The Impact of Prognostic Uncertainty on Condition-Based Maintenance Scheduling: an Integrated Approach

Tseremoglou, I.; Bieber, M.T.; Verhagen, W.J.C.; Santos, Bruno F.; Freeman, F.C.; van Kessel, P.J.

DOI
10.2514/6.2022-3967

Publication date
2022

Document Version
Final published version

Published in
AIAA AVIATION 2022 Forum

Citation (APA)
Tseremoglou, I., Bieber, M. T., Verhagen, W. J. C., Santos, B. F., Freeman, F. C., & van Kessel, P. J. (2022). The Impact of Prognostic Uncertainty on Condition-Based Maintenance Scheduling: an Integrated Approach. In *AIAA AVIATION 2022 Forum: June 27-July 1, 2022, Chicago, IL & Virtual* Article AIAA 2022-3967 (AIAA AVIATION 2022 Forum). American Institute of Aeronautics and Astronautics Inc. (AIAA). https://doi.org/10.2514/6.2022-3967

AIAA AVIATION Forum
June 27-July 1, 2022, Chicago, IL & Virtual
AIAA AVIATION 2022 Forum

10.2514/6.2022-3967

The Impact of Prognostic Uncertainty on Condition-Based Maintenance Scheduling: an Integrated Approach

* Iordanis Tseremoglou and Marie Bieber.†
Delft University of Technology, Delft, 2629HS, The Netherlands

Wim J.C. Verhagen‡
RMIT University, Carlton, Victoria, 3053, Australia

Bruno F. Santos§
Delft University of Technology, Delft, 2629HS, The Netherlands

Floris C. Freeman and Paul J. van Kessel¶
KLM Royal Dutch Airlines, Schiphol, 1117ZL, The Netherlands

One of the challenges of Condition-Based Maintenance (CBM) is to combine health monitoring and predictions with efficient scheduling tools. However, the majority of literature is focusing on the assessment of prognostics algorithms performance. In fact, the added value of these algorithms can only be assessed when considering their impact on maintenance decision process. Furthermore, in practice, when considering the scenario of an aircraft fleet with multiple monitored components, it is hard for a human decision-maker to translate and identify the effect of probabilistic results from all prognostics models from all systems on the maintenance schedule. Therefore, to support the implementation of CBM, the prognostics algorithms have to be integrated within a scheduling framework. Our paper proposes this integration in order to evaluate the impact of different level of prognostics accuracy and uncertainty on the aircraft fleet maintenance scheduling level. First, a Support Vector Regression (SVR) model is used to predict the Remaining Useful Life (RUL) distributions of the monitored components. Second, the maintenance scheduling problem is solved within a Reinforcement Learning (RL) approach incorporating a state-of-the-art Partially Observable Monte Carlo algorithm. Implementing a rolling horizon approach, our proposed framework is applied to a fleet of 10 aircraft, each equipped with multiple monitored systems. A case study with multiple different prediction accuracy and uncertainty scenarios is performed to assess the impact of prognostics uncertainty on optimal maintenance scheduling. The performed analysis aims to guide the development and assessment of prognostic models in terms of accuracy and uncertainty in the context of CBM.

I. Introduction

With the increased number of sensors installed in the modern aircraft and the necessity for cost savings due to Covid-19 impact, aircraft maintenance is currently shifting from the corrective and preventive approach towards Condition-Based Maintenance (CBM) [1], [2]. In the CBM framework, maintenance decisions are based on health diagnostics and prognostics, i.e, the component's health is monitored through sensors on a permanent basis and the collected information are processed by prognostic algorithm on the ground to predict the Remaining Useful Life (RUL) [3]. Those RUL estimates are then used to facilitate the decision making process of aircraft maintenance scheduling.

Regardless of the choice of the prognostics algorithm, RUL predictions are subject to uncertainty resulting from a variety of factors, such as the model inherent uncertainty or noise and disturbances in the sensor signals themselves

*The first two authors contributed equally.
†PhD Candidates, Air Transport and Operations, Faculty of Aerospace Engineering, Delft University of Technology, Delft, 2629HS, The Netherlands.
‡Senior Lecturer, Aerospace Engineering and Aviation, RMIT University, Carlton, Victoria, 3053, Australia.
§Associate Professor, Air Transport and Operations, Faculty of Aerospace Engineering, Delft University of Technology, Delft, 2629HS, The Netherlands
¶KLM Engineering and Maintenance,KLM Royal Dutch Airlines, Schiphol, 1117ZL, The Netherlands

[4]. It is therefore crucial to take into account the uncertainty in the prognostic output, especially when using them for decision making [5]. The decision-making process is further complicated by the fact that RUL predictions are updated on a continuous basis. This means that the maintenance schedule has to be adapted continuously using updated information from the prognostics. At the same time, the operational availability of the airline fleet must be ensured and last-minute schedule changes must be prevented.

Most of the existing CBM research focuses mainly on developing RUL prognostics algorithms. With respect to maintenance scheduling in a CBM context, the vast majority of published papers on this area focuses on phased missions, i.e. missions with multiple, consecutive, non-overlapping operational phases [6], with short scheduled breaks and fleet sizes up to 15 aircraft operating in military environments. What is more, there are relatively few studies that develop prognostics models and integrate them in a maintenance scheduling model, the most recent developed by [7]. However, all the aforementioned studies focus on defining the optimal maintenance schedule with respect to a single type of component, yet they lack the consideration and understanding of the impact of uncertainty related to multiple different system health predictions on the feasibility of the estimated maintenance plan.

