Stochastic discrete event simulation of airline network and maintenance operations

Varenna, Sara; Li, Haonan; Ribeiro, Marta; Santos, Bruno F.

DOI
10.1016/j.jairtraman.2025.102789

Publication date
2025

Document Version
Final published version

Published in
Journal of Air Transport Management

Citation (APA)
Varenna, S., Li, H., Ribeiro, M., & Santos, B. F. (2025). Stochastic discrete event simulation of airline network and maintenance operations. *Journal of Air Transport Management*, *125*, Article 102789. https://doi.org/10.1016/j.jairtraman.2025.102789

Journal of Air Transport Management 125 (2025) 102789

Contents lists available at ScienceDirect

Journal of Air Transport Management

journal homepage: www.elsevier.com/locate/jairtraman

Stochastic discrete event simulation of airline network and maintenance operations

Sara Varenna, Haonan Li *, Marta Ribeiro, Bruno F. Santos

Operations & Environment, Faculty of Aerospace Engineering Delft University of Technology, The Netherlands

ARTICLE INFO

Keywords:
Decision support
Stochastic discrete event simulation
Airline operations
Maintenance planning

ABSTRACT

The complexity of airline operations requires operations planning to be divided into multiple problems solved sequentially by the respective departments. This is particularly the case for (1) network planning and (2) maintenance planning. Despite the close interaction of these two departments, airlines typically evaluate plans from both domains separately. However, an integrated perspective is necessary to develop robust plans and effective recovery policies in this intrinsically uncertain environment. This paper presents a new modular, stochastic, discrete event simulation model of airline operations named ANEMOS (Airline Network and Maintenance Operations Simulation). ANEMOS contains both network and maintenance dynamics, allowing the evaluation of plans, policies, and scenarios from both domains. The model is validated using data from a major European airline. We show that the simulated results closely resemble the airline's historical operational performance. ANEMOS is tested with a use-case investigating the effects of adding a second reserve aircraft to a fleet of fifty wide-body aircraft. The results show that the second reserve is capable of reducing cancellations by 55%. However, such does not cover the lost revenue associated with keeping an aircraft non-operational for a part of the time.

1. Introduction

Airlines try to make the best use of their fleet of aircraft while applying strict maintenance regulations to ensure the airworthiness of their aircraft fleet. The choices about when to schedule maintenance for each aircraft and which aircraft to assign to each flight significantly affect the airlines' profitability. Ideally, optimizing both maintenance and network operations simultaneously would yield the best outcome. However, due to the complexity of each operation type, varying requirements, decision time frames, and objectives, planning is typically divided into several steps carried out by different departments within the airline.

In practice, the maintenance department sets high-level fleet requirements while the network department determines the flight schedule. Major aircraft maintenance checks, which can last from a day to several weeks in the hangar, are planned based on this flight schedule, influencing aircraft availability to flights. The fleet management department finalizes the aircraft routes a few days before operations start. Any subsequent changes to these routes, whether due to maintenance or other needs, must be approved by the fleet management department or the Operations Control Center (OCC) in the days before the operation. Concurrently, the maintenance department defines the allocation of aircraft to the maintenance checks and minor inspections, defining the task packages that make up each check or inspection for the fleet in the upcoming days. The maintenance department aims to ensure the aircraft remains airworthy and healthy for the days ahead while considering routine tasks and maintenance needs from reported technical issues.

The interconnection between maintenance and network operations is also very clear in the uncertain environment of the day-to-day operations. Disruptions caused by bad weather conditions, airspace and airport congestion, or technical problems can easily spread through the network and compromise maintenance plans. Operations must be planned robustly so that disruptions can be avoided or mitigated. When disruptions occur, the airline must be capable of applying effective disruption recovery policies to restore the undisrupted plans efficiently. However, these policies are often based only on experience. Given the high uncertainty in this environment, it is hard to evaluate how policies used in each domain will affect the other domains.

Airlines can greatly benefit from a model capable of investigating how decisions made in a domain would affect operations as a whole. The literature on this subject, however, is scarce. The proposed works tend to focus on either network (Rosenberger et al., 2002) or maintenance (Duffuaa and Andijani, 1999; Öhman et al., 2020; Iwata and

* Corresponding author.
E-mail address: George.Li@tudelft.nl (H. Li).

https://doi.org/10.1016/j.jairtraman.2025.102789
Received 17 September 2024; Received in revised form 28 February 2025; Accepted 21 March 2025
Available online 5 April 2025

Mavris, 2013) operations separately, simulating the other domain in a simplified manner. As a result, the full picture of airline dynamics is not correctly represented.

To overcome this limitation, this paper presents a new modular, stochastic discrete event simulation model of airline operations named ANEMOS (Airline Network and Maintenance Operations Simulation). ANEMOS simulates network and maintenance operations of the intercontinental fleet of a hub-and-spoke carrier. It may be used as a framework to test policies, plans, and scenarios involving both network and maintenance operations to understand the system's performance as a whole. Its dynamic structure allows the evaluation of plans and policies at the strategic, tactical, and operational levels. In addition, ANEMOS also allows the evaluation of the effects of multiple external factors (e.g. increased hub congestion, (departure) delays, unexpected maintenance repairs) on the airline's performance.

This paper is structured as follows. First, Section 2 presents an overview of comparable works found in the literature. The problem tackled by this research is detailed in Section 3, with an overview of the decisions made at strategic, tactical, and operational levels by ANEMOS. Section 4 defines the structure of ANEMOS, explaining each module in detail. ANEMOS is applied to a case study with a major European airline, as described in Section 5. The results of this case study are then presented in Section 6. How ANEMOS translates to other case studies, its limitations, and future improvements are discussed in Section 7. Finally, Section 8 concludes this work.

2. Literature review

Simulation has been previously used in literature in the field of airline maintenance and network operations planning and optimization. Most works use a simulation framework only as an instrument for testing their models (Barnhart et al., 2002; Aloulou et al., 2010; Vos et al., 2015). A few papers focus on the simulation models, with the objective of using them to assess scenarios and to support decision-making (Duffuaa and Andijani, 1999; Iwata and Mavris, 2013; Nisse et al., 2023; Xu et al., 2024; Geske et al., 2024; van Schilt et al., 2024). While some of these works are developed by airlines that are interested in evaluating what-if scenarios in their operations (Duffuaa and Andijani, 1999; Öhman et al., 2020; Nisse et al., 2023), other models are mainly developed for research purposes to allow model testing and comparison (Rosenberger et al., 2002; Xu et al., 2024; Geske et al., 2024; van Schilt et al., 2024).

Several simulation models for airline operations have been proposed in the literature. Among the simulation models presented in the literature, two noteworthy mentions are made. First, the SimAir (Lee et al., 2003) academic simulation tool was developed for simulating airline operations and recovery strategies. It has been used in a few cases (Lan et al., 2006; Ben Ahmed et al., 2017; Rosenberger et al., 2004). It includes turnaround and block time, weather, influences from other airlines, and crew and passenger flow. However, maintenance is simulated in a simplified manner by only considering regular maintenance stops and unscheduled maintenance in between flights with a certain probability. Second, the Discrete Event Simulation (DES) framework presented in Pohya et al. (2021) is capable of evaluating the effects of using specific products, technologies, and policies in the long run throughout the life cycle of an aircraft or fleet. It simulates the complete lifetime of an aircraft, from its purchase to the flights and maintenance executed on it, up until its retirement. The long-term perspective used makes this a very useful model for evaluating the effects of high-level, strategic policies on the overall life cycle of an aircraft. Due to this wide perspective, however, both network and maintenance operations are simulated in a simple manner: maintenance slots are not scheduled but rather executed at fixed intervals or when pre-defined degradation levels of components are reached.

Works (Jacobs et al., 2005; Duffuaa and Andijani, 1999; Öhman et al., 2020) make clear that airlines value the insights that operations simulation can provide. Jacobs et al. (2005) describes an operations simulation by an international airline. The model assesses the schedule by simulating disruptions and recovery, allowing aircraft swaps, the use of a reserve aircraft, reducing maintenance time, and cancelling flights. The goal of Duffuaa and Andijani (1999) is to evaluate the impact of different maintenance policies on airline operations. The presented framework is modular and includes interactive modules such as a planning and scheduling module for maintenance planning. In turn, Iwata and Mavris (2013) can be used for assessing maintenance policies such as postponing task execution and parts logistics. More recently, works have been directed at scheduling maintenance tasks under uncertainty. Tseremoglou and Santos (2024) used a reinforcement learning for maintenance planning under uncertainty. Villafranca et al. (2025) developed a heuristic model to design a daily aircraft maintenance schedule under uncertain task times across bases with limited technicians. Both cases show that considering uncertainties can reduce maintenance costs. Finally, Zhang et al. (2024) provides each maintenance package the appropriate buffer time based on a risk assessment thus reducing the impact of flight delays in maintenance scheduling.

2.1. Research gap

Overall, current simulation efforts focus on either network or maintenance operations while modeling the other aspect in a simplified manner. As a consequence, these models are not capable of evaluating the integrated performance of airline operations, in which network and maintenance plans and policies closely interact with each other. A gap exists in the form of a simulation model of airline operations that includes the simulation of both network and maintenance operations. For airlines, such a model could be used to facilitate the negotiations between departments when defining constraints and requirements for planning and scheduling, testing the obtained plans, and evaluating the effects of disruption scenarios and specific recovery policies. In academia, this model would allow testing the effectiveness of proposed optimization models in a stochastic and network-maintenance integrated environment.

3. Problem definition

As already introduced, ANEMOS is a modular, stochastic, discrete event simulation model of the network and maintenance operations of hub-and-spoke carriers. It is developed with a modular structure, allowing for changing and adapting the single modules to a simulation's needs. The model simulates the dynamics of each aircraft, including flights and maintenance slots. The input to the model comprises a list of aircraft with their subtypes, a flight schedule with fleet assignment, a list of the maintenance slots available for each aircraft subtype, and a list of requirements and deferred defects (DDs) for each subtype. For the discrete event simulation, the input must describe a deterministic or stochastic measure of each simulated activity. Note that passenger connections and crew rosters are not considered.

Fig. 1 displays the operations levels, and respective actions, covered by ANEMOS. By utilizing this framework, we can systematically evaluate decisions across three distinct levels: strategic, tactical, and operational. Strategic decisions, such as fleet planning and network development, shape the long-term direction of operations. Tactical decisions, including fleet assignment and aircraft rotations, optimize resource allocation within the established strategy. Operational decisions, such as line maintenance planning, ensure the smooth execution of daily operations. Since these decisions are integrated within different modules, the framework provides a comprehensive approach to assessing their interactions and overall impact.

The following sections provide more detail on the actions of ANEMOS at each operational level.

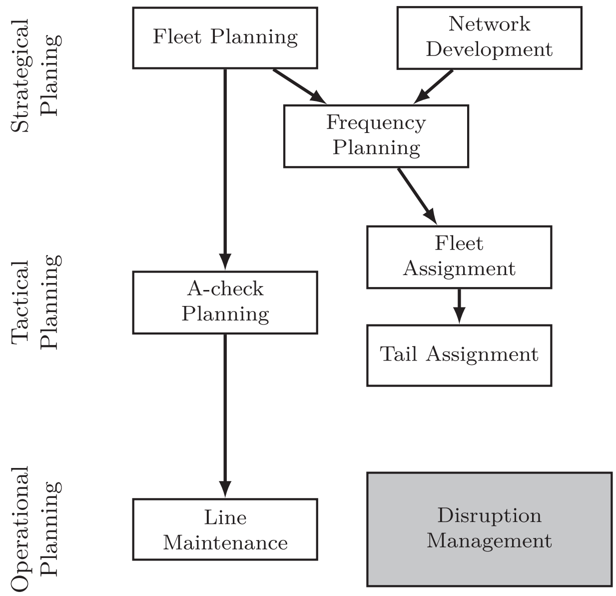

Fig. 1. The operations levels, and respective actions, covered by ANEMOS.

3.1. Actions at strategic level

At the strategical level, airlines plan their fleet, determining both the aircraft type and quantity—based on the origin–destination pairs they aim to serve. The combination of these two factors dictates the number of flights an airline can operate, shaping its frequency planning. During the model construction phase of ANEMOS, fleet planning and network development are integrated. By modifying these parameters, we can systematically assess the impact of such decisions through subsequent simulations.

