?" abstract analysis, as presented by Munzner [5], which aims to pick apart and categorise every aspect of a visualisation problem to facilitate the comparison and borrowing of methods and techniques used in diverse fields and contexts.

The increasing development of AI algorithms and the need for humans to interact with them increase the need to providing guidelines for human–AI interaction as [6,7]. Wright et al. [8] present a detailed comparative analysis of industry human–AI Interaction Guidelines. Guzdial et al. [9] noted, however, that it is not yet clear how best to design AI interfaces that focus on explainability or co-creativity. Abdul et al. [10] and Wang et al. [11] agree that explainable, accountable and intelligible systems remain key challenges. In this line of thought, much research has been performed on explainability: Zhou et al. [12] presents a comprehensive overview of methods for the evaluation of ML explanations and Linardatos et al. [13], a review of ML interpretability methods. In a more user-centred perspective, Bhatt et al. [14] synthesise the limitations of current explainability techniques that hamper their use by end users.

Reinforcement Learning (RL) has been used for maintenance planning optimisation in multiple domains. Knowles et al. [15] used a basic Q-Learning in a maintenance scheduling problem to decide, at each step, if a maintenance job should be performed or not. An RL solution to optimise maintenance scheduling in a flow line system was proposed by Wang et al. [16]. Barde et al. [17] uses an on-policy first visit Monte Carlo to obtain the optimal replacement policy that minimises the downtime of military trucks, composed of different types of components with random time-to-failure. In the aviation domain, Hu et al. [18] propose the Q-learning algorithm for solving the problem of aircraft long-term maintenance decision optimisation.

## 3. Related Work

We describe ways of visualising the indicators needed to conduct CBM and we describe methods that can be used to carry out maintenance planning using ML and interacting methods with ML planning algorithms in Section 2. However, how should the user exploration of the CBM planning methods be enabled, helping to develop the participant perspective and appropriation of a new tool?

Cultural Probes were proposed by Gaver et al., as "An approach of user-centred design for understanding human phenomena and exploring design opportunities" [19] and "Probes are collections of evocative tasks meant to elicit inspirational responses from people—not comprehensive information about them, but fragmentary clues about their lives and thoughts" [20]. J. Wallace et al. argue that the process of mediating both the relationship between participant and researcher and participant and her own feelings about a question can be achieved with design probes that provide more than just inspiration for the design [21]. Furthermore, cultural probes can be a tool for designers to understand users [22]. F. Lange-Nielsen shows some studies in which probes are used as a scientific method or a design tool and [23,24] show how technology probes can be a promising new design tool in the design of new technologies. Using support artefacts, cultural probes allow participants to document their activities and experiences, to be used as research material. While collecting the perspective of the participants in the process, this method allows them to explore new things beyond the expected.

The role of playfulness in cultural development has been recognised at least since Huizinga [25]. Since then, there has been extensive work on this topic in the scientific community. Playful probing approach uses games designed specifically for the study, and these games are tailored to the research area and purpose [26]. Sjovoll and Gulden [27] suggests that a game designed for playful Probing "opens up for a playful and autonomous environment for data-gathering which involves learning about individual and shared social practices". The Playful Probes technique uses similar principles to those of Cultural Probes while exploiting games as a research tool to enable learning and data collection. Playful

Probes could also potentially enable the exploration of the CBM planning methods, helping to develop the participant perspective and appropriation of new tools [28]. A preliminary study [29] showed that Playful Probing artefacts can be used to design new ML algorithms in a critical and highly regulated operational context.

## 4. Design Case

In the context of the development of a prototype planning tool in a new aircraft maintenance CBM paradigm, we studied how the decision support tool should be designed to (1) give the user a reading and understanding of the indicators of prognostic aircraft components/systems obtained by external systems and stored in Database and (2) to integrate the reinforcement learning (RL) agent to human collaboration in planning maintenance as can be seen in Figure 1.
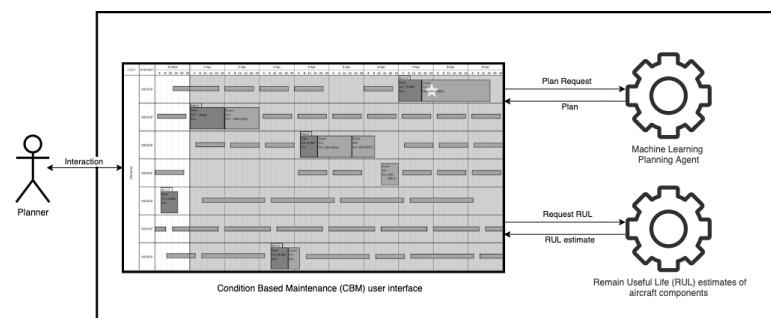


**Figure 1.** Condition-Based Maintenance (CBM) planning aircraft tool architecture. The user interacts with a Graphical User interface (GUI) that requests plans to the Reinforcement Learning (RL) automatic planning agent and requests Remaining Useful Life (RUL) visualisations for a certain component/system. Both the AutoPlan agent and the RUL visualisation obtain the historical information from the Database.

Figure 1 represents the three components we propose for a CBM Maintenance Planning Aircraft Tool. We can see that it is composed by the automatic planning agent, by the graphical user interface that allows the interaction and refinement of the plan by the planner, and by visualisation of the RUL indicators that gives the planner the confidence of a particular maintenance plan.

We did not have any information about what indicators we could use for RUL and how they could be visualised. We started to use the RUL of the component/system in a simple way, just as a value in Flight Hours (FH). Regarding the automatic planning algorithm, we have already developed a first version of a planning algorithm for long-term routine maintenance when we started this study.

## 5. Method

Playful Probing approach proposed by Bernhaupt et al. [26] uses games designed specifically for the study tailored to the research area and purpose of the study. It is the approach we used to study how to design the planning tool for a new CBM paradigm since it was intended to explore the design of the new tool by provoking inspirational responses from participants. From a practical point of view, planners play with the artefacts in paper prototype form to solve a maintenance challenge previously prepared for the workshop (see Figure 2). In the game scenario, the planners are confronted with a new RUL in a maintenance in the maintenance plan. To solve this game planners have to move the maintenance blocks in time and move flight plans to other airplanes to respect the new RUL. Including planners in this process, we bring participants to the development of the tool from an early stage, letting them appropriate the new tool. To research the design of the Playful Probes we adopted the Design Science Research (DSR) approach [30]. This work reports lessons from a DSR iteration, with useful results to refine future DSR cycles, but also to inform similar efforts by researchers and practitioners on similar research processes.

In the context of the ReMap project, we recruited two domain experts in aircraft maintenance management to conduct the design experiment. The participants, both male and between 20 and 40 years old have a good background in aviation and practical knowledge of planning tools, although they were not daily practitioners. The workshop was facilitated by the researcher, who ensured the application of the protocol and clarification of doubts on play scenarios and materials, e.g., role playing the gamemaster role, and the designer researcher who assisted in the discussion. With this experiment, lasting 74 min, we expected to open relevant questions about how ML-based Remaining Useful Life (RUL) estimates could be instrumented as part of the Playful Probe simulation.