The former is reflected in the absence of a framework detailing the implications and limitations of different prognostics predictions accuracy and uncertainty levels in a CBM context. Doing so, however, would both help in the evaluation of prognostic outputs as well as in the development of scheduling tools. The objective of this work is to bridge this gap. In this paper, we therefore present an integrated prognostic and scheduling framework which explicitly considers uncertainty in the output of the prognostics algorithms. To assess the prognostics accuracy and uncertainty impact on optimal maintenance scheduling, a case study is conducted on a fleet of 10 aircraft using prognostic outputs from different systems. Several maintenance scenarios with different uncertainty levels are tested to evaluate their impact on the maintenance planning process.

The main contributions of our research are summarized as follows:

- A novel integrated prognostic and scheduling framework is proposed, addressing both RUL predictions and scheduling objectives, while considering uncertainty in the RUL predictions.
- Using the developed model, we perform an extensive evaluation of the impact of different levels of accuracy and uncertainty, achieved by the prognostics algorithms for different monitored systems, on optimal maintenance scheduling.
- Finally, we demonstrate how the framework applied to a fleet of aircraft can help deriving thresholds in terms of uncertainty and accuracy to guide stakeholders in the development and implementation of CBM.

II. Methodology

The methodology consists of a prognostic and a scheduling module as shown in Figure 1. In the prognostics module, introduced in Section II.A, the prognostic models are trained on historical system data and their uncertainty is calculated. In the scheduling module, introduced in Section II.B, the aircraft maintenance schedule is produced and updated based on predictions. To be more precise, the RL scheduling tool take as an input the predicted RUL distributions of the aforementioned models for every day of the planning horizon together with the current maintenance schedule and available maintenance slots. It outputs the revised maintenance schedule, which forms the starting solution for the next day of the planning horizon.

A. Prognostics

In the prognostics module the RUL predictions are produced for each system and the uncertainty of the predicted RULs is calculated. In the following, we introduce the prognostics model and explain how the uncertainty is quantified.

1. Prognostic Model

The prognostic model used in this study is a Support Vector Regression (SVR) trained on the system input data with the capability of predicting RUL distributions. Support Vector Machines are AI techniques based on the statistical learning theory proposed by Vapnik [8]. Although there exist many kinds of AI techniques as highlighted in Section I,

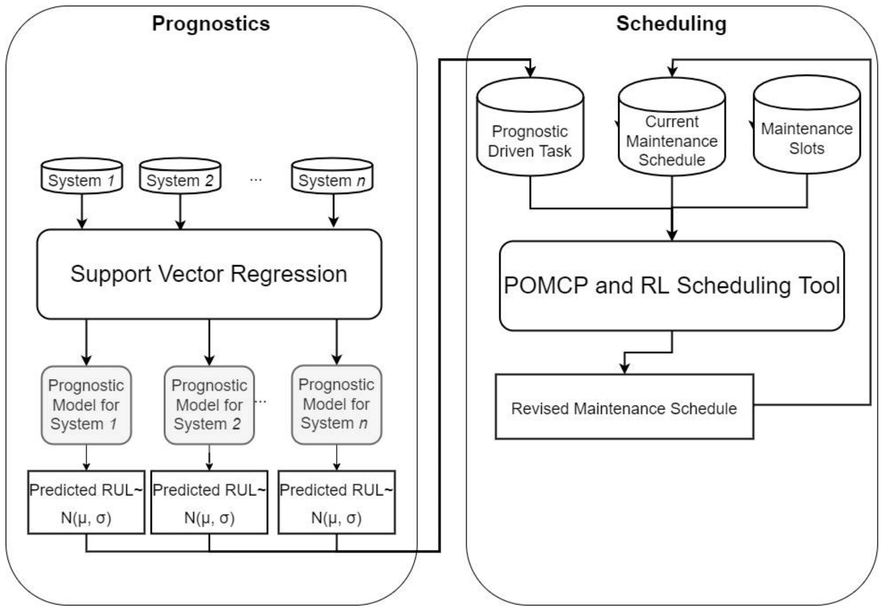

Fig. 1 The prognostic and scheduling module of the integrated prognostic scheduling framework.

we chose to use SVR in this study for several reasons: First, the study aims at investigating the effect of uncertainty on maintenance scheduling. Uncertainty in predictions can result from several sources, amongst them model inherent uncertainty. Nearly all AI models are subject to this type of uncertainty, therefore the technique of choice does not matter so much when demonstrating the effect of uncertainty. Second, SVR not only is easily adaptable to different data, but is also excellent in processing multi-dimensional data, such as data collected from different sensors [9]. As this is exactly the type of data that we use in our case study, presented in Section III, SVR is a suitable method. The SVR model is trained on the system input data and has the capability of predicting RUL distributions. Note that when we assume the following points regarding the system data:

- The system of consideration is operated until failure.
- According system data, which includes all sensor measurements related to system health, as well as operating conditions, is available from the begin of operations until the failure.
- The remaining useful life of the system is known at any time of operations.

2. Prediction Uncertainty

To quantify the uncertainty of RUL predictions, a commonly used metric is the estimation of the prediction interval [10]. Similarly as presented in [11] given the variance obtained from the model we fit a Gaussian distribution to outputs at each time step and calculate the probability cumulative distribution function. In this way we translate our RUL point estimates to prediction distributions. Usually the assumption is that the RUL estimates follow a Gaussian distribution and using the predictions at each time step, the mean is directly calculated using the RUL estimates. However, the RULs are decreasing over time, so the assumption that they follow a Gaussian distribution is not necessarily true. Therefore, we instead convert the RUL to the time between failure (TBF), as follows:

$$TBF = RUL + t_c, \tag{1}$$

where RUL ist the estimated RUL and t_c is the current time. We assume that the time between failures over all components follows a Gaussian distribution $\mathcal{N}(\mu, \sigma^2)$ centered around the mean TBF and calculate the according parameters for each time step. Finally, the mean TBF is shift back to the RUL, by subtracting t_c, the current time. Furthermore we calculate the upper and lower bounds at 3σ, i.e. retaining 99.7% of the data. Those together with the

Gaussian distribution and the predictions are further used as input parameters for the scheduling model, as explained in the next Section II.B.