3.2. Actions at tactical level

At the tactical level, ANEMOS helps airlines carry out network and maintenance planning. The former includes fleet and tail assignment. ANEMOS simulates the operations of hub-and-spoke carriers with one hub. Thus, the assignment of flights to aircraft is done at the rotation level, meaning each aircraft follows a sequence of flights that depart from and return to the hub. A rotation comprises at least two flights, with one or multiple intermediate stops at outstations. ANEMOS takes as input a weekly schedule of rotations, including flight departure and arrival times as well as assigned fleet types.

Alongside network planning, ANEMOS will schedule maintenance taks. The duration of the maintenance slot is calculating by summing the total labor hours required for the assigned tasks (including both routine and non-routine maintenance) and dividing by the available workforce during the slot. Non-routine maintenance refers to unscheduled repairs identified during inspections and can only occur in maintenance slots performed in a hangar. The probability of encountering non-routine maintenance is derived from historical data. At the start of each hangar slot, ANEMOS uses historical probabilities to determine whether non-routine issues will arise within the assigned work package. Notably, ANEMOS assumes that the likelihood of non-routine findings is independent of the number, type, or duration of scheduled tasks.

Finally, note that the scope of maintenance slots included in ANEMOS is limited to slots with a duration comparable to A-checks, which for wide-body aircraft is around 24 h. Furthermore, it is assumed that all maintenance is carried out at the hub. Requirements that are included within letter-checks work packages are also excluded since the scope of the simulation is limited to maintenance up to A-checks.

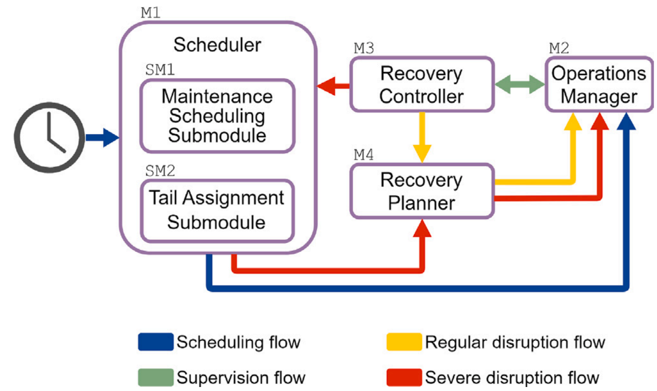

Fig. 2. The four modules of ANEMOS, and the interaction flows between them.

3.3. Actions at operational level

At the operational level, ANEMOS considers two key decisions: line maintenance (LM) and disruption management. Line maintenance includes all maintenance tasks that can be performed on the aircraft while it is on the ground between flights. These tasks ensure continued airworthiness and operational readiness without significant downtime.

Regarding disruptions, ANEMOS models two types: Aircraft on Ground (AOG) events, and delays. An AOG event occurs when an aircraft is deemed unairworthy and requires maintenance before it can resume operations. In ANEMOS, AOG events are modeled independently for each aircraft . Note that the simulation is limited to AOG scenarios occurring at the hub.

A model of aggregated primary delays is present in the ANEMOS. This model represents all delays with the exception of propagated delays and technical delays, which should emerge from the simulation dynamics (e.g., delays caused by factors such as crew availability, airport congestion, and weather disruptions). This delay is included in the form of an action the aircraft goes through after its scheduled departure time, or, in the case of delayed departure due to technical disruptions or propagated delay, after the turnaround activities have ended.

Two different models of aggregated primary delay are developed, one for outstations and one for the hub. While at outstations flights often depart hours or days apart, at the hub many flights depart within a short period of time. And, given that primary departure delays often appear to be linked to airport congestion, the departure delay of flights departing in short time window cannot be assumed to be independent of one another. To capture this dependency, the disruption state of the hub airport is modeled as a discrete-state, discrete-time process, and a different distribution for each disruption state is then used to determine the primary delay that a departing flight should experience. This approach ensures a more accurate representation of primary departure delays at congested hub airports.

4. Methodology

Fig. 2 shows the four modules that make up the simulation, along with the interaction flows that connect them.

A simulation clock regularly calls the *Scheduler* (M1), which assigns to each aircraft a feasible sequence of maintenance slots and flights to be flown. This module has two submodules: the *Maintenance Scheduling* Submodule (MSS.SM1), and the *Tail Assignment* Submodule (MSS.SM2), which are called sequentially.Within these two submodules, several tactical decisions are implemented, including fleet assignment, aircraft rotations, and maintenance slot allocation. Additionally, certain operational decisions, such as line maintenance planning, are also incorporated. The output of the *Scheduler* is the input to the second module, the *Operations Manager* (M2), which includes separate discrete event processes for each of the simulated aircraft and manages the dynamics of the simulation. The *Recovery Controller* (M3) monitors the aircraft processes and intervenes whenever a disruption in the original schedule is found. Generally, the *Recovery Controller* calls the *Recovery Planner* (M4) to find an optimal solution for the disruption at hand. However, if the occurring disruption impacts flights falling after the end of the recovery window, then the *Recovery Controller* calls the *Tail Assignment* Submodule first to provide a long-term solution. This solution is then given as input to the *Recovery Planner*, which optimally solves the disruption within its considered recovery window. Modules M2, M3, and M4 primarily focus on exploring operational decisions. The MSS receives a list of available maintenance slots for each aircraft subtype and a list of tasks for each registration, assigns slots to specific aircraft, and schedules the execution of tasks within these slots. The scheduling window of the MSS covers a fixed number of weeks. The following sections describe each module and submodule in detail.

4.1. Maintenance scheduling submodule (MSS)

The MSS receives a list of available maintenance slots for each aircraft subtype and a list of tasks for each registration, assigns slots to specific aircraft, and schedules the execution of tasks within these slots. The scheduling window of the MSS covers a fixed number of weeks.

4.1.1. Maintenance tasks

Each aircraft has a set of tasks that need to be scheduled and executed. Each task is characterized by its arrival date (i.e. the date when a DD is found or when the previous instance of a requirement is completed), ready date (i.e. the day from which the task can be executed), and due date (i.e. the date before which the task must be performed). Each task has an estimated duration and labor hours for its execution. We assume that a set of tasks can be scheduled within a maintenance slot if (1) the duration of each task is shorter than that of the maintenance slot and (2) the total number of labor hours associated with the set of tasks does not exceed the slot's maximum allowed labor hours. Each slot is characterized by its start and end date, its assigned aircraft type, the number of labor hours that can be scheduled within it, and the maximum duration and labor hours of each task that can be scheduled in it. Finally, the task is characterized by a location where it can be executed. Hangar tasks can only be executed in the hangar; platform tasks can either be executed within the hangar or on the platform.

The tasks to be scheduled may be requirements or DDs, which differ in nature. Requirements are routine tasks characterized by an interval defined in calendar days. When requirements are characterized by an interval expressed in flight hours or flight cycles, this is translated into calendar days assuming a fixed number of flight hours and flight cycles flown per day. The arrival date of the first instance of a requirement is randomized. Note that the simulation does not allow requirements due at the beginning of the simulation. A new requirement is generated whenever the previous one is executed. The scope of the considered requirements is limited based on their interval. Requirements with an interval shorter than the time intercurring between two MSS calls are excluded since it would be impossible to schedule them before their due dates.

In turn, DDs are one-off tasks and not independent of other tasks—these often occur at the same time. ANEMOS considers a separate task inter-arrival process for each simulated aircraft and assumes that DDs only arrive at the beginning of each day. A weighted choice is used to determine the number of days between DDs' arrival. A second weighted choice is used to determine how many DDs arrive on an arrival day. The specific DDs arriving are sampled from historically arrived DDs. With each call of the MSS, the DDs arriving before the next call are disclosed.

Note that, in real-life operations, the ready date of a task is generally limited by the availability of material. For requirements that are scheduled, repetitive tasks, it is assumed that the required material is always available, and their ready date coincides with their arrival date. For DDs, we consider that the necessary material may not be immediately available. If the DD requires material, then the historical date of part availability is used to compute its ready date. If not, then the DD can be executed from its arrival date. Finally, in real life, when a task exceeds its due date, the aircraft is grounded until the task is executed. However, the event of a task going due is extremely rare. Thus, given the limited flexibility available to the model, a task going due does not have direct effects on the simulation dynamics and does not lead to the grounding of aircraft. Grounding is modeled independently by the Aircraft on Group process, defined in Section 4.3.2. Additionally, requirements are always executed at a fixed fraction of the interval of the requirement, independently of when the previous requirements task was due.

4.1.2. Scheduled maintenance slots

Three types of scheduled maintenance slots are included in the simulation: Line Maintenance (LM) slots, Flexible (Flex) slots, and Mandatory Hangar (MH) slots. LM are modeled as weekly 'bins' that include all the tasks scheduled in line maintenance for the week. An LM slot is defined for each aircraft each week, and tasks can be scheduled within it if their ready date and due date allow scheduling between its start and end date. When a requirement is executed within an LM slot, it is assumed to be executed at a fixed fraction of its interval. For DDs, the interval is defined as the time between the task's arrival and the due date. When the computed execution date falls out of the start and end date of the LM slot, the date is moved to the closest time boundary.

Flex slots are part of the input of the model. They can be defined for one or more weeks, and they are repeated over the simulated time window. The input must also specify the location in which they are executed, i.e. on the platform or in a hangar. These slots are not necessarily used, but they are only simulated when the Scheduler assigns them to an aircraft, meaning that at least one task is scheduled within them. Differently from LM slots, these slots are simulated within the aircraft processes of the Operations Manager.

MH slots represent all maintenance slots scheduled to execute a set of routine tasks (i.e. A-checks or slots scheduled to execute cabin modifications). These are characterized by a nominal duration independent of the tasks scheduled within them. It is assumed that a certain number of labor hours, defined for each slot, can be executed during this time. Additionally, they are mandatory, i.e. they should always be assigned to an aircraft, independently of the fact that any additional task is scheduled in their work package. The cancellation of MH slots by the MSS and Recovery Planner is not prohibited, but it is strongly disincentivized.

4.1.3. Mathematical formulation

This mathematical formulation is derived from the work of van Kessel et al. (2022). The sets, variables, and parameters are described in Table 1. The objective function described in Eq. (1a) minimizes the costs associated with assigning or unassigning a task to a specific slot.

The cost of assigning a task to a specific maintenance slot varies per task type and slot. Requirements should be scheduled as close as possible to their due date in order to minimize the wasted interval, i.e. the fraction of the required interval that is lost due to the anticipation of its execution. On the other hand, DDs should be executed as soon as possible. Regardless, all tasks should be executed with some anticipation with respect to their due date, in order to add a buffer for maintenance postponement in case of disruptions. The weight of assigning a task to a maintenance slot follows the function displayed in Fig. 3, where the cost of assigning a requirement to a slot generally decreases for later slots, while the cost of assigning a DD increases for slots starting later in time. After the preferred anticipation is reached, the cost of the assignment for all tasks then increases at a higher rate. Not having a task assigned to a slot is penalized if the task is due within a week after the end of the scheduling window.

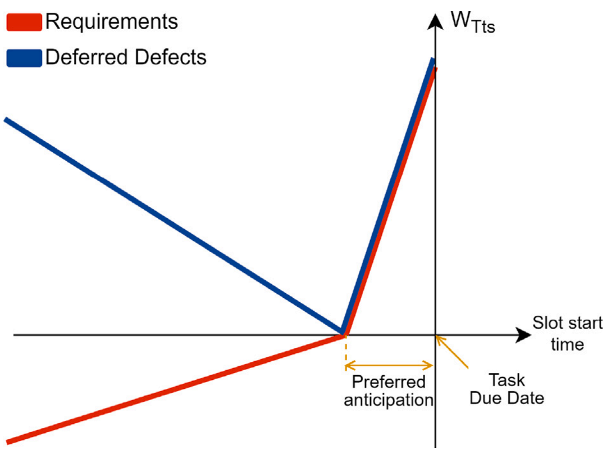

Fig. 3. Cost of assigning a task to a slot, based on task type and slot date.

The second term of the objective represents the cost of activating a maintenance slot with the exception of LM slots, which are not penalized as LM tasks can always be executed in this context. This cost assumes different values for different types of slots. Mandatory Hangar (MH) slots should, in theory, be mandatory. However, in order to avoid infeasible situations, the use of MH slots is strongly incentivized by giving a high negative value to their activation weight. Flex slots should only be used when necessary, and the use of shorter slots should be preferred. Therefore, the cost of assigning a Flex slot to an aircraft is made up of two components: a fixed value connected to the activation of a slot dependent on its location, and a value proportional to the duration of the slot.