Affected by COVID-19 measures, these experiments were conducted online with the figma web-based collaborative design platform [31], simulating the paper prototype [32] exercise. For this experiment, we played with simplified maintenance scenarios featuring work packages and only one RUL indicator for each aircraft maintenance.

Next, we will describe the steps of this probing methodology, from workshop preparation to email interviews.

1. Materials. The basic elements of a maintenance schedule (Figure 2):
   Flight: blue ribbons represent flights that some aircraft (row) can do at some time.
   Block: A predefined routine maintenance (A-checks).
   Cluster: A non-routine maintenance composed by other tasks not included in Blocks. This could have Due Date, RUL, both, or none.
   Canvas: The "background" of all artefacts, the scenario with information about the fleet, time representations and restrictions.
2. Solving the Problem Path. The beginning of the resolution was linear, only possible in one direction. Participants would be faced with the simplest concepts of the flight plan and maintenance. Subsequently, the resolution would lead to a path where users would necessarily be faced with more complex issues such as conflicting conditions and 90% confidence RUL.
   This probe was instrumented by placing visual artefacts RUL estimates with 90% confidence to confront participants with situations that could lead to debate and the generation of insights. The questions that we want to be raised by the participants are: Does it make sense to have a large degree of uncertainty? How do we represent it to enable decisions?
3. Test specification. To prepare the workshop, all artefacts were designed digitally but printed and pre-tested manually as in a common paper prototype exercise. After testing multiple approaches to instrument the probes with visual artefacts, and adjustments in size and complexity, the exercise was migrated to a digital collaboration tool (Figure 2).
4. Briefing. In an initial part of the experiment/workshop, an introduction was made explaining what the basic maintenance elements of the game were and demonstrating how to solve a simple problem.
5. Playful procedure. In this part of the experimental session, artifacts were presented to participants with a non-trivial maintenance scheduling problem to be solved, i.e., a problem that needs several plays both in the artifacts of the maintenance and artifacts of flights to respect the new RUL (Figure 2). The participant's voice and the collaborative canvas were recorded presenting their ideas and playing with the representations to solve the maintenance problem. The facilitator answered participants' questions about whether they could take a certain action or not. Furthermore, he alerted when they were ignoring some important conditions while trying to explore the problem.
6. Debriefing. After participants solved the scheduling problem, a wider discussion space opened, namely on the role of RUL visualisation and the use of an ML planning agent in the planning process.
7. Email Interview. After viewing the recording, some questions were sent to the participants. The intention was to clarify or deepen the reflections that they expressed during phases 5 and 6.

**Figure 2.** Maintenance scheduling problem presented to the participants in the experimental session. Flights : blue rectangles; Blocks: red rectangles; Clusters: orange rectangles; Star: the yellow star inside Cluster represent a new maintenance problem presenting a restrictive RUL.

This PDW generated audio and video recordings. The conversation between participants and the video with the manipulation of game artefacts took place in phase 5 and discussion in phase 6. Data were analysed by splitting into small time segments, coded into groups according to the first analysis categories in the conversation.

## 6. Workshop Results

In this section the results obtained in the Participatory Design Workshop (PDW) will be presented, first illustrating the temporal coding of the topics covered, and in a second part an analysis of the conversation obtained.

### 6.1. Content Coding

The workshop was started by explaining the basic maintenance elements in the game and illustrating how to solve a linear problem. In this part, lasting 10 min, the participants cleared some doubts about the game mechanics but did not interact with artefacts. The experiment followed with a non-trivial maintenance scheduling problem to be solved. Participants' dialogue and collaborative canvas were recorded, while discussing ideas and manipulating plan artefacts to solve the maintenance problem.

The facilitator intervened to: (a) answer questions about whether or not some actions were possible; (b) alert participants when they were missing relevant information; (c) try to encourage them to further explore aspects of the problem, to assess informational or action needs. The session developed freely following the problem to be solved, with no constraints regarding ordering of participants' actions or managing concurrency among open explorations, favouring dialogue while supporting each other's exploration. After participants solved the scheduling problem, a wider discussion focused on the role of ML in the planning process. This part lasted 74 min.

When we look at the focus of the conversation during stage 5 in Figure 3 we can see that at the beginning the participants talk about the representation of planning artefacts and deal with technical issues related with technology unfamiliar to participants before the workshop. Immediately after the problem has been placed, participants start talking about maintenance-related meta-speech (practices, procedures or regulations but not directly related to the resolution of the maintenance challenge). Only at 5:30, they change the focus to solve the problem, and only at 8 min, they start to manipulate the artefacts. From this moment onwards, the participants do not lose focus on solving the problem until the end of the exercise. This problem resolution is accompanied alternately by moments of artefact manipulation or maintenance-related meta-speech.
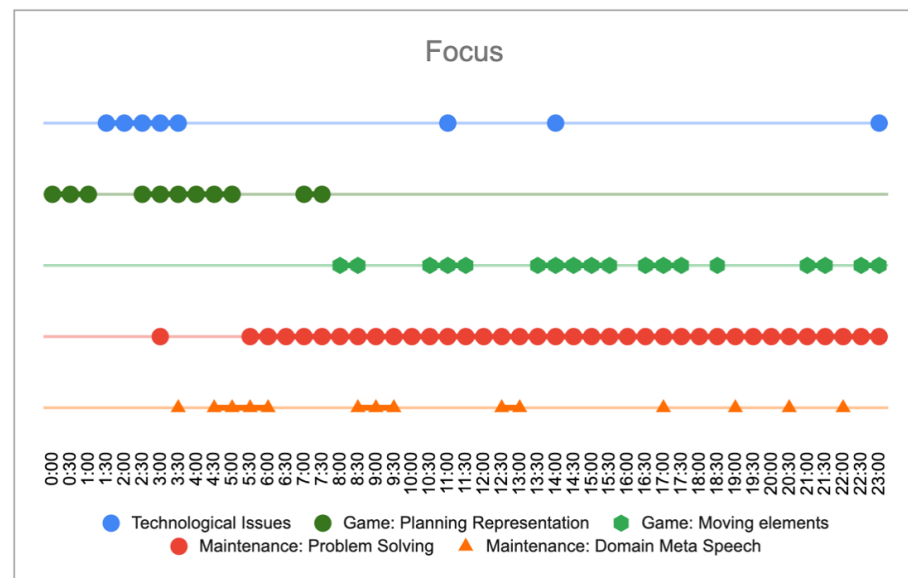


**Figure 3.** Focus of conversation during stage 5.

Reflections (on ML-tools and practises) mainly take place during stage 6, after minute 23, immediately after the problem has been solved, as can be seen in Figure 4. During this phase, it is noteworthy a quite intense discussion about maintenance planning practises. The discourse alternates between current practises and speculation on what future practises will look like. The reflective discourse in this phase is divided into three major blocks. Between 23 and 40 min, we find a speech oscillating between current and future practises; then, between 42 and 58 min, the conversation is focused on future practises and between 58 and 70 min, only current practises are discussed.