B. Scheduling

For every day of the planning horizon, a new RUL prediction is generated for every system, and the corresponding prognostics-driven task is updated. For the purposes of this study we consider components pertaining to Non Safety Related Equipment (NSRE). These specific tasks have no impact on the operating safety of the aircraft.

However, in a CBM context, maintenance planning for this type of tasks becomes more challenging than the preventive and corrective maintenance tasks because of the uncertainty included in the RUL predictions. More specifically, the maintenance planner, instead of planning tasks by considering a deterministic task due date, as in the case of preventive maintenance tasks, has to derive an optimal maintenance plan based on an uncertain outcome, which is captured by the probability distribution of the predicted RUL. An example of health predictions for an aircraft monitored component is depicted below. The uncertainty of the prognostic model on the (future) condition of the considered component leads to a range in RULs, which raises great ambiguity regarding when to schedule the maintenance of the specific component.

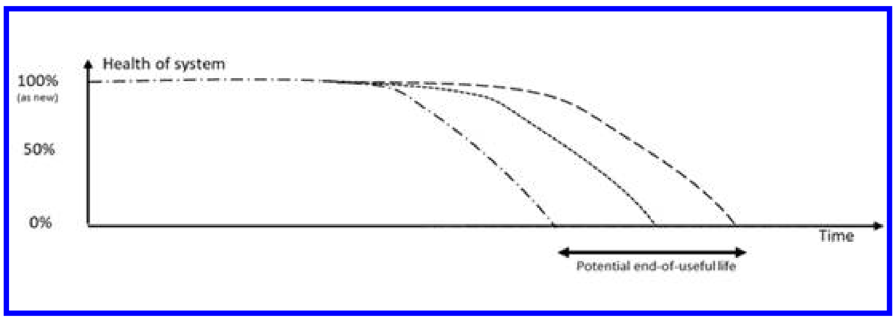

Fig. 2 Example of prognostics output.

Preventively maintaining such a component before the end of the RUL induces less maintenance costs that repairing it after a failure has been observed. Nevertheless, replacing the component too early can potentially lead to a high number of replacements, which could outweigh the benefit of less costly repairs. As such, there is a time window in which the maintenance action is beneficial (see figure below)

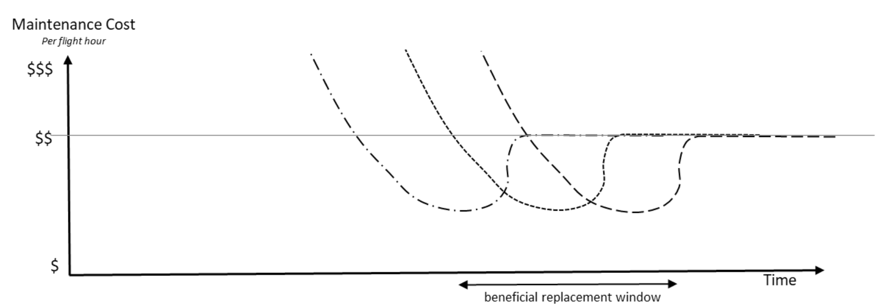

Fig. 3 Beneficial replacement window.

The optimal maintenance action for every task within this window is defined by the Partially-observable Monte Carlo Planning (POMCP) algorithm, which is used to create belief states of the health degradation of the corresponding systems. Having these belief states, knowing the maintenance slots available to each aircraft in the foreseen future, and having information about the maintenance costs associated with preventive and reactive maintenance, the POMCP prescribes the recommended dates for the execution of replacement tasks associated with the systems in an estimated degradation stage.

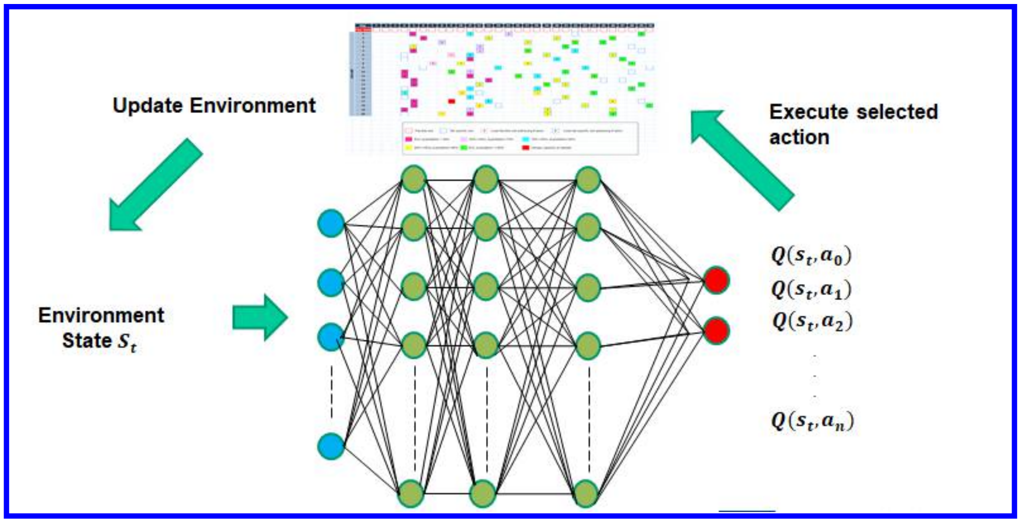

Fig. 4 RL scheduling algorithm

When the optimal maintenance action for each component has been defined, the next step is to create a suitable maintenance opportunity in the airline's network to execute the maintenance action. However, not all maintenance opportunities share the same cost. In case a maintenance buffer is selected, the aircraft occupies a slot that is no longer available to the rest of the fleet. A less costly option would be to execute maintenance on already planned maintenance, but this maintenance slot may not align well with the beneficial time window defined by the POMCP.