The *fixed scheduling window* is defined as the time from the beginning of the scheduling window to the next scheduled call of the MSS. When MH slots and Flex slots fall within this time window and they are assigned to an aircraft different from the one they were assigned to in the previous call of the MSS, they should receive an additional penalty. The purpose is twofold: first, in real life, rescheduling maintenance slots close to their start date can cause a waste of resources. Second, the MSS works in a close relationship with the TAS, which, in previous calls of the Scheduler, had found a tail assignment solution based on the previous input of the MSS. Changing the assignment of maintenance slots within the fixed scheduling window could cause incompatibility with the previous plans of the TAS, leading to unnecessary flight cancellations.

Constraints (1b) impose that all tasks are scheduled within a slot or unassigned. Constraints (1c) ensure that only one aircraft is assigned to a slot. Constraints (1d) allow a task to be assigned to a slot only if the slot is assigned to the task's aircraft. Constraints (1e) impose that a slot can be activated only if at least one task is assigned to it. This set of constraints does not apply to MH slots since they should always be assigned. Constraints (1f) restrict the assignment of a slot to an aircraft to one per week. Constraints (1g) limits the total labor scheduled within a slot's work package to the slot's maximum allowed labor. This set of constraints, along with Constraints (1b) are the only two sets of constraints that interest LM slots in addition to Flex and MH slots. This is because each aircraft has a pre-assigned weekly LM slot which does not require activation. Eqs. (1h)–(1j) describe the decision variables' domain.

A maintenance slot can be assigned to an aircraft if their subtype matches. A task can be assigned to a maintenance slot when its duration is shorter than the maximum allowed duration of a slot, when it is within the maximum labor hours that a slot allows per task, when their location matches, when the slot falls between the task's ready and due date, and when the task's aircraft is compatible with the slot in terms of subtype. These constraints are imposed through the use of subsets.

Table 1
Mathematical formulation of the MSS.

Sets and subsets		
A		Aircraft
T		Tasks
S		Flex slots and MH Slots
L		Line maintenance slots
C		Weeks included within the scheduling window
A^s	$\subseteq A$	Aircraft that can be assigned to slot s
S_F	$\subseteq S$	Flex slots
S^t	$\subseteq S$	Slots in which task t can be executed
S^c	$\subseteq S$	Slots in week c
L^t	$\subseteq L$	Line Maintenance slots in which task t can be executed
T^a	$\subseteq T$	Tasks of aircraft a
T^s	$\subseteq T$	Tasks that can be executed in slot s
Decision variables		
$\delta_{A_{as}}$	$\in \{0,1\}$	1 if slot s is assigned to aircraft a, 0 otherwise
$\delta_{T_{ts}}$	$\in \{0,1\}$	1 if task t is scheduled in slot s, 0 otherwise
δ_{U_t}	$\in \{0,1\}$	1 if task t is not scheduled in any slot, 0 otherwise
Parameters		
$W_{S_{sa}}$	Cost of assigning slot s to aircraft a	
$W_{T_{ts}}$	Cost of scheduling task t in slot s	
W_{U_t}	Cost of leaving task t unscheduled	
P_{TL_t}	Labor hours required to execute task t	
P_{SL_s}	Maximum labor hours that can be assigned to slot s	
M	Large constant	

Minimize:

$$\sum_{t\in T}\left(\sum_{s\in S^t\cup L^t} W_{T_{ts}}\delta_{T_{ts}} + W_{U_t}\delta_{U_t}\right) + \sum_{s\in S}\sum_{a\in A^S} W_{S_{sa}}\delta_{A_{as}} \tag{1a}$$

Subject to:

$$\sum_{s\in S^t\cup L^t}\delta_{T_{ts}} + \delta_{U_t} = 1 \qquad \forall t\in T \tag{1b}$$

$$\sum_{a\in A^s}\delta_{A_{as}} \le 1 \qquad \forall s\in S \tag{1c}$$

$$\sum_{t\in T^a\cap T^s}\delta_{T_{ts}} \le M\delta_{A_{as}} \qquad \forall s\in S, \forall a\in A^S \tag{1d}$$

$$\sum_{a\in A^s}\delta_{A_{as}} \le \sum_{t\in T^s}\delta_{T_{ts}} \qquad \forall s\in S_F \tag{1e}$$

$$\sum_{s\in S^c}\delta_{A_{as}} \le 1 \qquad \forall c\in C \forall a\in A^s \tag{1f}$$

$$\sum_{t\in t^s} P_{\mathrm{TL}_t}\delta_{T_{ts}} \le P_{\mathrm{SL}_s} \qquad \forall s\in S\cup L \tag{1g}$$

$$\delta_{A_{as}} \in \{0,1\} \qquad \forall s\in S, \forall a\in A^s \tag{1h}$$

$$\delta_{T_{ts}} \in \{0,1\} \qquad \forall t\in T, \forall s\in S^t \tag{1i}$$

$$\delta_{U_t} \in \{0,1\} \qquad \forall t\in T \tag{1j}$$

Table 2
Mathematical formulation of the TAS.

Sets and subsets		
R		Rotations and reserve slots (segments)
A		Aircraft
A^r	$\subseteq A$	Aircraft that can be assigned rotation or reserve slot r
OV		Set of unordered sets $(r,t), r \in R, t \in R$ where r and t overlap in time
Decision variables		
$\delta_{R_{ra}}$	$\in \{0,1\}$	1 if rotation or reserve slot r is assigned to aircraft a, 0 otherwise
δ_{U_r}	$\in \{0,1\}$	1 if rotation or reserve slot r remains unassigned, 0 otherwise
Parameters		
$W_{R_{ra}}$	Cost of assigning rotation or reserve slot r to aircraft a	
W_{U_r}	Cost of leaving rotation or reserve slot r unassigned	

4.2. The Tail Assignment Submodule (TAS)

The TAS takes the output of the MSS as input and assigns a feasible sequence of flights to each aircraft, considering the pre-assigned maintenance slots. In addition to flights, the TAS is also capable of assigning a certain number of reserve slots, i.e. time slots during which an aircraft is scheduled to act as a reserve aircraft. Similarly to the MSS, the TAS considers a scheduling window that goes from the end of the recovery window for the Recovery Planner to a fixed number of weeks after the call day. The TAS also assigns reserve slots to aircraft, identifying the aircraft that are acting as reserves so that specific recovery policies involving the reserve aircraft can be implemented by the Recovery Planner. The number of daily reserve slots and their start and end times are inputs of the model.

4.2.1. Mathematical formulation

Since rotations and reserve slots are modeled in the same way, the term *segment* will be used to refer to either one of these entities. Eq. (2a) minimizes the costs of assigning a segment to a specific aircraft and leaving segments unassigned. In particular, the cost of unassigning segments is the highest since cancelling rotations and reserving slots should always be avoided.

The cost of assigning a segment to an aircraft depends on the nature of the segment. While the cost of assigning a reserve slot is constant, the cost of assigning a rotation to an aircraft is dependent on the aircraft type. Categories of preferred subtypes are defined, so that if a feasible assignment cannot be done within the originally assigned subtype, then a rotation can be assigned to other subtypes according to the preference.

Constraints (2b) is the cover constraints that impose that each segment is either assigned to one aircraft or unassigned. An unassigned rotation from the TAS solution is deemed cancelled in the simulation only if it falls before the next scheduled call of the maintenance scheduler, as changes in the slots assignment can lead to changes in the rotations assignments and to cancellations. Constraints (2c) prevent two overlapping segments from being assigned to the same aircraft - a buffer is considered before and after each rotation. The feasible assignment of a segment to an aircraft while considering the aircraft's pre-assigned maintenance slots is achieved by reducing the feasible subsets A^r. Reductions of these subsets can also be used to reduce the feasibility of aircraft-route assignments. Constraints (2d) and Constraints (2e) define the domain of the decision variables (see Table 2).

$$\text{Minimize:} \sum_{\substack{r\in R\\ a\in A^r}} W_{R_{ra}}\delta_{R_{ra}} + \sum_{r\in R} W_{U_r}\delta_{U_r} \tag{2a}$$

Subject to:

$$\sum_{a\in A^r} \delta_{R_{ra}} + \delta_{U_r} = 1 \qquad \forall r \in R \tag{2b}$$

$$\delta_{R_{ra}} + \delta_{R_{ta}} \leq 1 \qquad \forall (r,t) \in \text{OV}, \forall a \in A^r \cap A^t \tag{2c}$$

$$\delta_{R_{ra}} \in \{0,1\} \qquad \forall r \in R, \forall a \in A^r \tag{2d}$$

$$\delta_{U_r} \in \{0,1\} \qquad \forall r \in R \tag{2e}$$

4.3. The operations manager

The Operations Manager in Fig. 4 is responsible for the discrete event simulation dynamics. It includes three types of discrete event processes: the aircraft process, the Aircraft On the Ground (AOG) process, and the hub disruption process.

4.3.1. The aircraft process

One aircraft process is built for each of the simulated aircraft, describing the sequence of activities that the aircraft goes through. At the beginning of the simulation, all aircraft are located at the hub, ready to execute the next assigned rotation or maintenance slot, which will be generally defined as *duty* (Block AP1 in Fig. 4). If the next duty scheduled for the aircraft (AP2) is a rotation, then the aircraft waits for its scheduled departure time. Once this is reached (AP3), the rotation becomes the aircraft's current duty and its assignment cannot be changed anymore. At this point, the aircraft can experience a primary departure delay, which is summed to the propagated delay the aircraft is experiencing from previously executed duties. This delay represents a combination of all delays that are not technical or propagated delays, including delays related to crew, weather, and congestion, to cite some. When the delay time has passed, the aircraft takes off (AP4), and it reaches its destination after flying for a certain amount of time (AP5).

The aircraft undergoes the turnaround activities at an outer station (AP7) and then waits for the scheduled departure time of the next flight in the rotation (AP3). The flight activities are then repeated until the aircraft lands back in the hub (AP5) after executing the last flight in the rotation (AP6). At this point, or during the duration of its ground time at the hub, the aircraft can experience a grounding (AP8), as determined by the Aircraft On Ground (AOG) process of the corresponding aircraft (AP9, AP10). The aircraft then undergoes turnaround activities at the hub, and it is again ready for the next duty (AP1).

If the next duty of the aircraft (AP2) is a maintenance slot, the aircraft must wait for the scheduled start time. In particular, if it is a hangar maintenance slot the aircraft must wait for the scheduled start of towing (AP11), after which it is towed to the hangar, while if it is a platform slot, then the aircraft simply waits for the scheduled start time of the slot (AP12). In the slot, both scheduled tasks and non-routines are executed, and when the slot ends (AP13), the aircraft is towed back to the platform, if not already there. Once on the platform (AP14), the aircraft must wait for the turnaround time to elapse before it can start flying again (AP1). Before the next duty starts, the aircraft can be grounded, in accordance with the AOG process (AP9, AP10).

The duration of all cited activities, i.e. the time elapsed between two subsequent events, can assume a stochastic or deterministic value based on the simulation's needs. The arrows in Fig. 4 define an activity as deterministic or stochastic as implemented in the case study proposed in Section 5.