**Figure 4.** Reflection utterances in participants' conversation along the experiment.

Concerning the introduction of the Remaining Useful Life concept, we could verify that whenever there is a dialogue about time or the confidence interval, it comes with a discussion on RUL meaning and implications it may have. This took place mainly in the first block of the mixed discourse between current and future practices (23–44). The ML debates took place during this second block. It is interspersed between the form of interaction and reflection on the functioning of the algorithms and appears at several points in time simultaneously. The third block was exclusively a reflection on current practices. Between the first and the second block, a moment of reflection on the game (Playful Probe) itself takes place, but only for 2 min.

*6.2. Conversation Analysis*

The discussion is based on the analysis of the meanings expressed in conversation during the experiment and on the comparison with the feedback from the participants' interview answers. The conversation was very extensive; we will only focus on the discussion related to the visualisation of the RUL and the interaction with a machine learning algorithm.

These participants did not start immediately solving the problem. They begin by addressing the problem using meta-speech, suggesting that they were "reading" the problem first and obtain the right connection between artefacts and the maintenance language that they are familiar with. They took about 5 min between the moment that the problem was placed, until they started moving the elements in a very intricate collaboration process, such as analysing and negotiating the movements as if they were learning to play a game of chess. At minute 13:30, they decided to each take a different role "you do the flight and I do the maintenance"(P1) perhaps as a form of collaboration, but something that made sense later in terms of reassessing practices.

The participants found it easy and clear to understand what needed to be done. However, they found the RUL was not easy to interpret and considered it as a fixed due date. P2 said "was quite tricky estimate what risk you took when you interpreted the RUL", while P1 said the representation of RUL requires some mental effort to visualise,

"was a bit challenging to determine the due dates for the tasks, it required some mental efforts" P1 adding during the exercise that "the difference between 95 and 99 in my head is not playing a role". Despite the difficulty of seeing the impact of the confidence level during the exercise, they have made an effort to understand it, e.g., P1 said "I won't to risk, because 90% is quite high". During the exercise, P2 suggested a RUL of 60H with a confidence level of 90%, "it would be nice if we could see (…) 65 ± 6 h , than you kind have an idea of how close the edge you are", and when asked if a boxplot could fit, P1 answered "Yeh, I'm thinking out aloud now, but perhaps instead a square box, it could be a kind of distribution".

P1 said that the planner wouldn't matter too much in a preventive task of a component that usually fails with no impact (maintenance consequences), but if it is a component that must not fail because of the risk of an aircraft on the ground situation or a flight being cancelled; then it would make a big difference. They suggested automatically visualising the RUL on the timeline, and P1 also suggested it would be good to "visualise operation impact" as costs, availability and the maintenance components, asking P2 "But it could actually depend on what's these 65 h based on, right? What kind of components we are talking about!?".

At some point, P1 considered scheduling two hours maintenance over the limit, and wondered "what is the consequences of not making the exact Due Date? what's the consequences of having the component filled before the preventive removal?" and "How critical is it if we don't respect a RUL?", suggesting that due dates maybe could be more flexible, if the return is large enough. At the end of the exercise, P1 took a co-constructive move and started using the collaboration tool to make some design proposals. They started to draw how this kind of distribution can occur, as shown in Figure 5, using as a reference the representation of "Trends, Rul, & Uncertainty" by [33]. This representation can also be seen as a visual analogy based on how the arrival time is modeled, but in this case as a view of the risk.
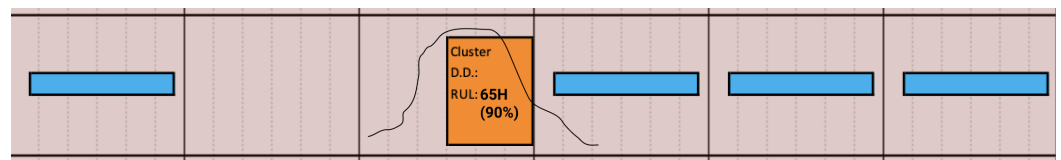


**Figure 5.** The participant proposed visualizing the RUL with a "probability distribution curve", showing how it could be represented by roughly drawing the thin black curve.

P1 complemented how this model can work when there is a situation with two maintenance needs partially overlapping, "the planner can for example choose ever when the arrival time is that close to each other respecting this potential overlap (…) we can just wide to the left and the right as much as we can" as shown in Figure 6.
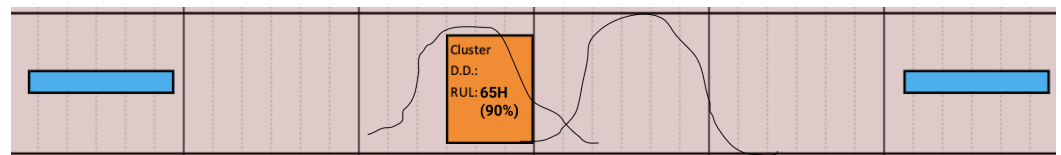


**Figure 6.** The participant proposed visualizing two RUL, showing how it could be represented by roughly drawing the two overlap thin black curves.

P2 was cautious, saying that with current practices "we don't want our planner to access the technical state of the plane (…) I wouldn't be very comfortable with letting him decide whether it's an acceptable risk to take". P1 complemented "that's how we work now, so if there is a prognostic alert, then somebody makes the due date, and the scheduler respects that due date, and the guy that makes the due date, doesn't know about the schedule, he just looks at the task, looks at the criticality and then he says, ok, this needs

to be done in 10 days". In some situations, instead of accessing the data itself, the scheduler should have a specialist's assessment to sign an agreement to deviate from the due date and then act on that assessment to take that risk.

The participants agreed that the future state should be different as they can use predictive values such as RUL to represent the uncertainty of the aircraft condition and also use maintenance opportunities to solve the problems. P2 presented their vision: "we should have a kind of class of component or class of consequences, and depending on that class, it must not run the risk, or it can run the risk of exhausting the RUL". P1 agreed "the decision on whether to schedule something, should not be just dependent on the description of the task but should be also dependent of the maintenance opportunities and the state of the fleet", and " take in consideration the probability that's something might fill with the large or small impact". A task with low probability and very high impact can trigger a discussion about whether it should be planned, and they should simply accept this schedule if they "have a spare aircraft stand bay or have some buffer in the network", otherwise, they will not take this risk, which may lead to cancellation.

Based on the knowledge and experience of the participants, the discussion took very interesting paths through existing planning practices as well as speculations on what future practices might look like with machine learning algorithms that both enable predictive indicators for RUL and help the planner and scheduler in decision making.

In the next section we describe how automatic planning has been designed and integrated into the maintenance planning tool, considering the characteristics and needs of the planners.