The decision which aircraft to schedule in which maintenance opportunity and which aircraft tasks is assigned to that maintenance opportunity is determined by the RL algorithm. For this research, the Deep Q-Learning algorithm was used. The input of the neural network is the state of the environment S_t, and the output is the Q function values of the different action that can be taken, where each action corresponds to selecting (or not) an aircraft for maintenance on that specific maintenance opportunity The design of the algorithm is summarized in Figure4.

The state of the environment is defined by the following features, as illustrated in Figure 1:

1. **Prognostic Driven Tasks:** The prognostics tasks refer to the maintenance of systems of components that are monitored through sensors permanently. The prognostics tasks are driven by the RUL predictions produced by the SVR prognostics model. Maintenance actions are triggered only when there is strong evidence of failure risk, hence decreasing the number of unnecessary maintenance actions and, at the same time, avoiding unforeseen failures (and corresponding unscheduled maintenance events).

2. **Maintenance Slots:** Execution of a maintenance task requires the availability of workforce. For each maintenance slot, an availability of eight workhours is considered.Each slot is assumed to have a start date, an end date and a designated aircraft type. We considered two types of maintenance slots:

   Fixed slots: An assigned aircraft is predefined for fixed slots, and only additional maintenance tasks from that aircraft can be assigned to the maintenance slots. Fixed maintenance slots are usually scheduled several weeks in advance and consist of more extensive maintenance operations such as letter checks.

   Flexible slots: For flexible maintenance slots, the aircraft registration is variable, and a maintenance scheduler is free to decide which aircraft to allocate to the slot, provided that the aircraft type matches the slot type. These slots are generally used for last-minute maintenance operations to execute corrective maintenance tasks.

3. **Current maintenance schedule:** A feasible schedule before new RUL predictions are received. In a feasible schedule, prognostics-driven tasks are assigned to maintenance slots.The maintenance slots schedule outlines ground time reserved at the hangar to execute maintenance.

III. Case Study

In order to test and illustrate the performance of the proposed framework, we apply the rolling horizon approach over a planning period of 70 days for a fleet of 10 aircraft with 250 CBM tasks. For every day of the planning horizon, the prognostic models predict a RUL under uncertainty. Based on the RUL prediction, the belief about the state of the component is updated and thereby impacts the maintenance scheduling.

As the aim is to understand the impact of uncertainty in predictions, we run the following three scenarios:

- **Increased Accuracy Scenario:** High prediction accuracy with low uncertainty,
- **Baseline Scenario:** Medium prediction accuracy with medium uncertainty,
- **Decreased Accuracy Scenario:** Low prediction accuracy with high uncertainty.

To analyze the impact of prognostics uncertainty on optimal maintenance planning, the scheduling framework presented in Section II is solved for the above scenarios. An overview of the simulation setup is presented in Figure 5, whereas an analysis of the simulation steps is presented in Sections III.A-III.D.

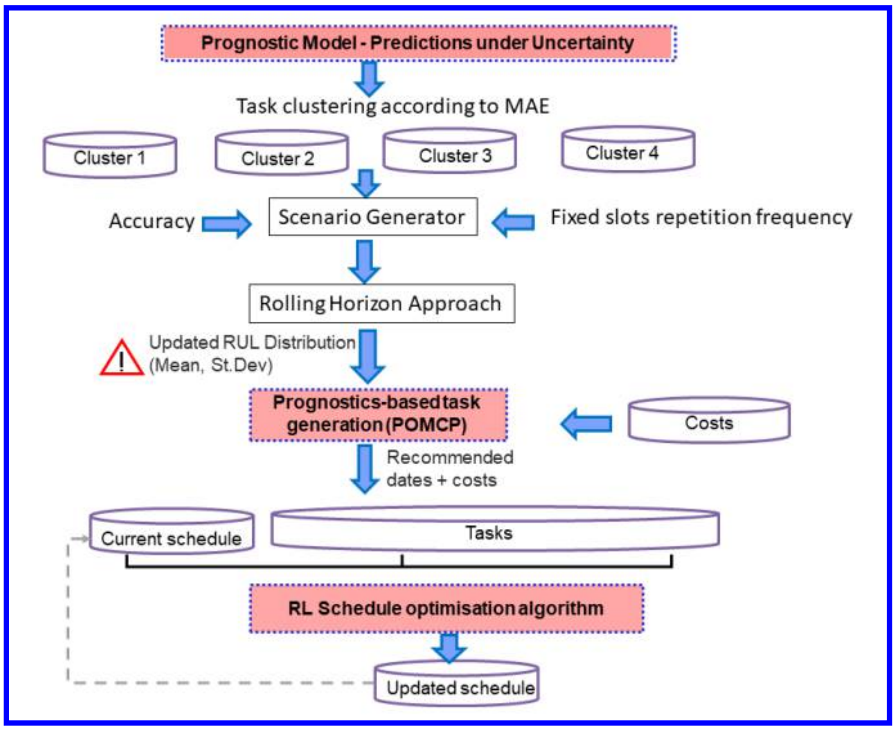

Fig. 5 Architecture - online stage simulation framework.

A. System data

We use the C-MAPSS (Commercial Modular Aero-Propulsion System Simulation) data set, which consists of four data sets, each containing simulated run-to-failure data for turbofan engines [12] [13]. The data sets differ mainly in the number of fault modes ('modes') and operating conditions ('conditions') as listed in Table 1.

Each engine is considered to be from a fleet of engines of the same type and each time series, also often referred to as trajectory, is from a single unit. The engines are operated until failure, i.e. the time series capture the operations of each unit until it fails. In the test set, the time series ends at some point before the failure and the objective is to estimate the RUL, or in other words the number of remaining operational cycles before failure. There are 21 sensor measurements and each row in the data contains the measurements corresponding to operations during one time cycle for a certain unit. From the above listed data set properties, it becomes clear that the data set is suiting for our case study as it follows the assumptions presented in Section II.A.1.