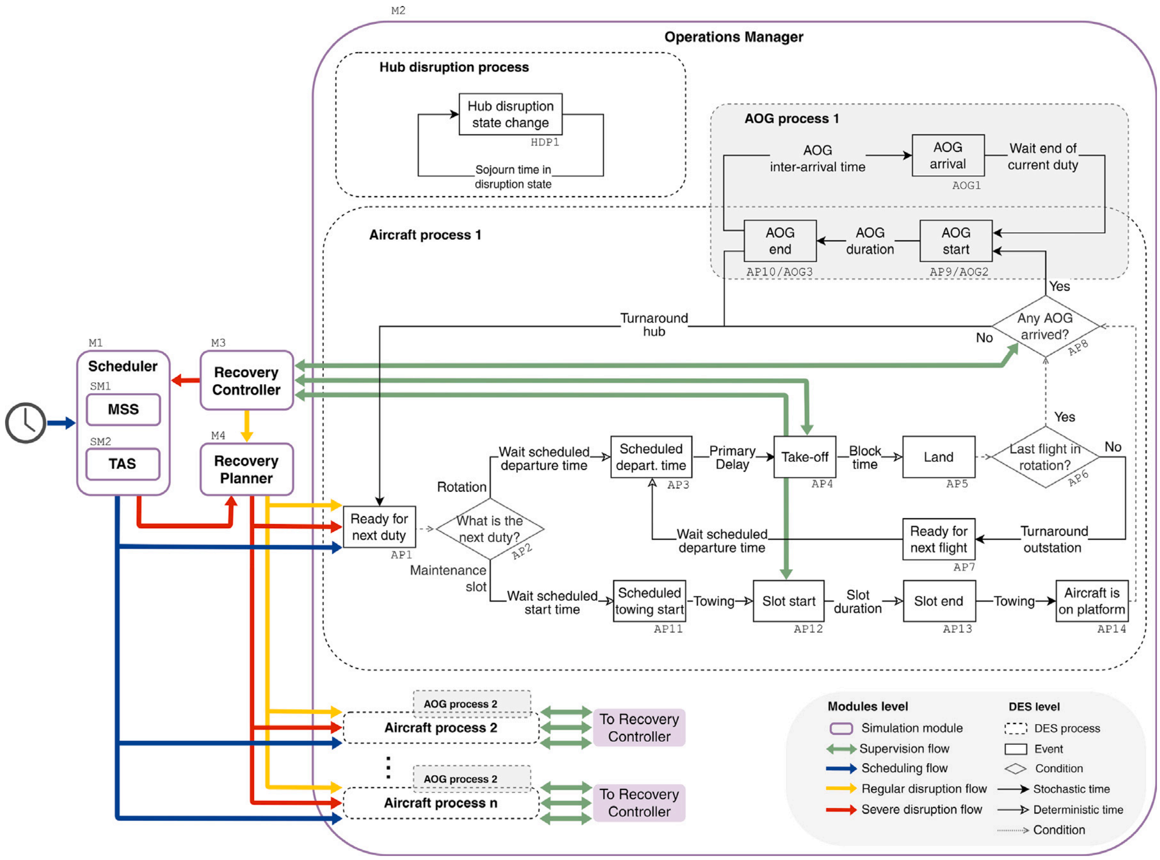

Fig. 4. Expansion of Fig. 2 that details the Operations Manager and its discrete event processes: the aircraft processes, i.e. the sequences of activities and events each aircraft goes through within the discrete event simulation, the AOG processes, which manage the grounding of the aircraft, and the hub disruption process, which keeps track of the disruption state at the hub.

4.3.2. The aircraft on the ground (AOG) process

AOG situations are modeled as an independent process, so that AOG situations happen with exponentially distributed inter-arrival time, and have a duration that follows a log-normal distribution. Given that the recovery module is rotation based, as opposed to flight based, the AOGs are assumed only to happen when the aircraft is located at the hub. When an AOG arrives during the execution of a rotation, it is postponed to when the aircraft reaches the hub. In some cases AOGs can require days to be solved. It can happen that an AOG slot overlaps with a scheduled maintenance slot and, in these cases, a call to the recovery module would generally lead to the cancellation of the maintenance slot. However, AOGs are opportunities in which the aircraft is on the ground available for receiving maintenance, and there is therefore no reason why a work package should not be executed as scheduled. Therefore, it is assumed that maintenance slots whose duration is shorter than that of the AOG by a certain multiplicative factor can be executed within the AOG slot. If this condition does not apply, then the maintenance slot is postponed to after the AOG time has elapsed.

An AOG process is defined for each aircraft, and it interacts with the corresponding aircraft process as shown in Fig. 4. When an AOG arrives for an aircraft(AOG1), i.e. when there is a finding that requires the grounding of the aircraft, the aircraft should stop executing duties. If the aircraft is on the ground at the hub at the time of the arrival, then the AOG starts right away. However, if the aircraft is currently executing a rotation or a maintenance slot, then the start of the AOG is postponed until the end of the execution of the aircraft's current duty. This is necessary to avoid having big disruptions within a rotation while both the Scheduler and the Recovery Planner are rotation-based rather than flight-based. Once the aircraft reaches the hub or it is back on the platform after undergoing scheduled maintenance, the AOG can start (AOG2), and end (AOG3) after a stochastically determined duration.

During the duration of the AOG, the aircraft cannot execute any duty. Once an AOG ends, a new one arrives after a certain time named the *AOG inter-arrival time* has elapsed.

Finally, AOGs can require days to be solved. AOGs are opportunities in which the aircraft is on the ground available for receiving maintenance, and there is, therefore, no reason why a work package should not be executed as scheduled during a grounding. Therefore, it can be assumed that maintenance slots whose duration is shorter than that of the AOG by a certain multiplicative factor can be executed within the AOG slot. If this condition does not apply, then the maintenance slot is postponed to after the AOG time has elapsed.

4.3.3. The hub disruption state process

Given that primary departure delays often appear to be linked to airport congestion, the departure delays of flights departing in a short time window cannot be assumed to be independent of one another. In order to account for this, the disruption state of the hub airport is modeled as a discrete-state, discrete-time process, and the primary delay of the departing flights is expressed as a function of the current disruption state.

To describe this process, two parameters must be set: the number of categorical disruption states considered, and the time steps, or time brackets that should be used to discretize time. The disruption state of the hub is initialized at the lowest level at the beginning of the day of operations. Then, the sojourn time in this state, i.e. the number of brackets during which the disruption state at the hub remains unvaried, is determined by sampling from an exponential distribution and by rounding the obtained value to the nearest integer. After the sojourn time has elapsed, the new state is determined by means of a transition

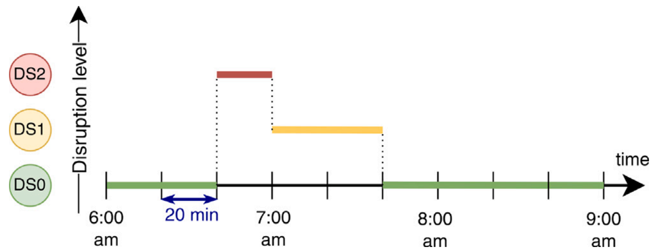

Fig. 5. Evolution of the hub disruption state over three hours, for three disruption states, and 20 min long brackets.

probability matrix, which describes the probability of transitioning from each state to each of the other considered states. The new sojourn time can now be computed, and so on. The process continues until it is initialized again at the beginning of the next morning, where a daily initialization is necessary because it is uncommon for an airport to be congested at night or at the early hours of the morning.

Fig. 5 shows the evolution of the hub disruption state over three hours, when three disruption states are considered and time is discretized in 20-min brackets. The disruption state is initialized at the minimum disruption level (DS0) at 6:00. The state changes after two-time brackets and transitions to disruption level DS2, where it remains until 7:00. After that, there is a transition to disruption level DS1, followed by a sojourn time of two-time brackets, and so on.

This stochastic process can be easily translated into a discrete event process with one recurring event of state change (HDP1), separated by an activity of duration corresponding to the sojourn time. This makes the process easily integrable within the Operations Manager in the form of the hub disruption process.

4.4. The recovery controller

The Recovery Controller supervises the aircraft processes of the Operations Manager to detect when disruptions occur, and when this happens, it requests a recovery action. The Recovery Controller interacts with each aircraft process in correspondence with three events: when a flight takes off (AP4 in Fig. 4), when a maintenance slot starts (AP12), and when an AOG arrives (AP8). At these points in time, it estimates the time at which the aircraft will be ready to start its next scheduled duty, given the current state. For a flight, this estimate is done by summing average turnaround times and flight duration, while the duration of maintenance slots and AOGs is assumed to be known.

This estimate of the next ready time of the aircraft is then compared to the currently expected departure time of the next duty: if the estimated ready time falls after the expected departure time by a minimum defined value, then a recovery action is deemed necessary. Notice that the expected, instead of the scheduled departure time of a duty, is considered due to the fact that both maintenance slots and rotations can be delayed by the Recovery Planner. If previous recovery calls have already delayed a duty, and it is expected that that delay will not be increased, then there is no need to call the Recovery Planner again.

When the Recovery Controller detects that a recovery action is needed, the procedure involves calling the Recovery Planner to find a solution within a relatively short recovery window, lasting for a time that is in the order of days. However, in some cases, disruption can be so severe that it affects duties not included within the recovery window of the Recovery Planner. These disruptions generally occur due to the arrival of AOGs that last for days. In these cases, it is necessary to call the TAS to define a long-term solution over the TAS' scheduling window, before the Recovery Planner can be called. Note that a call to the MSS is not necessary, because if any maintenance slot overlaps with such long AOGs, it is automatically included within them.

4.5. The recovery planner

The Recovery Planner is called whenever a disruption occurs, to find a short-term feasible solution. The model acts on a recovery window with a duration in the order of days, and it must produce a solution that is compatible with the assignment of rotations and slots that do not fall within the recovery window. The implemented Recovery Planner models one parallel time-space network for each aircraft in the considered fleet, as defined in Vink et al. (2020). Since the Recovery Planner proposed for ANEMOS is rotation-based, the time-space network is collapsed into a timeline, where the only airport from which arcs generate and terminate is the hub. The allowed recovery options include delaying or cancelling a rotation or a maintenance slot, changing the appointed aircraft to fly a rotation, using a reserve aircraft, swapping maintenance slots, or postponing maintenance slots to a future opportunity.

Eq. (3a) describes the objective of the ILP, which is a cost-minimization of all considered recovery options. The cost of assigning a rotation to an aircraft is dependent on the aircraft type and, in particular, on aircraft subtype preference groups, similar to what is done in the TAS. Delayed rotations are modeled as copies of the original rotations, departing and arriving at later nodes, in an approach that was initially proposed by Levin (1971). For this reason, a rotation can only be delayed by pre-defined discrete amounts of time. The cost of delaying a rotation and assigning it to an aircraft depends both on the aircraft subtype and on the duration of the delay. The latter component is assumed to vary linearly with the delay duration. In addition to the base cost of the assignment, the weight should be increased whenever the assignment does not correspond to the original assignment of the rotation in order to favor keeping the plan as it is.

Similarly to rotations, slots can be executed as originally planned, delayed, or cancelled. However, differently from rotations, slots cannot be freely reassigned, but they can only follow simple swap patterns. This means that if three aircraft A, B, and C are respectively assigned maintenance slots a, b, and c, it is possible to do a swap such as $A \rightarrow b, B \rightarrow a$, but not a swap such as $A \rightarrow b, B \rightarrow c, C \rightarrow a$. Furthermore, cancelling a slot should, in principle, not be allowed. This is because leaving some tasks un-executed would lead to the grounding of the aircraft. However, in order to avoid infeasible situations, cancelling a slot is allowed, at a very high cost. Another recovery option included in the model is the possibility of postponing maintenance slots to *flexible maintenance arcs*, which are arcs generated any time an aircraft is on the ground at the hub for a time that allows fitting a maintenance slot. Note that, especially for what concerns hangar slots, the assumption of always having the available resources to provide maintenance to an aircraft is a strong assumption, that does not necessarily represent actual operations. However, this option can be used to investigate scenarios of maximum maintenance flexibility. In general, ground arcs can be assigned a cost of zero, unless particular conditions require otherwise.

Finally, a binary slack variable is used to avoid unnecessary involvement of an aircraft in the recovery strategy. If any of the rotations previously assigned to an aircraft is reassigned to a different one, then the slack variable is activated and a penalty applies. This term acts in parallel with respect to the additional weight incurred by each re-assigned rotation, and it allows for a reduction of the number of aircraft involved in the recovery solution.

Constraints (3b) ensure that a rotation is executed as scheduled, delayed, or cancelled. Constraints (3c) impose that a slot is either executed as planned, delayed, cancelled, swapped, or postponed to a flexible maintenance arc. Constraints (3d) ensures the balance of the network. For each aircraft, the origin and termination nodes are defined based on their assigned rotations and slots before any recovery action is taken. In particular, the origin node corresponds to the time at which their current (or last) duty is expected to end, or, if an AOG has arrived, it corresponds to the expected end time of the AOG. The termination node is imposed on the time when the first duty not included within the recovery window is scheduled to start. All central nodes are put in correspondence with duties arriving or departing, always considering a possible buffer before and after the duty for either turnaround or towing operations. At each node, the sum of entering and exiting arcs must be equal to the balance of the node, i.e. 1 for origin nodes, −1 for termination nodes, and 0 for central nodes. Constraints (3e) activate the slack variables that avoid the re-assignment of rotations for each aircraft. Constraints (3f) is required given the formulation of the slot swaps. In fact, the decision variables associated with slot swaps refer to ordered pairs of slots (s,t) that can be swapped. This constraint imposes that if slot s is swapped with slot t, then slot t is also swapped with slot s. Finally, Constraints (3g)–(3p) define the domain for the included decision variables.