## 7. Design Outcomes

From the results presented in the previous Section 6, it is possible to elect the set of important maintenance decision requirements not only for the future design of the runnable version of the tool (this study will be conducted in the future), but in particular for the components that incorporate the ML algorithms, the automatic planning agent and the RUL visualisation shown in the Table 1

**Table 1.** List of main design outcomes that can be incorporated into the design of the Automatic Planning Agent and RUL visualisation. The letter "x" address the design outcomes to the ML components.

| Design Outcomes | Automatic Planning Scheduling | RUL Visualisation |
|---|:---:|:---:|
| RUL only applies to short-term maintenance. | x | x |
| Different RUL should have the same confidence level. | x | x |
| RUL should be visualised as a distribution of probabilities instead of fixed value or Boxplot. | | x |
| RUL estimates should be flexible (not mandatory as due dates). | x | x |
| It should be possible to plan maintenances after RUL estimates. | x | |
| The planned schedule should take into account maintenance opportunities. | x | |
| The planned schedule should take into account the global state of the fleet. | x | |
| The planned schedule should take into account the impact of every maintenance necessary (operation/costs). | x | |
| The planner is not responsible for taking decision related to the technical state of the plane. | x | |

Using RUL indicators in predictive maintenance means working with a short-term planning, focusing only on the next few days or weeks. The core of CBM maintenance is in small maintenance issues that arise, for instance with a 25 FH RUL, with a small quantity of tasks and which has to be scheduled for the following days. The planning agent can not have a holistic view of the plan, at least compared to a human, since much information is not in the system. However, it can be designed to take into account the information of the current fleet state (ensuring there are no collisions), the availability of hangars and the criticality of the tasks to be planned. This can be done by allowing the user to conduct particular maintenance as planned, not letting the agent change their position. Another important aspect is the impact of conducting or not conducting certain maintenance. If we are talking about replacing a component on the coffee machine, it may be preferable to run the risk of it breaking, rather than incurring costs of stopping the plane to conduct this maintenance, or take into account the working time that could be wasted to conduct small maintenance that can be conducted with other planned maintenance, which encompasses the opening of the same panels that give access to that component. These issues will be detailed in the Section 7.1.

The main design outcome for RUL visualisation is that it should be represented through a probability density function (PDF) rather than a fixed value or boxplot. When fixing the same confidence level to the whole of RUL, it is possible to visually analyse and compare different PDFs through the drawing of the respective curves. It was important to understand that planners are responsible for planning with respect to the various maintenance domains (Manpower, Materials, Machinery, and Method), not by evaluating the status of the components or by creating maintenance tasks to solve a problem. However, it may be useful to show information that complements RUL estimates of a task or component, or even the RUL history associated with this task or component. This feature to used visualise RUL with models created from historical data will be detailed in Section 7.2.

### 7.1. Automatic Planning Agent

One of the main challenges for a maintenance planner is dealing with unexpected maintenance events efficiently. The planning usually aims to balance minimising ground time and maximising task utilisation. However, the conflicting nature of these two metrics means the optimal decision is not always obvious. It is also not feasible to drastically change the current plan every time new information comes in.

The automatic planning algorithm went through several iterations. In early versions, the algorithm was designed for long-term maintenance planning for a time horizon of six months. However, after the workshop, it became apparent that prognostic-driven tasks only apply in the short term, and it is not feasible to employ CBM so far in the future. As a result, the time horizon was reduced to one month, meaning that heavy maintenance checks, which can last multiple weeks but are less frequent, were no longer considered. RUL prognostics share the same confidence level, and tasks originating from those prognostics have a flexible due date. The reason for this flexibility is that, in some non-critical systems, the cost of allowing the component to fail and replacing it with a new one might be lower than performing preventive maintenance on that component.

Aircraft operation also has significant relevance in the maintenance decision process. Therefore, the flight plan for the fleet was included in the algorithm, giving the user a complete and more accurate representation of the fleet status. The flight and maintenance plans being combined allows the RL agent to have a complete understanding of the fleet state and available maintenance opportunities when scheduling tasks, which are two important requirements. The requirement stating that the plan should account for the impact of maintenance decisions is not being considered yet and can be regarded as future work.

One more requirement that arrived from the workshop was the possibility for the human planner to have some control over the final solution, which is essential in cases where the RL agent does not have access to all the available information. Therefore, a new

feature introduced was the possibility to hold slots in place, meaning they are permanent and cannot be moved by the algorithm to another time. This feature helps the planner in cases where it is very advantageous or even required to perform maintenance at a specific time.

The Key Performance Indicators (KPIs) used to evaluate the quality of the maintenance plan also evolved over time. Initially, the two KPIs used were the ground time of the fleet (in hours) and the average utilisation percentage for all scheduled tasks. While the ground time remains the same, the utilisation KPI changed to what we are calling time slack, which is the difference (in hours) between a task due date and its scheduled date. These KPIs are provided after every algorithm run, allowing the planner to see the consequences of every proposed modification in the context of the entire fleet.

The goal of the automatic planning algorithm is to adapt the current maintenance plan for a fleet whenever newly available information becomes available. That includes new faults and updated RUL prognostics. The RL algorithm Deep Q-learning is used to optimise the generation of an updated maintenance plan for the fleet. The input for this adaptive planning algorithm contains the flight plan for the fleet and the scheduled maintenance slots and tasks. The output corresponds to a new maintenance plan, consisting of the updated maintenance slots and tasks.

Before the planning begins, a simulation of new maintenance information is performed. On each day of the plan, there is a probability that new maintenance events are discovered for every aircraft, namely faults and RUL prognostics. This simulation provides the required data to train and test the RL agent. Historical data provide relevant metrics such as the average number of faults per week and their criticality, which is used to define the simulation parameters. Faults are simulated with a certain urgency level based on the "Rectification Interval" categories specified in the operators' Master Minimum Equipment List (MMEL) [34]:

- Category A: No standard interval is specified. However, items in this category shall be repaired within the time interval specified in the "Remarks and Exceptions" column of the operator's approved MMEL.
- Category B: Items in this category shall be rectified within 3 calendar days (excluding the day of discovery).
- Category C: Items in this category shall be rectified within 10 calendar days (excluding the day of discovery).
- Category D: Items in this category shall be rectified within 120 calendar days (excluding the day of discovery).

A new task is created for every fault, and its due date is set based on the respective urgency class. The generation of updated RUL prognostics results in updating of the due date of already existing tasks.

A scheduling priority is assigned to all simulated tasks. The ones created to rectify faults have a higher priority and are arranged based on the respective fault urgency. The ones updated after the simulation of new RUL prognostics have the lowest priority. Next, the planning phase of the algorithm begins. Tasks are ordered based on their priority and are scheduled individually by the RL agent. When multiple tasks have the same priority, their due date defines the scheduling order. At each step, the RL agent acts by choosing a slot to schedule the next task. The task is scheduled in the chosen slot, and the remaining workload available in that slot is updated. This process is repeated until all tasks have been scheduled. The automatic planning algorithm is illustrated in Figure 7.

The RL agent has two possible actions when choosing the maintenance slot to schedule a task. It can opt for the slot that minimises the time slack or the one that minimises the aircraft ground time. In the first case, the chosen slot will be the closest to the task due date. The second case takes into account the task duration and access cost. Most times, to access a component and perform the task, it is required to open multiple panels, which may substantially increase the ground time. Therefore, it would be better to group tasks with common access requirements and perform them in the same slot.