Table 1 Characteristics of the four turbofan engine data sets, note that the difference between the four data sets lies within the number of fault modes ('modes') and operating conditions ('conditions')

Data set	#modes	#conditions	#Train units	#Test units
#1	1	1	100	100
#2	1	6	260	259
#3	2	1	100	100
#4	2	6	249	248

B. Prognostic Tasks

The SVR prognostic model is trained on all four data sets, predicting the RUL at each time step for each of the test trajectories as defined in Table 1.

For the purpose of our case study, we assume that the time cycles correspond to flight cycles (FC) and assume an aircraft usage of 4 FCs per day. The predictions are evaluated for each trajectory in terms of the Mean Absolute Error (MAE) as defined in [14]. Then, all the trajectories are pooled together and organized into four clusters based on their MAE as listed below:

- Cluster 1: MAE < 20FCs
- Cluster 2: 20FCs <= MAE < 30FCs
- Cluster 3: 30FCs <= MAE < 50FCs
- Cluster 4: 50FCs <= MAE

This results in four groups/ systems for each scenario, containing data from 270 time cycles, i.e 270 FC, corresponding to 67.5 days before failure until 0 time cycles. The resulting prediction means and uncertainty bounds as introduced in Section II.A.2 are visualized in Figure 6.

For every system, two states are defined by partitioning its true RUL, RUL_{true}, in two intervals depending on FCs before failure, $I_1 = [270, 40)$ and $I_2 = [40, 0)$. The first state is referred to as the healthy state, and we consider that the component is in the healthy state when the $RUL_{true} \in I_1$. The second state is the degrading state, and the component is considered to be in that state when its $RUL_{true} \in I_2$. As such, the discrete state space for every component at flight cycle t is defined as $S_t = \{1 : Healthy, 2 : Degrading, 3 : Fail\}$.

The true state of the component cannot be directly inferred from the prognostics since at every decision epoch, the predicted RUL follows the Gaussian distribution $\mathcal{N}(\mu, \sigma^2)$. Based on the mean value μ and intervals I_1 and I_2 we define the observed state O_t. This observed state stochastically relates to the true underlying but hidden state of the component, which is either healthy or degrading. This relationship is described by the state observation matrix, which captures the probability that the decision-maker observes that the component is in state $S_t = y$ while the component is in true unobservable state $S_t = i$. The corresponding state observation matrices for every cluster are presented in Table 2.

C. Simulation Scenarios

As mentioned above, we simulate three scenarios with different prediction accuracy and uncertainties. In the baseline scenario, the predictions from the SVR prognostic model as depicted in Table 5 are used. To capture the effect of varying prediction uncertainty, we simulate two additional sets of clusters,where in each set the prediction accuracy of the components belonging in the clusters described in Section III.B has been increased and decreased respectively by 40%. The corresponding state observation matrices are presented in Table 3 & 4.
Moreover, each scenario is tested on two different fixed slot repetition policies. According to the first policy, the fixed maintenance slots are repeated every 7 days resulting to a total of 92 fixed maintenance slots and 15 flexible slots, whereas in the second policy the fixed maintenance slots are repeated every 15 days resulting to a total of 50 fixed maintenance slots and 50 flexible slots. It is noted that in both cases the same amount of workhours is distributed among slots.

