Aircraft Maintenance Planning

Genetic Algorithm Optimization of Aircraft Hangar Maintenance Planning Under Uncertainty

Ties Hollander

MSc Thesis Aerospace Engineering

TUDelft

Aircraft Maintenance Planning

Genetic Algorithm Optimization of Aircraft Hangar Maintenance Planning Under Uncertainty

Thesis report

by

Ties Hollander

to obtain the degree of Master of Science
at the Delft University of Technology
to be defended publicly on July 23, 2025 at 13:30.

Thesis committee:	
Chair:	Dr.ir. M.F.M. Hoogreef
Supervisors:	Dr.ir. M.J. Ribeiro
	Ir. M.M.D. Witteman
External examiner:	Dr.ir. M. Popović
Place:	Faculty of Aerospace Engineering, Delft
Project Duration:	September, 2024 - July, 2025
Student name:	T.J.A. Hollander
Student number:	4850939

An electronic version of this thesis is available at `https://repository.tudelft.nl/`.

Cover image by Jinally
(https://mydecorative.com/5-tips-for-aircraft-hangar-maintenance/)

Faculty of Aerospace Engineering · Delft University of Technology

Preface

Handing in this report means the end of a seven-year adventure in Delft. It was an adventure that has shaped me not only academically but also personally. One memory that stands out during the thesis is my visit to an aircraft hangar, where I had the opportunity to observe aircraft maintenance up close. This experience helped me connect my thesis work to real-world practice and strengthened my motivation for the field.

The process of this thesis began over a year ago, when I met for the first time with my supervisor, Marta Ribeiro, to discuss a topic. We quickly decided to work on hangar maintenance planning. I want to thank my daily supervisors, Marta Ribeiro and Max Witteman, for their continuous support throughout the project. You were always available when I had questions or needed to discuss ideas. Your quick replies, whether by e-mail or Microsoft Teams, made a real difference. Your clear and concrete feedback was very valuable and I have learnt a lot from it.

In addition to my supervisors, I want to thank my friends who were also working on their theses at the same time. The shared experiences helped to make the process enjoyable. I also grateful to my fellow board members of the VSV, and the study society as a whole, for the warm environment and the support I received. Finally, I would like to thank my parents and sisters for their support and love from home. And a special mention goes to the university library of the TU Delft: without that building, I doubt I would have made it past the first year.

To conclude, I would like to continue a small personal tradition by ending with a limerick:

> This thesis took a year,
> With guidance kind and clear.
> Through effort and stride,
> With peers by my side,
> We've reached the end we hold so dear.

Ties Hollander
Delft, July 2025

Contents

Nomenclature

List of Abbreviations

AMOS Aircraft Maintenance Operations Simulation

AMP Aircraft Maintenance Program

CS Change Score

DP Dynamic Programming

EASA European Union Aviation Safety Agency

ETA Estimated Time of Arrival

FAA Federal Aviation Administration

FC Flight Cycles

FH Flight Hours

GA Genetic Algorithm

H-AMCS Hangar Aircraft Maintenance Scheduling

HMP Hangar Maintenance Planning

IP Integer Programming

KPI Key Performance Indicator

LLN Law of Large Numbers

MCS Monte Carlo Simulation

MILP Mixed-integer Linear Programming

MIP Mixed Integer Programming

ML Machine Learning

MP Mathematical Programming

MPD Maintenance Planning Document

MRO Maintenance, Repair, and Overhaul

RL Reinforcement Learning

RUL Remaining-Useful-Life

RWS Roulette-wheel Selection

SI Similarity Index

SSS Steady-state Selection

SUS Stochastic Universal Selection

DY Calendar Days

List of Figures

List of Tables

1 Introduction

Aviation is one of the safest modes of travel in the world, due to strict regulations. One of the important factors in aviation safety is maintenance. On average, aircraft maintenance forms around 10% of an airline's operational cost [1]. Heavy maintenance, or maintenance that is executed in the hangar, is a significant part of that. It can account for more than 70% of the maintenance costs and requires a large amount of resources [2]. In present days, airlines operate in environments with small margins and it is beneficial for them to keep the costs low by improving their maintenance scheduling and efficiency.

Initially, maintenance started as a simple and straightforward process. However, due to a more dynamic environment where both costs and the complexity of aircraft continued to increase, manual maintenance planning became more and more impracticable [3]. This resulted in airlines organizing their maintenance in a more systematic way to save costs and achieve a higher efficiency, with Air Canada being one of the first in 1977 [4]. Nowadays, maintenance is usually planned by maintenance planners aided by computer tools.

This thesis proposes a stochastic optimisation-simulation model, where a Genetic Algorithm (GA) is applied to generate an initial maintenance schedule for aircraft maintenance. The planning is subjected to a Monte Carlo simulation to analyse feasibility across different scenarios to improve its robustness. The model is validated in a case study of a European airline with a large, heterogeneous fleet. The framework can be used in the maintenance planner decision process to evaluate various schedules with a feasibility assessment, giving insight into the performance of the planning and possibilities for robust improvement.

This thesis report is divided into three parts. Part I provides an introduction to the topics treated in this work; the principles of aircraft maintenance are introduced, a literature review is performed and the research scope is defined. In Part II, the work performed in this research is presented in the form of a scientific paper, where the appendix is included. Finally Part III, provides the bibliography used for this work.

Part I

Literature Review & Research Definition

2

Problem Definition

The proper execution of aircraft maintenance is important for three reasons: safety, time, and costs. Firstly, the aviation industry adheres to very high safety standards. Aircraft need regular maintenance to ensure compliance with safety and airworthiness requirements from institutions such as the European Union Aviation Safety Agency (EASA) and the Federal Aviation Administration (FAA). These are aviation safety institutions with the mission of ensuring safe air operations by formulating rules, standards, and guidance and by certifying aircraft, parts, and equipment. An example of such a regulation is that a certificate of release of an aircraft to service after maintenance can only be issued by a subcontractor who has received a certification authorisation from the Part-145 organisation. Next to that, the Maintenance Planning Document (MPD) exists. These documents are provided by aircraft manufacturers to explain the repetitive tasks that are required to maintain their aircraft. Maintenance planning engineers use the MPD information to develop operator maintenance programs that are then submitted to the relevant aviation authority for approval.

Secondly, aircraft maintenance is labour- and material-intensive. It is constrained by the availability of material, machinery, method, and manpower (4M). These are explained in Table 2.1. This requires a lot of planning, and maintenance planners can take up to several weeks to create an annual schedule.

Combining all the time and labour spent to adhere to safety regulations, it is not surprising that maintenance expenses are a major contributor to the total operational costs of an airline. In 2022, airlines spent around 10.9% of all operational expenses on Maintenance, Repair, and Overhaul (MRO) [6]. This chapter first discusses different types of maintenance in Section 2.1, then goes more into one of those types in Section 2.2 and lastly, maintenance planning is discussed in Section 2.3.

2.1. Types of Maintenance

Aircraft maintenance comprises all aspects of keeping an aircraft airworthy and in serviceable condition at a minimum cost. This is similar for both civil and military aircraft. It involves overhaul, repair, inspection

Table 2.1: Formulation of **4M** requirements within the analysis [5].

4M’s	Explanation
Method	The estimated execution duration of a task needs to fit within the scheduled duration of a maintenance slot.
Machinery	The scheduled execution date needs to be after the ETA of the required machinery.
Material	The scheduled execution date needs to be after the ETA of the required material.
Manpower	Workforce which satisfy the skill requirements of the task, need to be scheduled to the corresponding maintenance slot for the required amount of workforce hours.

or modification of aircraft and aircraft components [7].

Maintenance is usually classified into three categories:

- **Corrective maintenance:** Maintenance tasks are performed after failure to restore a component to a satisfactory condition by providing correction of a known or suspected malfunction and/or defect [7].
- **Predictive maintenance:** Maintenance tasks are performed shortly before when data-driven analytics predict equipment failures, by direct monitoring of the condition of the equipment in service [8].
- **Preventive or scheduled maintenance:** Maintenance tasks are performed at defined intervals to retain an item in a serviceable condition by systematic inspection, detection, replacement of worn-out items, adjustment, calibration, cleaning, etc. [7].

Figure 2.1 shows a simplified flowchart of the steps an aircraft goes through in hangar maintenance. It starts in the green box on the left at operation. Then, depending on the airline, a positioning flight might be necessary to bring the aircraft to the maintenance facility, where it goes to the hangar. They are coloured in orange, as this step is not always necessary. For some checks, access panels need to be opened on the aircraft. These panels allow access for maintenance or inspection of specific aircraft systems and structures. Simultaneously, they protect underneath parts and components when closed [9]. When the right panels are open, the actual maintenance operation can take place, such as lubrication, repair, or visual inspection. After successful maintenance execution, the panels can be closed, and the aircraft can leave the hangar and fly back to its base to continue operations. As said before, it is desired that this cycle be as efficient as possible so that the aircraft has maximum time in operation and can create revenue for the airline by carrying passengers. At the same time, it is beneficial to maximize the number of maintenance slots to have as many opportunities to plan maintenance as possible. This creates a conflict in optimisation and should be balanced: enough time on the ground to execute the necessary maintenance checks, while having enough operational availability for the aircraft to generate revenue by carrying cargo and passengers.

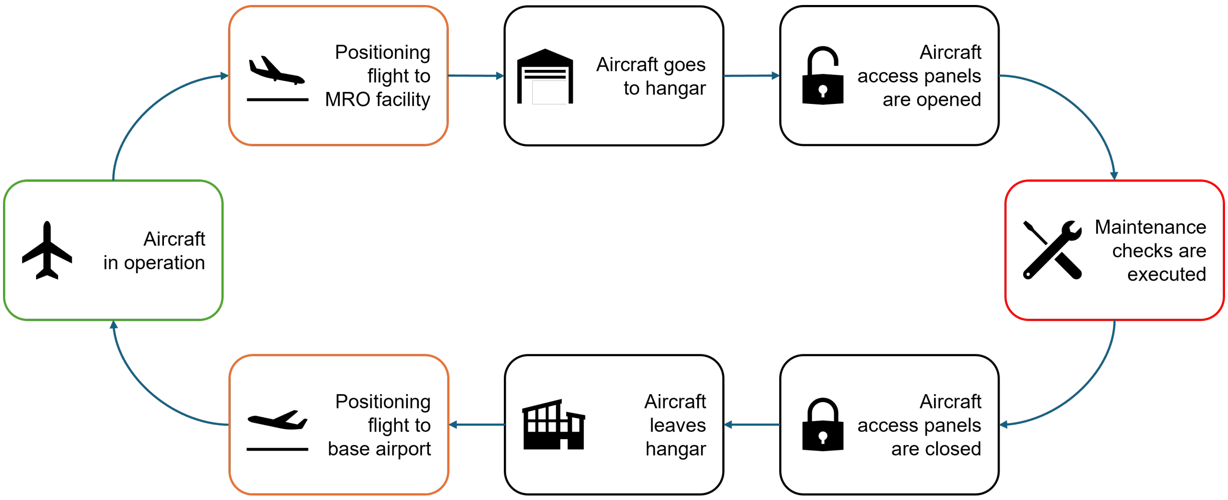


Figure 2.1: Simplified hangar maintenance flow.

Growth of Predictive Maintenance

Predictive maintenance is a growing research topic, with several companies offering products that can give insight into sensor data, such as Prognos from Airfrance-KLM and Aviatar from Lufthansa Technik. However, the implementation into practice is still small. There are several reasons for that. Examples are safety considerations in a conservative industry, insufficiently formalized decision-making processes and methods and the extensive list of different aircraft systems and failure modes requiring a specialized and dedicated approach [10]. Hence, preventive maintenance remains the dominant and standardized method of planning maintenance. Thus, it also remains relevant to investigate the optimization of preventive maintenance schedules. This will be discussed in the following section.

2.2. Preventive Maintenance

Many different requirements need to be met to execute preventive aircraft maintenance. The requirements state that different maintenance tasks must be carried out after a certain interval, which can be a number of Flight Hours (FH), Flight Cycles (FC), or Calendar Days (DY). The interval that is exceeded first determines the due date of this specific check. All three parameters are set back to zero once the task is executed. If a task exceeds the due date, an aircraft will lose its airworthiness, resulting in an Aircraft-on-Ground (AOG) scenario. This is not desired by airlines, as it is very expensive. A rough estimate for the daily potential revenue from a wide body aircraft in commercial operations equals around €18,000 [5]. The MPD or maintenance programme contains the maintenance information: the intervals and all maintenance tasks to be carried out for each aircraft type, approved by the competent authority [11]. This strategy takes care of the safety and reliability of aviation today, but because of the statistical generalizations on which those intervals are based, it can lead to component replacements long before their actual due date is reached or to component failures before the assigned maintenance date. In both cases, additional operational costs are caused [12].

The maintenance can be performed at either line or base maintenance, depending on the amount of maintenance and type of tasks to be executed. Line maintenance means it can be done at the gate or apron. Base maintenance is executed at a hangar, and they are traditionally divided into different letter checks named A, B, C, and D. Table 2.2 categorises the different aircraft maintenance tasks. An A-check is conducted every 400–600 flying hours and lasts no more than 24 hours for a narrow-body aircraft, while the B-check should be conducted every 6–8 months. C- and D-checks are carried out much less frequently, as these take an aircraft out of flight service for several weeks and are more extensive and complex [13]. The overview can be seen in Table 2.3, which also indicates the type of maintenance tasks executed per letter check.

Table 2.2: Taxonomy for aircraft maintenance (Modified from: [3]).

<table>
<tr><th rowspan="2"></th><th colspan="3">Lay-over or light maintenance</th><th colspan="2">Heavy maintenance</th></tr>
<tr><td>Line maintenance</td><td colspan="2">Line or hangar maintenance</td><td colspan="2">Hangar maintenance</td></tr>
<tr><td rowspan="2">Preventive or routine</td><td>Short-term</td><td colspan="2">Mid-term or regular checks</td><td colspan="2">Long-term</td></tr>
<tr><td>Pre-flight, transit, daily checks</td><td>A-check</td><td>B-check</td><td>C-check</td><td>D-check</td></tr>
<tr><td rowspan="2">Unscheduled or non-routine</td><td>Predictive or on-condition maintenance</td><td colspan="2">Predictive or on-condition maintenance</td><td colspan="2">Predictive or on-condition maintenance</td></tr>
<tr><td>Corrective or emergency maintenance</td><td colspan="2">Corrective or emergency maintenance</td><td colspan="2">Corrective or emergency maintenance</td></tr>
</table>

2.3. Maintenance Planning

Maintenance planning is considered to consist of all activities required to ensure maintenance can be executed. This ranges from determining the work to be planned and estimating the required manhours to the development of schedules for the planned maintenance jobs, bays, equipment, aircraft, etc. In this document, maintenance planning and maintenance scheduling are used interchangeably. Figure 2.2

Table 2.3: Aircraft letter check and corresponding inspection interval [14].

Check	Maintenance Type	Interval	Maintenance Tasks
A-check	Light maintenance	2-3 months	External visual inspection, filter replacement, lubrication etc.
B-check	Light maintenance	Rarely mentioned	Tasks are commonly incorporated into successive A-checks
C-check	Heavy maintenance	18-24 months	Thorough inspection of the individual systems and components
D-check	Heavy maintenance	6-10 years	Thorough inspection of most structurally significant items

shows an overview of the most important steps in planning. It flows from left to right getting closer to the maintenance execution in time and more detailed in planning. Of the 4M (Manpower, Machine, Material and Method), three are present in this figure. These are the important building blocks for maintenance.

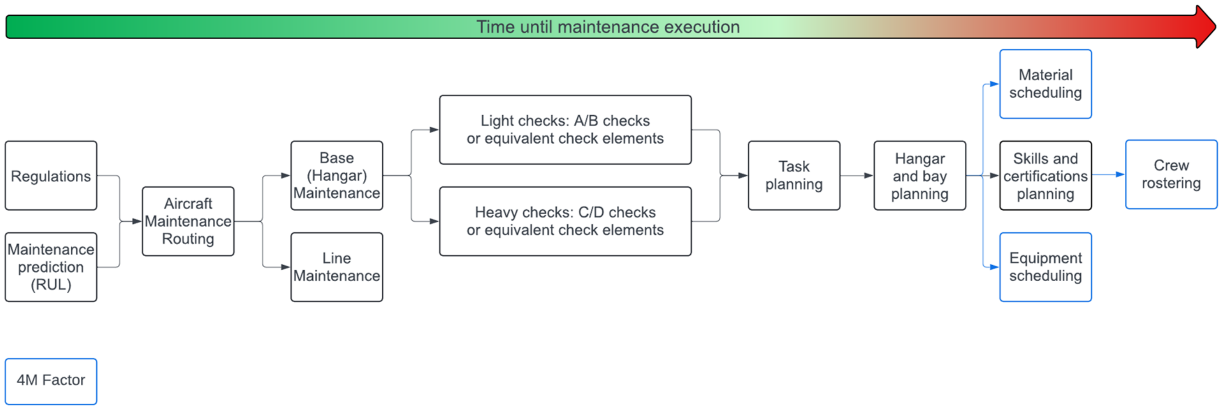


Figure 2.2: Overview of Aircraft Maintenance Planning.

According to Samaranayake and Kiridena (2012) [15], there are two large streams of literature regarding the heavy maintenance problem. One focuses on the overall scheduling of a fleet of aircraft at specific hangars and the other on more detailed planning aspects, such as tasks, materials, resources, and personnel. The first stream can be found more on the left side of Figure 2.2, whereas the second stream is merely found on the right side. The first stream is more extensive and uses approaches such as linear programming, heuristics, integer programming, etc. to optimise the overall performance. The second stream is more sparse and uses mathematical models, decision support systems, and expert systems focusing mostly on workforce allocation.

Going through Figure 2.2, one would first encounter regulations and maintenance prediction, or the prediction of the Remaining-Useful-Life (RUL). Regulations, as described earlier, are mandated by aviation authorities and determine when maintenance should take place. At the same time, the prediction of the RUL is used in predictive maintenance to determine when an aircraft would need service. Following, the aircraft maintenance routing problem aims to determine the assignment of available aircraft to cover all the flights in an airline network with the objective of maximizing aircraft utilisation with full coverage of the necessary maintenance. Many of the studied aircraft maintenance optimization problems concern this optimization (see, e.g., Sriram and Haghani, (2003) [16]) [17].

The overview then splits into base or hangar maintenance and line maintenance. The focus of this research is on hangar maintenance, hence this was further investigated and divided into light and heavy checks as described earlier in Section 2.2. The following blocks are discussed in Chapter 3 to go into more detail on the literature. This is however not done for skills and certifications planning, which is why it is explained underneath.

Skills and Certifications
Skills and certifications are also important in aircraft maintenance, as they help to ensure the quality and traceability of the maintenance performed. Personnel and engineers need to have valid licenses to work. When the maintenance is finished and the airplane is approved to fly again, it can only be released by licensed personnel. In an ideal scenario, every engineer has a certificate for every aircraft type. The regulations per country and airline where the maintenance is performed, however, determine how many licenses the engineers and technicians may have [3]. For example, KLM's internal safety rules prohibit engineers from holding licenses for more than two aircraft types and one skill [18]. Engineers in Taiwan, on the other hand, can obtain licences for three aircraft types [19].

The maintenance planning problem examined in literature rarely includes certification, despite its importance. Although required for technicians before they may perform maintenance, it varies greatly across different countries and airlines. That can be a reason why certification may be frequently mentioned in operations research studies but it is seldom used as a constraint in the workforce scheduling problems for maintenance. Skill levels are more often included in planning problems. Usually, this skill requirement is merely integrated via manpower or other constraints. Some do mention specific skill constraints of the personnel, such as Dijkstra et al. [18] did. In general, many different skills are mentioned in research. All of those specific skills, however, are not incorporated in workforce scheduling or planning problems. This could be explored further and connected to particular training procedures and costs [3].

To conclude, it is clear that it is important to plan aircraft maintenance well. Airlines spend a lot of effort to properly schedule all their maintenance tasks to adhere to safety regulations while staying within the limits of their resource availability, providing high aircraft availability, and keeping costs as low as possible. An optimization model that can create an efficient and realistic maintenance schedule is desired.

3 Literature Review

Extensive research has already been done on maintenance and aircraft maintenance planning and scheduling specifically. They comprise many different approaches aimed at improving the operations around aircraft maintenance. The problem is complex due to many factors, including resource limitations, compliance with regulations, and the need to minimize disruptions and delays while maintaining operational safety. This chapter discusses the state-of-the-art, an overview of the methodologies used in maintenance planning, which also assesses how uncertainty is incorporated in literature.

3.1. State-of-the-art

Scheduling of airline maintenance has a long history, which started in 1977 with the development of Aircraft Maintenance Operations Simulation (AMOS) [4]. Over time, aircraft maintenance scheduling has changed significantly. Deterministic methods for scheduling maintenance checks were the main focus of early research. For example, a complete model for maintenance scheduling was developed by Sriram and Haghani (2003) [16] and solved heuristically. Later researches improved this by using dynamic programming to optimise long-term check schedules for heterogeneous fleets, such as Deng, Santos and Curran (2020) [20].

Recent studies, such as Van Kessel et al. (2023) [5] and Tseremoglou et al. (2023) [12], examine disruption management and condition-based maintenance, addressing the challenges of task arrival unpredictability and resource allocation. This section goes into the developments in task scheduling, resource allocation and crew scheduling.

3.1.1. Task Scheduling

Task planning in aircraft maintenance scheduling involves organising and prioritising maintenance activities to ensure aircraft readiness while minimising operational disruptions. This process requires aligning scheduled tasks, such as routine checks, with unscheduled or non-routine tasks that arise during inspections or due to unexpected failures. Effective task planning accounts for resource constraints, such as manpower, bays, and materials, while adhering to strict regulatory requirements. Tasks must be scheduled in the correct sequences [21].