The assignment of a rotation or of a delayed rotation to an aircraft is not permitted when the rotation falls out of the aircraft's origin and termination nodes. Also, the re-assignment of a rotation or of a delayed rotation to a different aircraft is not permitted when a minimum anticipation, i.e. the time intercurring between the current time and the time when the rotation is expected to depart, is not guaranteed. The reduction of subsets can also be used to limit the types of aircraft that can fly specific routes (see Table 3).

Two Flex slots can be swapped when their scheduled work package would fit in the destination slot in terms of duration and scheduled labor, and when no task in either work package would go due before the new assigned slot's scheduled start. Also, there should be a match in aircraft subtype, slot location, and slot type for the swap to be possible. Flexible maintenance arcs are specifically generated for each maintenance slot assigned to each aircraft so that the expected duration of the slot fits within the maintenance arc. Finally, the objective function weights and the reduced subsets can easily be used to impose airline-specific policies. For example, the assignment of a rotation to an aircraft when the latter has a reserved slot scheduled can be disincentivized by assigning a higher cost of the specific rotation-aircraft assignment weight. If an airline policy requires a reserve aircraft to be available at the beginning of each day of operations, even at the cost of cancelling rotations scheduled in the coming days, the subset A^r can be reduced to prevent the assignment of rotations to aircraft when this would cause the overlap with reserve slots scheduled on the coming days.

Minimize:

$$\begin{aligned}&\sum_{r\in R}\sum_{a\in A^r} W_{R_{ra}}\delta_{R_{ra}}+\sum_{r\in R} W_{CR_r}\delta_{CR_r}\\&+\sum_{(r,a)\in DR}\sum_{a\in A^{rd}} +W_{DR_{rda}}\delta_{DR_{rda}}+\sum_{s\in S} W_{S_s}\delta_{S_s}\\&+\sum_{s\in S} W_{CS_s}\delta_{CS_s}+\sum_{(s,d)\in SD} W_{DS_{sd}}\delta_{DS_{sd}}\\&+\sum_{(s,t)\in SW}\frac{1}{2}W_{SW_{st}}\delta_{SW_{st}}+\sum_{(s,m)\in M} W_{M_{sm}}\delta_{M_{sm}}\\&+\sum_{g\in G^a}\sum_{a\in A} W_{G_{ga}}\delta_{G_{ga}}+\sum_{a\in A} W_{\Gamma_a} z_{\Gamma_a}\end{aligned}\tag{3a}$$

Subject to:

$$\sum_{a\in A^r}\delta_{R_{ra}}+\sum_{a\in A^{rd}}\sum_{\substack{(r,d)\\ \in DR^r}}\delta_{DR_{rda}}+\delta_{CR_r}=1\forall r\in R\tag{3b}$$

$$\begin{aligned}&\delta_{S_s}+\sum_{\substack{(s,d)\\ \in DS^s}}\delta_{DS_{sd}}+\delta_{CS_s}+\sum_{\substack{(s,m)\\ \in M^s}}\delta_{M_{sm}}\\&\quad+\sum_{\substack{(s,t)\\ \in SW^s}}\delta_{SW_{st}}=1\qquad\forall s\in S\end{aligned}\tag{3c}$$

$$\begin{aligned}&\sum_{\substack{r\in\\ R\cap OR^{na}}}\delta_{R_{ra}}-\sum_{\substack{r\in\\ R\cap TE^{na}}}\delta_{R_{ra}}+\sum_{\substack{(r,d)\in\\ DR\cap OR^{na}}}\delta_{DR_{rda}}\\&\quad-\sum_{\substack{(r,d)\in\\ DR\cap TE^{na}}}\delta_{DR_{rda}}+\sum_{\substack{s\in\\ S\cap OR^{na}}}\delta_{S_s}-\sum_{\substack{s\in\\ S\cap TE^{na}}}\delta_{S_s}\\&\quad+\sum_{\substack{(s,d)\in\\ DS\cap OR^{na}}}\delta_{DS_{sd}}-\sum_{\substack{(s,d)\in\\ DS\cap TE^{na}}}\delta_{DS_{sd}}+\sum_{\substack{(s,m)\in\\ M\cap OR^{na}}}\delta_{M_{sm}}\\&\quad-\sum_{\substack{(s,m)\in\\ M\cap TE^{na}}}\delta_{M_{sm}}+\sum_{s\in S^a}\sum_{\substack{(s,t)\in SW^s\\ t\in OR^{na}}}\delta_{SW_{st}}\\&\quad-\sum_{s\in S^a}\sum_{\substack{(s,t)\in SW^s\\ t\in TE^{na}}}\delta_{SW_{st}}+\sum_{\substack{g\in\\ G\cap OR^{na}}}\delta_{G_{ga}}-\sum_{\substack{g\in\\ G\cap TE^{na}}}\delta_{G_{ga}}\\&\quad=P_{B_n a}\qquad\forall a\in A,\forall n\in N^a\end{aligned}\tag{3d}$$

$$\sum_{r\in R^a_{\text{orig}}}\left(\delta_{R_{ra}}+\delta_{CR_r}\right)+\sum_{\substack{(r,d)\\ \in DR^r}}\delta_{DR_{rd}}\geq\left|R^a_{\text{orig}}\right|\left(1-z_{\Gamma_a}\right)\qquad\forall a\in A\tag{3e}$$

$$\delta_{SW_{st}}=\delta_{SW_{ts}}\qquad\forall(s,t)\in SW\tag{3f}$$

$$\delta_{R_{ra}}\in\{0,1\}\qquad\forall r\in R,\forall A\in A^r\tag{3g}$$

$$\delta_{DR_{rda}}\in\{0,1\}\qquad\forall(r,d)\in DR,\forall A\in A^d r\tag{3h}$$

$$\delta_{CR_r}\in\{0,1\}\qquad\forall r\in R\tag{3i}$$

$$\delta_{S_s}\in\{0,1\}\qquad\forall s\in S\tag{3j}$$

$$\delta_{DS_{sd}}\in\{0,1\}\qquad\forall(s,d)\in DS\tag{3k}$$

$$\delta_{CS_s}\in\{0,1\}\qquad\forall s\in S\tag{3l}$$

$$\delta_{SW_{st}}\in\{0,1\}\qquad\forall(s,t)\in SW\tag{3m}$$

$$\delta_{G_{ga}}\in\{0,1\}\qquad\forall a\in A,\forall g\in G^a\tag{3n}$$

$$\delta_{M_{sm}}\in\{0,1\}\qquad\forall(s,m)\in M\tag{3o}$$

$$z_{\Gamma_a}\in\{0,1\}\qquad\forall a\in A\tag{3p}$$

5. Case study

An implementation of ANEMOS was developed in collaboration with a major European airline, which provided the historical data necessary to define all input parameters and distributions described in this section. The proposed case study investigates the effects of adding a reserve aircraft to the partner airline's fleet in different operational disruption scenarios. The fleets considered in the implementation include 4 different aircraft models.

5.1. Maintenance scheduling submodule (MSS)

The MSS is called once a week, and its scheduling window goes from 3 days to three weeks after its call day. The weights of the parameters of the MSS are determined following the logic in Section 4.1. The values to be given to each parameter of the objective function are determined using a hierarchical logic (van Kessel et al., 2022), which involves defining a hierarchy of importance associated with each scheduling decision. The hierarchy is defined as follows, from highest to lowest importance: (1) assignment of MH slots, (2) assignment of maintenance tasks, (3) maintaining the assignment of slots in the fixed scheduling window unchanged, (4) reducing the use of ground time, i.e. avoiding the unnecessary activation of maintenance slots, (5) scheduling tasks with the preferred anticipation (see Table 4).

Requirements are limited to those having an interval between 15 days and three months, as this is the interval of A-checks for the considered fleets. The simulation of the arrival of DDs requires the

Table 3
Mathematical formulation of the recovery planner.

Sets and subsets		
A		Aircraft
R		Rotations
DR		Delayed rotations. Pair (r,d) denotes rotation r being delayed by d time units
S		Maintenance slots (Flex & MH slots)
DS		Delayed slots. Pair (s,d) denotes maintenance slot s being delayed by d time units
SW		Aircraft swap. Ordered pair (s,t) denotes the feasible swap of maintenance slots s and t
M		Free maintenance arcs. Pair (s,m) denotes the postponement of slot s to free maintenance arc m
G N		Ground arcs Nodes
OR^{na}		Arcs originating in node n that interest aircraft a
TE^{na}		Arcs terminating in node n that interest aircraft a
A^r	$\subseteq A$	Aircraft a that can be assigned rotation r
A^{rd}	$\subseteq A$	Aircraft a can be assigned delayed rotation (r,d)
DR^r	$\subseteq DR$	Delayed rotation derived from rotation r
R^a_{orig}	$\subseteq R$	Rotations currently assigned to aircraft a
S^a	$\subseteq S$	Maintenance slots of aircraft a
DS^s	$\subseteq DS$	Delayed maintenance slots derived from slot s
SW^s	$\subseteq SW$	Ordered pairs of maintenance slots that can be swapped where the first slot is s
M^s	$\subseteq M$	Free maintenance arcs may be used for postponing slot s
G^a	$\subseteq G$	Ground arcs tmay used by aircraft a
N^a	$\subseteq N$	Nodes that interest aircraft a
Decision variables		
$\delta_{R_{ra}}$	$\in \{0,1\}$	1 if rotation r is assigned to aircraft a, 0 otherwise
$\delta_{DR_{rda}}$	$\in \{0,1\}$	1 if delayed rotation (r,d) is assigned to aircraft a, 0 otherwise
δ_{CR_r}	$\in \{0,1\}$	1 if rotation r is cancelled, 0 otherwise
δ_{S_s}	$\in \{0,1\}$	1 if slot s is kept active, 0 otherwise
$\delta_{DS_{sd}}$	$\in \{0,1\}$	1 if delayed slot (s,d), 0 otherwise
δ_{CS_s}	$\in \{0,1\}$	1 if slot s is cancelled, 0 otherwise
$\delta_{SW_{st}}$	$\in \{0,1\}$	1 if slot s and t are swapped, 0 otherwise
$\delta_{G_{ga}}$	$\in \{0,1\}$	1 if aircraft a uses ground arc g, 0 otherwise

(*continued on next page*)

Table 3 (*continued*).

$\delta_{M_{sm}}$	$\in \{0,1\}$	1 if slot s is moved to flexible maintenance arc m, 0 otherwise
z_{Γ_a}	$\in \{0,1\}$	1 if any rotation initially assigned to aircraft a is reassigned, 0 otherwise
Parameters		
$W_{R_{ra}}$		Cost of assigning rotation r to aircraft a
$W_{DR_{rda}}$		Cost of delaying rotation r by d time units and assigning it to aircraft a
W_{CR_r}		Cost of cancelling rotation r
W_{S_s}		Cost of keeping slot s active
$W_{DS_{sd}}$		Cost of delaying slot s by d time units
W_{CS_s}		Cost of cancelling slot s
$W_{SW_{st}}$		Cost of swapping slot s and t
$W_{M_{sm}}$		Cost of postponing slot s to flex maintenance arc m
$W_{G_{ga}}$		Cost of aircraft a using ground arc g
W_{Γ_a}		Cost of changing rotation assignment of aircraft a
$P_{B_{na}}$		Node balance at node n of aircraft a. Equal to 1 if n is an origin node, to -1 if n is a termination node, and to 0 if n is a central node.

Table 4
Decision variables for the MSS.