When an agent acts it will receive a reward signal that determines how good that decision was and is used in the learning process. This reward is defined with the following function:

$$R = -u - (d + a) \tag{1}$$

It contains a factor ($u$) corresponding to the task utilisation and a factor corresponding to the sum of the task duration ($d$) and the access cost ($a$), which represents the added ground time. The access cost is zero if the task is scheduled in a slot that already shares the same access requirements. Both are negative factors because the goal is to minimise them.
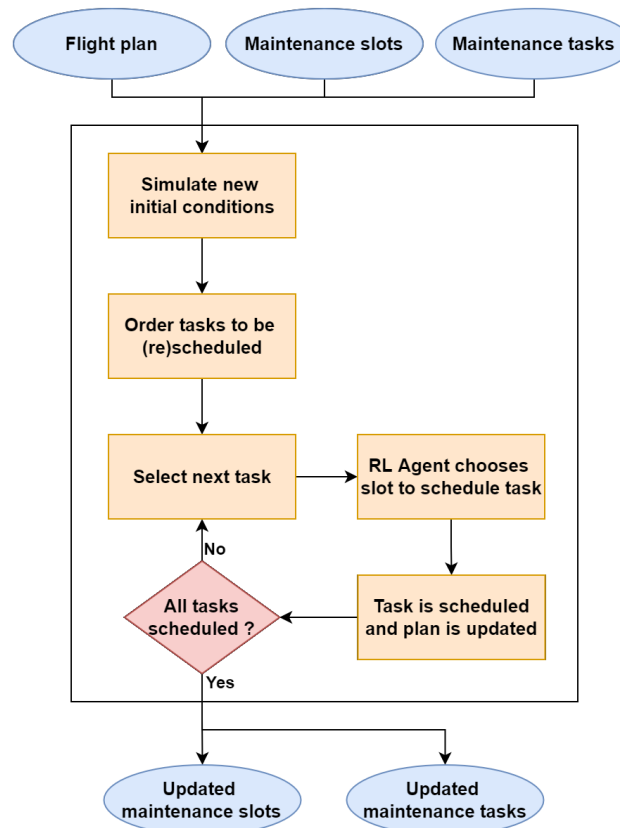


**Figure 7.** Activity diagram representing the automatic planning algorithm.

The initial maintenance plan served as a baseline for the training and testing of the RL algorithm, and it was created based on real maintenance data for a 16-aircraft fleet. Additionally, three maintenance scenarios (mild, medium, and aggressive) were created to validate the algorithm's solution. The number of new maintenance events simulated along with their severity varies in each scenario.

The time horizon considered was 4 weeks and the number of training episodes was 200. At the end of each episode, a new solution is obtained and evaluated according to the predefined KPIs. Figures 8 and 9 show the evolution of the ground time and time slack KPIs, respectively, throughout training. We can see that both are being minimised, meaning the quality of the updated maintenance plan is improving. These results indicate that the agent is able to improve its decision-making capabilities over time and learn to produce better maintenance plans from past experiences. Furthermore, RL is able to produce a feasible maintenance plan that complies with all the problem constraints.
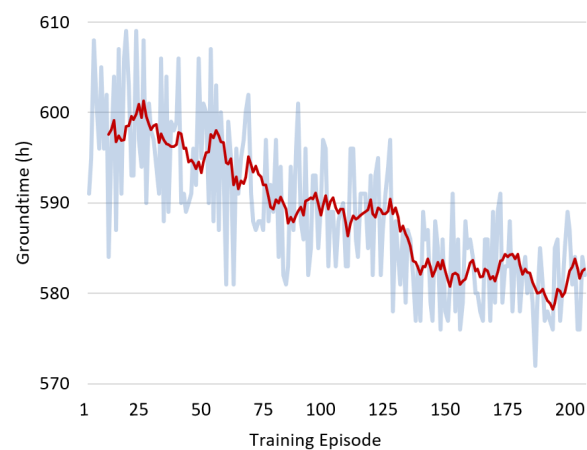
**Figure 8.** Evolution of the ground time KPI during training (in blue) with a moving average trend line (in red).
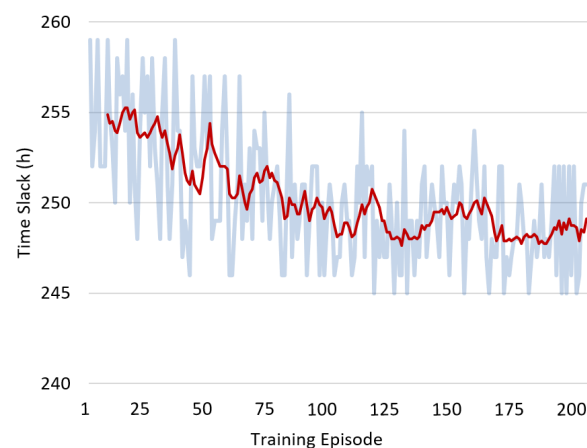


**Figure 9.** Evolution of the time slack KPI during training (in blue) with a moving average trend line (in red).

*7.2. RUL Visualisation*

This section describes how RUL visualisation was developed, after receiving the inputs obtained in the Workshop. The development of the RUL visualisation started with a set of guidelines resulting from the workshop and a set of graphs and visualisations related with RUL prognostics. These concepts were then contextualised within the InfoVis field and we implemented a usable visualisation prototype where we could experiment with and test our ideas. The visualisation tool was developed iteratively, as new data and prognostic algorithms were made available, and was interleaved with recurrent feedback sessions with members that accompanied the workshop process and also with one of the participants.

As referenced by the P1 participant during the playful procedure, Figure 5, the visual representations of RUL prognostics can rely on some of the classical methods of representing time-based data, such as the ones shown by Goebel et al. [33] relative to "Trends, Rul, & Uncertainty". From this perspective, we expanded on the use of the line and area charts and tried to keep the visual idiom within the established visual aspect, such as by using the same coordinate mapping, combining the two types of plot and by allowing the comparison of different trajectories on the same graph. We use the term idiom, within the visualisation context to refer to the set of choices that determine the meaning of the visualisation elements, in line with its use in [5].

The outputs resulting from the Playful Probe included thoughts and guidelines related with RUL visualisation and these were also used as a starting point, or better, as possible goals for the visualisation. One of the most relevant concepts received pointed

towards visualising the RUL by presenting its current (newest) prognostic as a PDF curve. Nonetheless, initial analysis and experiments with RUL prognostic results showed that any instance of RUL prognostics, devoid of context, was not very meaningful. In the sense that, at the end of a given RUL trajectory, most individual predictions could be or would be inaccurate. Moreover, the accuracy of the predictions would start low and increase towards the end of the trajectories. This was a strong indication that representing only the most recent prognostics might not be enough to help the user to understand the real condition of the components.

Additionally, depending on the prognostic model used, the resulting trajectory might look different and might need different interpretations. For instance, if the RUL is modeled with an Elbow Point, the earlier predictions are expected to be understood as "more than X hours", and the trajectory will look like a bent arm, starting horizontally and then rising (or falling) after the elbow point.