3.1.2. Resource Allocation: Material and Equipment

Robust aircraft maintenance planning also depends on efficient scheduling of materials and equipment. Spare parts can be unavailable, with high lead times leading to delays in maintenance. However, having a large inventory can be expensive and space-inefficient. Qin et al. (2020) [22] and Oenzil and Ishak (2021) [23] provide two examples of stochastic models that optimise spare parts inventory by forecasting demand using probabilistic scenarios and the reliability of components. This can minimise excess inventory costs and guarantee timely resource availability, reducing disruptions and leading to better maintenance performance. Further research in this area can focus on the extension of uncertainties or integration of other planning components to increase the quality of the model [22] or review the optimal cost analysis for parts procurement [23].

3.1.3. Crew Scheduling

With crew scheduling for aircraft maintenance, one has to make sure that sufficient staffing levels are maintained while minimising costs. De Bruecker et al. (2015) [17] used a heuristic approach to Mixed-Integer-Linear Programming (MILP) creating robust rosters, handling stochastic aircraft arrival times based on simulation results. Pereira et al. (2021) [24] proposed a support information system for the planning of aircraft maintenance teams. They allocate, using nonlinear integer programming and Monte Carlo simulation, available technicians by skills to maintenance teams under the uncertainty of workload. De Bruecker et al. (2015) [17] recommend future work on extending models to include uncertainty in workforce capacity and workload. Next to that, they suggest improving the diversification strategy, by for example using a metaheuristic mechanism.

3.2. Uncertainties and Disruptions

Hangar Aircraft Maintenance Scheduling (H-AMCS) has multiple sources of uncertainty that influence the execution and creation of a maintenance schedule. Because of that, it is essential for H-AMCS to identify these uncertainties and their effect on the maintenance planning [2].

The consideration of uncertainties in research has been increasing in the last few years. Where first mostly deterministic models were created, more and more stochasticity is taken into account. An older paper mentioning uncertainty in maintenance scheduling is Vassiliadis & Pistikopoulos (2001) [25], which describes an optimization framework for determining the best maintenance policies in continuous process operations in the presence of parametric uncertainty and assessing and measuring how uncertainty affects optimal maintenance schedules. This research has not yet been applied to aircraft maintenance.

There are many uncertainties that can occur in aircraft maintenance A list of examples is shown below. It must be noted that some uncertainties are far more significant and impactful for an airline than others, e.g., access panels have less impact than available manhours.

- Aircraft utilization
- Failure rates
- Duration of routine tasks
- Number of findings and/or non-routine tasks
- Duration of non-routine tasks or repair times
- Duration of opening and closing of access panels
- Aircraft availability or flight arrival
- Available bays
- Available manhours
- Available skills and certificates
- Availability of material: spare parts (inventory)
- Delivery time of spare parts
- Availability of equipment

The uncertainty in duration can increase the duration of maintenance checks and cause delays. These delays affect the start and due dates of subsequent maintenance checks, which leads to adjustments to the initial schedule. This effect can cause a prolongation in the long term as small deviations accumulate.

There are several researches that have looked at uncertainty and robustness in aircraft maintenance planning. These papers are listed in Table 3.1. The most common uncertainty is the duration or workload of maintenance. Of those, failure rate or failure times were the most common. Aircraft utilization, available skills, or material and equipment are however little covered.

3.2.1. Robust Planning

The robustness of a planning is usually defined as a plan that is prepared for the uncertainties of future unknown events. This is useful in dynamic and stochastic environments. The uncertainties may be from changes in the environment or from inaccurate execution of the solution itself. This can, however, be very broadly interpreted. This makes it hard to measure or define the concept of robustness [35].

Robust plans have multiple advantages [14], [2]:

- Stable resource allocation: as the planning is revised less often or with minimal changes, the resources and crew can also be planned consistently.
- Cost reduction: robust planning can minimise disruptions during operational peaks, thereby reducing

Table 3.1: Uncertainties in stochastic aircraft maintenance planning research.

Author	Uncertainty	Method or Framework
Masmoudi and Haït (2012) [26]	Task duration, procurement delays	Fuzzy GA
Samaranayake and Kiridena (2012) [15]	Unplanned maintenance activities	Unitary Structuring Technique
Rosales et al. (2014) [21]	Non-routine task variability	System Dynamics
De Bruecker et al. (2015) [17]	Flight arrivals	MILP with heuristic enhancement
Dinis et al. (2019a) [27]	Workloads	Bayesian networks with Expectation-Maximization algorithm
Dinis et al. (2019b) [28]	Unscheduled workloads	Space-time-skill coordinate system
Semaan and Yehia (2019) [29]	Task duration and probability of breakdown	Monte Carlo Simulation with cyclic operation network
Qin et al. (2020) [22]	Spare parts demand	Benders decomposition
Oenzil and Ishak (2021) [23]	Demand of spare parts	Component Reliability Analysis
Pereira et al. (2021) [24]	Workloads	Non-linear integer programming, Monte Carlo Simulation
Shahmoradi-Moghadam et al. (2021) [30]	Task duration	e-Conservative, Monte Carlo Simulation
Deng and Santos (2022) [14]	Aircraft daily utilization, Maintenance elapsed time	Approximate Dynamic Programming, Monte Carlo Simulation
Hu et al. (2022) [31]	Maintenance performance and system degradation	Markov Decision Process, Reinforcement Learning, Linear Programming
Van der Weide et al. (2022) [2]	Check duration, aircraft utilisation rates	GA, Monte Carlo Simulation
He et al. (2023) [13]	Task duration	Column generation, integer programming
Tseremoglou et al. (2023) [12]	RUL prediction	MILP, Deep Reinforcement Learning
Van Kessel et al. (2023) [5]	Stochastic task arrival, resource availability, flight arrivals	MILP
Zhang et al. (2023) [32]	Check duration, personnel transfer	Non-dominated sorting GA
Li et al. (2024) [33]	Non-routine task Workloads	Supervised learning
Tseremoglou et al. (2024) [34]	RUL prediction, task arrival	Support Vector Regression, Rolling horizon, Deep Reinforcement Learning

labour and inventory costs and increasing revenue (i.e., by not having to execute maintenance in summer periods).

- Operational efficiency: the increased reliability can reduce the buildup of maintenance delays over time and avoid 'contamination' of other delayed aircraft due to excessive rescheduling.

3.3. Methods used for scheduling

Often, research on maintenance scheduling uses methodologies to tackle specific difficulties, such as resource allocation, workforce allocation, or the arrival of ad hoc maintenance. These solution methods can generally be categorised into four main categories [3]:

- **Mathematical Programming (MP):** These methods use mathematical models to solve decision problems optimally, usually within a set of constraints [36]. Examples are Mixed-integer Linear Programming (MILP) and Dynamic Programming (DP).
- **(Meta)heuristics:** Heuristic methods are procedures that can probably discover a very good feasible solution, but not necessarily the optimal one. They do so in a computer-efficient manner, suitable for very large problems. Metaheuristics are general solution methods that provide a set of guidelines to develop heuristic optimization algorithms [37] [38]. Examples are Genetic Algorithm (GA), Simulated Annealing or Tabu Search.
- **Simulation:** Simulation-based approaches analyse and refine scheduling strategies by modelling real-world events, which frequently take into account stochastic components such as demand variations or equipment failures. A common example is the use of Monte Carlo Simulation.
- **Machine Learning (ML):** Machine learning is a subset of artificial intelligence and refers to the ability of machines to learn without explicitly being programmed. A common example is Reinforcement Learning (RL).

Van den Bergh et al. (2013) [3] has created a more elaborate overview of the state-of-art and goes into the different solution methods for both airline scheduling and maintenance planning. It can be noted that the majority of research used mathematical programming methods, such as IP or MILP. These are often used, because they can generate optimal solutions for deterministic problems and they are easy to set up, with commercial solvers, like Gurobi, widely available. Several other papers use heuristics or metaheuristics, of which GA is the most common one. Simulated annealing and tabu search are only addressed once, according to [3]. The paper that uses simulated annealing is a deterministic model, but the tabu search incorporates uncertainty in the timing of the workload.

The use of simulation is common as well, as one can generate many scenarios. It has the ability to show the different possibilities of a model or tool and can visualise the results of all simulations well. This can be especially interesting for testing robustness by simulating many different disruptions. Machine learning is currently a growing topic, with RL being applied more and more in optimisation frameworks. RL can reach very good results in a short time but is heavily dependent on the training it is subjected to. Metaheuristics, as stated by Hillier and Lieberman (2015) [37], are good at combining the heuristic procedure of discovering a nearly optimal solution while remaining sufficiently efficient in large problems with higher-level strategies. They can thereby escape local optima and perform a robust search of a feasible region.

In Table 3.1, the models or the framework applied in the research have also been indicated. It can be seen that the Genetic Algorithm, Mixed-integer Linear Programming, Monte Carlo simulations, and Reinforcement Learning are recurring multiple times. Some other methods or frameworks are only mentioned once and have thus not been explained in detail in this proposal.

Ma et al. (2022) [39] also discussed the emerging technologies that tackle uncertainties but do so in the context of aircraft maintenance routing. The advantages and disadvantages of these methods should be investigated to determine what is best for robust aircraft maintenance planning. This could also be taken from other industries that do maintenance planning, but was however not yet done in this research proposal.

3.4. Scale and Planning Horizon

The papers reviewed usually include a case study with an airline with heterogeneous fleets of around 40 to 70 aircraft. The planning horizon varies a lot. Van der Weide et al. (2022) [2] creates a long-term schedule around four years into the future, only focusing on C-checks. This is similar to Deng and Santos (2022) [14], although they focus on all letter checks with a heterogenous Airbus A320 fleet of 40-50 aircraft. Van Kessel et al. (2023) [5] has a horizon of both 10 and 120 days, but includes reactive scheduling, just like Silva et al. (2023) [40]. However, the latter also has a planning horizon of only one month. It depends on the paper whether or not the processing time of the model is mentioned. The numbers found can range from just a couple of seconds ([14]), to 15-30 minutes ([2]). De Bruecker et al. (2015) [17] limits the MILP optimisation to only one minute, but allows fifteen iterations of the model enhancement heuristic. They plan only one week into the future. The scale was unfortunately not found in this paper.

4
Research Proposal

This chapter discusses the research proposal that follows from the problem definition as presented in Chapter 2 and the literature review as presented in Chapter 3. It first discusses the research gap in Section 4.1, then the objective in Section 4.2 and lastly the research questions in Section 4.3.

4.1. Research Gap

As written in Chapter 3, there is a lot of research already done in the field of airline maintenance planning. Many different components can be investigated and optimized via multiple optimization methods.

Uncertainty in maintenance is investigated quite often, but there are only few papers available that include it in a scheduling optimisation framework and apply it to aircraft maintenance. Of those, maintenance duration is the most investigated type of uncertainty, but manhour, material, or equipment availability is less common. The use of prognostics of the RUL is growing in research but is mostly useful for predictive maintenance only, whereas this research will focus on preventive maintenance. The papers that discuss an optimisation framework only sometimes incorporate rescheduling, such as [40], [5] and [32].

The common approach to hangar maintenance planning is using the four different letter checks and allocating them to predetermined maintenance slots for certain aircraft types. Several papers schedule full checks, such as Van der Weide (2022) [2], and apply a probabilistic distribution of maintenance duration on the full check as well to include uncertainty. Witteman et al. (2021) [41] also mentioned that the task allocation problem could be expanded by considering stochasticity or exploring uncertainty related to aircraft utilization or the emergence of non-routine tasks.

From the literature, it appears that genetic algorithms are relatively often used, but the use of other metaheuristics is scarce. Besides that, although metaheuristic methods provide the opportunity to include uncertainties and find a solution that is good enough in a computing time that is small enough, they are, to the best of my knowledge, not applied to stochastic airline maintenance planning scenarios considering uncertainty in repair times/task durations [3] [38]. Hence, it is concluded that there is a gap in the literature where metaheuristics other than GA are hardly applied to stochastic airline maintenance planning scenarios. It should be investigated why this is the case. Other often used methods are mathematical programming and simulations. There is a growth in the use of machine learning, specifically of RL. The methods have different advantages; hence, their performance depends on the type of problem that needs to be optimised.

To my knowledge, there has not yet been a study that optimises a maintenance schedule incorporating uncertainty in check duration and assesses feasibility. This could be investigated further in this master's thesis. However, if combined with a large heterogeneous fleet of around 120 aircraft, this might be computationally too expensive, which should be taken into account when selecting the optimisation framework.

4.2. Research Objective

The research objective of this thesis project is defined to be:

> *To develop a stochastic optimization model for the planning of hangar maintenance, that optimizes aircraft bay assignment, interval usage and feasibility, under uncertainty of non-routine tasks.*

The research objective will focus on these three aspects (optimized bay assignmeent, interval usage, and feasibility) to create an efficient and stable schedule. The interval usage for preventive maintenance should be maximized to make as much use as possible of the three intervals DY, FH, and FC, which can save one or more maintenance checks over the lifetime of an aircraft and thus save costs. Feasibility should be high, to be able to complete maintenance schedules with as little schedule changes as possible due to disruptions. Schedule changes need to be replanned manually, and shifts workforce and resources, costing time and money. Robust planning can avoid this as it needs fewer changes in case of disruptions as explained in Section 3.2.1. Hence, feasibility is desired.

4.3. Research Questions

The research objective as presented in Section 4.2 can be transformed into the following research question:

> *How does a stochastic maintenance planning model perform taking into account uncertain non-routine tasks considering bay assignment, interval usage and feasibility?*

Uncertainty in task durations was chosen as the uncertain parameter, as it was indicated that this is the most important factor for the airline to cause delays in their maintenance schedule. As described by Rosales et al. (2014) [21], this uncertainty mainly stems from unplanned and unscheduled events, such as discrepancies, damages or something broken. This is around 40 to 60% of all maintenance activities. The discrepancies and damages are usually found during the inspection stage that need to be corrected by programming non-routine activities. Unscheduled tasks might require additional resources and activities, forcing to adjust and change an initial plan, causing delays and disruptions within a whole process, again for which a robust planning can be a solution.

Several subquestions were determined to together answer this research question and grouped in four topics. They are presented below:

1. **Uncertainty in non-routine maintenance tasks:**
   How can non-routine tasks be incorporated into the maintenance planning?
   (a) How can the arrival of non-routine tasks be defined?
   (b) How can the durations of non-routine tasks be defined?
   - An answer to this question will be formulated using the data provided by the stakeholder airline, using historical data to determine probability distribution for the arrival and duration of non-routine tasks, per aircraft type and if possible also by aircraft age and task card type.
2. **Maintenance model:**
   How can the stochastic maintenance model be built?
   (a) What optimisation method is applicable to create an optimized maintenance schedule under uncertainty?
   (b) What are the desired capabilities of the model?
   (c) What constraints should be adhered to?
   - These questions will be answered using articles that give an overview of the used methods, such as Van den Bergh et al. (2013) [3], as well as research online about capabilities of different optimisation methods. The advantages and drawbacks of the methods described in Section 3.3 should be compared and see which methods can best incorporate uncertainty while satisfying

other demands. The second and third subquestion will be answered in coordination with the stakeholder airline as well as the academics to determine capabilities both useful to academia and the stakeholder.

3. **Definitions and performance:**
   How can the performance of the model be analysed?
   (a) How can ground time, interval usage, and feasibility be defined?
   (b) How does the model's performance compare to an exact method regarding aircraft ground time, interval usage, feasibility, and required planning time?
   (c) How can simulation be applied to examine the feasibility of a planning?
   - The first subquestion can be answered by using definitions from articles such as Silva et al. (2023) [40]. The other two questions can be answered in two steps. The first step is to compare the planning of the optimization with an adapted version of the deterministic, exact method from the stakeholder airline that is created in Gurobi. Secondly, different simulation scenarios can be generated to determine how feasible the created planning will be. This can be fed back to the Genetic Algorithm to increase the confidence intervals.
4. **Validation:**
   How can the validity and practical applicability of the model be assessed?
   (a) What is the model's sensitivity to changes in the input variables and the objective function?
   (b) Does the model produce the required output for the case study?
   (c) Are the model's outcomes satisfactory when considering computation time, assumptions validity, and optimization quality?
   - This question serves as validation to see if the model responds as expected to, for example, longer maintenance durations. It can be answered by following the same steps for assessing the performance but changing the inputs in the model and looking how the created planning shifts. The last question discusses whether the created model is actually useful for its intended use.

5
Methodology

In this chapter, the research as proposed in Chapter 4 is transformed into a planning and explained. Section 5.1 discusses the planning of the thesis and Section 5.2 briefly addresses the data used during the project.

5.1. Research Planning

The overall research planning is shown in Figure 5.1 and is explained afterwards. The weeks added to the diagram correspond to the duration of the thesis, which was started on the 23rd of September. This is referred to as week 1.

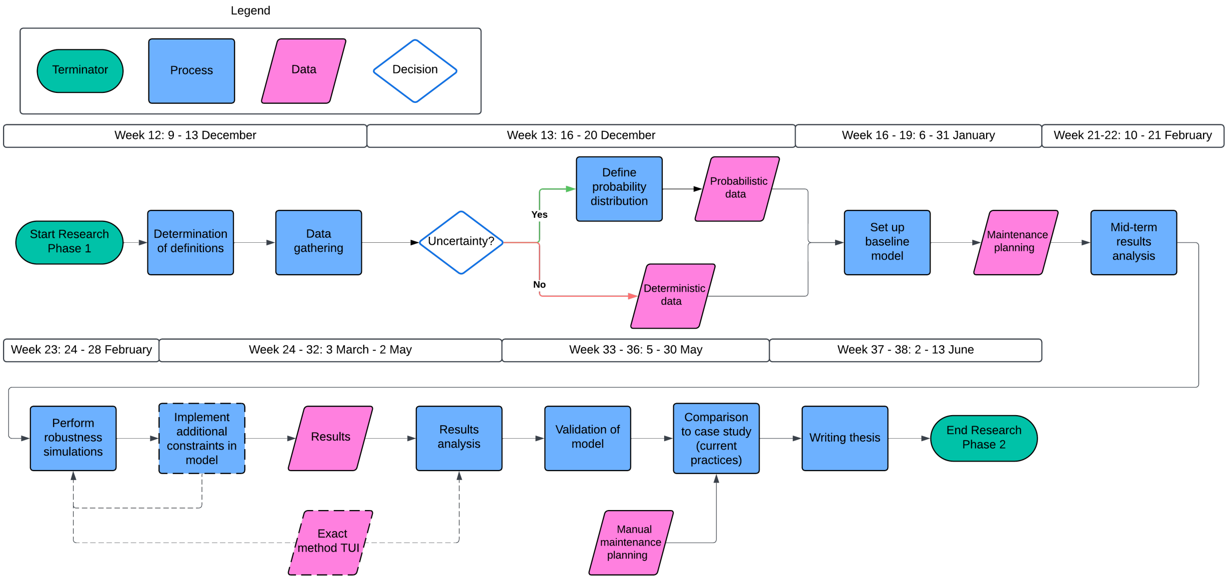


Figure 5.1: Flowchart of research planning.

Determination of Definitions

In this phase, the key definitions for the hangar maintenance planning will be defined, such as the measurement of ground time, interval usage and schedule changes as well as robustness. This gives an answer to research subquestions 1 and 2.

Data Gathering

In this phase, the necessary data, including historical maintenance records, task durations, and current maintenance schedules will be gathered.

Uncertainty and Probability Distribution Definition
This phase analyses the uncertainty in task durations and determines what probability distribution the data follows. Here, it is also determined which checks exactly will be used in the model, so only distributions need to be found for those. This gives an answer to research subquestion 3.

Set up Baseline Model
This phase sets up the baseline model. The first important step here is to determine which (metaheuristic) optimisation method will be used to answer research subquestion 4. Then, a simple first model that optimises the hangar maintenance planning will be created that adheres to the following constraints:

- Operator constraint: the operator decides on a maximum number of aircraft that can have maintenance per period of time
- Black-out window: no hangar checks are scheduled during the summer period
- Number of manhours available
- Number of bays available
- Due date of the task cards/check

A first analysis of the results will be done to get an idea of the answer to research subquestion 5. After this phase, the midterm review will take place on the **7th of March, 2025**.

Maintenance Planning Simulations
After that, the following phase performs robustness simulations to test how the baseline model handles uncertainty and in how many cases the planning needs rescheduling or can remain the same by changing input variables and the probability distribution of uncertain parameters. The robustness will be evaluated based on the number of feasible cases and the number of schedule changes necessary to return to a feasible solution. This will answer research subquestion 2 and 6.

If that works well, other constraints and implications can be added to the model to make it more realistic. These additions include at least increasing the number of bays, addressing bay limitations at the aircraft type level, and adding multiple hangars. Other constraints can be added as well if time allows.

Results Analysis
This phase will analyse the planning and simulation results to understand the impact of uncertainty on aircraft ground time, interval usage, and schedule changes. If possible and time allows, the results will be compared with an exact MILP method that has the same operating conditions and constraints. This will complete the answers to questions 5 and 6.

Model Validation and Comparison
This phase will validate the model by comparing the results to a case study of current maintenance practices. It is compared on ground time, interval usage and the time needed to create the plan. This will answer question 7.

Writing the Thesis
Lastly, the methodology, findings, simulations and analyses will be documented. The thesis is concluded with a defence mid- to end-July.

The time planning with dates is shown in Table 5.1. The midterm review is planned for the **7th of March 2025**, and the green light review is planned for the **27th of June 2025**. The gathering of data is excluded from the planning as it will be provided by the airline.

Table 5.1: Time planning of thesis.