Decision variables		
C_{MH}	-5×10^{12}	Activating an MH slot.
W_{U_t}	5×10^7 10^7	Unassigning task t, if t is a dd Unassigning task t, if t is a requirement
C_{fix_s}	10^6	Changing the assignment of a slot s in the fixed scheduling window. 0 if s is not in the fixed window.
C_{S_s}	10^5 10^4	Activating flex slot s, if s is a hangar slot Activating flex slot s, if s is a platform slot
C_{sd}	10^3	Additional cost per hour of slot duration, when a slot is activated
C_{ant_t}	10^2 -4×10^1	Anticipating the execution of requirement t with respect to its due date by one day Anticipating the execution of dd t by one day

definition of the limits of two weighted choices. In the case at hand, the value of some objective function weights can be defined by one component, such as the cost of leaving a task unassigned (W_{U_t}). In other cases, the value of these weights derives from the combination of multiple cost components, such as in the case of the weight of assigning a slot to an aircraft, which depends both on slot duration and location. For the latter cases, some *bridging costs* (C) are defined in the hierarchy and later assembled into the final objective function weights (W). The hierarchy is translated as follows:

The cost W_{U_t} is higher for DDs, since the scheduling of requirements is less critical. The remainder of the weights, which are compound weights, can now be defined. $W_{S_{sa}}$ can be defined as:

$$W_{S_{sa}} = \begin{cases} C_{MH} + C_{fix_s}, & \text{if } s \text{ is a MH slot} \\ C_{S_s} + C_{sd} \cdot d_s + C_{fix_s}, & \text{if } s \text{ is a flex slot,} \end{cases} \tag{4}$$

where d_s is the duration of slot s. The cost of scheduling a task t in maintenance slot s, $W_{T_{ts}}$, depends on the number of anticipation days with which the task would be executed. Calling this anticipation, a_{ts}, the parameter can be determined as:

$$W_{T_{ts}} = C_{ant_t} \cdot a_{ts} \tag{5}$$

5.2. Tail assignment submodule (TAS)

The recovery window of the TAS goes from three days to two weeks after the call of the TAS. Its parameters are determined using the same

Table 5
Decision variables for the TAS.

W_{U_r}	10^5	Unassigning segment r, if r is a reserve slot
	10^4	Unassigning segment r, if r is a rotation
$W_{R_{ra}}$	10^2	Assigning rotation r to aircraft a, if r is a rotation and a is not of its originally assigned subtype. 0 if r s a reserve slot or a is of the preferred subtype.

hierarchical logic used for the MSS. The hierarchy is defined as follows: (1) assignment of segments, (2) assignment of rotations to the preferred aircraft type. Subtype preference groups are introduced to determine which aircraft subtype should be used in the case of the unavailability of the originally assigned subtype. The following groups are defined:

- Group 1 (lower passenger capacity, range): Types I, II
- Group 2 (higher passenger capacity, range): Types III, IV

Rotations assignment is not allowed outside of the preference group, meaning that for example a rotation originally scheduled for a Type I cannot be assigned to a Type IV. This is done by reducing the model subsets. The hierarchy is translated into the following weights:

5.3. Operations manager

This model assumes flight time to be fixed for each route. Primary delays are assumed to be positive, meaning that a flight cannot depart before its scheduled departure time. These delays can be described by a simple probability of experiencing a delay, and by an analytical distribution limited at values greater than zero describing the duration of the delay. These delays are assumed to be independent of the fleet type, but dependent on the station at which they occur. The same simple probability and analytical distribution are used for all outstations due to limited data availability. Four other sets of parameters are used for the hub, one for each of the considered hub disruption states (see Table 5).

It is a challenge to isolate the duration of a turnaround in historical data. In order to make a distinction, only flights that departed with a delay due to IATA delay code 93, i.e. propagated delay, are considered in the analysis. For these flights, it is assumed that the time needed for turnaround operations corresponds to the ground time between arrival and departure of the limiting flights. Turnaround activities are assumed to be independent of the aircraft type. Three empirical distributions are obtained: one for turnaround time at the hub, and two for turnaround time at outstations. Two separate turnaround time distributions are used for outstations because, at certain airports, only shorter technical turnaround activities are executed, as suggested by the bimodal empirical probability density function of the full dataset. Outstations are categorized into short or regular turnaround stations based on the mean registered turnaround time.

The towing time is only considered when an aircraft needs to undergo maintenance in the hangar. Towing time to and from the hangar is fixed to one hour.

The duration of a maintenance slot depends on its work package and on the NR labor originating from findings happening during the slot. It is assumed that NRs can only happen within slots executed in the hangar. The probability of NRs coming up in a hangar work package is determined from historical data. The total NR labor hours executed within the historical maintenance slots are also computed and used to fit an analytical curve for each aircraft subtype. At the beginning of each hangar slot, the historical probability of experiencing non-routines is used to determine whether there will be some findings in the work package or not. If there should be some findings, the total required NR labor is sampled from the reference distribution and added to the work package. Once the NR labor hours are sampled, the duration of the maintenance slot can be computed as the maximum of the following two values:

- The maximum duration of the tasks included in the slot's work package.
- The sum of the labor hours associated with the tasks included in the work package (including non-routine labor) is divided by the available workforce in the maintenance slot.

The AOG inter-arrival time and duration are defined by stochastic distributions fitted on historical data by minimization of the RSS. Different distributions are obtained for each aircraft type.

The hub disruption process is characterized by four disruption states, and time is discretized in 20-min brackets. The process is initialized every day at 6:00 UTC. Both the exponential distribution describing the sojourn time in a state and the transition probability matrix are derived from historical data. To determine the duration of the disruption state, historical flights are grouped in twenty-minute brackets based on their actual departure time. Each bracket is then characterized by a disruption state based on the mean departure delay observed within that bracket. To do so, arbitrarily defined minimum and maximum mean delays are defined for each state. Adjacent brackets characterized by the same state are then counted, and this measure of sojourn time is used to determine the best-fitting exponential distribution by RSS minimization. The transition probability matrix can easily be determined by computing the empirical probability of transitioning from one state to another.

5.4. Recovery controller

The recovery controller requires a recovery action when it computes that the next duty of an aircraft will experience an increase in expected delay of at least ten minutes.

5.5. Recovery planner

The Recovery Planner works on a recovery window that covers three and a half days. At the end of the recovery window operations must be resumed as originally scheduled, and for this reason, duties are included in the recovery space if their *arrival* time falls within the recovery window. In order to better resemble our partner airline's operations, the possibility of postponing maintenance to free maintenance arcs is excluded from the solution of the Recovery Planner. Also, with the objective of reducing passenger disruption on the day of operation, it is imposed that the designated reserve aircraft should always be available at the start of each day. As a consequence, the Recovery Planner cannot assign a rotation to an aircraft if this overlaps with a reserve slot assigned to the aircraft over the coming days. Finally, rotations are allowed to be reassigned to aircraft of any subtype, and the preference group logic described for the TAS applies (see Table 6).

The parameters of the Recovery Planner are defined based on a hierarchy, this time formulated in terms of avoidance preference, from the recovery action that should be avoided the most, to the recovery action that is considered most acceptable. The following order is defined: (1) cancelling maintenance slots, (2) cancelling rotations (3) changing the aircraft subtype assignment of a rotation (4) swapping maintenance slots, (5) changing the aircraft assignment of a rotation, (6) delaying the start time of a maintenance slot, and (7) delaying the start time of a rotation. The following costs are derived from the hierarchy:

The cost of assigning a rotation to an aircraft can be defined as follows:

$$W_{R_{ra}} = C_{\text{type}_{ra}} + C_{\Delta ra} \tag{6}$$

Concerning delaying duties, copies of the original rotation arcs are generated with a delay of 5, 10, 20, 40, 60, 120, 180, and 240 min. For maintenance slots, copies are created with a delay of 5, 10, 20, 40, 60, 120, and 180 min. The chosen values are denser for shorter delays because these values are more commonly observed, and because they allow the avoidance of more drastic recovery interventions that are

Table 6
Decision variables for the recovery planner.

W_{CS_s}	3×10^6 10^6	Cancelling slot s, if s if an MH slot Cancelling slot s, if s if a flex slot
W_{CR_r}	10^5	Cancelling a rotation
$C_{\text{type}_{ra}}$	2×10^4 10^4	Assigning rotation r to aircraft a, if aircraft a's subtype is not included in r's preference group Assigning rotation r to aircraft a, if aircraft a's subtype is within to rotation r's preference group. 0 if a is of the originally assigned subtype
$W_{SW_{st}}$	10^3	Swapping two maintenance slots
W_{Γ_a}	2×10^2	Involving aircraft a in the recovery solution
$C_{\Delta_{ra}}$	10^2	Assigning rotation r to aircraft a, if the assignment is different from what was previously planned
C_{DS}	2.5×10^1	Delaying a slot by one minute
C_{DR}	2×10^1	Delaying a rotation by one minute

often required for significant delays that are in the range of the hours. Also, the maximum delay allowed for a maintenance slot is shorter than that allowed for a rotation given the lower resource flexibility associated with a maintenance slot in terms of, for instance, manpower and hangar space. For the same reason, delaying a maintenance slot is more expensive than delaying a rotation. Calling d_d the delay imposed on a slot or rotation, the weights of assigning delayed duty arcs are defined as:

$$\begin{aligned} W_{DS_{sd}} &= C_{DS} \cdot d_d \\ W_{DR_{rda}} &= W_{R_{ra}} + C_{DR} \cdot d_d \end{aligned} \tag{7}$$

The costs of using a maintenance slot as planned (W_{S_s}) and of using a ground arc ($W_{G_{ga}}$) are set to zero.

5.6. Scenarios generation

Scenarios consider a fleet of 50 aircraft (or 51 when a reserve is added) including the already cited four aircraft subtypes. The simulated schedule is a weekly schedule that includes 263 rotations. The simulated maintenance slots are the slots that were historically available during that week. Each scenario is simulated a hundred times over 180 days. Two variables model this case study: the number of reserve aircraft available (i.e. one or two) and the distribution describing the duration of AOGs. In order to generate different scenarios to evaluate, different distributions describing AOG duration must be chosen. In the baseline scenario, AOGs for all aircraft types are modeled as having a duration described by a lognormal distribution. Using the generally accepted parametrization of the lognormal distribution in μ and σ, the expected value is as follows:

$$E[X] = e^{\mu + \frac{1}{2}\sigma^2} \tag{8}$$

Given this property, new scenarios are generated by multiplying the scale parameter e^{μ} by a constant, so that the new distributions are characterized by an expected value that is equal to the original expected value multiplied by the same constant. The multiplicative constant, which will be called the AOG duration factor, is set to 1, 1.2, 1.4, 1.6, 1.8, and 2.

6. Results

Here, we present the results of ANEMOS applied to the case study described in Section 5. The results are validated in Section 6.1. Additionally, Section 6.2 xplores a special case involving the inclusion of reserve aircraft, demonstrating how ANEMOS can be utilized to evaluate different operational policies.

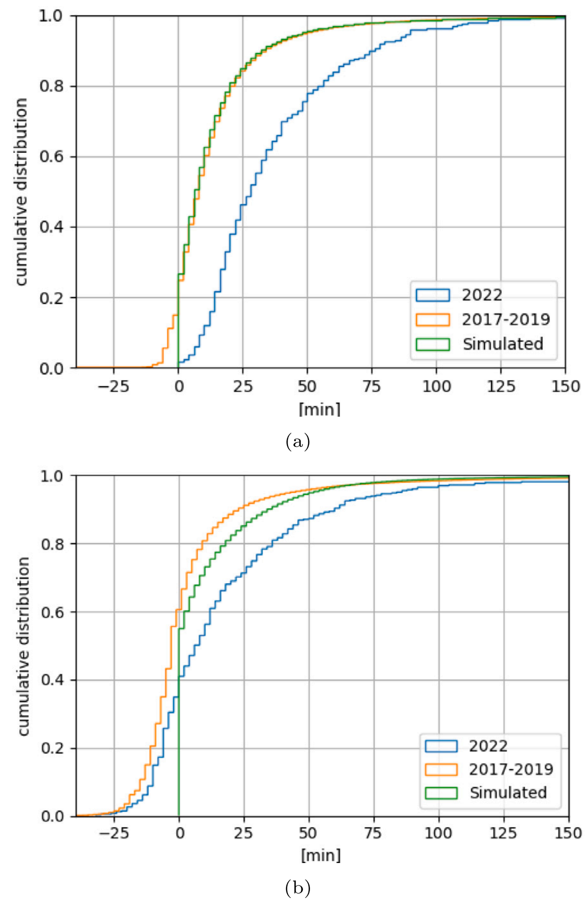

Fig. 6. Empirical CDFs of the departure delays of simulated and historical flights departing from the hub 6(a) and from outstations 6(b). Two historical curves are shown. The 2017–2019 curve corresponds to data used for building the delay model, while the 2022 curve refers to flights included in the simulated schedule, corresponding to a specific week of operations.