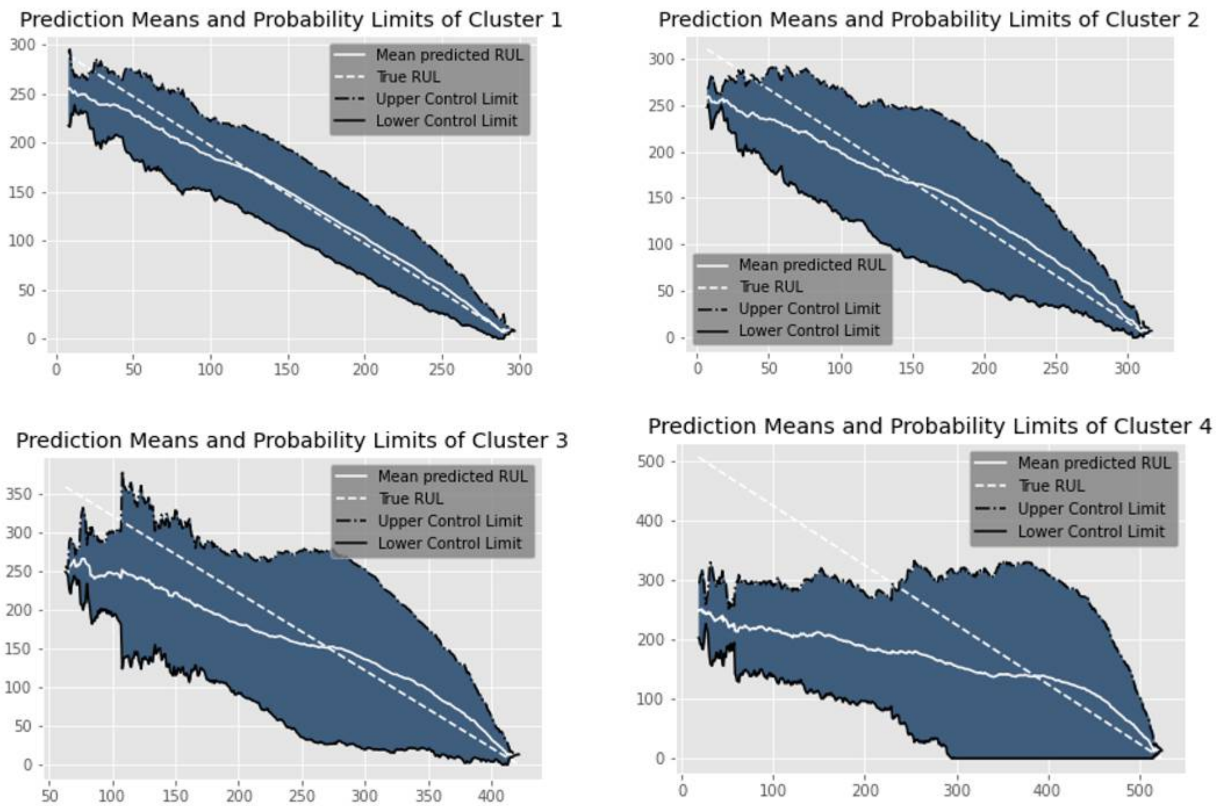

Fig. 6 The predicted mean RUL and uncertainty bounds for each cluster.

Table 2 State Observation Matrices - Baseline Scenario.

CLUSTER 1 (MAE~8.12)	Healthy	Degrading
Heathy	0.99	0.01
Degrading	0.14	0.86

CLUSTER 2 (MAE~13.89)	Healthy	Degrading
Healthy	1	0
Degrading	0.22	0.78

CLUSTER 3 (MAE~23.52)	Healthy	Degrading
Healthy	0.99	0.01
Degrading	0.31	0.69

CLUSTER 4 (MAE~37.28)	Healthy	Degrading
Healthy	0.96	0.04
Degrading	0.39	0.61

D. Scheduling

The aircraft maintenance scheduling algorithm is solved using a rolling horizon approach for a period of 70 days. In each iteration of the planning horizon, the POMCP algorithm calculates the optimal maintenance date for each component based on the minimization of the maintenance cost and the RL agent devises the maintenance schedule accordingly. Afterwards, the planning horizon shifts one day ahead, at which a new RUL prediction for every component is obtained and a new maintenance schedule is devised. The same process continues until all prognostics-driven tasks have been scheduled.

Table 3 State Observation Matrices - Increased Accuracy Scenario.

CLUSTER 1 (MAE~4.87)	Healthy	Degrading
Healthy	0.99	0.01
Degrading	0.11	0.89

CLUSTER 2 (MAE~8.33)	Healthy	Degrading
Healthy	0.99	0.01
Degrading	0.16	0.84

CLUSTER 3 (MAE~14.11)	Healthy	Degrading
Healthy	0.98	0.02
Degrading	0.23	0.77

CLUSTER 4 (MAE~22.37)	Healthy	Degrading
Healthy	0.95	0.05
Degrading	0.29	0.71

Table 4 State Observation Matrices - Decreased Accuracy Scenario.

CLUSTER 1 (MAE~11.39)	Healthy	Degrading
Healthy	1	0
Degrading	0.17	0.83

CLUSTER 2 (MAE~19.44)	Healthy	Degrading
Healthy	1	0
Degrading	0.28	0.72

CLUSTER 3 (MAE~32.93)	Healthy	Degrading
Healthy	1	0
Degrading	0.39	0.61

CLUSTER 4 (MAE~52.20)	Healthy	Degrading
Healthy	1	0
Degrading	0.48	0.52

IV. Results

The goal of the analysis is to evaluate the impact of the different uncertainty levels on the optimal schedule devised by the RL agent and identify promising maintenance policies. The uncertainty impact is going to be evaluated on the basis of the following four objectives:

1. The first objective is **to execute tasks ahead of their due date**. When exceeding the task due date, the aircraft is no longer airworthy and would have to be kept on the ground until the task is performed, inducing major costs to the airline.

2. Secondly, after guaranteeing the execution of tasks in time, the goal of an MRO should be **to provide fleet availability towards the airline**. In general, the airline prefers to assign tasks to the already defined fixed maintenance slots and limit the use of flexible slots to cases where it is absolutely necessary. This will be reflected to the number of flexible slots that will be used when running each scenario.

3. A schedule change is defined to be a change in the aircraft registration assigned to a flexible maintenance slot, when compared with the original schedule. Reallocation of tasks between fixed maintenance slots is not considered as a schedule change since it does not affect the availability of the aircraft for operations. In the case of a schedule change, the number of days of notice is of high importance. An aircraft allocation change on the day of operations is more costly than an aircraft change one week ahead. Therefore, **the number of schedule changes in the last three days before the day of operations are assessed**.

4. The last objective is **to plan individual tasks at the optimal moment in time**. As described above task execution

is the most important objective. For the prognostics-driven task the goal is to schedule close to the date that minimizes the maintenance cost and subsequently minimize the waste of the RUL of the component.

In order to have a low bound reference, the RL agent is run with perfect information, i.e., assuming that all tasks have deterministic due days and are known in advance. Thus the schedule is created in one iteration. Moreover, each scenario is tested on two different fixed slot repetition policies. According to the first policy, the fixed maintenance slots are repeated every 7 days resulting to a total of 98 fixed maintenance slots and 25 flexible slots, whereas in the second policy the fixed maintenance slots are repeated every 15 days resulting to a total of 50 fixed maintenance slots and 50 flexible slots. The results of the model with perfect information are summarized in Table 5.

Table 5 Scheduling with perfect information

Perfect Information		Fixed Slot Frequency	
		7 Days	15 Days
RL agent	Failed tasks	0	0
	Flexible slots used	1	3
	Task utilization	82.20%	70.6%

The results of the six simulated scenarios are summarized in Table 6.

Table 6 Resulting failed tasks, flexible slots used, schedule changes and task utilization for the different simulated scenarios.