Research Phase 1	**Start Date**	**End Date**	**Duration (working days)**
Review research proposal	27-11-24	06-12-24	6
Determination of definitions	09-12-24	13-12-24	4
Definition of probability distribution	16-12-24	20-12-24	4
Christmas holiday	*23-12-24*	*03-01-25*	-
Set up baseline model	06-01-25	31-01-25	16
Skiing holiday	*03-02-25*	*08-02-25*	-
Perform mid-term results analysis and prepare mid-term deliverable	10-02-25	21-02-25	8
Set up simulations	24-02-25	28-02-25	4
Research Phase 2	**Start Date**	**End Date**	**Duration (working days)**
Implement additional constraints	03-03-25	21-03-25	12
Perform simulations	24-03-25	04-04-25	8
Results analysis	07-04-25	02-05-25	16
Validation	05-05-25	30-05-25	16
Writing draft thesis	02-06-25	13-06-25	8
Research Dissemination	**Start Date**	**End Date**	**Duration (working days)**
Work on final thesis	16-06-25	27-06-25	8
Implement feedback green light review	02-07-25	11-07-25	6
Prepare thesis defence	14-07-25	25-07-25	8

5.2. Data Management

This section briefly touches upon the data that will be used during the thesis. Table 5.2 lists data that is required from the airline to include that in the model. Some of the entries in the table will only be used if time allows, such as the opening and closing times of access panels. It must be noted that this table is very similar to the list provided in Section 3.2, but now, deterministic values will be used for the other parameters apart from task duration.

The data will be provided anonymously by the company supervisor, i.e. the aircraft ID is unknown to the author to make publication of the master thesis possible at the end of the research. All data will not be stored locally but only on cloud-based platforms such as Snowflake or Gitlab. During the project, only the thesis student and the two supervisors will have access to the data.

Table 5.2: Useful data for research

Data
Aircraft fleet
Aircraft checks (elements) with due dates, task cards, tail registration
# of findings and non-routine tasks
Predicted duration per routine task
Actual duration of routine tasks
Actual duration of non-routine tasks
Access panels that need to be opened per task and time it takes
Available manhours

Part II

Scientific Paper

Genetic Algorithm Optimisation of Aircraft Hangar Maintenance Planning under Uncertainty

Ties Hollander

Faculty of Aerospace Engineering, Delft University of Technology

ABSTRACT

Aircraft maintenance planning plays a large role in ensuring operational efficiency and safety while minimising costs. Hangar maintenance scheduling can be trivial due to various uncertainties, such as non-routine tasks, resource availability, and unforeseen delays. Deterministic methods might struggle to account for these complexities and do not scale well with large, heterogeneous fleets, causing frequent and costly adjustments to the schedule. Previous research has focused on incorporating different uncertainties using robust scheduling methods. This research aims to develop and assess a stochastic and scalable aircraft hangar maintenance planning model that can provide insight in the robustness of the planning, next to incorporated uncertainties, to reduce the need for frequent planning revisions.

The proposed method creates a schedule using a Genetic Algorithm (GA) that minimises maintenance costs and interval losses while adhering to operational constraints. After that, a Monte Carlo simulation is applied to assess the feasibility of the schedule under randomly generated check duration scenarios. Critical checks that can cause grounding of aircraft due to exceeded due dates are modified in a feedback loop to improve the robustness of the schedule. The maintenance optimisation is tested in a case study, provided by a European airline and discusses the trade-off between maintenance cost, interval loss, run time, and feasibility in hangar maintenance planning under uncertainty.

The schedule is compared to a Mixed-Integer Linear Programming (MILP) benchmark model. Results show that the MILP outperforms the MILP in terms of cost and run time, but the GA might be useful in more complex scenarios. The simulations give insight in the robustness of the planning and show that delay propagation and grounding probabilities can be decreased by adjusting critical checks during re-optimisation. The overall grounding probability can go down by 9 to 40%, with 5 to 10% of checks fixed in time, respectively. This can lead to a more robust schedule, minimising revisions. An airline can use this framework as a decision-support tool to create variations on the planning and assess the impact of its decisions on the robustness of the schedule.

***Keywords*: Genetic Algorithm, Monte Carlo Simulation, Maintenance Scheduling, Aircraft Maintenance, Robustness**

1. Introduction

Aviation is one of the safest modes of travel in the world, due to strict regulations. One of the most important factors in aviation safety is maintenance. On average, aircraft maintenance forms around 10% of an airline's operational cost [1]. Heavy maintenance, or maintenance that is executed in the hangar, is a significant part of that. It can account for more than 70% of the maintenance costs and requires a large amount of resources [2]. In present days, airlines operate in environments with small margins and it is beneficial for them to keep the costs low by improving their maintenance scheduling and efficiency.

Initially, maintenance started as a simple and straightforward process. However, due to a more dynamic environment where both costs and the complexity of aircraft continued to increase, manual maintenance planning became more and more impracticable [3]. This resulted in airlines organising their maintenance in a more systematic way to save costs and achieve a higher efficiency, with Air Canada being one of the first in 1977 [4]. Nowadays, maintenance is usually planned by maintenance planners aided by computer tools.

This paper proposes a stochastic optimisation-simulation model, where a Genetic Algorithm (GA) is used to generate an initial maintenance schedule for aircraft maintenance. The planning is subjected to a Monte Carlo simulation to analyse feasibility across different scenarios to improve its robustness. The model is validated in a case study of a European airline with a large, heterogeneous fleet.

The framework can be used in the maintenance planner decision process to evaluate various schedules with a feasibility assessment, giving insight into the performance of the planning and possibilities for robust improvement.

This article is structured as follows. Section 2 gives more background information on aircraft maintenance planning and the problem with uncertainty. Section 3 gives an overview of related work, uncertainties in aircraft maintenance and the use of metaheuristics in maintenance research. Next, methodology with assumptions, mathematical model and constraints for the maintenance scheduling problem are introduced Section 4. The Genetic Algorithm and Monte Carlo simulation framework are also explained in this section. The airline case study and hypotheses are introduced in Section 5 and Section 6, respectively. They are followed by the results in Section 7 and validation in Section 8. The results, the validation, and the framework are discussed with recommendations for future work in Section 9. The paper is concluded in Section 10.

2. Problem Definition

This section discusses what aircraft maintenance entails and how the aircraft hangar maintenance check scheduling problem can be described. Aircraft maintenance is regulated strictly to ensure safety. It is usually executed as preventive maintenance, meaning tasks are performed on an interval-basis to prevent failure.

Despite the rise of Predictive Maintenance (PdM) in literature, preventive maintenance remains the dominant form of maintenance in aviation due to regulatory requirements [5]. This makes it still a relevant and important area of investigation. Airlines generally divide their maintenance into line maintenance and base maintenance. This research focuses on base maintenance, which is also known as hangar maintenance. Base maintenance is often divided into different letter checks: A-, B-, C- and D-checks [3]. A- and B-checks are considered light maintenance taking usually a couple of days, while C- and D-checks are heavy maintenance events, taking multiple days to several weeks to complete [2]. They require severe planning and coordination of parts, tools, bays and certified personnel [6].

The Maintenance Planning Document (MPD) defines maximum allowed intervals for scheduled tasks based on Flight Hours (FH), Flight Cycles (FC) or Calendar Days (DY). These intervals are determined by the manufacturer [7]. Maintenance must be completed before any of these parameters is exceeded. After execution, the parameters are set back to zero. If a task exceeds its deadline, an aircraft loses its airworthiness and must be grounded. This leads to revenue losses for airlines. According to Knotts (1999) [8], there are two goals in maintenance scheduling: maximising fleet availability and minimising maintenance cost. These goals translate into minimising interval loss or slack, the unused lifespan between required checks. Performing checks prematurely wastes usable interval, reduces overall utilisation and increases indirect costs [9]. Proper aircraft maintenance scheduling can avoid grounding and avoid lost interval, hence saving costs. Airlines also aim to plan maintenance outside of commercial peak periods to provide more revenue. These periods are called black-out windows. No hangar maintenance can take place as all aircraft are needed for operation or the hangar is closed. This might be during the summer holiday or during Christmas [2]. However, there are many other factors that can play a role in the planning of maintenance. Aircraft maintenance is labour- and material-intensive and is constrained by the availability of Material, Machinery, Methods and Manpower, also known as the 4Ms [10]. Planning must take into account the necessary spare parts and equipment, hangar bay availability and technicians with appropriate certifications to execute the maintenance. Dinis et al. (2019) [11] described that the maintenance workload can be divided into scheduled and unscheduled tasks. The first component is the deterministic part and the second is the stochastic part resulting from the execution of scheduled tasks and depending on probabilistic failure patterns [12]. These non-routine tasks result in a high degree of uncertainty [13]. It is unknown when something is expected to occur, or what technical skills non-routine tasks will require [11]. It can cause planned resources to be insufficient, or excessive.

The uncertainty in workload due to unscheduled maintenance tasks or activities might prolong maintenance checks, cause delay propagation and compromise the efficiency and feasibility of the initial maintenance schedule [10]. Deviations in aircraft utilisation can also impact the due dates of upcoming maintenance checks. Despite having an efficient maintenance schedule before operations have started, it is very common to have disruptions that result in frequent adjustments to the initial schedule. The rescheduling is often done manually, which can last up to a few days or a week, and impacts maintenance costs and quality of service. Robust schedules that require fewer adjustments under disruptions can reduce costs and inefficiencies in maintenance execution. [2]

It can be concluded that an optimal schedule on paper, i.e. one with minimal maintenance cost or lost interval, might not actually be optimal in practice when uncertainty comes into play. The additional costs associated with disruptions and adjustments can outweigh the initial benefits and highlight the value of incorporating uncertainty into the scheduling process. By doing so, airlines can reduce the likelihood of groundings or large modifications, resulting in plans that are not only efficient but also resilient across different scenarios.

3. Literature Review

Aircraft maintenance is a complex part of airline operations, with multiple layers of planning, such as routing, scheduling, crew rostering and resource allocation. As it takes place in a dynamic environment with strong cost pressures, extensive research has been done on aircraft maintenance. This section reviews the historical development of maintenance scheduling, optimisation objectives (cost and wasted interval), non-routine tasks, uncertainty, and approaches for robust scheduling.

While Aircraft Maintenance Routing (AMR) is a critical aspect of airline operations and closely connected to Hangar Maintenance Planning (HMP) and has been widely studied, (e.g. He et al. (2023) [14] or Es Yurek (2024) [15]), this research focuses specifically on HMP. AMR is assumed to be resolved separately and is hence not discussed here.

3.1. Historical development

Historically, aircraft maintenance scheduling has relied on manual scheduling based on the experience of maintenance planners taking weeks of planning [2]). One of the earliest attempts to improve this was Air Canada's Aircraft Maintenance Operations Simulations model (AMOS), developed in the 1970s to generate an accurate maintenance schedule in a short time (Boere (1977) [4]). Currently, several planning tools exist and are used by airlines to create a computer-aided maintenance planning, but final schedules still rely on the experience of maintenance planners [16].

Early approaches to aircraft maintenance scheduling often focused on deterministic problems, minimising costs or wasted interval. Sriram and Haghani (2003) [17] applied heuristic optimisation to minimise maintenance cost on maintenance routing problems. Later, in 2012, Jiang (2012) [18] minimised aircraft-on-ground losses with an Artificial Bee Colony algorithm. Qin et al. (2019) [19] expanded aircraft maintenance scheduling with hangar layout planning. Lastly, Deng et al. (2020) [9] used dynamic programming to optimise long-term maintenance check schedules by reducing wasted interval time.

Over time, the literature has shifted toward stochastic models, simulation, and hybrid approaches that account for real-world uncertainties as described before in Section 2.

3.2. Uncertainty in Aircraft Maintenance

Samaranayake (2006) [6] recognised that this uncertainty mainly comes from non-routine and unscheduled maintenance. It was found that approximately 50% of the maintenance workload is unplanned and identified during inspection. Dinis et al. (2019) [11] found that the unscheduled maintenance can even be up to 198% of the scheduled workload, increasing with aircraft age. This creates significant uncertainty in workload, resource needs, and duration, often leading to cascading delays and resource conflicts. Hence, recent studies have focused on predicting non-routine tasks. The authors proposed a framework to characterise maintenance work to manage uncertainty and improve capacity planning. Li et al. (2024) [20] applied supervised learning to forecast future non-routine task workloads. To handle uncertainty, recent research turned to machine learning and metaheuristics. Reinforcement Learning (RL) has been used by Andrade et al. (2021) [21] to optimise long-term maintenance scheduling and compare it to a dynamic programming (DP) based approach, minimising interval loss. RL is also applied by Lee and Mitici (2023) [22] to trigger predictive maintenance actions based on Remaining-Useful Life (RUL) distributions. Metaheuristics, such as Genetic Algorithms (GA), have been proven effective in discovering nearly optimal solutions while remaining sufficiently efficient in large problems [23]. It has often been used for scheduling problems, such as by Kleeman and Lamont (2005) [24] who solved the multi-objective scheduling problem for aircraft engine maintenance using a genetic algorithm.

3.3. Robust Scheduling

To mitigate the impact of uncertainty, recent studies have introduced robust scheduling methods, creating more stable schedules [10]. Zhang et al. (2023) [25] applied a non-sorted dominating GA to optimise personnel and equipment allocation under uncertain conditions. Deng and Santos (2022) proposed approximate dynamic programming, considering uncertainty of aircraft daily utilisation and maintenance check elapsed time, minimising wasted interval.

Monte Carlo simulation (MCS) is often used to model and evaluate stochastic maintenance plans, for example by Semaan and Yehia (2019) [26] and Pereira et al. (2021) [27]. Shahmoradi-Moghadam et al. (2021) [28] combined simulation and optimisation to generate robust solutions for a fighter aircraft fleet for different conservatism levels. Van der Weide et al. (2022) [2] integrated a GA with MCS to create robust long-term heavy check schedules, incorporating uncertainty in check duration and aircraft utilisation. Marseguerra and Zio (2000) [29] used a similar approach to optimise maintenance and repair strategies of an industrial plant. Later, Van Kessel et al. (2023) [10] addressed disruption management by developing a reactive rescheduling tool to ensure plan stability when disruptions occur. An overview of research that discusses uncertainties in a stochastic aircraft maintenance planning context is shown in Table 1. Several uncertainties are investigated using different methods or frameworks.

3.4. Research Gap

Despite growing interest in robust and stochastic maintenance planning, several research gaps remain in literature.

1. Feasibility assessment under disruption: While Van der Weide et al. (2022) [2] used MCS to assess feasible schedules and Van Kessel et al. (2023) [10] developed rescheduling tool, there is no research that evaluates the feasibility and stability of a schedule by analysing knock-on effects, delay propagation and maintenance bay conflicts.
2. Full-year planning for large, heterogeneous fleets: Most studies apply their framework to case studies of European airlines with small-to-medium fleet sizes (typically 30-60 aircraft) and focus on a single check type (e.g. C-checks). Larger, more complex planning problems for a full annual cycle are under-represented.
3. Confidence-based robustness: Few models incorporate confidence levels on the on-time execution of maintenance checks under uncertainty. Combining feasibility analysis with optimisation allows planners to quantify the robustness of schedules.

Given this context, this research addresses the following question: *How does a stochastic maintenance planning model perform taking into account uncertain task durations considering aircraft maintenance cost, interval usage and feasibility?*

To answer this question, this research develops an optimisation-simulation approach that:

- Uses a Genetic Algorithm to optimise hangar maintenance planning divided over internal and external hangars,
- Applies Monte Carlo simulation to assess the feasibility of the schedule under uncertainty,
- Evaluates knock-on conflicts and grounding and applies that to the confidence interval of planning the checks,

The approach is validated on a full-year planning of a large, heterogeneous airline fleet.

4. Methodology

In this section, the methodology is explained for the development and analysis of an aircraft hangar maintenance planning under uncertainty for a heterogeneous fleet, using optimisation and simulation models. To begin, the assumptions used in this research are listed in Section 4.1. It continues with the explanation of the mathematical formulation

Table 1. Uncertainties in stochastic aircraft maintenance planning research.

Author	Uncertainty	Method or Framework
De Bruecker et al. (2015) [30]	Flight arrivals	MILP with heuristic enhancement
Deng and Santos (2022) [31]	Aircraft daily utilisation, Maintenance elapsed time	Approximate Dynamic Programming, Monte Carlo Simulation
Dinis et al. (2019a) [11]	Unscheduled workloads	Space-time-skill coordinate system
Dinis et al. (2019b) [32]	Workloads	Bayesian networks with Expectation-Maximisation algorithm
Hu et al. (2022) [33]	Maintenance performance and system degradation	Markov Decision Process, Reinforcement Learning, Linear Programming
Van Kessel et al. (2023) [10]	Stochastic task arrival, resource availability, flight arrivals	MILP
Li et al. (2024) [20]	Non-routine task Workloads	Supervised learning
Masmoudi and Haït (2012) [34]	Task duration, procurement delays	Fuzzy GA
Mattila and Virtanen (2011) [35]	Failure rates and maintenance duration	Gamma distribution and Reinforcement Learning
Oenzil and Ishak (2021) [36]	Demand of spare parts	Component Reliability Analysis
Pereira et al. (2021) [27]	Workloads	Non-linear integer programming, Monte Carlo Simulation
Qin et al. (2020) [37]	Spare parts demand	Benders decomposition
Rosales et al. (2014) [38]	Non-routine task variability	System Dynamics
Samaranayake and Kiridena (2012) [39]	Unplanned maintenance activities	Unitary Structuring Technique
Semaan and Yehia (2019) [26]	Task duration and probability of breakdown	Monte Carlo Simulation with cyclic operation network
Shahmoradi-Moghadam et al. (2021) [28]	Task duration	e-Conservative, Monte Carlo Simulation
Sohn and Yoon (2010) [40]	Mean time between failure (MTBF) and mean time between repair (MTTR)	Random effects Weibull regression model
Tseremoglou et al. (2024) [41]	RUL prediction, task arrival	Support Vector Regression, Rolling horizon, Deep Reinforcement Learning
Tseremoglou et al. (2023) [42]	RUL prediction	MILP, Deep Reinforcement Learning
Van der Weide et al. (2022) [2]	Check duration, aircraft utilisation rates	GA, Monte Carlo Simulation
Zhang et al. (2023) [25]	Check duration, personnel transfer	Non-dominated sorting GA

of the problem in Section 4.2. After that, the working and design of the genetic algorithm is explained in Section 4.3. Next, the approach and use of the Monte Carlo simulations is discussed in Section 4.4.

4.1. Assumptions

Several assumptions were made during this research. They are written in A.1 - A.10 and are based on Van der Weide et al. (2022) [2], Deng et al. (2020) [9] and on real-life practice from the case study.

- A.1 Only the Calendar Day maintenance interval is taken into account. It is assumed that the Flight Hours (FH) and Flight Cycles (FC) will not be exceeded.
- A.2 The minimum time step in the maintenance schedule is one calendar day.
- A.3 Due dates and interval losses of check elements inside work packages are not considered. Only the due date of the entire work package will be used for the calculation of interval loss.
- A.4 Both internal and external hangars are assumed to be fully flexible and no predetermined slots for specific check types are assigned.
- A.5 The daily man-hour capacity CAP_b per bay is assumed to be constant every day, including weekends.
- A.6 The employees in the hangar are assumed to have the necessary skills to execute every task planned for the aircraft and check type.
- A.7 The location of a hangar does not influence check possibility. It is assumed that aircraft routing is flexible.
- A.8 The checks that need to be planned are assumed to not have a maximum number of lost interval days.
- A.9 It is assumed that delays or problems in one bay do not affect other bays. The bays are assumed to be independent from each other, even if they are in the same hangar.
- A.10 It is assumed that the actual maintenance duration follows a normal distribution with a standard deviation between 5% - 20% of the estimated mean.

These assumptions were made to simplify the problem and to focus on only one uncertainty. A.1 and A.4 are specific to the airline of the case study and can thus be different for other airlines.

Next to these, the research scope is limited to MPD tasks only. Ad hoc checks and airworthiness directives are not considered as part of the maintenance planning.

4.2. Mathematical Model

This subsection shows the mathematical model formulation of the Aircraft Hangar Maintenance Planning. Firstly, the sets, parameters and decision variables are introduced. Secondly, the objective function and constraints are described.

4.2.1. *Nomenclature*

In the following tables, the sets, parameters and decision variables are presented that will be used in the mathematical model.

Table 2. Sets

W	Set of work packages
B	Set of bays
T	Time planning period
T_{blackout}	Black-out time periods (subset of T)

Table 3. Parameters

CAP_b	Capacity in man-hours per day for bay b.
C_{internal}	Cost per man-hour for executing a maintenance work package in an internal hangar.
C_{external}	Cost per man-hour for executing a maintenance work package in an external hangar.
D_w	Deadline to start maintenance check for work package w.
$e_{\text{a/c type}(w),b}$	Equal to 1 if aircraft type of work package w can be handled by bay b; 0 otherwise.
$f_{w,b}$	Equal to 1 if check type of work package w can be handled by bay b; 0 otherwise.
m_w	Number of required man-hours for work package w.
W_{ground}	Weight of cost of man-hours on the objective function.
W_{interval}	Weight of anticipating a maintenance work package one day (i.e., lost from not using the full interval).

Table 4. Decision Variables

s_w	Starting time of maintenance work package w
$x_{w,b,t}$	Equal to 1 if work package w is being executed in bay b at time t; 0 otherwise.
$y_{w,b,t}$	Equal to 1 if work package w is planned in bay b, starting at time t; 0 otherwise. $y_{w,b,t}$ can only be equal to 1 for one time step.
$i_{w,b}$	Equal to 1 if work package w is allocated to an internal bay b; 0 otherwise.
$e_{w,b}$	Equal to 1 if work package w is allocated to an external bay b; 0 otherwise.