Another concern is related to noisy results. If there are up and down fluctuations in the prediction results, it is very difficult to understand the trajectory trend only by assessing individual points.

Concerns expressed in the playful procedure, regarding the lack of confidence in a single value of RUL prognostics as well as a lack of objectivity when comparing similar RUL values, further cemented the decision to represent the whole trajectory. By representing the whole trajectory, the user can see if the component is degrading normally or, for instance, if the current prognostic is happening in a steep descent or within a noisy period. This might change how the planner takes into account the proposed maintenance plan. Therefore, in order to not stray too far from the workshop considerations, we present the current prognostic as a PDF curve and we also provide a separated detailed visualisation where the user can see the components' prognostic history and how the prediction has been changing.

The visualisation idiom developed, shown in Figure 10, uses line and area charts to represent the RUL prognostic. The information is mapped on a horizontal axis, marking time in flight hours, which is divided into past and future and is centred on the present time. The vertical axis shows RUL in flight hours, that is the amount of time we expect the component to be usable.
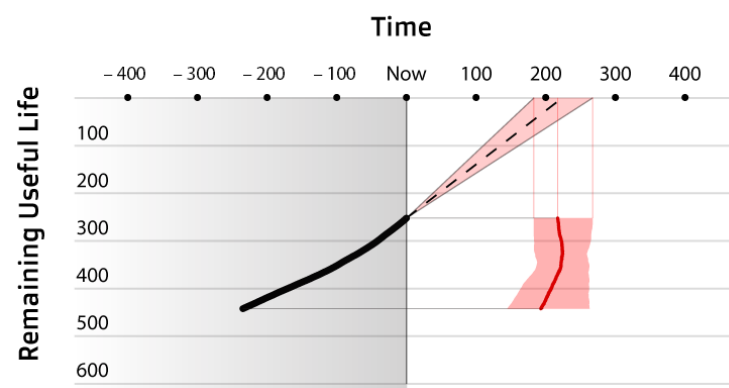


**Figure 10.** RUL visualisation: RUL trajectory as a black thick line; the current prognostic is shown by the red triangle and the dashed line starting at the end of the black trajectory. The dashed line can show RUL in single value or the mean of a PDF. The width of the triangle marks one standard deviation away from the mean or a minimum and maximum RUL; the red area in the lower right zone represents the Expected End of Life (EEoL) according to each RUL prognostic.

In the Past (left) side of the visualisation, we represent the RUL trajectory with a line plot. This line shows all the available prognostics since the component was installed. This side is shaded darker. On the Future (right) side we present the expected behaviour based on the current prognostic. If the RUL is calculated as a PDF we represent the mean and the standard deviation with an area and line plots. If the RUL is a single number then we represent the expected behaviour as a dashed line. In addition, on the Future side, we

represent the Expected End of Life (EEoL) of the component mapped in relation to the RUL prognostic. That is for any point in the line plot of the trajectory on the past side, we mark the place in time that we expect the trajectory to end. If the RUL is a PDF we mark the minimum, mean and maximum, otherwise just one number. This results in an area or a line plot, respectively. The area resembles a triangle or an inverted tornado, pointing to the end of the trajectory.

Figure 11 presents a rough 3D sketch of this relation, where we can see how the PDF curves form the red area.

We can see an example of a complete trajectory in Figure 12, showing the complete red shape.

By analysing the vertical deviation or the slant of the resulting area or line it is much easier to see if the prognostic model predictions were consistent or if there were lots of corrections over time. This is because the ground truth of a trajectory, if represented in this mapping would be a vertical line, centred on the end of life.

If there are parts of the red area that overlap with the grey background it means that the corresponding RUL prognostics expected the component to fail before our current time. If the overall shape is leaning to the right, it means that the component is lasting longer than expected. If the overall shape is leaning to the left it means that the component is deteriorating faster than expected. This artefact may help to infer how planning can relate to future behaviour. For instance, a low RUL can lead to a flight plan that makes the component last longer than expected as it was used on less damaging routes—this would be represented by the red line pointing to the right side.
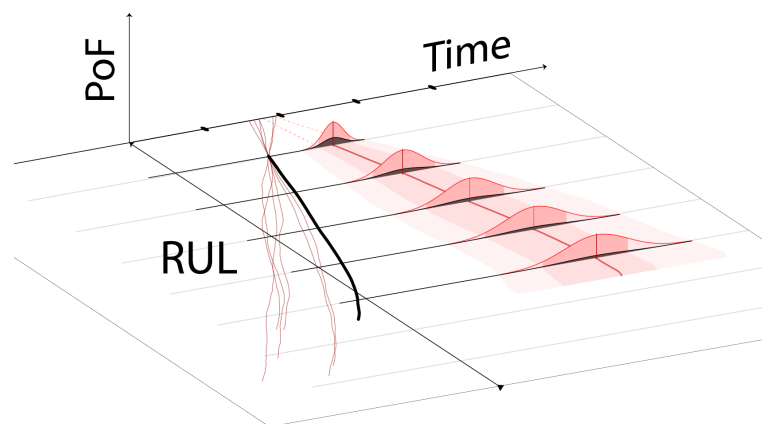


**Figure 11.** Sketch representation of the RUL trajectories and the PDF curves in 3D. The black line shows the current trajectory and the red lines are other past trajectories of the same component, all in Flight Hours. Here we can see that the red area, mentioned previously, is formed by successive PDF curves. In this figure, the PDF curves represented result from the trajectory comparison, and not from the model, as these were single value RUL prognostics. As such there are two red areas, the darker area shows the observed interval and the lighter one marks the 95% confidence interval, that is 2 standard deviations away from the mean.

We also implemented a simple technique to compare a current RUL trajectory with past trajectories of components of the same type, to help contextualise its behaviour and make deviations from a "normal" behaviour much more explicit. If there is available data with past RUL trajectories of other components of the same type, we can also represent the past trajectories as line plots. Each past trajectory is mapped relative to the time when their RUL prognostic was closest to the reference RUL, as shown in Figure 13. This results in a distribution, showing how long other components of the same type lasted after having a similar prognostic. The comparison method can be applied to single value RUL trajectories, or by using the mean value of RUL PDF trajectories.
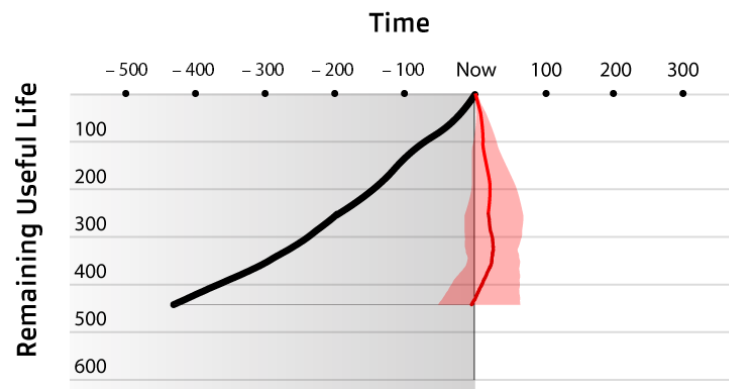
**Figure 12.** RUL visualisation at the end of life, with the resulting triangular shape showing that all predictions were very close to the real end of life, specially at the start and end of the trajectory. The black line represents the current trajectory. The red line marks all the Expected Ends of Life according to the RUL PDF mean, and the width of the area marks one standard deviation away from the mean.