		Base Scenario		Scenario with increased accuracy		Scenario with decreased accuracy	
		Fixed Slot Frequency		Fixed Slot Frequency		Fixed Slot Frequency	
		7 Days	15 Days	7 Days	15 Days	7 Days	15 Days
RL agent	Failed tasks	25	9	21	4	32	15
	Flexible slots used	18	27	15	19	23	45
	Schedule Changes	34	43	21	37	36	25
	Task utilization	81.11%	71.57%	77.5%	68.2%	83.6%	73.6%

From the table, it can be observed that a maintenance strategy using less fixed and more flexible slots results, in all scenarios, in less failures. Moreover, increasing the accuracy of the predictions results to less components failing, less flexible slots used and less schedule changes. However, a lower task utilization is achieved, due to the fact that the RL agent schedules more tasks in comparison with the baseline and the decreased accuracy scenario. To shed further light into the effect of uncertainty into the number of failures and the stability of the schedule, we analyze the effect of each cluster on the former objectives. The results are presented in Figures 7 & 8.

From Figure 7, it can be concluded that the highest amount of schedule changes and failed tasks are due to tasks belonging to cluster 3 and 4 which almost in all scenarios (except for the cluster 3 in scenario with increased accuracy) has a $MAE > 20$. To gain insights in the uncertainty impact, we also plot the total number of schedule changes and failed tasks observed across all scenarios for all tasks having $MAE > 20$ and tasks having $MAE < 20$ in Figure 9.

It can be observed that out of a total of 196 schedule changes and 106 failed tasks, approximately 71.4% (140) and 77% (82) are due to tasks having $MAE > 20$. Translating this to uncertainty from the observation matrices, it can be concluded that having an uncertainty greater than approximately 29% in the identification of the degrading state can have a severe impact on the efficiency and stability of the schedule. As such, having components with historical values of predictions falling within this interval of uncertainty introduces the need for the airline to include more flexible slots in the schedule as maintenance planning strategy. On the other hand, an order of uncertainty for the degrading state below 29% corresponds to a 'low-threat' planning state, as relatively high values are achieved for the four target features of planning.

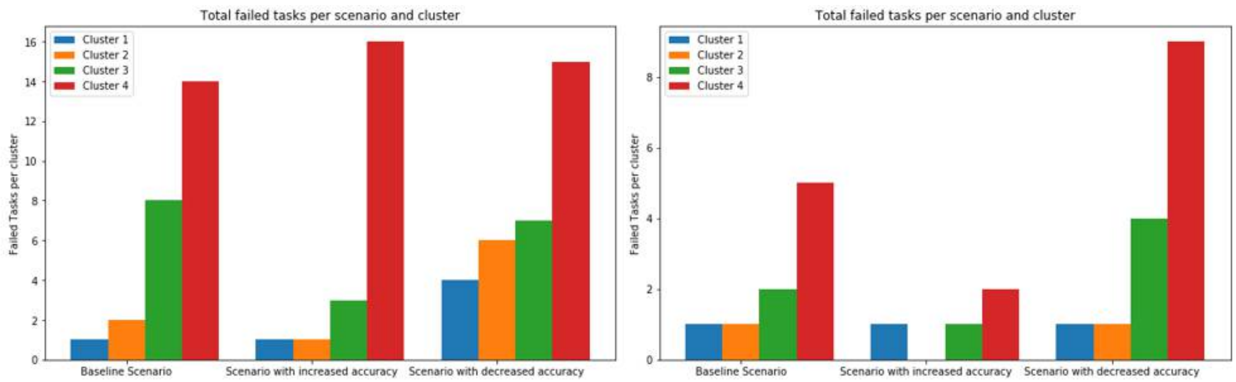

Fig. 7 Failed tasks per cluster – 7 days(left figure) vs 15 days (right figure).

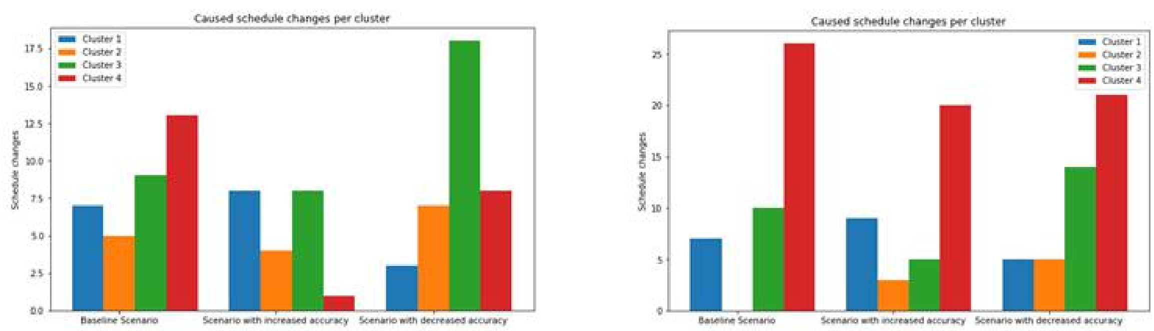

Fig. 8 Caused schedule changes per cluster – 7 days(left figure) vs 15 days (right figure).

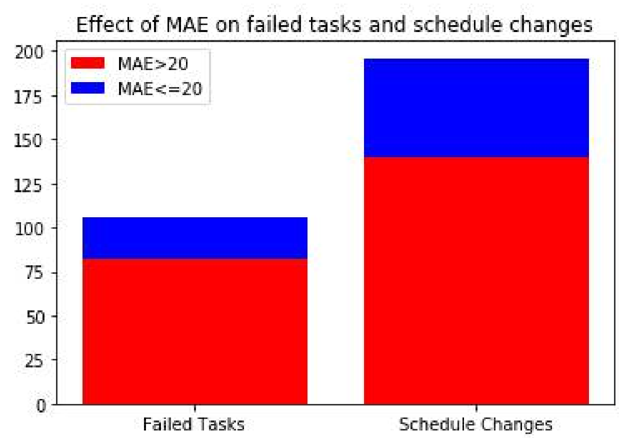

Fig. 9 Total failed tasks and schedule changes in relation with MAE.