4.2.2. Objective Function (Fitness Function)

The objective function is defined as follows:

Minimise

$$\sum_{w\in W}\Bigg[\underbrace{m_w\cdot\left(C_{\text{internal}}\cdot a_{w,b_i}+C_{\text{external}}\cdot a_{w,b_e}\right)\cdot W_{\text{ground}}}_{\text{Maintenance Cost}} + \underbrace{(D_w-s_w)\cdot W_{\text{interval}}}_{\text{Loss of Interval}}\Bigg] \tag{Obj}$$

This objective function is also used as the fitness function for the Genetic Algorithm (GA). It consists of two parts, maintenance cost and lost interval. The first part looks at the man-hours of checks that are planned in an internal bay (i.e. property of the airline) or if it the check is outsourced to an external MRO. External MROs are more expensive than the internal hangars, which is reflected in the cost factors C_{internal} and C_{external}. Hence, it is desired that as many checks are planned internally to minimise this part of the objective function. The second part is the minimisation of the lost interval, which was discussed in Section 1. It is desired to plan as close as possible to the due date and minimise the lost interval, because that might save the airline an additional check over the lifespan of the aircraft, hence the operational time of the aircraft increases and the cost decrease which is desirable. This part of the objective function tries to obtain that. The weights W_{ground} and W_{interval} can be used to determine the importance of the two parts of the objective function and depends on the airline's preferences.

4.2.3. Constraints

The optimisation problem is subject to eight constraints, that are described here. First, constraint C1 ensures that every work package is planned exactly once in one bay and starts at one time step t. It is required that every work package starts before they exceed their interval, as explained before. C2 ensures that by constraining the start date to be equal or lower than the due date. As an aircraft covers a full bay, there cannot be multiple work packages and thus aircraft assigned to a certain bay at a certain time step as dictated by C3. When a maintenance check starts, all tasks must be finished. Hence, the bay should be occupied by that check for the entire duration. The duration is dependent on the capacity in the bay and the number of required man-hours for the check as also explained later in Section 4.3. C5 and C6 ensure that only aircraft types and check types are planned in bays that have the skills, space and capacity to do so, respectively. C7 ensures that nothing is planned during black-out windows. Lastly, constraint C8 ensures that the correct weighting of internal or external bays is used in the objective function.

- C1: Plan all work packages with a unique start time

$$\sum_{b\in B}\sum_{t\in T} y_{w,b,t} = 1 \qquad \forall \quad w\in W \tag{1}$$

- C2: Plan all work packages on or before their due date

$$s_w \le D_w \qquad w\in W \tag{2}$$

- C3: Only one work package per bay per moment of time

$$\sum_{w\in W} x_{w,b,t} \le 1 \qquad \forall \quad b\in B, t\in T \tag{3}$$

- C4: Duration enforcement

$$x_{w,b,t'} \ge y_{w,b,t} \qquad \forall \quad w\in W, b\in B, t\in T, \quad t'\in\left[t, t+\frac{m_w}{\text{CAP}_b}-1\right] \tag{4}$$

- C5: Aircraft type

$$x_{w,b,t} \le e_{\text{aircraft type}(w),b} \qquad \forall \quad w\in W, b\in B, t\in T \tag{5}$$

- C6: Check type

$$x_{w,b,t} \le f_{w,b} \qquad \forall \quad w\in W, b\in B, t\in T \tag{6}$$

- C7: Black-out windows

$$x_{w,b,t} = 0 \qquad \forall \quad w \in W, b \in B, t \in T_{\text{blackout}} \quad (7)$$

- C8: Mutual exclusivity of internal and external bay assignments

$$a_{w,b_i} + a_{w,b_e} = 1 \qquad \forall \quad w \in W, b \in B \quad (8)$$

$s_w = \sum_{b\in B} \sum_{t\in T} t \cdot y_{w,b,t}$
$x_{w,b,t}, y_{w,b,t} \in [0,1] \quad \forall \quad w \in W, b \in B, t \in T$ and $a_{w,b_i}, a_{w,b_e} \in [0,1] \quad \forall \quad w \in W$

4.3. Genetic Algorithm

In this section, it is explained why the GA was chosen for the optimisation, how the mathematical model is created in the Genetic Algorithm and how a Genetic Algorithm works.

4.3.1. Trade-off Different Optimisation Techniques

Genetic algorithm (GA) is an evolutionary algorithm inspired by natural selection and the reproduction of the fittest individuals [43]. It was chosen from seven techniques including Mixed-Integer Linear Programming, Reinforcement Learning, Simulated Annealing, Particle Swarm Optimisation, Ant Colony Optimisation and Tabu Search.

Compared to these methods, GA offers a promising balance of advantages. For instance, while Mixed-Integer Linear Programming provides exact solutions, it struggles with scalability [42]. Reinforcement Learning, although powerful in dynamic environments, often requires extensive training data and tuning [33]. The ease of implementation, low reliance on training data, ability to handle uncertainty, and relatively fast computing times made GA particularly well-suited to this context. Despite other methods showing strengths in specific areas, GA provided the best combination of advantages for this optimisation problem.

Albadr et al. (2020) [44] described the steps of a GA. First, an initial population is generated consisting of a certain number of candidate solutions for the aircraft maintenance planning. The fitness of every individual is calculated using the fitness function. Using a parent-selection technique, parents are chosen for mating to form a new generation. New individuals are created with the combination of parents and through crossover and mutation. The fitness of the new population is assessed and the cycle will start again until a termination criterion is satisfied, for example a certain fitness value that is reached.

The following subsections explain how the different components of GA are used in this optimisation model.

4.3.2. Chromosome Representation

The maintenance schedules have to be represented in such a way that they can be implemented into the GA. Hence, they have to be encoded into a chromosome. In this model, every work package is represented by two genes: the bay assigned and the priority ranking. Both genes are integers. An example can be: [6, 77]. The specific work package is assigned to bay 6 and has a priority of 77. All checks that are in the same bay with a priority ranking lower than 77 will be planned earlier than this check and hence will have a higher chance to be planned close to their due date. The gene space for the bay assignment is restricted by the capacity of every hangar and bay. If the work package involves a check or aircraft type that the hangar cannot support, this bay is excluded in the gene space. The priority ranking is limited only by the number of work packages requiring scheduling.

All the work packages together form a candidate solution (individual), which is a possible maintenance schedule. All candidate solutions form the population.

By trying different bay assignments for the work packages that need to be planned, the model aims to minimise the cost of maintenance by allocating as many checks as possible to internal hangars. This corresponds to the first term in the objective function (Obj). The priority ranking supports the minimisation of lost interval, which is the second part in the objective function. Additionally, it prevents conflicts in bay usage, in line with constraint 3. For each bay, the model starts with placing the black-out windows in the schedule. It then plans the work packages, starting with the highest priority checks (i.e. those with the lowest numbers) of assigned to that bay. The duration of each work package, expressed in days, is calculated by dividing the required man-hours by the number of man-hours available per day in the assigned bay, as defined in Equation 9.

$$d_w = \left\lceil \frac{m_w}{CAP_b} \right\rceil \quad (9)$$

The algorithm initially tries to schedule the work package to start on its due date. It then verifies if this placement creates a conflict with a black-out window or another work package already planned in that time frame. If a conflict is detected, the work package will be shifted one day earlier and re-evaluated. This process repeats until a valid, conflict-free time slot is found. This procedure continues sequentially for all work packages, following their priority ranking.

4.3.3. Initial Population

The initial population can have a large impact on the performance of the GA, as it determines the starting point in the solution space. Its quality depends on the size and diversity of the population. A diverse population enables exploration and can prevent premature convergence. A small population may also limit performance as it quickly loses diversity, but an overly large population can become computationally inefficient [43]. The initial population can be turned off or on. When disabled, the population is randomly generated based on the gene space of every work package. When enabled, a custom initial population can be implemented. In this case, an initial population was created by strongly favouring internal and high-capacity bays (i.e. a high number of available man-hours per day). To amplify differences, bays were weighted using a power-weighting method. Priority rankings were assigned randomly. The approach can be found in Appendix A, Algorithm 1. The preference for internal bays may be too strong, leading to too many work packages being assigned to internal bays and resulting in infeasible schedules. These infeasible schedules are heavily penalised in the fitness function with a conflict penalty, which discourages the algorithm to use these further in evolution.

4.3.4. Parent Selection

The population is assessed using the fitness function, where the fitness of every schedule is calculated using Equation Obj. Based on these fitness scores, schedules are selected as parents to produce offspring for the next generation. Parent selection plays a critical role in the convergence rate of the GA as fit parents tend to produce fitter individuals, speeding up convergence. However, it is also important to maintain diversity in the population. Without it, the GA risks premature convergence. It occurs when a single very fit solution dominates the population in an early stage, preventing the discovery of better alternatives [45].

Multiple parent selection methods were tested in this research, which were steady state, rank, random, tournament (for $n = 2, 3, 4, 5$), roulette wheel, and stochastic universal selection, to see which gave the best convergence and computational time [43]. Steady-state selection (sss) and tournament-2 led to the most optimal results. The parameter selection is further explained in Subsection 4.3.7.

4.3.5. Crossover and Mutation

Crossover is the process of recombination of the genes of two parents to create offspring. The most common methods are single-point, two-point, and uniform crossover [45]. All three were tested and the uniform crossover led to the best result, as discussed later in Subsection 4.3.7.

After generating a new offspring population, mutation is applied to maintain diversity. It can affect both bay assignments and priority rankings and its influence depends on the mutation rate. Two mutation strategies were explored in this study: inversion and a custom method. The custom mutation followed the same power-weighting logic used in the initial population (Subsection 4.3.3). It is important that the bay assignments of the mutated individual still comply with the gene space. The results of the mutation parameter selection are provided in Subsection 4.3.7.

4.3.6. Termination

The process described above is repeated until the termination criterion is met: if the fitness value remains constant for 50 generations, the optimisation terminates, assuming convergence. The chromosome representation is translated into a Gantt chart of the schedule, and evaluated. Verification checks confirmed that no schedule overlaps occurred and that bays were only assigned aircraft and check types matching their capabilities.

4.3.7. Parameter Selection

After setting up the model, the GA parameters were selected. This was done for two maintenance seasons ('24-'25 and '25-'26). The different parameters are summarised in Table 5, for parent selection, crossover, mutation and mutation percentage. Combined, 320 parameter configurations were tested, both with and without a custom initial population. Runs without the initial population consistently generated less optimal results and are therefore excluded from further discussion.

4.4. Maintenance Planning with Uncertainty

As discussed in Section 2, it is important to develop robust schedules, that can withstand different sources of uncertainty. This section introduces a framework to assess the feasibility and robustness of the schedules created by the GA. A feasible schedule is defined as one that can be executed without any changes.

These insights allow for a trade-off between having of a schedule that remains feasible across various scenarios, and keeping the optimality of the initial plan in terms of minimal maintenance costs or minimal wasted interval. Feasibility might ask for more buffers and gaps to absorb delays, whereas optimality is expected to prefer a tightly packed planning to maximise the internal hangar usage and the maintenance intervals.

Monte Carlo simulations (MCS) can give insight into this. A Monte Carlo simulation uses random sampling and statistical modelling to predict a range of possible outcomes for uncertain variables, providing probabilities for each result. It can help to predict the impact of uncertainty in scheduling problems. [46]

4.4.1. Scenario Generation

The first step in simulating the behaviour and uncertainty in the maintenance schedule is the generation of scenarios. The planning of the GA was created assuming a fixed maintenance workload. In the simulation, uncertainty must be incorporated to be able to assess the feasibility across various scenarios. This means that the total number of man-hours needed per aircraft check should fluctuate following probability distributions. Ideally, these distributions are based on historical maintenance data and expertise.

By generating a diverse and large set of scenarios based on the uncertainty in maintenance duration, the simulation provides an insight of how the schedule might perform under different conditions. This can help maintenance planners to identify potential bottlenecks or risk of infeasibility. The generation of different scenarios is an important first step for a robust assessment of the maintenance planning created.

In this research, the distributions were not based on historical maintenance data. It was assumed that the probability distributions would follow a normal distribution with the mean being the estimated number of required man-hours and the standard deviation a percentage of this value. Standard deviations from 5% to 20% were tested, with steps of 5% to be able to compare the differences and influences of the magnitude of the standard deviation on the feasibility of the maintenance planning. It is assumed that real standard deviations would lie in this range.

4.4.2. Conflicts

In the simulation, four things are interesting regarding feasibility: the number of conflicts, knock-on conflicts, the cascade length of the knock-on conflicts, and the grounding of aircraft. First, conflicts are explained.

Conflicts are defined as an overlap of one check with another check or a black-out window. This occurs if the previous check experiences a delay and the next check or black-out window is starting shortly behind the previous check. If an overlap occurs in the maintenance duration, a conflict is raised. This is the first part of the simulation. After the generation of scenarios, the new duration for each work package is calculated with the same formula as before, Equation 9. Due to the variation of the maintenance workload, it might happen that some checks take additional days,

Table 5. Genetic Algorithm parameter selection.

Parent Selection	Crossover	Mutation	Mutation Percentage
Steady-state selection (sss)	Single-point	Paired inversion	5%
Tournament ($n = 2, 3, 4$)	Two-points	Custom mutation	10%
Stochastic universal sampling (sus)	Uniform		15%
Roulette wheel selection (rws)	Scattered		20%
Rank	Paired two-points		
Random			

whereas others are finished earlier than planned.

4.4.3. *Knock-on Delays and Cascading Lengths*

Subsequently, the knock-on delays and cascading lengths play a key role in the feasibility of a schedule. These delays occur when a conflict causes one check to be postponed, forcing subsequent checks to start and finish later, even if they were not directly delayed themselves. When checks are scheduled closely together, this effect can amplify the knock-on delay and it is propagates further in the schedule, . A visual example of knock-on delays is shown in Figure 1. In this scenario, three checks (C-1, C-3, and C-5) are delayed, shown with the red elongations. The checks C-2 to C-5 suffer from the propagation of these delays. This is made visible with the orange marks. Checks C-6 and C-7, however, are unaffected due to the sufficiently large gap in the schedule. The cascading length, or knock-on length, can also be derived from this picture: the delay from check C-1 propagates through four additional checks, despite the gap between C-4 and C-5. The delay in C-3 affects only two additional checks and the delay in C-5 does not affect later checks. This makes the delay of C-1 the most disruptive. Analysing these patterns can help determine which checks are critical and should receive extra care to maintain schedule stability.

The probabilities for conflicts and knock-on delays were calculated by dividing the number of conflicts and knock-on delays per check by the number of simulations run.

4.4.4. *Grounding*

The knock-on delays might cause some work packages to start after their due date. If this is the case, an aircraft has to be grounded. Grounding of an aircraft occurs when a maintenance task is due before the task starts and must be avoided as explained before in Section 2, which is why the probability of occurrence for grounding is also investigated with this simulation.

4.4.5. *Number of Simulations*

The law of large numbers (LLN) states that the average of observations gets closer to the expected value as the number of observations or simulations increases. The average values obtained should converge to certain values when the number of simulations increases [47]. This law is considered when determining a suitable number of simulations, such that the outcomes have a converging result. For this, the average number of conflicts per scenario and the average number of knock-on conflicts were analysed. Simulations were run for maintenance season 2024-2025, with a standard deviation of 5%, with the number of simulations ranging from 100 to 10,000. Although the outcomes did not show large variations, it was concluded that the values for average number of conflicts and knock-on conflicts were converging after around 9,000 simulations. This was based on a trendline of a two-point moving average and the deviation from the total average at 9,000 simulations being less than 0.5% and 0.2% for the average number of conflicts and knock-on conflicts, respectively.

4.4.6. *Robust Optimisation*

Insights from the feasibility assessment can be used in a second optimisation of the maintenance schedule with the GA to create a more robust schedule. Critical checks, those causing knock-on conflicts and groundings, were identified and fed back into the genetic algorithm planner, with two important differences:

1. The current best solution from the previous run of the GA, is now included in the initial population of the re-optimisation,
2. The estimated number of required man-hours of the selected critical checks will be increased by one standard deviation in the re-optimisation and are fixed in time and bay allocation.

With the increase of the hours with one standard deviation, the confidence interval (the certainty that the maintenance check can be executed within the assigned time) rises from 50% to 84.1%. In the initial planning, the number of man-hours that is planned for was equal to the mean μ. Because of the assumed normal distribution, increasing to a value of $\mu + \sigma$ adds 34.1% of possibilities to fall within the confidence interval of the check. It was decided to increase by only one standard deviation length, because this has the highest impact on the confidence interval. Further increasing the interval by another standard deviation length, would increase the confidence interval with only 13.6%, which does not weigh up to the increase in maintenance time. If critical checks are still causing a lot of conflicts and grounding, this might be reconsidered.

A new solution will be generated by the Genetic Algorithm, with the updated parameters. By including the previous optimal solution in the initial population, it is expected that the optimisation will be very fast and the outcome will be close to the first schedule.

With that, the approach followed in this research comes to an end. The complete framework of the model can be revisited in Figure 2. It shows the inputs for the optimisation and simulation models, the intermediate and final outputs

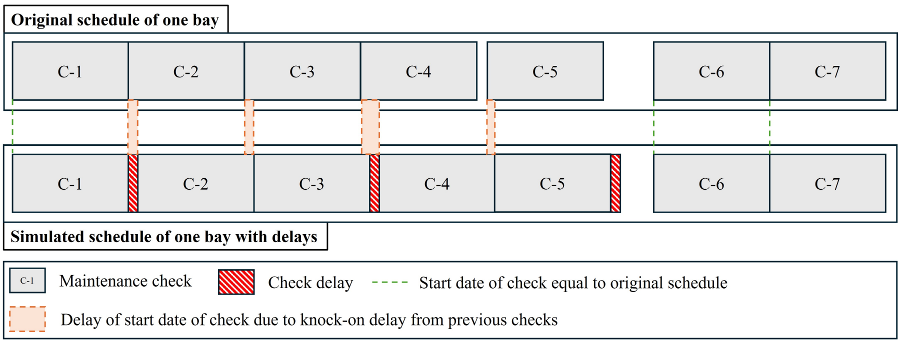


Fig. 1. Visualisation of knock-on delay.

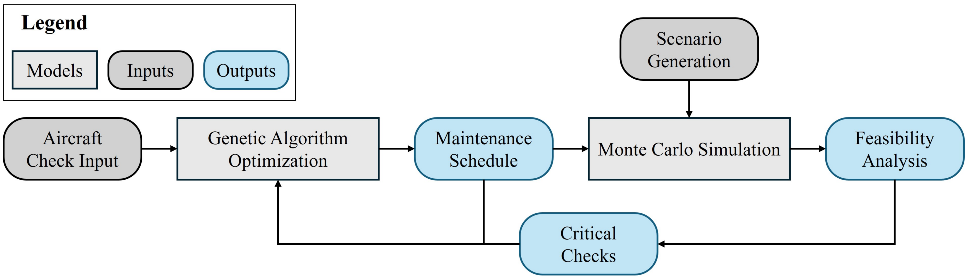


Fig. 2. Overview of the optimisation-simulation framework.

as well as the feedback loop from the Monte Carlo simulation to the genetic algorithm. In the following sections, this is applied to a case study and its results are analysed and discussed.

5. Case Study

To analyse the performance of the created model and apply the Monte Carlo simulations, a case study is performed with maintenance data from a European airline, that partially performs her own maintenance and partially outsources it to external MROs. The GA model will be compared with an exact Mixed-Integer Linear Programming model. This model was developed in-house by the airline.

This MILP model can function as a good benchmark for determining the optimality gap of the genetic algorithm. It has the same objective function and is set to adhere to the same constraints as the genetic algorithm model. It was also run on the same online platform, to allow for a fair comparison in computational time too. The KPIs and the heat maps from the simulation are compared for both models, as the exact method will be subjected to the same simulation scenarios.

5.1. Scope and Season Scenarios

The scope of the case study is to create a planning for a complete maintenance season, with heterogeneous checks and aircraft types about four to six months before the start of the season. The model was tested on maintenance seasons 2024-2025 and 2025-2026. The number of checks per dataset is presented in Table 6, together with the number of aircraft types and check types. The check types are broken down using the abbreviations as mentioned in the caption. All GA parameter combinations are tested for both seasons to see how sensitive the selection of GA parameters is to the input scenario.

Table 6. Different tested scenarios, where L=Light, H=Heavy, E=EOL and P=Paint for the check types.

Scenario	# of checks	# of A/C	A/C types	Check types
2024-2025	193	122	5	4 (L: 107, H: 73, E: 2, P: 11)
2025-2026	145	121	5	4 (L: 91, H: 45, E: 7, P: 2)

5.2. Determining Uncertainty

Throughout the research, multiple methods for determining uncertainty in maintenance duration and non-routine tasks were tried. Due to scarce data, this could not be put to practice and the scope was refined to focus on the creation of the genetic algorithm model.

Hence, it was decided to assume that the maintenance durations follow a normal distribution, with the mean being the estimated required man-hours and the standard deviation expressed as a percentage of the mean.

6. Hypotheses

Before the execution of the research, several hypotheses were set regarding the optimisation and the optimisation-simulation framework. They describe the expected comparisons between the GA and MILP planning.

6.1. Optimisation

- H_{O_1}: The Genetic Algorithm will have a lower computation time than the MILP in large scenarios. In smaller scenarios, it is expected to be similar or reversed. This is expected because MILP solvers are very fast in small scenarios but can explode for larger scenarios when it is difficult to find the optimum. The GA is not expected to explode as much, because it will create small iterations to every planning and can be stopped at multiple moments without an optimal solution.

- H_{O_2}: The Genetic Algorithm will find a less optimal solution than the MILP in large scenarios. By design, MILP solvers will return the optimal solution or be infeasible. The larger the scenario, the more likely the GA gets stuck in a local optimum and returns that as the best solution.
- H_{O_3}: The optimality gap between GA and MILP will increase when the scenario size increases. This from the previous hypothesis that GA gets stuck in a local optimum more easily in a larger scenario, hence also the optimality gap will increase.