With this representation we try to convey more information about the behaviour of each component and also about the prognostics accuracy, so that the user can have a more objective notion about the intrinsic uncertainty of any prediction.
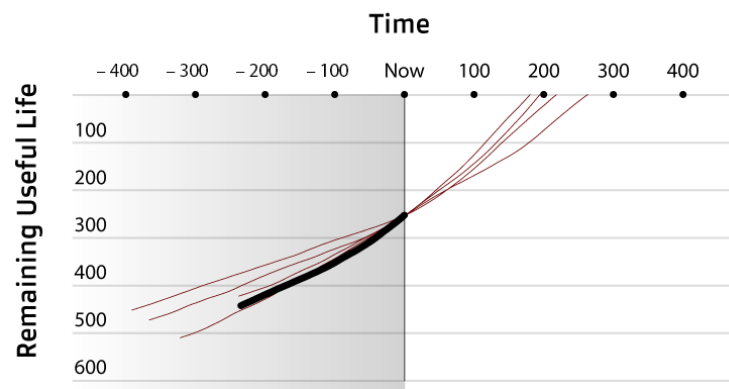


**Figure 13.** RUL trajectory comparison showing some past trajectories aligned with a current trajectory. The current trajectory is represented with the black line and the past trajectories are represented with the thin red lines.The past trajectories are aligned by the time they had a prognostic similar to the last prognostic of the current trajectory. With this representation we can see how the other trajectories behaved after this point. When there is a new prognostic, the trajectories are re-aligned using the new prognostic value.

## 8. Discussion and Reflection

The Playful Probe approach, materialised in a shared digital paper prototype, enabled an exploratory environment in which researchers and domain experts were able to explore diverse aspects of ML adoption in aircraft maintenance. By playing with probe artefacts to solve a concrete scenario, participants can focus and reflect on changes to their domain practices and open productive dialogue on how CBM maintenance could be designed, as evidenced by our content and speech analysis relative to the exercise and materialised in the main design outcomes. By using the design outcomes as a list of design requirements for the ML components of our design case, it was possible to evolve the RL agent for Human–AI collaboration in maintenance planning and propose the visualisation of CBM indicators.

Regarding the RL component, a planning solution capable of optimising maintenance decision-making when facing new unexpected events was developed. This solution evolved over time to a more realistic representation of a maintenance planning scenario by reducing the time horizon and including aircraft flights in the problem. The introduction of the slot

hold feature also allows the user to have more control over the resulting maintenance plan. The automatic planning algorithm can deal with faults of different urgency classes and prognostics information by optimising the scheduling of the respective maintenance tasks. The goal is to achieve a balance between fleet ground time and time slack. Results indicate that the RL agent is able to produce better maintenance plans during training by improving both KPIs. Furthermore, by taking into consideration the current plan as a baseline, we avoid creating a completely new plan whenever new information becomes available, which would be unfeasible in real scenarios.

The reflection generated in this exercise allowed the participants to speculatively imagine how the planner could use this tool in the future by incorporating the predictive RUL indicators of aircraft components. The RUL indicator triggered an extensive speech that required a constant reflective discourse on current and future practices. The workshop allowed the interpretation of the RUL indicator and pointed to a possible form of representation through PDF curves. Although box plots and PDF curves are reliable ways of representing probability distributions, in order to better visualise the degradation trends it was necessary to contextualise the instance prognostics in their time frame, i.e., through their trajectory. Therefore, the proposed RUL detailed visualisation was primarily focused on contextualising the prognostic instances within their current and past behaviours. It can be divided into three parts: the RUL trajectory (Past); the current RUL prognostic (Present); and the Expected End of Life (Future). In order to provide in-depth knowledge of the prognostics, we present the component's current condition, contextualise it in its past behaviour and show how it can translate into the future. By representing this information visually, the user can quickly grasp if the proposed plan is fit for the component condition or take action if some modifications are required. Additionally, the visualisation idiom might help to infer how planning can relate to future behaviour.

All things considered, it was possible to speculatively but explicitly generate requirements for a runnable version of the planning tool that includes the automatic ML planning agent and generates the detailed RUL visualisation increasing the planner's confidence in the maintenance plan. However, other forms of visualisation and interaction with the ML agent should be explored in future runs. The current study does not yet inform the acceptance by planners of a runnable Playful Probe as either a training or final use device, which remains on our project agenda.

**Author Contributions:** Conceptualisation, J.R. and L.R.; investigation, J.R.; methodology, J.R.; supervision, C.S., B.R. and L.R.; validation, L.R.; writing—original draft, J.R., P.A. and M.C.; writing—review and editing, C.S., B.R. and L.R. All authors have read and agreed to the published version of the manuscript.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Informed consent was obtained from all subjects involved in the study.

**Data Availability Statement:** Not applicable.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Andrade, P.; Silva, C.; Ribeiro, B.; Santos, B.F. Aircraft maintenance check scheduling using reinforcement learning. *Aerospace* **2021**, *8*, 113. [CrossRef]
2. Bødker, S.; Kyng, M. Participatory design that matters—Facing the big issues. *ACM Trans. Comput. Hum. Interact.* **2018**, *25*, 1–31. [CrossRef]
3. Bødker, S.; Roque, L.; Larsen-Ledet, I.; Thomas, V. Taming a Run-Away Object: How to Maintain and Extend Human Control in Human-Computer Interaction? 2018. Available online: https://pure.au.dk/ws/files/135965176/bodker_taming_a_runaway_object.pdf (accessed on 15 March 2022).
4. Aigner, W.; Miksch, S.; Schumann, H.; Al, E. *Visualization of Time-Oriented Data*; Springer: London, UK, 2011.

5.    Munzner, T. *Visualization Analysis & Design*; CRC Press, Taylor & Francis Group: Boca Raton, FL, USA, 2015.

6.    Amershi, S.; Inkpen, K.; Teevan, J.; Kikin-Gil, R.; Horvitz, E.; Weld, D.; Vorvoreanu, M.; Fourney, A.; Nushi, B.; Collisson, P.; et al. Guidelines for Human-AI Interaction. In Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems—CHI '19, Glasgow, UK, 4–9 May 2019; ACM Press: New York, NY, USA, 2019; pp. 1–13. [CrossRef]

7.    Holbrook, J. Human-Centered Machine Learning . 2020. Available online: https://medium.com/google-design/human-centered-machine-learning-a770d10562cd (accessed on 16 April 2020).