6.1. Model validation

The model is validated by comparing the results obtained for the baseline case of the case study (AOG duration factor of 1, 1 reserve aircraft), with historically operational performance comprises data on flights flown between 2017 and 2019. However, the proposed case study refers to a schedule and maintenance plan implemented in 2022. For this reason, the simulated delays are compared to historical data of two sets of flights: the flights flown between 2017 and 2019, and the flights flown during the week of the simulated schedule in 2022.

Figs. 6(a) and 6(b) show the empirical cumulative distribution functions (CDF) of the departure delay of flights departing from the hub and from outstations, respectively. The curve of the simulated data closely follows the 2017–2019 curve for departure delays from the hub, with the sole difference that there are no departures before the scheduled departure time since they are not allowed by the model. When considering departure delays from outstation, a similar trend is observed between simulated and 2017–2019 flights, although simulated departure delays tend to be higher than 2017–2019 ones. This is due to the assumption of fixed flight leg duration, which in some routes leads to a systematic accumulation of delay that propagates to the following flight legs.

A comparison of the simulated delays and delays from 2022 shows a significant underestimation of delays in the simulated results. This can be explained by the widespread ground personnel shortages in the historical period from 2022, leading to strong disruptions in numerous airports worldwide. This result, however, does not invalidate the proposed case study, since the case study focuses more on flight cancellations rather than delays, and historical data of our partner airline show that no cancellation was caused by delay propagation in 2022.

The second validated performance measure regards cancellations, for which the historical value to be used as a validating comparison should be discussed. As already mentioned, delay propagation does not cause any cancellations thanks to the buffer time scheduled at the hub for intercontinental operations — only historical cancellations of full rotations are considered. Several reasons for this underestimation of cancellations can be given. First, the model generally has more flexibility in recovery than it is available in reality. Second, simulated A-checks have more operational flexibility than in reality. Third, the scope of maintenance limited to A-checks excludes heavy maintenance which is more likely to exceed the scheduled time, causing disruptions.

The model is capable of executing between 99.3% and 99.5% of tasks for all aircraft subtypes apart from the Type II, for which an execution rate of 96.1% is simulated. The tasks that the model is incapable of scheduling are generally tasks that are longer than the available maintenance slots, or tasks for which a maintenance opportunity cannot be made available due to the short time between the task's ready date and due date. In the case of Type II, the reduced execution rate is due to the inclusion of recurring tasks that, in real-life operations, are executed in specifically designated maintenance slots that are longer than the ones included in the scope of the simulation.

The scheduling logic used by the simulator is validated by considering the *tasks' relative anticipation*, defined as the ratio between the number of days intercurring between a task execution date and due date and the number of days intercurring between the task arrival date and due date. Fig. 7 shows the empirical CDF of the relative task anticipation of the simulated and historical tasks. The simulated requirements in Fig. 7(a) follow the trend of historical requirements, as they tend to be executed close to their due date to minimize the requirements' lost interval. However, they are generally executed closer to their due date than in reality, which can be explained by setting the preferred anticipation for requirements execution of the MSS to zero.

Fig. 7(b) shows that DDs follow the correct trend of being executed early after their finding. When compared to the full dataset of historical DDs, a general postponement of simulated task execution is observed. This is because historically, many DDs are generated from crew complaints registered within the Aircraft Maintenance Log (AML), i.e., a book located on board each aircraft that can be used to report any Minimum Equipment List (MEL) problem detected on the aircraft during operations. When an aircraft undergoes maintenance, the AML is checked and the included tasks are often executed on the same day on which they are found, without needing to be scheduled. If these tasks are excluded from the historical dataset, the second historical curve shown in Fig. 7(b) is obtained (historically reduced), which closely resembles the curve of simulated tasks.

6.2. Adding reserve aircraft

Four indicators are used for the evaluation of the results of this case: (1) the cancellation factor (CF), i.e. the percentage of rotations that are cancelled, (2) the arrival delays, (3) the costs of disruptions, and (4) the avoided disruption costs when a second reserve is used.

The results for the cancellation factor for all scenarios are shown in Fig. 8 along with the 95% confidence interval computed using the bootstrap technique. The increase in the AOG average duration causes a significant increase in the CF, with the number of cancellations almost doubling for an AOG duration factor of 1.4 and becoming more than triple for an AOG duration factor of 2. Adding a second reserve aircraft, on the other hand, allows for a reduction of the expected cancellations. At the baseline, the second reserve halves the number of cancellations, bringing the CF from 0.11% to 0.05%. As the value of the AOG duration factor increases, the impact in terms of the number of cancellations

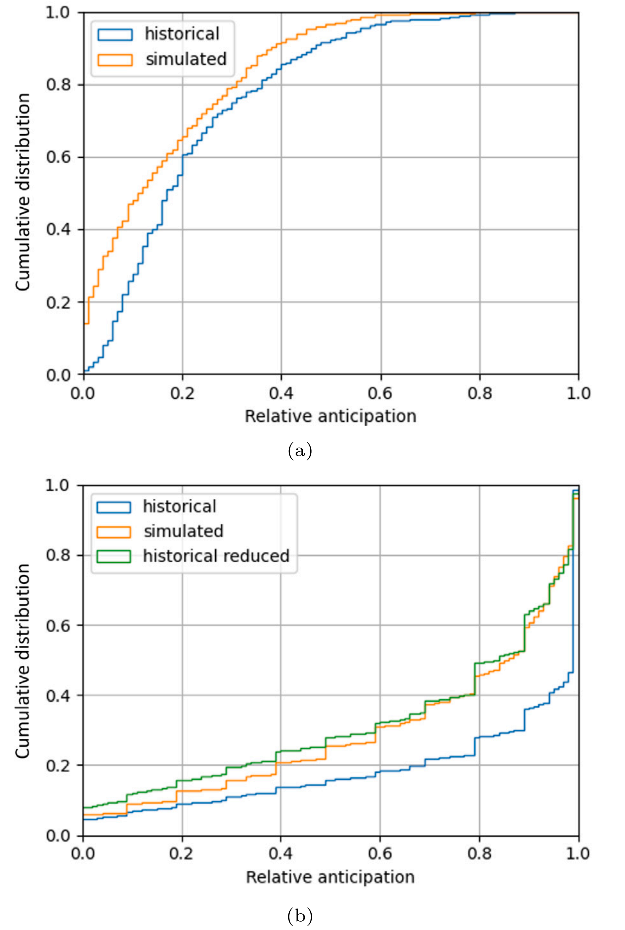

Fig. 7. Empirical CDFs of the relative task anticipation of simulated and historical requirements 7(a) and deferred defects 7(b). For DDs, a third curve showing the relative anticipation of historical tasks not found and executed on the same day is shown.

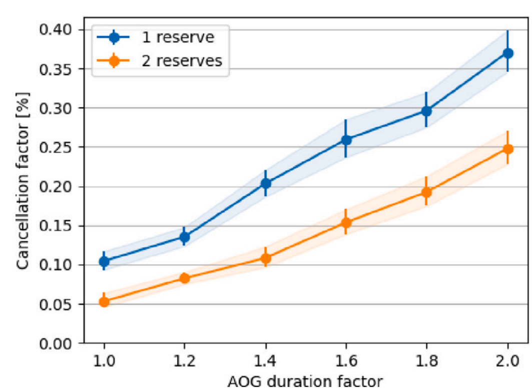

Fig. 8. Cancellation factor for different AOG duration factor and reserve aircraft scenarios.

that the reserve can avoid increases, and then it stabilizes at around 0.10%.

Fig. 9(a) shows the empirical exceedance probability curve of departure delays, which describes the probability of observing a delay greater than a specified value. The results shown are obtained with an AOG duration factor of 1, but other scenarios show a similar impact on the second reserve aircraft. A detail of delays between one and four hours is shown in Fig. 9(b). The use of a second reserve aircraft reduces the probability of observing a delay longer than one hour by 0.4% (from

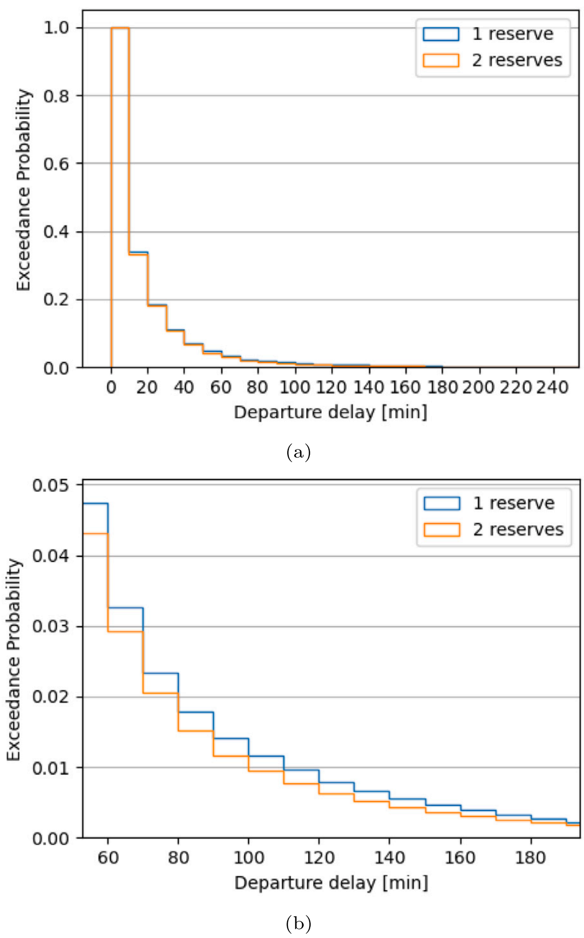

Fig. 9. Exceedance probability curve of departure delays, for an AOG duration factor of 1: full plot 9(a) and detail 9(b).

3.3% to 2.9%) and the probability of observing a delay longer than two hours by 0.1% (from 0.8% to 0.7%).

The economic impact is computed considering the costs of cancellations and delays, including costs associated with European regulations on passenger compensation and soft costs related to passenger satisfaction. The computation of delay costs disregards the cost of passengers' lost connections. Fig. 10 displays the disruption costs. The two columns on the right show the cost components, and the column on the left shows the total costs obtained from their summation. The effects of an increased AOG duration factor are significant, with a cost increase of +30% for an AOG duration factor of 2. The primary contributors to this cost increase are cancellation costs. This can be explained by the high costs associated with the cancellation of a flight, rather than with its delay, especially considering that cancellation costs, differently from delay costs, account for passenger misconnections.

Fig. 11 shows the avoided disruption costs, i.e. the difference between the costs incurred with one and two reserves, are computed. For low values of AOG duration factors, the impact of the second reserve aircraft is comparable for delay and cancellation costs. For higher values of the AOG duration factor, the avoided costs of delay remain stable, while the avoided cancellation costs increase significantly. This can be explained by the fact that as AOGs become longer, the probability of experiencing disruptions that could lead to both delays and cancellations increases, but it is more cost-efficient to use the reserve aircraft to avoid cancellations, rather than delays. Furthermore, the much higher costs associated with cancellations with respect to delays, lead to a higher impact on cancellation costs, rather than delay costs, when a comparable number of disruptions of the two types are avoided.

This result shows how using a reserve aircraft is an expensive measure, which turns out to be not economically advantageous, despite the benefits obtained in terms of operational performance.

7. Discussion

The integration of airline network and maintenance planning presents a significant challenge due to the inherent complexities and operational constraints faced by different departments. Traditional airline operations often suffer from inefficiencies stemming from a lack of coordination between network planning and maintenance scheduling, leading to suboptimal performance and increased operational disruptions. The proposed framework, ANEMOS, seeks to address these issues by providing an integrated approach to airline operations, enabling interdepartmental coordination to improve overall system performance. However, there are still limitations that must be addressed towards a large acceptance of ANEMOS by both industry and academy. The following sections explore these limitations in greater detail. Section 7.1 discusses how ANEMOS can mitigate challenges and practical constraints faced by airlines. The current assumptions made in the model and their implications are outlined in Section 7.2. Finally, Section 7.3 explores potential directions for future research.

7.1. Practical implementation

ANEMOS implements a combined optimization approach that integrates network and maintenance planning. Such differs from current airline operations, where these decisions are typically managed by separate departments. The adoption of ANEMOS would therefore require closer interdepartmental collaboration, posing a challenge to its acceptance within the industry.