6.2. Combined Optimisation-Simulation

- H_{OS_1}: The loss of interval and maintenance cost will increase when the confidence interval is raised for critical checks. This is expected as elongating the time for critical checks will create a tighter planning, causing some checks to be replanned to external hangars or with more interval loss.
- H_{OS_2}: Feasibility will increase when the confidence interval is raised for critical checks. This is expected as the critical checks will have more on-time execution, due to more time per check, and thus lead to less grounding events.

7. Results

In this section, the results of the research are presented. Firstly, Section 7.1 discusses the outcomes of the GA optimisation. Secondly, the results of the Monte Carlo simulations are discussed in Section 7.2. Lastly, the re-optimisation with fixed critical checks is presented in Section 7.3.

7.1. Genetic Algorithm Optimisation

This section discusses the outcomes of the genetic algorithm for both seasons. They are assessed using several KPIs, which are shown in Table 7. They are set in priority order. The maintenance cost is the largest contributor to fitness value, hence it is considered the most important.

Table 7. Key Performance Indicators (KPIs) for the Genetic Algorithm schedule.

Key Performance Indicators
Estimated maintenance cost
Run time
Total lost interval
Average lost interval
Number of internal checks
Number of external checks

Before evaluating the outcomes, the results of the parameter selection are presented in Table 8. For both maintenance seasons, two combinations were chosen: one with the highest fitness score, and another with a trade-off that provides a better balance between fitness score and computation time.

7.1.1. Comparison with MILP

The optimisation is run on an online cloud-based platform in a Kubernetes cluster, with the limits set on 7 CPU cores and a maximum RAM memory of 20 GB, for both the MILP and the GA. The optimisation resulted in a Gantt chart with the planning, a fitness evolution over the generations, and the KPIs. The fitness evolution for both GA-scenarios of season 2024-2025 is shown in Figure 3a, where it can be seen that the fitness value is rising and eventually stagnates. The outcome of the highest fitness scenario is shown in Figure 3b, which shows all the internal and external bays with planned checks as coloured boxes. The other fitness evolution plot and maintenance schedules can be found in Appendix A, in Section A.2.

The KPIs of both models were compared with each other. They are shown in Table 9 and Table 10 for seasons '24 - '25 and '25 - '26, respectively. For season '24 - '25, the genetic algorithm performs usually worse than the MILP for all KPIs. The only exception is the run time. Although the best fitness scenario of the GA has a longer run time than the MILP, the run time for the traded-off scenario is only 246.1 seconds. The GA plans more checks externally, which indicates that the MILP makes better use of the space available in the internal hangars, which reduces costs.

For season '25 - '26, the genetic algorithm does reach the same estimated maintenance cost as the MILP and a similar division of internally and externally planned checks. On the other KPIs, the GA performs worse than the MILP. Especially the computational time is a lot higher in these runs.

7.1.2. Scalability

Based on the results above, the MILP would outperform the GA. Hence, scalability is assessed. To do so, the number of man-hours available per bay was reduced using a capacity factor. The results, shown in Table 11, focus on the differences in run time and maintenance cost. Other KPIs were considered less relevant.

The MILP consistently outperforms the GA in maintenance cost, as expected from an exact optimisation approach. Regarding feasibility, the GA began returning infeasible solutions from a man-hour capacity reduction factor of 0.4, whereas the MILP only became infeasible below 0.3. While the GA remains relatively constant in computational time, the MILP run time explodes for factors 0.8 and 0.9. Interestingly, the computational time recovers at lower reduction factors, eventually outperforming the GA again from 0.7 for the highest fitness solution, and from 0.5 for the traded-off solution.

It was expected that the computation time for the MILP would continue to increase at lower reduction factors because the problem becomes more complex, but it seems that this is actually only the case for factors 0.8 and 0.9 and after that the problem becomes easier again.

7.2. Feasibility Analysis

The schedules created by the two models are subjected to Monte Carlo simulations to analyse their feasibility.

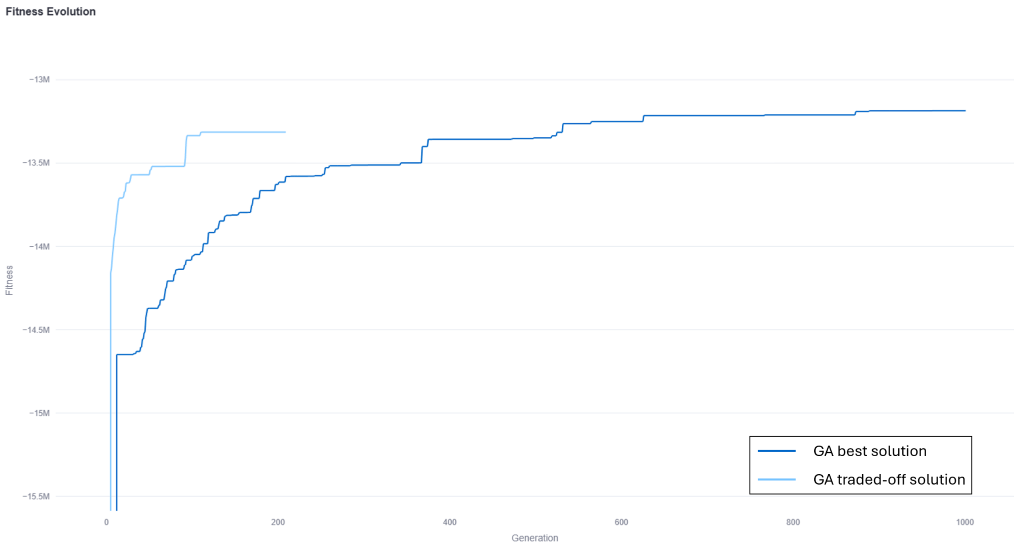
7.2.1. Average number of conflicts

Many different statistics could be defined from the feasibility analysis, but the most important ones were average number of conflicts, knock-on delays and groundings.

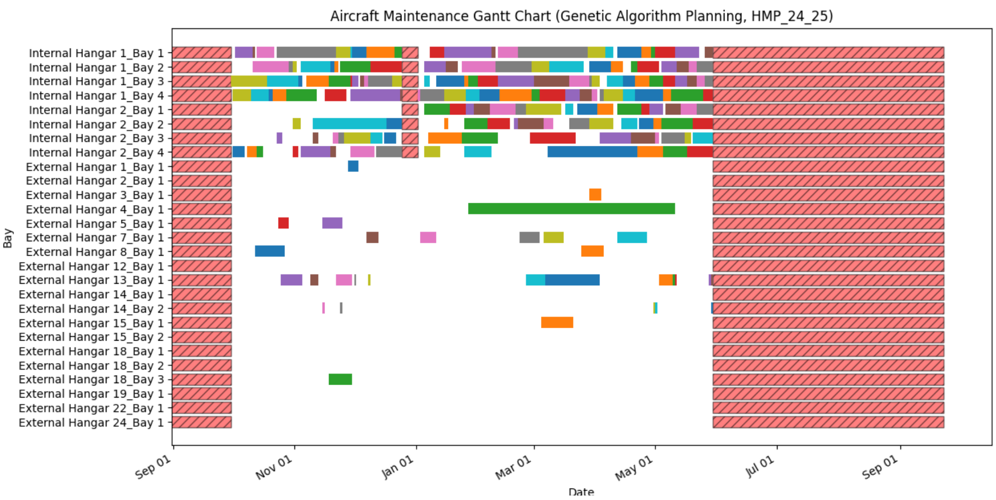
In season 2024 - 2025, the MILP tends to have more

Table 8. Parameter selection results.

Season	Variant	Parent selection	Crossover	Mutation	Mutation percentage	Run time (s)
2024-2025	Best solution	Tournament (n = 2)	Uniform	Paired inversion	20%	1791.5
	Traded-off solution	Steady-state selection	Uniform	Paired inversion	15%	246.1
2025-2026	Best solution	Steady-state selection	Paired two-points	Custom	5%	678.6
	Traded-off solution	Steady-state selection	Two-points	Custom	5%	442.9



(a) Fitness evolution plot of season 2024 - 2025.



(b) Planning with highest fitness (tournament-2, uniform crossover, paired inversion mutation, 20% mutation percentage).

Fig. 3. Genetic Algorithm fitness evolution and planning for scenario with highest fitness, season 2024 - 2025.

Table 9. KPI comparison MILP and GA optimisation, season 2024-2025.

Scenario	Est. cost optimality gap	Run time (seconds)	Run time optimality gap	Total slack (days)	Slack optimality gap	Avg slack (days)	# of internal checks	# of external checks
MILP	-	1060.4	-	2,064	-	11	175	18
GA (best fitness)	6.1%	1757.1	65.1%	2,323	12.5%	12	162	31
GA (trade-off fitness-computation time)	7.2%	246.1	-76.8%	2,955	43.2%	12	159	34

Table 10. KPI comparison MILP and GA optimisation, season 2025-2026.

Scenario	Est. cost optimality gap	Run time (seconds)	Run time optimality gap	Total slack (days)	Slack optimality gap	Avg slack (days)	# of internal checks	# of external checks
MILP	-	7.056	-	1,612	-	11	130	15
GA (best fitness)	0%	678.6	9517.3%	2,010	24.7%	14	130	15
GA (trade-off fitness-computation time)	0%	442.9	6176.9%	2,407	49.3%	17	130	15

Table 11. Man-hour capacity reduction results.

Capacity factor	GA variant	Run time GA (s)	Run time MILP (s)	Run time difference (s)	Run time gap	Maintenance cost difference	Maintenance cost optimality gap
1	Traded-off solution	425.6	2662.7	2237.1	84.0%	-891243	-7.2%
	Best solution	2982.2		-319.5	-12.0%	-763420	-6.1%
0.9	Traded-off solution	580.6	29673.4	29092.8	98.0%	-1165617	-9.2%
	Best solution	3451.1		26222.3	88.4%	-1108701	-8.7%
0.8	Traded-off solution	228	36008.2	35780.2	99.4%	-1456998	-11.0%
	Best solution	2826.3		33181.9	92.2%	-797840	-6.0%
0.7	Traded-off solution	517.2	1048.1	530.9	50.7%	-828176	-6.0%
	Best solution	3511.4		-2463.3	-235.0%	-575250	-4.1%
0.6	Traded-off solution	1033.3	1864.2	830.9	44.6%	-460492	-3.2%
	Best solution	3849.2		-1985.0	-106.5%	-405361	-2.8%
0.5	Traded-off solution	633.1	344.3	-288.8	-83.9%	-412322	-2.7%
	Best solution	2323.1		-1978.8	-574.7%	-341381	-2.3%
0.4	Traded-off solution	940.8	72.5	-868.3	-1197.7%	-10.00×10^9	Infeasible GA
	Best solution	3815.1		-3742.6	-5162.2%	-3.000×10^9	Infeasible GA
0.3	Traded-off solution	1442.3	561.1	-881.2	-157.0%	-40.00×10^9	Infeasible GA
	Best solution	6075.8		-5514.7	-982.8%	-9.000×10^9	Infeasible GA

conflicts than the GA plannings, which can be seen in Figure 4a. The box plots show the three planning variants on the x-axis and the occurring number of conflicts per scenario on the y-axis. The best fitness GA and MILP schedules are also visualised by the coloured heat maps in Figure 5. It can be observed that the GA schedule is less tightly packed than the MILP, with more checks being allocated to external hangars, which could be a logical cause of the difference in conflicts.

For season 2025-2026, the outcomes were almost similar, as shown in Figure 4b. This was also checked for other standard deviation percentages, where similar results were observed. More information on the influence of standard deviation percentages can be found in Appendix A.

7.2.2. *Knock-on Conflicts and Groundings*

Next, the knock-on conflicts were simulated. On average, for season '24-'25, 21% of the checks cause a knock-on conflict in the best-fitness GA, versus 44% in the MILP with $\sigma = 5\%$. The average length of the knock-on delay in days is also higher for the MILP. The difference in the probability of occurrence of knock-on conflicts is well illustrated by Figure 6. The MILP creates a cascading effect due to the back-to-back scheduling which increases the probability of knock-on conflicts. The GA is influenced less by that cascading effect, because of more gaps. This also has an effect on the grounding, illustrated in the heat maps in Figure 7, where the checks causing groundings of aircraft are coloured based on their probabilities. Here, it can be seen that there are more checks in the MILP that have a high probability to cause grounding, visible from the red blocks in the Gantt chart. The performance of the models is further discussed in Section 9. All three average parameters tend to increase when the standard deviation percentage increases. Interestingly however, the probability of a knock-on conflict becomes more spread across all checks, instead of a few checks that severely suffer from knock-on conflicts. A probable cause for this is that a higher standard deviation also allows for earlier completion, mitigating severe knock-on conflicts that propagate far in the schedule. This is not usual in practice hence it is addressed in Section 9.

For season 2025-2026, the results are very similar between the three schedules. This happens because there is no optimality gap in maintenance cost in this scenario as discussed in Section 7.1.

Based on the above results, it can be concluded that the Genetic Algorithm causes less grounding in season 2024-2025 than the MILP, because of fewer and smaller conflicts and knock-on conflicts, but differences are small in season 2025-2026. This result is further discussed in Section 9.

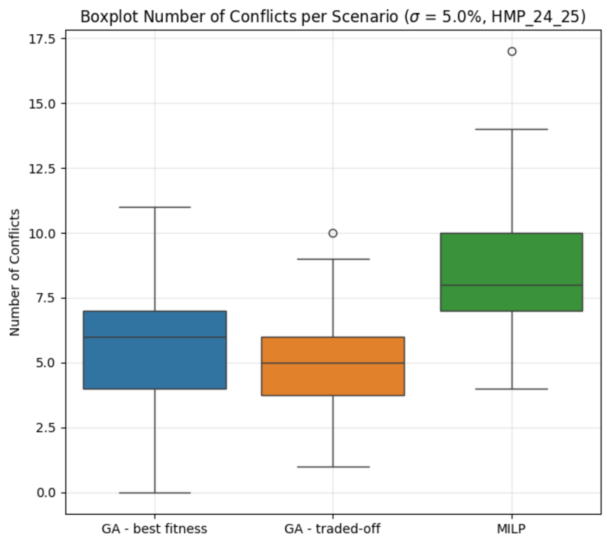
7.3. Feedback Monte Carlo simulations to Genetic Algorithm

To create a feedback loop to the Genetic Algorithm, the checks that endanger the feasibility of the schedule must be identified. These critical checks were selected based on the probability of grounding, since this was considered to be the most costly impact for an airline.

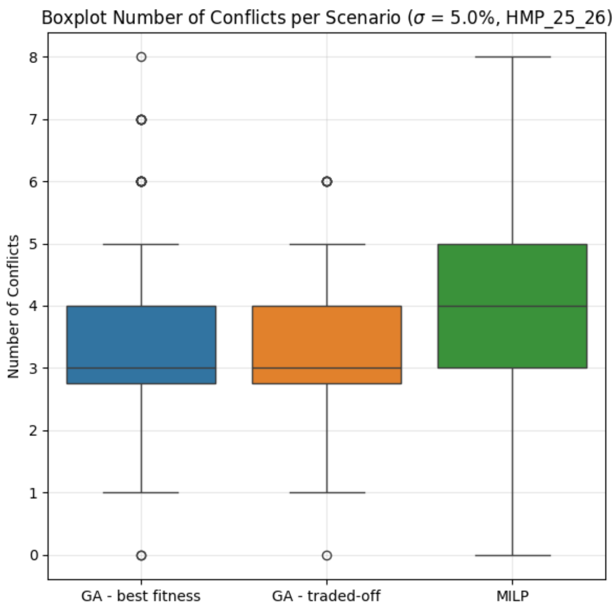
It was decided to test 5% and 10% of the total number of work packages being marked as critical, based on their grounding probability. This was deemed a relatively small number for which the confidence intervals need to be raised, hence not stepping too far from the Monte Carlo simulation on the Genetic Algorithm. Hence, for season 2024-2025, this resulted in 10 and 19 checks, respectively. The traded-off solution of the GA was used in this case. The reason for that being that the optimisation of the best fitness variant was not saturated but ran until the cut-off number of generations. During re-optimisation, this scenario would just continue its initial optimisation, which makes it unsuitable for testing the feedback framework and re-optimisation. The results of four different configurations is shown in Table 12. It can be observed that for the first three configurations, the run time is short and that the results do not change that much from the original optimisation. When we change the standard deviation to a higher percentage, and thus the allocation of extra time as well, the run time goes up. Note that the termination criterion is still set at 50 generations, which is a large share of the total number of generations run, while the improvement from the original is small.

Now, these configurations must run through the exact same simulation scenarios to analyse if the grounding probabilities have decreased and by how much. If so, the schedule is more robust, being able to remain feasible in more scenarios. The results for the first three configurations are shown in Table 13 and compared with the scenario where no checks are fixed. For every configuration, the sum is taken for the knock-on and grounding probabilities and the number of work packages that cause knock-on conflicts or groundings are counted. The fraction is determined by dividing the value by the total number of checks in the season. It can be observed that the overall probability of grounding goes down with around 9 to 40% when more checks are fixed and receive additional time, but the number of knock-on conflicts goes up in the scenario with 5% fixed checks, and for 5% fixed checks with 10% extra time with around 3-4%. This could be due to the schedule becoming tighter when the checks are elongated and remain in the same bay. This is not posing a problem for the configuration where 10% of checks are fixed, as the knock-on probability is lowered by 7.3%. When the fixed checks are inspected individually, it can be observed that the probability of grounding other checks goes down in all cases. There is an exception if a fixed checks is planned right before a black-out window. The check cannot propagate a knock-on conflict and cannot ground another check because of this, but it can be grounded itself. This probability also goes down when checks receive additional time. These results indicate that the number of fixed checks has a higher influence on the knock-on and grounding probabilities than the amount of time added, as the first and third configuration obtained very similar values. The configuration with different standard deviation is omitted here as it must be compared to a different base scenario.

The adapted planning can be compared with the initial planning as the new schedule might come with changes. As it is desirable to keep these as changes as small as possibles, it is interesting to quantify them using a change score (CS) or a similarity index (SI), where $SI = 100\% - CS$. The Change Score represents the proportion of altered checks



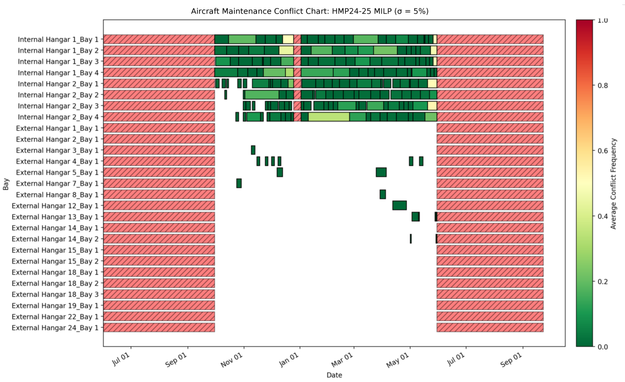
(a) Season '24-'25, $\sigma = 5\%$.



(b) Season '25-'26, $\sigma = 5\%$.

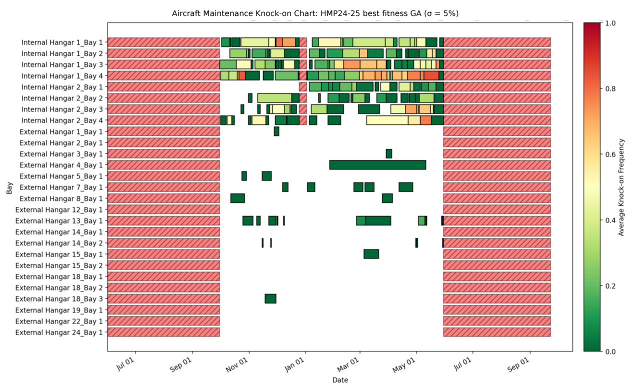
Fig. 4. Box plots of number of conflicts.

(a) GA best fitness planning.

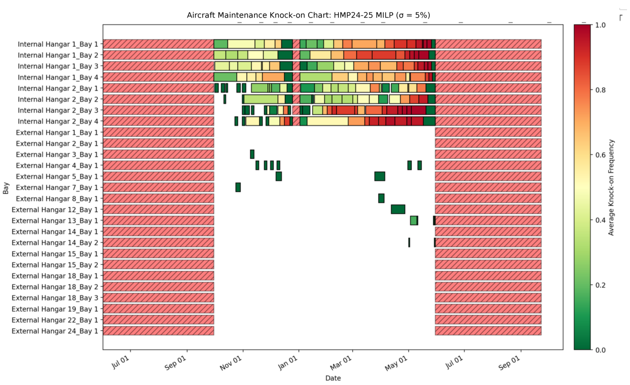


(b) MILP planning.

Fig. 5. Gantt chart heat maps of the probability of conflicts with $\sigma = 5\%$, season 2024-2025.

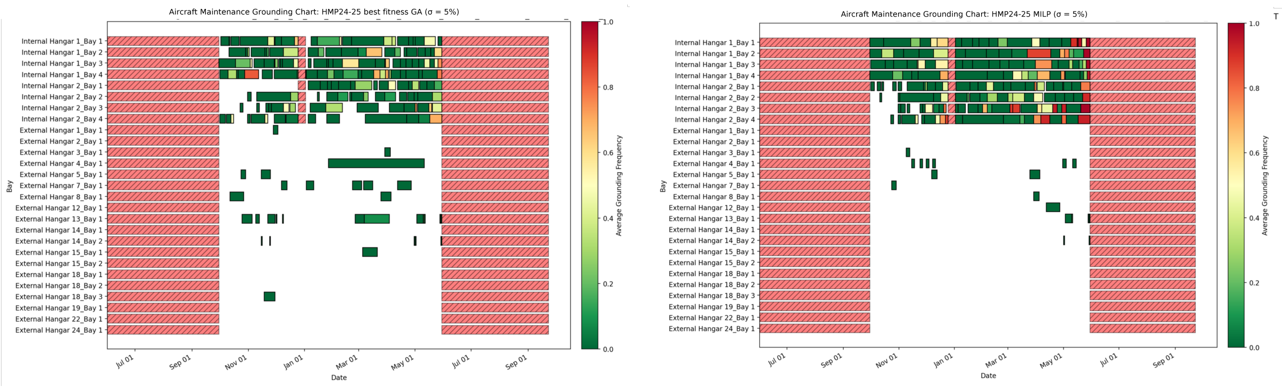


(a) GA best fitness planning.