V. Conclusion

In this paper we presented an analysis of the uncertainty impact on optimal maintenance planning. A SVR prognostics model was used to predict the RUL distributions for a set of monitored components. A set of maintenance scenarios were considered in order to identify problematic and promising regions of the uncertainty space of the predictions from prognostics models. The maintenance scenarios were solved using a RL scheduling algorithm. The results highlight that having prognostics tasks with prediction uncertainty within the degrading state higher than 29% corresponds to potentially high-risk maintenance decisions. A maintenance strategy to limit the influence of these tasks over the optimal maintenance planning would be to include more flexible slots in the upcoming planning horizon. Finally, the results indicate that a having open prognostics tasks having prediction uncertainty lower than 29% is a low-risk planning state with respect to optimal maintenance planning.

Further research steps involve:

- extending the model to account for more aircraft and more monitored systems,
- including routine and non-routine tasks in the scheduling process
- extending the analysis to include different prognostic models with various uncertainty and accuracy levels.

Acknowledgement

The work presented in this paper was financially supported by the European Union's Horizon 2020 research and innovation program ReMAP (Grant Agreement Number:769288). The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Contribution Statement

Marie Bieber: Conceptualization; Data curation; Formal analysis; Investigation; Methodology; Software; Validation; Visualization; Writing - original draft; Writing - review & editing.

Iordanis Tseremoglou: Conceptualization; Data curation; Formal analysis; Investigation; Methodology; Software; Validation; Visualization; Writing - original draft; Writing - review & editing.

Wim J.C. Verhagen: Conceptualization; Supervision; Writing - review & editing.

Bruno F. Santos: Conceptualization; Methodology; Supervision; Writing - review & editing.

Floris C. Freeeman: Conceptualization; Writing - review & editing.

Paul J. van Kessel: Conceptualization; Writing - review & editing.

References

[1] Peng, Y., Dong, M., and Zuo, M. J., "Current status of machine prognostics in condition-based maintenance: a review," *The International Journal of Advanced Manufacturing Technology*, Vol. 50, No. 1-4, 2010, pp. 297–313.

[2] Zhang, J., Wang, P., Yan, R., and Gao, R. X., "Deep Learning for Improved System Remaining Life Prediction," *Procedia CIRP*, Vol. 72, 2018, pp. 1033–1038. https://doi.org/10.1016/j.procir.2018.03.262, URL https://doi.org/10.1016/j.procir.2018.03.262.

[3] Swearingen, K., Majkowski, W., Bruggeman, B., Gilbertson, D., Dunsdon, J., and Sykes, B., "An open system architecture for condition based maintenance overview," Tech. rep., 2007. URL http://www.mimosa.org/mimosa-osa-cbm/.

[4] Javed, K., Gouriveau, R., and Zerhouni, N., "State of the art and taxonomy of prognostics approaches, trends of prognostics applications and open issues towards maturity at different technology readiness levels," , 2017. https://doi.org/10.1016/j.ymssp.2017.01.050.

[5] Saxena, A., Sankararaman, S., and Goebel, K., "Performance Evaluation for Fleet-based and Unit-based Prognostic Methods," *European Conference of the Prognostics and Health Management Society*, 2014, pp. 1–12.

[6] Li, X. Y., Huang, H. Z., and Li, Y. F., "Reliability analysis of phased mission system with non-exponential and partially repairable components," *Reliability Engineering and System Safety*, Vol. 175, No. March, 2018, pp. 119–127. https://doi.org/10.1016/j.ress.2018.03.008.

[7] de Pater, I., and Mitici, M., "Predictive maintenance for multi-component systems of repairables with Remaining-Useful-Life prognostics and a limited stock of spare components," *Reliability Engineering and System Safety*, Vol. 214, 2021. https://doi.org/10.1016/j.ress.2021.107761.

[8] Vapnik, V. N., "An overview of statistical learning theory," *IEEE Transactions on Neural Networks*, Vol. 10, No. 5, 1999, pp. 988–999. https://doi.org/10.1109/72.788640.

[9] Huang, H. Z., Wang, H. K., Li, Y. F., Zhang, L., and Liu, Z., "Support vector machine based estimation of remaining useful life: current research status and future trends," *Journal of Mechanical Science and Technology*, Vol. 29, No. 1, 2015, pp. 151–163. https://doi.org/10.1007/s12206-014-1222-z.

[10] Liao, Y., Zhang, L., and Liu, C., "Uncertainty Prediction of Remaining Useful Life Using Long Short-Term Memory Network Based on Bootstrap Method," *2018 IEEE International Conference on Prognostics and Health Management, ICPHM 2018*, , No. August 2019, 2018, pp. 1–8. https://doi.org/10.1109/ICPHM.2018.8448804.

[11] Biggio, L., Chao, M. A., and Fink, O., "Uncertainty-aware Remaining Useful Life predictors," *NeurIPS 2020-Workshop on Machine Learning for Engineering Modeling, Simulation and Design*, , No. Dl, 2020, pp. 1–6.

[12] Frederick, D. K., DeCastro, J. A., and Litt, J. S., "User's guide for the commercial modular aero-propulsion system simulation (C-MAPSS)," 2007.

[13] Saxena, A., Goebel, K., Simon, D., and Eklund, N., "Damage Propagation Modeling for Aircraft Engine Prognostics," 2008.

[14] Saxena, A., Celaya, J., Balaban, E., Goebel, K., Saha, B., Saha, S., and Schwabacher, M., "Metrics for evaluating performance of prognostic techniques," *2008 International Conference on Prognostics and Health Management, PHM 2008*, 2008. https://doi.org/10.1109/PHM.2008.4711436.