Designed as a decision-support tool, ANEMOS provides valuable insights by allowing users to adjust key input elements, such as fleet planning and maintenance requirements, and analyze different solutions while accounting for uncertainties. This capability offers significant benefits, even when network and maintenance planning are handled by different departments. The OCC, responsible for network planning, may use ANEMOS to better assign flights to specific aircraft while incorporating real-time information from the maintenance department regarding available maintenance slots and outstanding maintenance requirements. In turn, the maintenance department, can leverage ANEMOS both to get insights on optimal maintenance planning to decrease wasted interval, and to evaluate the effects of non-routine labor on the overall maintenance scheduling.

Ultimately, ANEMOS showcases the potential for optimization through a holistic approach to network and maintenance planning. In practice, implementing such an approach may be challenging due to the complexity of human resource coordination across operations. However, ANEMOS can facilitate negotiations between network planning and maintenance departments by providing a shared simulation platform, it enables the comparison of various constraints, decisions, and uncertainties within a unified model.

7.2. Assumptions

For simplification, ANEMOS assumes that several events are independent. First, non-routine findings are considered independent of the number, type, or duration of scheduled maintenance tasks. Recent work by Li et al. (2024) has shown that this is not the case. In maintenance planning, ANEMOS also assumes that materials are always available and that future labor man-hours are known and constant. Additionally, ANEMOS does not account for reactionary delays. However, recent data indicates that reactionary delay is the major contributor to the average departure delay per flight (Walker, 2022). Finally, turnaround activities are estimated based on the average of previous historical values. Nevertheless, studies have showed that the turnaround times

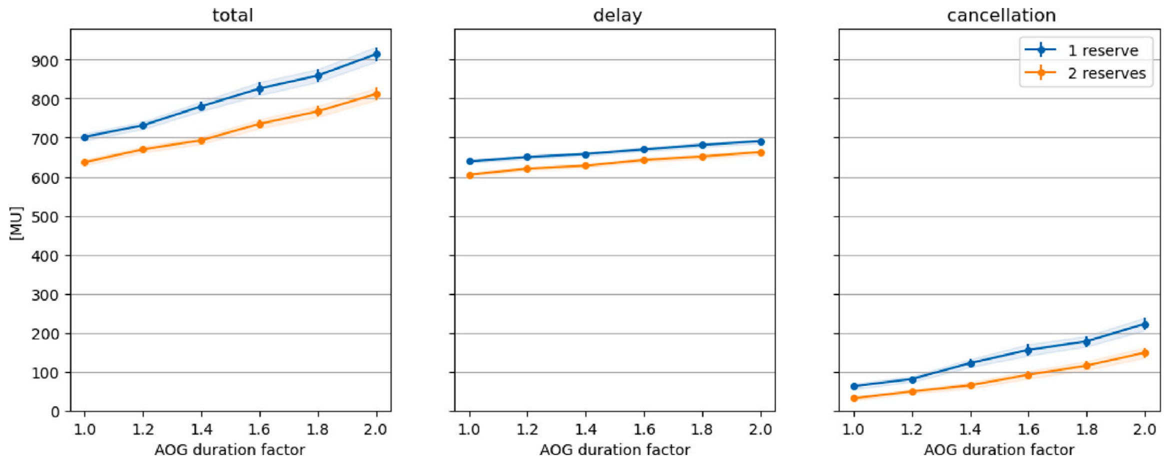

Fig. 10. Delay, cancellation, and total costs of disruptions for different AOG duration factors and reserve aircraft scenarios.

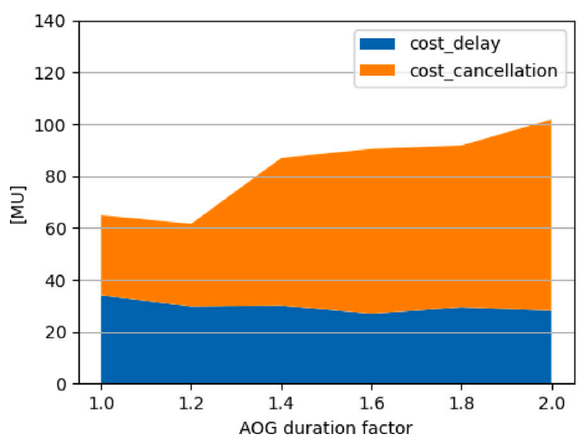

Fig. 11. Disruption costs are avoided through the use of a second reserve aircraft for different AOG duration factors.

vary significantly, particularly at larger airports, during peak hours, and when aircraft operate a high number of flights (Malighetti et al., 2023).

All these assumptions may limit the viability of the solutions presented by ANEMOS. However, accurately forecasting these factors remains challenging due to limited historical data and the large number of causal variables involved. In the future, integrating ANEMOS with predictive models could enhance its solutions by providing more reliable estimates for these elements. These external models could serve as valuable inputs, offering improved predictions for non-routine labor time, potential delays, and minimum turnaround times.

7.3. Future work

While the case study demonstrated the feasibility of the framework, it did not fully establish the extent to which ANEMOS can enhance interdepartmental coordination. Future studies should include empirical validation by analyzing real-world airline operations and assessing the improvements in decision-making effectiveness. Additionally, incorporating passenger and crew flows into the simulation would further enhance its applicability by providing a more comprehensive evaluation of operational performance. Finally, expanding the scope of the model to include multiple hubs and point-to-point carriers, allowing broader application across different airline operational structures. By addressing these aspects, ANEMOS can serve as a robust decision-support tool that enhances airline efficiency, improves interdepartmental collaboration, and mitigates the operational challenges that airlines frequently encounter.

8. Conclusions

This paper presented a stochastic discrete event simulation model of airline operations named ANEMOS (Airline Network and Maintenance Operations Simulation). This is the first approach in industry and research to combine network planning and maintenance in one single approach, considering several sources of uncertainty that airlines face in their operations. The capabilities of ANEMOS were validated through a case study developed in collaboration with a major European carrier, investigating the effects of using a second reserve aircraft for the simulated fleet. Direct comparison with historical data shows that the model closely resembles historically observed operational performance.

For future development, the model will be extended to allow the simulation of multiple hubs and point-to-point carriers. Secondly, considering passenger and crew flows would allow for testing a wider range of plans such as crew rosters, leading to a better quantification of airline performance in terms of passenger misconnections. Finally, widening the scope of the simulated maintenance slots to include maintenance heavier than A-checks would allow a better evaluation of maintenance-flights operations interaction.

CRediT authorship contribution statement

Sara Varenna: Writing – original draft, Validation, Methodology, Formal analysis, Conceptualization. **Haonan Li:** Writing – review & editing. **Marta Ribeiro:** Writing – review & editing. **Bruno F. Santos:** Writing – review & editing, Supervision, Methodology, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

Aloulou, M.A., Haouari, M., Zeghal Mansour, F., 2010. Robust aircraft routing and flight retiming. Electron. Notes Discrete Math. 36 (C), 367–374.

Barnhart, C., Kniker, T.S., Lohatepanont, M., 2002. Itinerary-based airline fleet assignment. Transp. Sci. 36 (2), 199–217.

Ben Ahmed, M., Ghroubi, W., Haouari, M., Sherali, H.D., 2017. A hybrid optimization-simulation approach for robust weekly aircraft routing and retiming. Transp. Res. Part C: Emerg. Technol. 84, 1–20.

Duffuaa, S.O., Andijani, A.A., 1999. An integrated simulation model for effective planning of maintenance operations for Saudi Arabian Airlines (SAUDIA). Prod. Plan. Control 10 (6), 579–584.

Geske, A.M., Herold, D.M., Kummer, S., 2024. Integrating AI support into a framework for collaborative decision-making (CDM) for airline disruption management. J. the Air Transp. Res. Soc. 3, 100026.

Iwata, C., Mavris, D., 2013. Object-oriented discrete event simulation modeling environment for aerospace vehicle maintenance and logistics process. In: Procedia Computer Science. Vol. 16, Elsevier B.V., pp. 187–196.

Jacobs, P.H., Verbraeck, A., Mulder, J.B., 2005. Flight scheduling at KLM. In: Proceedings - Winter Simulation Conference. Vol. 2005, pp. 299–306.

van Kessel, P.J., Freeman, F.C., Santos, B.F., 2022. Airline maintenance task rescheduling in a disruptive environment. European J. Oper. Res..

Lan, S., Clarke, J.P., Barnhart, C., 2006. Planning for robust airline operations: Optimizing aircraft routings and flight departure times to minimize passenger disruptions. Transp. Sci. 40 (1), 15–28.

Lee, L.H., Huei Chuen Huang, Chulung Lee, Ek Peng Chew, Jaruphongsa, W., Yean Yik Yong, Zhe Liang, Chun How Leong, Yen Ping Tan, Namburi, K., Johnson, E., Banks, J., 2003. Discrete event simulation model for airline operations: SIMAIR. In: Proceedings of the 2003 International Conference on Machine Learning and Cybernetics (IEEE Cat. No.03EX693). IEEE, pp. 1656–1662.

Levin, A., 1971. Scheduling and Fleet Routing Models for Transportation Systems. Transp. Sci. 5 (3), 232–255.

Li, H., Ribeiro, M., Santos, B., Tseremoglou, I., 2024. Prediction of non-routine tasks workload for aircraft maintenance with supervised learning. In: AIAA SCITECH 2024 Forum. American Institute of Aeronautics and Astronautics.

Malighetti, P., Morlotti, C., Redondi, R., Paleari, S., 2023. The turnaround tactic and on-time performance: Implications for airlines' efficiency. Res. Transp. Bus. Manag. 46, 100874.

Nisse, N., Salch, A., Weber, V., 2023. Recovery of disrupted airline operations using k-maximum matching in graphs. European J. Oper. Res. 309 (3), 1061–1072.

Öhman, M., Hiltunen, M., Virtanen, .K., Holmström, J., 2020. Frontlog scheduling in aircraft line maintenance: From explorative solution design to theoretical insight into buffer management.

Pohya, A.A., Wehrspohn, J., Meissner, R., Wicke, K., 2021. A modular framework for the life cycle based evaluation of aircraft technologies, maintenance strategies, and operational decision making using discrete event simulation. Aerospace 8 (7).

Rosenberger, J.M., Johnson, E.L., Nemhauser, G.L., 2004. A robust fleet-assignment model with hub isolation and short cycles. Transp. Sci. 38 (3), 357–368.

Rosenberger, J.M., Schaefer, A.J., Goldsman, D., Johnson, E.L., Kleywegt, A.J., Nemhauser, G.L., 2002. A stochastic model of airline operations. Transp. Sci. 36 (4), 357–377.

Tseremoglou, I., Santos, B.F., 2024. Condition-based maintenance scheduling of an aircraft fleet under partial observability: A deep reinforcement learning approach. Reliab. Eng. Syst. Saf. 241, 109582.

van Schilt, I.M., van Kalker, J., Lefter, I., Kwakkel, J.H., Verbraeck, A., 2024. Buffer scheduling for improving on-time performance and connectivity with a multi-objective simulation–optimization model: A proof of concept for the airline industry. J. Air Transp. Manag. 115, 102547.

Villafranca, M., Delgado, F., Klapp, M., 2025. Aircraft maintenance scheduling under uncertain task processing time. Transp. Res. Part E: Logist. Transp. Rev. 196, 104012.

Vink, J., Santos, B.F., Verhagen, W.J., Medeiros, I., Filho, R., 2020. Dynamic aircraft recovery problem - An operational decision support framework. Comput. Oper. Res. 117.

Vos, H.W.M., Santos, B.F., Omondi, T., 2015. Aircraft schedule recovery problem - A dynamic modeling framework for daily operations. In: Transportation Research Procedia. Vol. 10, Elsevier, pp. 931–940.

Walker, C., 2022. All-Causes Delays to Air Transport in Europe - Quarter 3 2022. Technical report, EUROCONTROL CODA Digest.

Xu, Y., Adler, N., Wandelt, S., Sun, X., 2024. Competitive integrated airline schedule design and fleet assignment. European J. Oper. Res. 314 (1), 32–50.

Zhang, Q., Chung, S.-H., Ma, H.-L., Sun, X., 2024. Robust aircraft maintenance routing with heterogeneous aircraft maintenance tasks. Transp. Res. Part C: Emerg. Technol. 160, 104518.