(b) MILP planning.

Fig. 6. Gantt chart heat maps of the probability of knock-on conflicts with $\sigma = 5\%$, season 2024-2025.



(a) GA best fitness planning.

(b) MILP planning.

Fig. 7. Gantt chart heat maps of the probability of groundings with σ = 5%, season 2024-2025.

Table 12. Results re-optimisation with different configurations.

Season	Variant	Standard deviation percentage	# of fixed checks	Extra allocated time	Percentual maintenance cost change from original	# of generations	Run time (seconds)	Lost interval	Percentual lost interval change from original
HMP24-25	Traded-off	5%	10 (5%)	5%	-0.00%	119	59.2	2910	-1.52%
HMP24-25	Traded-off	5%	19 (10%)	5%	+0.48%	99	59.6	3261	+10.36%
HMP24-25	Traded-off	5%	10 (5%)	10%	+0.00%	119	59.7	2957	+0.07%
HMP24-25	Traded-off	20%	10 (5%)	20%	+1.18%	384	273.0	3110	+5.25%

Table 13. Changes in knock-on and grounding probabilities after re-optimisation of GA traded-off (HMP24-25, σ = 5%)

Configuration	Parameter	Value	Fraction of all checks	% difference
0% fixed checks	Sum knock-on probabilities	44.83	0.23	-
	Count checks causing knock-on	117	0.61	-
	Sum grounding probabilities	21.39	0.11	-
	Count checks causing grounding	54	0.28	-
5% fixed checks	Sum knock-on probabilities	46.66	0.24	4.08%
	Count checks causing knock-on	112	0.58	-4.27%
	Sum grounding probabilities	19.36	0.10	-9.49%
	Count checks causing grounding	51	0.26	-5.56%
10% fixed checks	Sum knock-on probabilities	41.54	0.22	-7.34%
	Count checks causing knock-on	112	0.58	-4.27%
	Sum grounding probabilities	12.93	0.07	-39.55%
	Count checks causing grounding	43	0.22	-20.37%
5% fixed checks 10% extra time	Sum knock-on probabilities	46.45	0.24	3.47%
	Count checks causing knock-on	111	0.58	-5.36%
	Sum grounding probabilities	19.03	0.10	-12.19%
	Count checks causing grounding	51	0.26	-5.88%

out of the total checks in the schedule. This metric can quantify different types of change:

- Percentage of checks with changed man-hours (input)
- Percentage of checks with a changed duration
- Percentage of checks with a changed bay assignment
- Percentage of checks with a changed location type (internal or external hangar)
- Percentage of checks with a changed start or end date
- Percentage of checks with changed slack

Note that the percentages indicate the number of checks that remain the same, not, for example, the percentage change of the total duration or slack. The percentages can be combined to a composite similarity score. Table 14 shows the similarity index for three configurations compared to the original planning. The allocating 10% extra time to checks, instead of 5%, gives fewer changes than fixing more checks. Again, the configuration with changed standard deviation is omitted.

8. Validation

In this section, the approach of validating the genetic algorithm model is explained. The goal is to assess whether the hangar maintenance optimisation produces credible outcomes. Firstly, a sensitivity analysis of the parameter selection is discussed in Section 8.1. Secondly, a sensitivity analysis of the fitness function is discussed in Section 8.2.

8.1. Sensitivity Analysis - Parameter Selection

The GA-parameter selection sensitivity is analysed using a correlation matrix, to see which parameters have the largest influence on the outcomes of the algorithm. The matrices are shown in Figure 8 and Figure 9. It can be seen that the parent selection type is in both seasons the most influential on the fitness score. This happens especially because four selection types, namely rank, rws, sus and random, performed very badly in the optimisation. In the ten best parameter combinations of both seasons, tournament and sss occurred equally often. Of the other parameters, the mutation type was the most influential but results were mixed. In season '24-'25, the inversion mutation was more frequent and in season '25-'26, the custom mutation was dominant in the top 10. The results of the crossover type and mutation percentage were very mixed, with every possibility occurring frequently. This is reflected in the very low correlation scores of the crossover type.

It can be concluded that the parameter configuration strongly influences the result of the GA and is not always the same, based on the two tested seasons. However, tournament and sss usually give the most promising results.

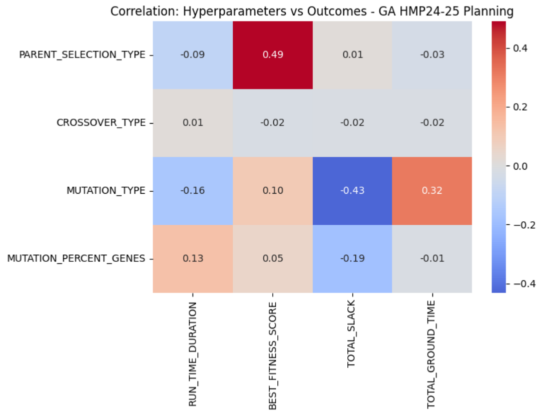


Fig. 8. Correlation matrix for season 2024-2025.

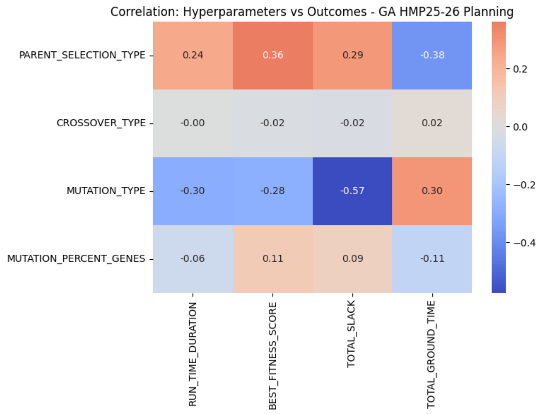


Fig. 9. Correlation matrix for season 2025-2026.

8.2. Sensitivity Analysis - Fitness Function

For the fitness function, several tests were executed to assess how it would influence the final solution. Firstly, the cost parameters for internal and external hangars were changed and made equal to each other. This resulted in a scattered planning with even use of internal and external hangars, as expected. With the cost parameters, an airline can influence how much they want to favour their internal hangars over external ones. Secondly, the weight of the lost interval component of the fitness function was increased. If increased with a factor of 1,000, the model sacrifices internal bay allocation for interval usage. In season '24 - '25, this change resulted in a contribution growth of wasted interval to the fitness function from 0.2% to 8.2%. It led to a maintenance cost increase of 7.2% and a decrease in wasted interval of 57.0%. For season '25 - '26, the contribution grew from 0.2% to 11.1%. The maintenance cost increased with 4.6% and the wasted interval decreased by 26.4%%.

Next to these two sensitivity analyses, the MILP model functions as a benchmark for the GA. The 0% optimality gap in season 2025-2026 for maintenance cost shows that the MILP and GA get to similar results. Smaller subsets of the scenarios were tested too. With half the checks in the data set but equal hangar availability, the GA will reach a completely optimal solution with multiple parameter combinations. Its run time decreases to less than two minutes.

The genetic algorithm can produce credible and near-

Table 14. Similarity index of adapted schedules, as compared to initial schedule of traded-off GA (season '24 - '25).

Similarity Index	5% fixed checks	10% fixed checks	10% extra time
# of checks with same man-hours	94.82%	90.16%	94.82%
# of checks with same duration	94.82%	93.26%	94.30%
# of checks with same start and end date	80.83%	67.36%	79.79%
# of checks with same bay	97.41%	97.93%	97.41%
# of checks with same location type	100.00%	98.45%	100.00%
# of checks with same slack	82.90%	72.54%	81.87%
Composite similarity	91.19%	85.91%	90.67%

optimal solutions, especially when parameters are well-tuned. The analyses demonstrate that the genetic algorithm's performance is sensitive to parameter configuration and the design of the fitness function, but the results show strong similarity with the MILP model. This makes the algorithm a suitable method for the optimisation of hangar maintenance planning.

9. Discussion

This section discusses the three components of this research: the optimisation of the maintenance planning using a Genetic Algorithm in Section 9.1, the feasibility assessment through simulation in Section 9.2, and the incorporation fo robustness improvements via the feedback framework in Section 9.3. If applicable, the parts include a validation of the relevant hypotheses.

9.1. Optimisation Performance

The comparison of the optimisation performance between the GA and MILP model showed multiple insights. While GAs are generally considered to have fast run times and better scalability in complex problems, these benefits were not consistently observed. In three of the four scenarios tested, the MILP outperformed the GA in both speed and solution quality. An important contributor to the longer run time of the GA is its termination criterion, which requires a large number of additional generations to confirm convergence. Lowering this criterion can reduce run time but increases the risk of premature convergence and suboptimal results. Airlines may need to balance this depending on their preferences.

The sensitivity of the GA to parameter tuning is also challenging, as the configurations varied between the two maintenance seasons, indicating that a universally optimal parameter combination may not exist. Looking at optimality, the GA showed a clear gap compared to the MILP. At this stage, the MILP delivers more optimal results with shorter runtimes in most tested scenarios. However, the GA may become more relevant in larger or more constrained problem settings where MILP models struggle to scale. This could be the case, for instance, with very large fleets, expanded planning horizons, or non-linear constraints. This was partly observed in the man-hour capacity reduction test, where GA outperformed the MILP for the scenarios where the complexity was increased, but the potential has not yet been conclusively demonstrated.

9.1.1. *Hypothesis Validation - Optimisation*

The results confirm hypothesis H_{O_2}. The GA achieved stable and decent results, but not optimal in larger scenarios. H_{O_1} cannot be confirmed as the MILP was faster in most cases. Although it is backed up by little diversity in the data sets, it could be that H_{O_3} can be confirmed too. The optimality gap is larger for the more complex season '24-'25, containing 193 work packages, than season '25-'26 containing 145 work packages. Statistical validation is needed to confirm this hypothesis.

9.2. Feasibility Assessment via Simulation

The feasibility of the created schedules was evaluated using Monte Carlo simulations to assess the plans performed under different scenarios. The results showed that schedules, which are more tightly packed, experienced more knock-on conflicts and grounded checks. This was the case for the exact planning in season 2024-2025. A more tightly packed schedule, with more checks in the internal hangars and closer to their due dates, leaves less room for potential disruptions. This aligns with the conclusion of Clarke (1998) [48], who noted the increasing severity of disruptions due to airline scheduling having increasingly less slack nowadays.

While the GA might appear to perform better in the simulation, this is not what the GA was designed for. Its initial shortcomings; reduced optimality and leaving more gaps in the schedule, actually prove beneficial by providing buffers. However, this robustness is not a feature of the GA, as it results from poor initial planning rather than a designated strategy. This is indeed observed in season 2025-2026, where the GA has a slightly higher probability for groundings than the MILP. An approach that might work as well in that case is systematically planning checks at least two days before its due date or black-out window to avoid groundings. In terms of robustness, it is unlikely that the current GA outperforms the MILP but we can investigate how we can accurately determine the checks that could benefit from additional time and for which checks it would be less useful.

9.3. Feedback Framework for Robustness Improvement

The final part of this study focuses on how adjustments could be fed back into the optimisation. This was done by fixing critical checks in time and adjusting man-hour capacity before re-running the GA. The potential benefit of the

GA lies in its ability to improve the initial planning in the optimiser. By reintroducing the previous schedule as part of the new initial population, we might decrease the run time, improve robustness, and keep the updated planning close to the original.

This feedback framework showed promising results for the different configurations proposed in Section 7.3. The new planning that was created had little changes, improved robustness and a short run time, thanks to the initial population approach. This framework might be useful as a quick decision-support tool for maintenance planners. Alternative planning options can be explored with parts of the planning fixed in time and with extra man-hours assigned to investigate the influence on robustness. For practical application, however, the feasibility assessment should be done quickly. The Monte Carlo simulations are currently very time-consuming.

Based on the findings of this study, the MILP model remains the preferred method for generating efficient and consistent maintenance plans, especially in cases with moderate complexity. The GA can however provide additional value when it that might outperform the MILP in complex scenarios and its feedback framework can be helpful as a quick decision support tool for a maintenance planner to assess and improve the robustness of maintenance schedules with alternative plannings while staying close to the original schedule.

9.3.1. Hypothesis Validation

Hypothesis H_{OS_1} is rejected. In the re-optimisation, one variant achieved a lower maintenance cost than the original optimisation. This can happen if the original optimisation did not find the optimal solution and the re-optimisation can find a solution that is cheaper. The lost interval also went down for this configuration. For the other configurations, the hypothesis could be confirmed, but it will not always be true. Hypothesis H_{OS_2} can be supported, since the grounding probabilities decreased in all re-optimisations. A statistical analysis should show whether these results are significant.

9.4. Limitations

There are several limitations to the current algorithm, that could be improved in future research.

- Maintenance is assumed to occur every day, including weekends, and every day has the same man-hour capacity. This may not be true in reality.
- No maximum lost interval was enforced. This allowed some checks to be planned with extremely large slack, up to 180 days, including annual checks. Such early scheduling can cause unnecessary extra checks.
- Maintenance durations were modelled with a normal distribution, while a right-skewed distribution might better reflect real-world variability. This tends to only extend tasks, while now work package can also be finished early.
- The combined framework and feasibility assessment were only evaluated in two maintenance seasons. To generalise the conclusions, broader application is needed.

9.5. Recommendations for Future Work

Further research should try to address the limitations discussed above and improve applicability on real-world environments.

- **Include uncertainty from historical data** to replace the assumed probability distributions for maintenance task durations with statistical distributions based on historical maintenance data. This would improve the practicality and accuracy of the Monte Carlo simulations.
- **Develop a real-time rescheduling tool** to create a dynamic rescheduling framework that can adjust the maintenance plan throughout the maintenance season in response to disruptions.
- **Explore combinations of multiple objectives** to look at different trade-offs, such as the benefit of shorter aircraft ground time versus cheaper internal maintenance.

10. Conclusion

This study investigated the development and analysis of a stochastic maintenance planning model, taking into account uncertainty of non-routine tasks. In effective maintenance planning, it is important to address potential disruptions to avoid additional costs and rescheduling. To do so, a stochastic optimisation framework was developed. A genetic algorithm was created to optimise he maintenance planning, and its outcomes were compared to an deterministic MILP model. A feedback framework was introduced to implement observations from a feasibility analysis through Monte Carlo simulations back into the optimisation with adjusted inputs. It was expected that the combination of these methods would improve the robustness of maintenance schedules, reducing the need for schedule changes due to disruptions.

Although the GA demonstrated potential in optimisation, with small optimality gaps for maintenance costs, the MILP model generally provides better and faster solutions. However, the GA could be useful in more complex scenarios. Next to that, the feedback mechanism showed that the robustness can be improved with limited run time and changes to the original schedule. This allows for a decision-support tool for maintenance planners to explore alternative schedules.

Overall, the framework gives insight into feasibility of schedules and provides an opportunity to improve the robustness of an aircraft hangar maintenance planning subject to uncertainties.

References

[1] G. Cros. *Airline Maintenance Cost Executive Commentary*. Tech. rep. IATA, 2022.

[2] T. van der Weide, Q.C. Deng, and B.F. Santos. "Robust long-term aircraft heavy maintenance check scheduling optimization under uncertainty". In: *Computers & Operations Research* 141 (May 2022). ISSN: 0305-0548. DOI: 10.1016/j.cor.2021.105667.

[3] J. van den Bergh et al. “Aircraft maintenance operations: state of the art”. In: (2013). URL: https://kuleuven.limo.libis.be/discovery/fulldisplay/lirias1677212/32KUL_KUL:Lirias (visited on 11/18/2024).
[4] N.J. Boere. “Air Canada Saves with Aircraft Maintenance Scheduling”. In: *Interfaces* 7.3 (May 1977), pp. 1–13. ISSN: 0092-2102. DOI: 10.1287/inte.7.3.1. URL: https://pubsonline-informs-org.tudelft.idm.oclc.org/doi/10.1287/inte.7.3.1 (visited on 11/05/2024).
[5] J. Lee. “Predictive Aircraft Maintenance: Integrating Remaining-Useful-Life Prognostics into Maintenance Optimization”. en. PhD thesis. Delft University of Technology, 2022. DOI: 10.4233/UUID:B219E3E7-285B-44CD-A0FF-47F03FB9F3AC. URL: http://resolver.tudelft.nl/uuid:b219e3e7-285b-44cd-a0ff-47f03fb9f3ac (visited on 07/06/2025).
[6] P. Samaranayake. “Current practices and problem areas in aircraft maintenance planning and scheduling – interfaced/integrated system perspective”. en. In: 2006.
[7] S.P. Ackert. “Basics of aircraft maintenance programs for financiers”. In: *Evaluation \& insights of commercial aircraft maintenance programs* 10 (2010).
[8] R.M.H. Knotts. “Civil aircraft maintenance and support Fault diagnosis from a business perspective”. In: *Journal of Quality in Maintenance Engineering* 5.4 (Jan. 1999). Publisher: MCB UP Ltd, pp. 335–348. ISSN: 1355-2511. DOI: 10.1108/13552519910298091. URL: https://doi.org/10.1108/13552519910298091 (visited on 07/07/2025).
[9] Q. Deng, B.F. Santos, and R. Curran. “A practical dynamic programming based methodology for aircraft maintenance check scheduling optimization”. In: *European Journal of Operational Research* 281.2 (Mar. 2020), pp. 256–273. ISSN: 0377-2217. DOI: 10.1016/j.ejor.2019.08.025. URL: https://www.sciencedirect.com/science/article/pii/S0377221719306782 (visited on 11/20/2024).
[10] P.J. van Kessel, F.C. Freeman, and B.F. Santos. “Airline maintenance task rescheduling in a disruptive environment”. In: *European Journal of Operational Research* 308.2 (July 2023), pp. 605–621. ISSN: 0377-2217. DOI: 10.1016/j.ejor.2022.11.017.
[11] D. Dinis, A. Barbosa-Póvoa, and A.P. Teixeira. “A supporting framework for maintenance capacity planning and scheduling: Development and application in the aircraft MRO industry”. In: *International Journal of Production Economics* 218 (Dec. 2019), pp. 1–15. ISSN: 0925-5273. DOI: 10.1016/j.ijpe.2019.04.029.
[12] H.K. Al-Fares and S.O. Duffuaa. “Maintenance Forecasting and Capacity Planning”. en. In: *Handbook of Maintenance Management and Engineering*. Ed. by M. Ben-Daya et al. London: Springer, 2009, pp. 157–190. ISBN: 978-1-84882-472-0. DOI: 10.1007/978-1-84882-472-0_8. (Visited on 07/07/2025).
[13] A.C. Márquez. “Models to deal with maintenance capacity planning.” In: *The Maintenance Management Framework: Models and Methods for Complex Systems Maintenance*. Springer, 2007, pp. 157–184.
[14] Y. He et al. “Maximizing robustness of aircraft routing with heterogeneous maintenance tasks”. In: *Transportation Research Part E: Logistics and Transportation Review* 177 (Sept. 2023), p. 103237. ISSN: 1366-5545. DOI: 10.1016/j.tre.2023.103237. URL: https://www.sciencedirect.com/science/article/pii/S1366554523002259.
[15] E. Es Yurek. “Combinatorial Benders decomposition for the operational aircraft maintenance routing problem”. In: *Computers & Operations Research* 164 (Apr. 2024), p. 106545. ISSN: 0305-0548. DOI: 10.1016/j.cor.2024.106545. URL: https://www.sciencedirect.com/science/article/pii/S0305054824000170 (visited on 11/05/2024).
[16] Q. Deng, B.F. Santos, and W.J.C. Verhagen. “A novel decision support system for optimizing aircraft maintenance check schedule and task allocation”. In: *Decision Support Systems* 146 (July 2021), p. 113545. ISSN: 0167-9236. DOI: 10.1016/j.dss.2021.113545. URL: https://www.sciencedirect.com/science/article/pii/S0167923621000555 (visited on 11/20/2024).
[17] C. Sriram and A. Haghani. “An optimization model for aircraft maintenance scheduling and reassignment”. In: *Transportation Research Part A: Policy and Practice* 37.1 (2003), pp. 29–48. ISSN: 09658564 (ISSN). DOI: 10.1016/S0965-8564(02)00004-6. URL: https://www.scopus.com/inward/record.uri?eid=2-s2.0-0037213213&doi=10.1016%2fS0965-8564%2802%2900004-6&partnerID=40&md5=483dbb99049d14ee4cf7d75b5dea7ef2.
[18] L.P. Jiang. “An Optimization Model for Aircraft Maintenance Scheduling Based on ABC Algorithm”. en. In: *Advanced Materials Research* 490-495 (Mar. 2012). Publisher: Trans Tech Publications, Ltd., pp. 147–151. ISSN: 1662-8985. DOI: 10.4028/www.scientific.net/amr.490-495.147. (Visited on 07/07/2025).
[19] Y. Qin et al. “A mathematical model and algorithms for the aircraft hangar maintenance scheduling problem”. In: *Applied Mathematical Modelling* 67 (Mar. 2019), pp. 491–509. ISSN: 0307-904X. DOI: 10.1016/j.apm.2018.11.008. URL: https://www.sciencedirect.com/science/article/pii/S0307904X18305353.
[20] H. Li et al. “Prediction of Non-Routine Tasks Workload for Aircraft Maintenance with Supervised Learning”. In: 2024. DOI: 10.2514/6.2024-2529.
[21] P. Andrade et al. “Aircraft Maintenance Check Scheduling Using Reinforcement Learning”. en. In: *Aerospace* 8.4 (Apr. 2021). Number: 4 Publisher: Multidisciplinary Digital Publishing Institute, p. 113. ISSN: 2226-4310. DOI: 10.3390/aerospace8040113. URL: https://www.mdpi.com/2226-4310/8/4/113 (visited on 11/20/2024).
[22] J. Lee and M. Mitici. “Deep reinforcement learning for predictive aircraft maintenance using probabilistic Remaining-Useful-Life prognostics”. In: *Reliability Engineering & System Safety* 230 (Feb. 2023),

p. 108908. ISSN: 0951-8320. DOI: 10.1016/j.ress.2022.108908. URL: https://www.sciencedirect.com/science/article/pii/S0951832022005233.