8.    Wright, A.P.; Wang, Z.J.; Park, H.; Guo, G.; Sperrle, F.; El-Assady, M.; Endert, A.; Keim, D.; Chau, D.H. A Comparative Analysis of Industry Human-AI Interaction Guidelines. 2020. Available online: http://xxx.lanl.gov/abs/2010.11761 (accessed on 22 October 2020).

9.    Guzdial, M.; Liao, N.; Chen, J.; Chen, S.Y.; Shah, S.; Shah, V.; Reno, J.; Smith, G.; Riedl, M.O. Friend, collaborator, student, manager: How design of an AI-driven game level editor affects creators. In Proceedings of the Conference on Human Factors in Computing Systems—Proceedings, Glasgow, UK, 4–9 May 2019; pp. 1–13. Available online: http://xxx.lanl.gov/abs/1901.06417 (accessed on 24 March 2020). [CrossRef]

10.   Abdul, A.; Vermeulen, J.; Wang, D.; Lim, B.Y.; Kankanhalli, M. Trends and trajectories for explainable, accountable and intelligible systems: An HCI research agenda. In Proceedings of the Conference on Human Factors in Computing Systems, Montreal, QC, Canada, 21–26 April 2018. [CrossRef]

11.   Wang, D.; Yang, Q.; Abdul, A.; Lim, B.Y.; States, U. Designing Theory-Driven User-Centric Explainable AI. In Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems, Glasgow, UK, 4–9 May 2019; pp. 1–15.

12.   Zhou, J.; Gandomi, A.H.; Chen, F.; Holzinger, A. Evaluating the quality of machine learning explanations: A survey on methods and metrics. *Electronics* **2021**, *10*, 593. [CrossRef]

13.   Linardatos, P.; Papastefanopoulos, V.; Kotsiantis, S. Explainable ai: A review of machine learning interpretability methods. *Entropy* **2021**, *23*, 18. [CrossRef]

14.   Bhatt, U.; Xiang, A.; Sharma, S.; Weller, A.; Taly, A.; Jia, Y.; Ghosh, J.; Puri, R.; Moura, J.M.; Eckersley, P. Explainable machine learning in deployment. In Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency, Barcelona, Spain, 27–30 January 2020; pp. 648–657. Available online: http://xxx.lanl.gov/abs/1909.06342 (accessed on 1 February 2022). [CrossRef]

15.   Knowles, M.; Baglee, D.; Wermter, S. Reinforcement Learning for Scheduling of Maintenance. In *Research and Development in Intelligent Systems XXVII*; Bramer, M., Petridis, M., Hopgood, A., Eds.; Springer: London, UK, 2011; pp. 409–422. [CrossRef]

16.   Wang, X.; Wang, H.; Qi, C. Multi-agent reinforcement learning based maintenance policy for a resource constrained flow line system. *J. Intell. Manuf.* **2016**, *27*, 325–333. [CrossRef]

17.   Barde, S.R.A.; Yacout, S.; Shin, H. Optimal preventive maintenance policy based on reinforcement learning of a fleet of military trucks. *J. Intell. Manuf.* **2019**, *30*, 147–161. [CrossRef]

18.   Hu, Y.; Miao, X.; Zhang, J.; Liu, J.; Pan, E. Reinforcement learning-driven maintenance strategy: A novel solution for long-term aircraft maintenance decision optimization. *Comput. Ind. Eng.* **2021**, *153*, 107056. [CrossRef]

19.   Mattelmaki, T.; Korkeakoulu, T. *Design Probes*; University of Art and Design: Helsinki, Finland, 2008.

20.   Gaver, W.W.; Boucher, A.; Pennington, S.; Walker, B. Cultural probes and the value of uncertainty. *Interactions* **2004**, *11*, 53. [CrossRef]

21.   Wallace, J.; McCarthy, J.; Wright, P.C.; Olivier, P. Making design probes work. In Proceedings of the Conference on Human Factors in Computing Systems, Paris, France, 27 April–2 May 2013; pp. 3441–3450. [CrossRef]

22.   Celikoglu, O.M.; Ogut, S.T.; Krippendorff, K. How Do User Stories Inspire Design? A Study of Cultural Probes. *Des. Issues* **2017**, *33*, 84–98. [CrossRef]

23.   Lange-Nielsen, F.; Lafont, X.V.; Cassar, B.; Khaled, R. Involving players earlier in the game design process using cultural probes. In Proceedings of the 4th International Conference on Fun and Games, Toulouse, France, 4–6 September 2012; ACM Press: New York, NY, USA, 2012; pp. 45–54. [CrossRef]

24.   Hutchinson, H.; Mackay, W.; Westerlund, B.; Bederson, B.B.; Druin, A.; Plaisant, C.; Beaudouin-Lafon, M.; Conversy, S.; Evans, H.; Hansen, H.; et al. Technology probes: Inspiring design for and with families. In Proceedings of the Conference on Human Factors in Computing Systems, Ft. Lauderdale, FL, USA, 5–10 April 2003; pp. 17–24.

25.   Huizinga, J. *Homo Ludens: A Study of the Play-Element in Culture*; Angelico Press: Brooklyn, NY, USA 2016.

26.   Bernhaupt, R.; Weiss, A.; Obrist, M.; Tscheligi, M. Playful probing: Making probing more fun. In *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*; 4662 LNCS; Springer: Berlin/Heidelberg, Germany, 2007; pp. 606–619._60. [CrossRef]

27.   Sjovoll, V.; Gulden, T. Play probes—As a productive space and source for information. In Proceedings of the 18th International Conference on Engineering and Product Design Education, Design Education: Collaboration and Cross-Disciplinarity, Aalborg, Denmark, 8–9 September 2016; The Design Society: Copenhagen, Denmark; 2016; pp. 342–347.

28.   Gaver, B.; Dunne, T.; Pacenti, E. Design: Cultural probes. *Interactions* **1999**, *6*, 21–29. [CrossRef]

29.   Ribeiro, J.; Roque, L. Playfully probing practice-automation dialectics in designing new ML-tools. In Proceedings of the VideoJogos 2020: 12th International Conference on Videogame Sciences and Arts, Mirandela, Portugal, 26–28 November 2020; pp. 1–9.

30.   Vaishnavi, V.K.; Purao, S., Eds. In Proceedings of the 4th International Conference on Design Science Research in Information Systems and Technology, Philadelphia, PA, USA, 7–8 May 2009.

31. Figma. Where Teams Design Together. 2020. Available online: https://www.figma.com (accessed on 13 July 2020).
32. Nielsen, J. The Usability Engineering Life Cycle. *Computer* **1992**, *25*, 12–22. [CrossRef]
33. Goebel, K.; Saxena, A.; Daigle, M.; Celaya, J.; Roychoudhury, I.; Clements, S. Introduction to prognostics. In Proceedings of the European PHM Conference, Dresden, Germany, 3–5 July 2012.
34. EASA. Certification Specifications and Guidance Material for Master Minimum Equipment List (CS-MMEL). 2021. Available online: https://www.easa.europa.eu/document-library/certification-specifications/cs-mmel-issue-3 (accessed on 1 May 2022).