[23] F.S. Hillier and G.J. Lieberman. *Introduction to Operations Research*. en. 10th ed. McGraw-Hill, 2015. ISBN: 978-1-259-25318-8.

[24] M.P. Kleeman and G.B. Lamont. "Solving the Aircraft Engine Maintenance Scheduling Problem Using a Multi-objective Evolutionary Algorithm". en. In: *Evolutionary Multi-Criterion Optimization*. Ed. by C.A. Coello Coello, A. Hernández Aguirre, and E. Zitzler. Berlin, Heidelberg: Springer, 2005, pp. 782–796. ISBN: 978-3-540-31880-4. DOI: 10.1007/978-3-540-31880-4_54.

[25] Y. Zhang et al. "A baseline-reactive scheduling method for carrier-based aircraft maintenance tasks". In: *Complex & Intelligent Systems* 9.1 (Feb. 2023), pp. 367–397. ISSN: 2199-4536. DOI: 10.1007/s40747-022-00784-9.

[26] N.M. Semaan and N. Yehia. "A stochastic detailed scheduling model for periodic maintenance of military rotorcraft". In: *Aircraft Engineering and Aerospace Technology* 91.9 (Oct. 2019), pp. 1195–1204. ISSN: 1748-8842. DOI: 10.1108/AEAT-09-2018-0254.

[27] D.P. Pereira et al. "Planning of Aircraft Fleet Maintenance Teams". In: *Aerospace* 8.5 (May 2021). ISSN: 2226-4310. DOI: 10.3390/aerospace8050140.

[28] H. Shahmoradi-Moghadam, N. Safaei, and S.J. Sadjadi. "Robust Maintenance Scheduling of Aircraft Fleet: A Hybrid Simulation-Optimization Approach". In: *IEEE Access* 9 (2021), pp. 17854–17865. ISSN: 2169-3536. DOI: 10.1109/ACCESS.2021.3053714.

[29] M. Marseguerra and E. Zio. "Optimizing maintenance and repair policies via a combination of genetic algorithms and Monte Carlo simulation". In: *Reliability Engineering & System Safety* 68.1 (Apr. 2000), pp. 69–83. ISSN: 0951-8320. DOI: 10.1016/S0951-8320(00)00007-7. URL: https://www.sciencedirect.com/science/article/pii/S0951832000000077 (visited on 09/27/2024).

[30] P. de Bruecker et al. "A model enhancement heuristic for building robust aircraft maintenance personnel rosters with stochastic constraints". en. In: *European Journal of Operational Research* 246.2 (Oct. 2015), pp. 661–673. ISSN: 03772217. DOI: 10.1016/j.ejor.2015.05.008. URL: https://linkinghub.elsevier.com/retrieve/pii/S037722171500380X (visited on 12/06/2024).

[31] Q.C. Deng and B.F. Santos. "Lookahead approximate dynamic programming for stochastic aircraft maintenance check scheduling optimization". In: *European Journal of Operational Research* 299.3 (June 2022), pp. 814–833. ISSN: 0377-2217. DOI: 10.1016/j.ejor.2021.09.019.

[32] D. Dinis, A. Barbosa-Póvoa, and A.P. Teixeira. "Valuing data in aircraft maintenance through big data analytics: A probabilistic approach for capacity planning using Bayesian networks". In: *Computers & Industrial Engineering* 128 (Feb. 2019), pp. 920–936. ISSN: 0360-8352. DOI: 10.1016/j.cie.2018.10.015.

[33] J.M. Hu et al. "Optimal maintenance scheduling under uncertainties using Linear Programming-enhanced Reinforcement Learning". In: *Engineering Applications of Artificial Intelligence* 109 (Mar. 2022). ISSN: 0952-1976. DOI: 10.1016/j.engappai.2021.104655.

[34] M. Masmoudi and A. Haït. "Fuzzy uncertainty modelling for project planning: application to helicopter maintenance". In: *International Journal of Production Research* 50.13 (2012), pp. 3594–3611. ISSN: 0020-7543. DOI: 10.1080/00207543.2012.670925.

[35] V. Mattila and K. Virtanen. "Scheduling fighter aircraft maintenance with reinforcement learning". In: *Proceedings of the 2011 Winter Simulation Conference (WSC)*. ISSN: 1558-4305. Dec. 2011, pp. 2535–2546. DOI: 10.1109/WSC.2011.6147962. URL: https://ieeexplore.ieee.org/abstract/document/6147962 (visited on 07/07/2025).

[36] F.R. Oenzil and D.P. Ishak. "Inventory Management based on Component Reliability toward Uncertainty of Demand in Aircraft Spare Parts". In: 2021, pp. 363–368. DOI: 10.1145/3468013.3468375. URL: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85143900280&doi=10.1145%2f3468013.3468375&partnerID=40&md5=3ffd1aa619615bced56963aec1c820d6.

[37] Y. Qin et al. "A scenario-based stochastic programming approach for aircraft expendable and rotable spare parts planning in MRO provider". In: *Industrial Management & Data Systems* 120.9 (Sept. 2020), pp. 1635–1657. ISSN: 0263-5577. DOI: 10.1108/IMDS-03-2020-0131. (Visited on 08/21/2020).

[38] L.J.S. Rosales, J.B. Yang, and Y.W. Chen. "Analysing delays and disruptions in Aircraft Heavy Maintenance". In: *The 32nd International Conference of the System Dynamics Society. Delft, Netherlands*. 2014. URL: https://proceedings.systemdynamics.org/2014/proceed/papers/P1326.pdf (visited on 11/27/2024).

[39] P. Samaranayake and S. Kiridena. "Aircraft maintenance planning and scheduling: an integrated framework". In: *Journal of Quality in Maintenance Engineering* 18.4 (2012), pp. 432–+. ISSN: 1355-2511. DOI: 10.1108/13552511211281598. (Visited on 11/08/2024).

[40] S.Y. Sohn and K.B. Yoon. "Dynamic preventive maintenance scheduling of the modules of fighter aircraft based on random effects regression model". In: *Journal of the Operational Research Society* 61.6 (June 2010). Publisher: Taylor & Francis _eprint: https://doi.org/10.1057/jors.2008.167, pp. 974–979. ISSN: 0160-5682. DOI: 10.1057/jors.2008.167. URL: https://doi.org/10.1057/jors.2008.167 (visited on 07/07/2025).

[41] I. Tseremoglou and B.F. Santos. "Condition-Based Maintenance scheduling of an aircraft fleet under partial observability: A Deep Reinforcement Learn-

ing approach". In: *Reliability Engineering & System Safety* 241 (Jan. 2024). ISSN: 0951-8320. DOI: 10.1016/j.ress.2023.109582. (Visited on 09/04/2024).

[42] I. Tseremoglou, P.J. van Kessel, and B.F. Santos. "A Comparative Study of Optimization Models for Condition-Based Maintenance Scheduling of an Aircraft Fleet". In: *Aerospace* 10.2 (Feb. 2023). ISSN: 2226-4310. DOI: 10.3390/aerospace10020120.

[43] B. Alhijawi and A. Awajan. "Genetic algorithms: theory, genetic operators, solutions, and applications". en. In: *Evolutionary Intelligence* 17.3 (June 2024), pp. 1245–1256. ISSN: 1864-5917. DOI: 10.1007/s12065-023-00822-6. URL: https://doi.org/10.1007/s12065-023-00822-6 (visited on 06/07/2025).

[44] M. Albadr et al. "Genetic Algorithm Based on Natural Selection Theory for Optimization Problems". en. In: *Symmetry* 12 (Oct. 2020), pp. 1–31. DOI: 10.3390/sym12111758. URL: https://www.researchgate.net/publication/344832850_Genetic_Algorithm_Based_on_Natural_Selection_Theory_for_Optimization_Problems (visited on 06/07/2025).

[45] J.E. Rowe. "Genetic Algorithms". en. In: *Springer Handbook of Computational Intelligence*. Ed. by J. Kacprzyk and W. Pedrycz. Berlin, Heidelberg: Springer, 2015, pp. 825–844. ISBN: 978-3-662-43505-2. DOI: 10.1007/978-3-662-43505-2_42. URL: https://doi.org/10.1007/978-3-662-43505-2_42 (visited on 07/07/2025).

[46] E. Zio. *The Monte Carlo Simulation Method for System Reliability and Risk Analysis*. en. Springer Series in Reliability Engineering. ISSN: 1614-7839. London: Springer, 2013. ISBN: 978-1-4471-4587-5 978-1-4471-4588-2. DOI: 10.1007/978-1-4471-4588-2. URL: https://link.springer.com/10.1007/978-1-4471-4588-2 (visited on 07/07/2025).

[47] F.M. Dekking et al. "The law of large numbers". en. In: *A Modern Introduction to Probability and Statistics*. ISSN: 2197-4136. Springer, London, 2005, pp. 181–194. ISBN: 978-1-84628-168-6. DOI: 10.1007/1-84628-168-7_13. URL: https://link-springer-com.tudelft.idm.oclc.org/chapter/10.1007/1-84628-168-7_13 (visited on 07/07/2025).

[48] M.D.D. Clarke. "Irregular airline operations: a review of the state-of-the-practice in airline operations control centers". In: *Journal of Air Transport Management* 4.2 (Apr. 1998), pp. 67–76. ISSN: 0969-6997. DOI: 10.1016/S0969-6997(98)00012-X. URL: https://www.sciencedirect.com/science/article/pii/S096969979800012X (visited on 07/09/2025).

A. Appendix

A.1. Algorithm for Initial Population

Algorithm 1 shows the algorithm used to generate the initial population of the genetic algorithm. It favours internal hangars and high-capacity bays. Each bay is assigned a base score equal to its daily man-hour capacity, with internal bays receiving an additional bonus of 1,000 hours. These scores are then used to sort the bays. A power-weighting selection is applied to randomly assign the work package to a bays in the initial population, increasing the probability of selecting options with higher scores.

Algorithm 1: Algorithm for initial population with internal and high-capacity bay preference

Input: Population size P, work packages W, bay capacity B, random seed s
Output: Initialised population matrix

Set random seed to s
Initialise empty population list
for $i \leftarrow 1$ **to** P **do**
    Initialise empty individual
    for work package w in W **do**
        Get aircraft type a_w and check type c_w from w
        Initialise empty list of compatible bays
        foreach bay b in B **do**
            if $a_w \in b$.aircraft types and $c \in b$.check_types **then**
                Add (b.number, b.location_type, b.capacity) to compatible bays
        Sort compatible bays: Internal bays first, then by descending man-hour capacity
        Select a bay from the list using power-weighted selection
        Append selected bay and a random sequence priority to individual
    Add individual to population
return population

A.2. Results Genetic Algorithm

In this section, the outcomes of the genetic algorithm can be found for the two different seasons, with two variants per season: one with the highest fitness and another where a relatively high fitness is combined with reasonable computation time. Tthe fitness evolution plot of season 2025-2026 is shown in Figure 10 in addition to the plot shown in the article. The plannings are shown in Figure 11 and Figure 12.

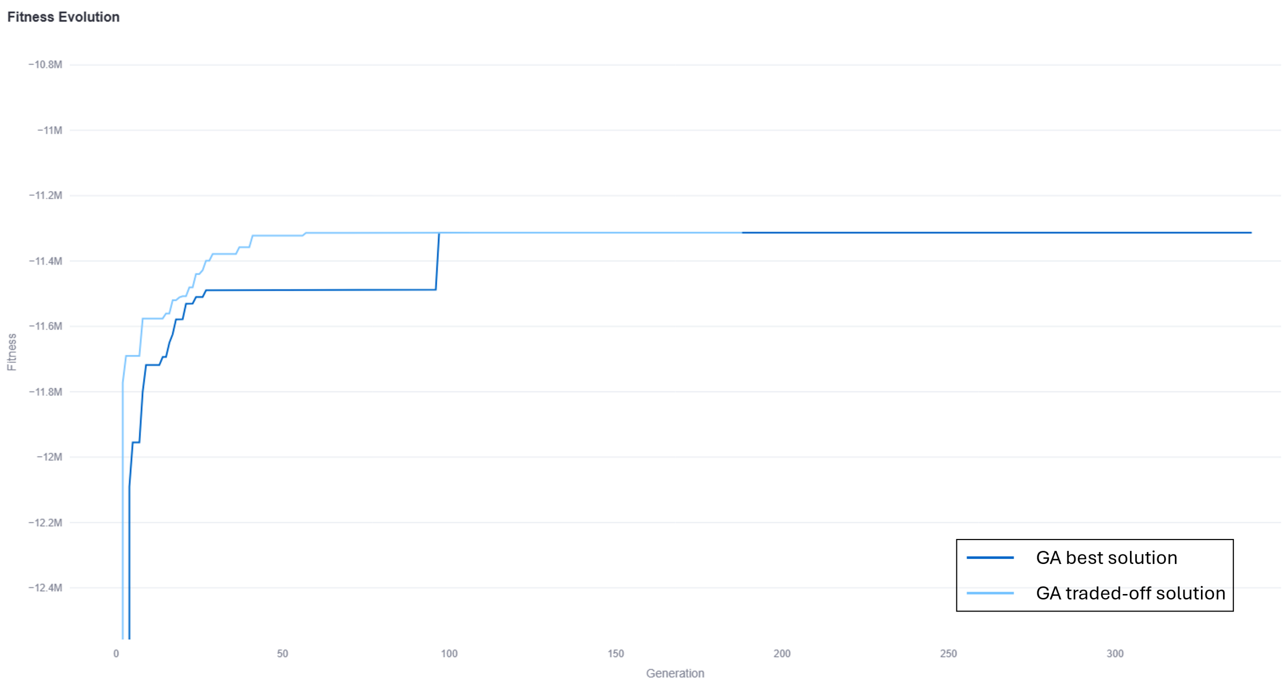
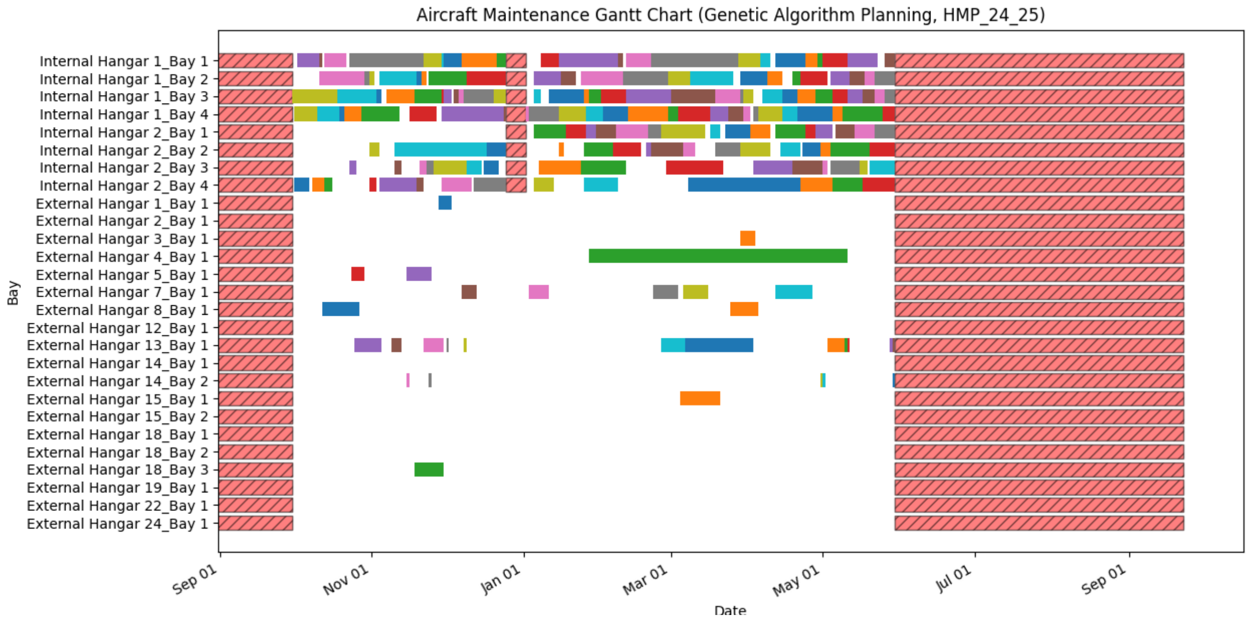
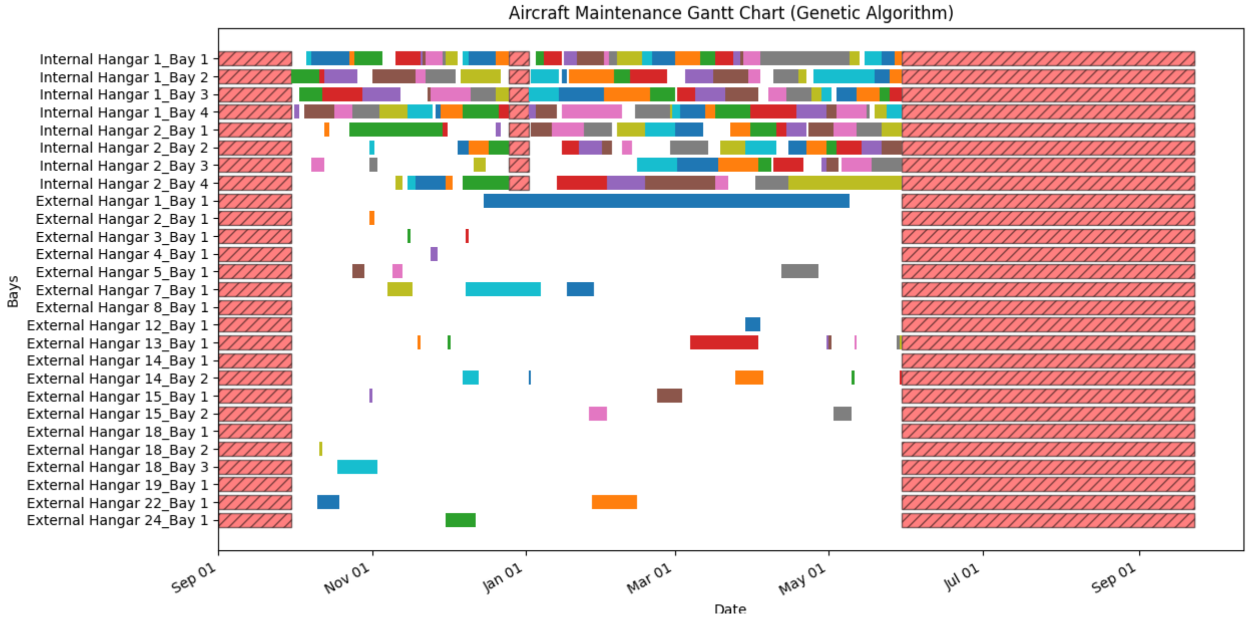


Fig. 10. Fitness evolution plots for season 2025-2026.

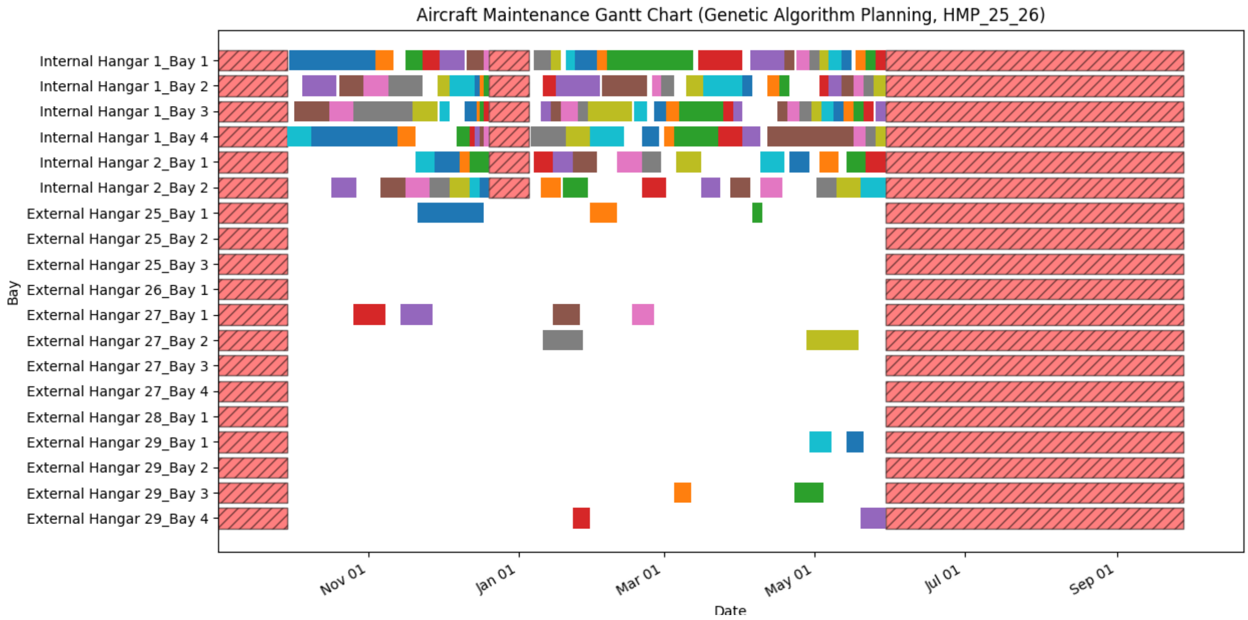


(a) Planning with highest fitness (tournament-2, uniform crossover, paired inversion mutation, 20% mutation percentage.

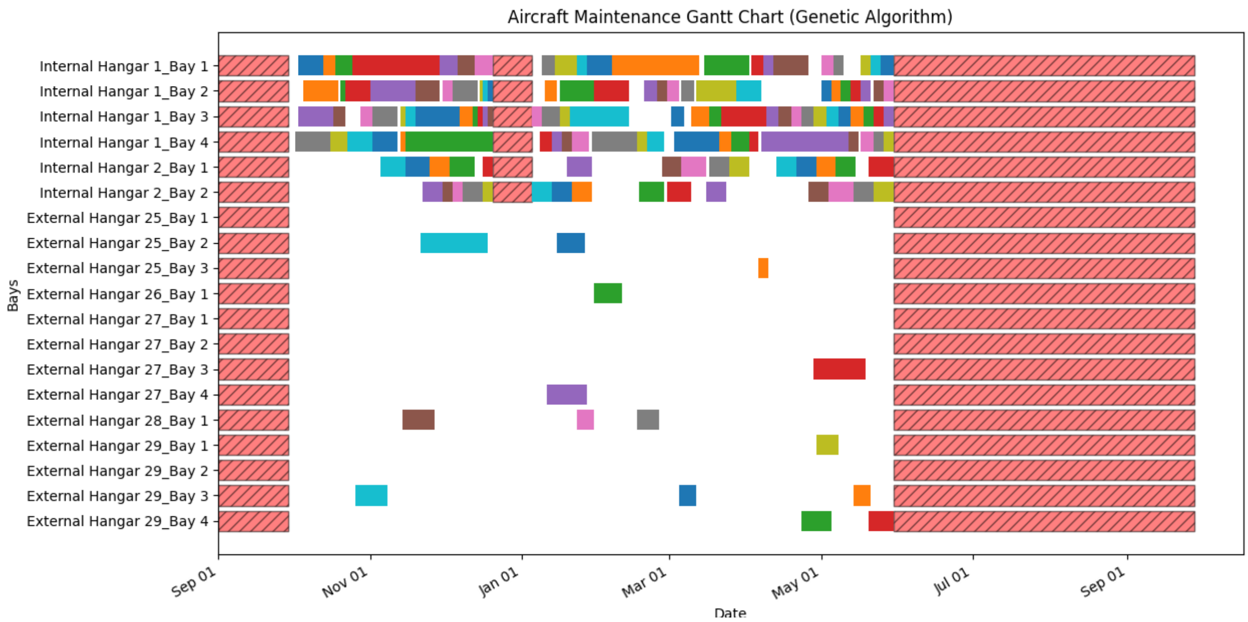


(b) Planning with high fitness combined with reasonable computation time (steady-state selection, uniform crossover, paired inversion mutation, 20% mutation percentage.

Fig. 11. Genetic Algorithm plannings for season '24 - '25.



(a) Planning with highest fitness (steady-state selection, paired two-points crossover, custom power-weighted mutation, 5% mutation percentage.



(b) Planning with high fitness combined with reasonable computation time (steady-state selection, two-points crossover, custom power-weighted mutation, 5% mutation percentage.

Fig. 12. Genetic Algorithm plannings for season '25 - '26.

A.3. Influence standard deviation percentage on simulations

In Figure 13 - Figure 15, the influence of the standard deviation on the three indicators is presented for both seasons. The percentage is calculated by dividing the number of conflicts by the number of simulations and by the number of checks in the schedule to normalise the values. The trend observed is relatively linear, with the MILP 2024-2025 having the highest probabilities in every graph. The three different variants for season 2025-2026 are very close to each other.

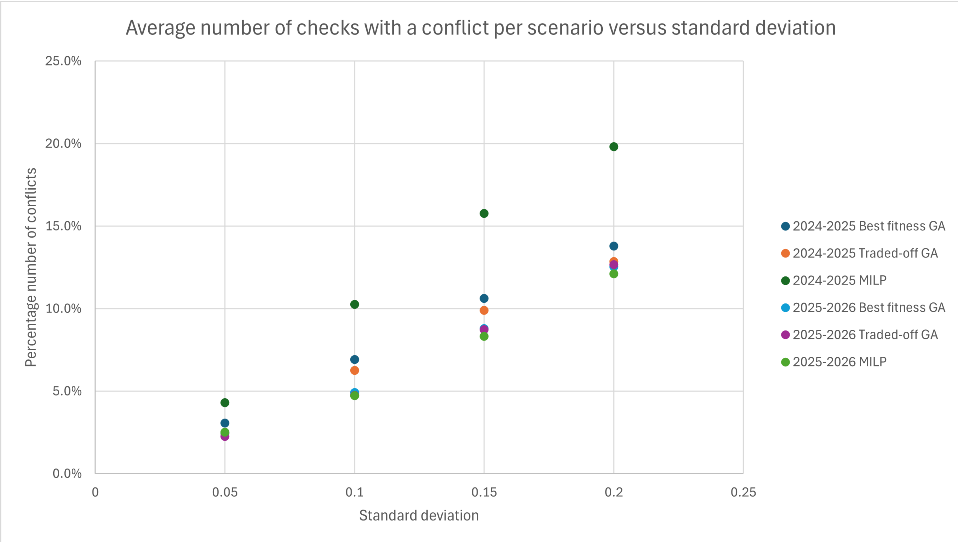


Fig. 13. Influence of standard deviation on percentage of checks with conflicts.

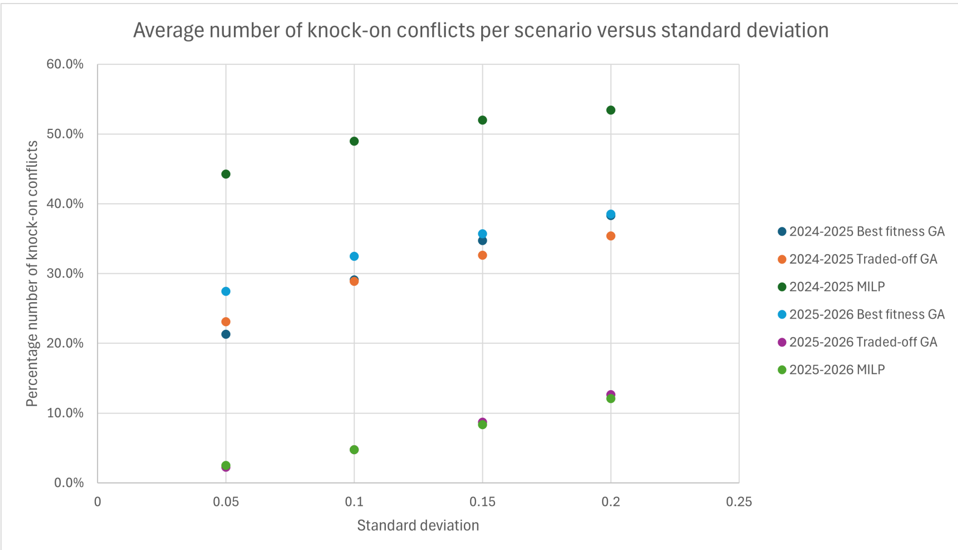


Fig. 14. Influence of standard deviation on percentage of checks with knock-on conflicts.

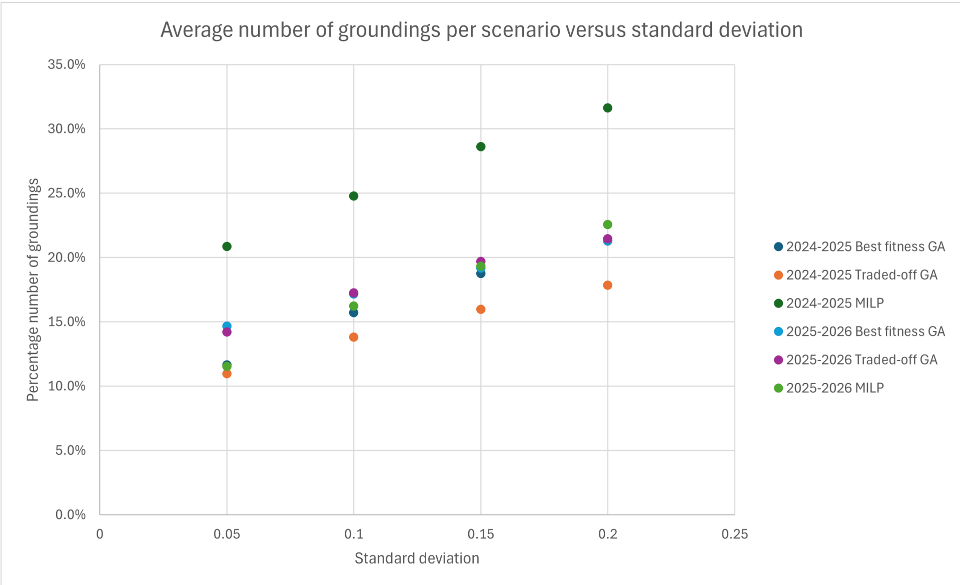


Fig. 15. Influence of standard deviation on percentage of checks with a grounding.

Part III

Closure

References

[1] G. Cros. *Airline Maintenance Cost Executive Commentary*. Tech. rep. IATA, 2022.

[2] T. van der Weide et al. “Robust long-term aircraft heavy maintenance check scheduling optimization under uncertainty”. In: *Computers & Operations Research* 141 (May 2022). DOI: 10.1016/j.cor.2021.105667.

[3] J. van den Bergh et al. “Aircraft maintenance operations: state of the art”. In: (2013). URL: https://kuleuven.limo.libis.be/discovery/fulldisplay/lirias1677212/32KUL_KUL:Lirias (visited on 11/18/2024).

[4] N.J. Boere. “Air Canada Saves with Aircraft Maintenance Scheduling”. In: *Interfaces* 7.3 (May 1977), pp. 1–13. DOI: 10.1287/inte.7.3.1. URL: https://pubsonline-informs-org.tudelft.idm.oclc.org/doi/10.1287/inte.7.3.1 (visited on 11/05/2024).

[5] P.J. van Kessel et al. “Airline maintenance task rescheduling in a disruptive environment”. In: *European Journal of Operational Research* 308.2 (July 2023), pp. 605–621. DOI: 10.1016/j.ejor.2022.11.017.

[6] *IATA - Unveiling the biggest airline costs*. URL: https://www.iata.org/en/publications/newsletters/iata-knowledge-hub/unveiling-the-biggest-airline-costs/ (visited on 11/14/2024).

[7] R.M.H. Knotts. “Civil aircraft maintenance and support Fault diagnosis from a business perspective”. In: *Journal of Quality in Maintenance Engineering* 5.4 (Jan. 1999). Publisher: MCB UP Ltd, pp. 335–348. DOI: 10.1108/13552519910298091. URL: https://doi.org/10.1108/13552519910298091 (visited on 07/07/2025).

[8] R.K. Mobley. *An introduction to predictive maintenance*. en. 2nd ed. Amsterdam New York: Butterworth-Heinemann, 2002.

[9] M. Kála et al. “Development of a tool for access worktime estimation in aircraft maintenance”. In: *Transportation Research Procedia*. INAIR 2020 - CHALLENGES OF AVIATION DEVELOPMENT 51 (Jan. 2020), pp. 46–55. DOI: 10.1016/j.trpro.2020.11.007. URL: https://www.sciencedirect.com/science/article/pii/S2352146520308577 (visited on 07/09/2025).

[10] R. Fedorov et al. “Taxonomy of candidate’s selection for prioritized predictive maintenance in maintenance, repairs and overhaul organizations”. en. In: *Journal of Quality in Maintenance Engineering* 29.3 (Oct. 2022). Publisher: Emerald Publishing Limited, pp. 589–605. DOI: 10.1108/JQME-04-2022-0022. URL: https://www.emerald.com/insight/content/doi/10.1108/jqme-04-2022-0022/full/html (visited on 11/19/2024).

[11] S.P. Ackert. “Basics of aircraft maintenance programs for financiers”. In: *Evaluation \& insights of commercial aircraft maintenance programs* 10 (2010).

[12] I. Tseremoglou et al. “A Comparative Study of Optimization Models for Condition-Based Maintenance Scheduling of an Aircraft Fleet”. In: *Aerospace* 10.2 (Feb. 2023). DOI: 10.3390/aerospace10020120.

[13] Y. He et al. “Maximizing robustness of aircraft routing with heterogeneous maintenance tasks”. In: *Transportation Research Part E: Logistics and Transportation Review* 177 (Sept. 2023), p. 103237. DOI: 10.1016/j.tre.2023.103237. URL: https://www.sciencedirect.com/science/article/pii/S1366554523002259.

[14] Q.C. Deng et al. “Lookahead approximate dynamic programming for stochastic aircraft maintenance check scheduling optimization”. In: *European Journal of Operational Research* 299.3 (June 2022), pp. 814–833. DOI: 10.1016/j.ejor.2021.09.019.

[15] P. Samaranayake et al. “Aircraft maintenance planning and scheduling: an integrated framework”. In: *Journal of Quality in Maintenance Engineering* 18.4 (2012), pp. 432–+. DOI: 10.1108/13552511211281598. (Visited on 11/08/2024).

[16] C. Sriram et al. “An optimization model for aircraft maintenance scheduling and re-assignment”. In: *Transportation Research Part A: Policy and Practice* 37.1 (2003), pp. 29–48. DOI: 10.1016/S0965-8564(02)00004-6. URL: https://www.scopus.com/inward/record.uri?eid=2-s2.0-0037213213&doi=10.1016%2fS0965-8564%2802%2900004-6&partnerID=40&md5=483dbb99049d14ee4cf7d75b5dea7ef2.

[17] P. de Bruecker et al. “A model enhancement heuristic for building robust aircraft maintenance personnel rosters with stochastic constraints”. en. In: *European Journal of Operational Research* 246.2 (Oct. 2015), pp. 661–673. DOI: 10.1016/j.ejor.2015.05.008. URL: https://linkinghub.elsevier.com/retrieve/pii/S037722171500380X (visited on 12/06/2024).

[18] M.C. Dijkstra et al. “Planning the Size and Organization of KLM’s Aircraft Maintenance Personnel”. en. In: *Interfaces* 24.6 (Dec. 1994), pp. 47–58. DOI: 10.1287/inte.24.6.47. URL: https://pubsonline.informs.org/doi/10.1287/inte.24.6.47 (visited on 12/04/2024).

[19] S. Yan et al. “A model and a solution algorithm for airline line maintenance manpower supply planning with multiple aircraft type maintenance certificates”. In: *Journal of the Chinese Institute of Engineers, Transactions of the Chinese Institute of Engineers,Series A/Chung-kuo Kung Ch’eng Hsuch K’an* 27.5 (2004), pp. 719–729. DOI: 10.1080/02533839.2004.9670919. URL: https://www.scopus.com/inward/record.uri?eid=2-s2.0-4544230157&doi=10.1080%2f02533839.2004.9670919&partnerID=40&md5=f6f4760b899e380c9227805af8135555.

[20] Q. Deng et al. “A practical dynamic programming based methodology for aircraft maintenance check scheduling optimization”. In: *European Journal of Operational Research* 281.2 (Mar. 2020), pp. 256–273. DOI: 10.1016/j.ejor.2019.08.025. URL: https://www.sciencedirect.com/science/article/pii/S0377221719306782 (visited on 11/20/2024).

[21] L.J.S. Rosales et al. “Analysing delays and disruptions in Aircraft Heavy Maintenance”. In: *The 32nd International Conference of the System Dynamics Society. Delft, Netherlands*. 2014. URL: https://proceedings.systemdynamics.org/2014/proceed/papers/P1326.pdf (visited on 11/27/2024).

[22] Y. Qin et al. “A scenario-based stochastic programming approach for aircraft expendable and rotable spare parts planning in MRO provider”. In: *Industrial Management & Data Systems* 120.9 (Sept. 2020), pp. 1635–1657. DOI: 10.1108/IMDS-03-2020-0131. (Visited on 08/21/2020).

[23] F.R. Oenzil et al. “Inventory Management based on Component Reliability toward Uncertainty of Demand in Aircraft Spare Parts”. In: 2021, pp. 363–368. DOI: 10.1145/3468013.3468375. URL: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85143900280&doi=10.1145%2f3468013.3468375&partnerID=40&md5=3ffd1aa619615bced56963aec1c820d6.

[24] D.P. Pereira et al. “Planning of Aircraft Fleet Maintenance Teams”. In: *Aerospace* 8.5 (May 2021). DOI: 10.3390/aerospace8050140.

[25] C.G. Vassiliadis et al. “Maintenance scheduling and process optimization under uncertainty”. In: *Computers & Chemical Engineering* 25.2-3 (Mar. 2001), pp. 217–236. DOI: 10.1016/S0098-1354(00)00647-5.

[26] M. Masmoudi et al. “Fuzzy uncertainty modelling for project planning: application to helicopter maintenance”. In: *International Journal of Production Research* 50.13 (2012), pp. 3594–3611. DOI: 10.1080/00207543.2012.670925.

[27] D. Dinis et al. “Valuing data in aircraft maintenance through big data analytics: A probabilistic approach for capacity planning using Bayesian networks”. In: *Computers & Industrial Engineering* 128 (Feb. 2019), pp. 920–936. DOI: 10.1016/j.cie.2018.10.015.

[28] D. Dinis et al. “A supporting framework for maintenance capacity planning and scheduling: Development and application in the aircraft MRO industry”. In: *International Journal of Production Economics* 218 (Dec. 2019), pp. 1–15. DOI: 10.1016/j.ijpe.2019.04.029.

[29] N.M. Semaan et al. "A stochastic detailed scheduling model for periodic maintenance of military rotorcraft". In: *Aircraft Engineering and Aerospace Technology* 91.9 (Oct. 2019), pp. 1195–1204. DOI: 10.1108/AEAT-09-2018-0254.

[30] H. Shahmoradi-Moghadam et al. "Robust Maintenance Scheduling of Aircraft Fleet: A Hybrid Simulation-Optimization Approach". In: *IEEE Access* 9 (2021), pp. 17854–17865. DOI: 10.1109/ACCESS.2021.3053714.

[31] J.M. Hu et al. "Optimal maintenance scheduling under uncertainties using Linear Programming-enhanced Reinforcement Learning". In: *Engineering Applications of Artificial Intelligence* 109 (Mar. 2022). DOI: 10.1016/j.engappai.2021.104655.

[32] Y. Zhang et al. "A baseline-reactive scheduling method for carrier-based aircraft maintenance tasks". In: *Complex & Intelligent Systems* 9.1 (Feb. 2023), pp. 367–397. DOI: 10.1007/s40747-022-00784-9.

[33] H. Li et al. "Prediction of Non-Routine Tasks Workload for Aircraft Maintenance with Supervised Learning". In: 2024. DOI: 10.2514/6.2024-2529.

[34] I. Tseremoglou et al. "Condition-Based Maintenance scheduling of an aircraft fleet under partial observability: A Deep Reinforcement Learning approach". In: *Reliability Engineering & System Safety* 241 (Jan. 2024). DOI: 10.1016/j.ress.2023.109582. (Visited on 09/04/2024).

[35] A Boudjelida. "On the robustness of joint production and maintenance scheduling in presence of uncertainties". In: *JOURNAL OF IN□LIGENT MANUFACTURING* 30.4 (Apr. 2019), pp. 1515–1530. DOI: 10.1007/s10845-017-1303-9.

[36] O. Naud et al. "Chapter 4 - Support to decision-making". In: *Agricultural Internet of Things and Decision Support for Precision Smart Farming*. Ed. by Annamaria Castrignanò et al. Academic Press, Jan. 2020, pp. 183–224. DOI: 10.1016/B978-0-12-818373-1.00004-4. URL: https://www.sciencedirect.com/science/article/pii/B9780128183731000044 (visited on 11/19/2024).

[37] F.S. Hillier et al. *Introduction to Operations Research*. en. 10th ed. McGraw-Hill, 2015.

[38] K. Sörensen et al. "Metaheuristics". en. In: *Encyclopedia of Operations Research and Management Science*. Ed. by S.I. Gass et al. Boston, MA: Springer US, 2013, pp. 960–970. DOI: 10.1007/978-1-4419-1153-7_1167. URL: https://doi.org/10.1007/978-1-4419-1153-7_1167 (visited on 11/20/2024).

[39] H.-L. Ma et al. "Tackling uncertainties in aircraft maintenance routing: A review of emerging technologies". In: *Transportation Research Part E: Logistics and Transportation Review* 164 (2022). DOI: 10.1016/j.tre.2022.102805. URL: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85132754027&doi=10.1016%2fj.tre.2022.102805&partnerID=40&md5=86d6e43c6b2f93aa80e5cbed0b2c0d32.

[40] C. Silva et al. "Adaptive reinforcement learning for task scheduling in aircraft maintenance". In: *Scientific Reports* 13.1 (Oct. 2023), p. 16605. DOI: 10.1038/s41598-023-41169-3. URL: https://doi.org/10.1038/s41598-023-41169-3.

[41] M. Witteman et al. "A bin packing approach to solve the aircraft maintenance task allocation problem". In: *European Journal of Operational Research* 294.1 (Oct. 2021), pp. 365–376. DOI: 10.1016/j.ejor.2021.01.027. URL: https://www.sciencedirect.com/science/article/pii/S0377221721000576 (visited on 11/20/2024).