Reserve Fleet Optimisation

Assessing airline operational performance impact of capacity changes on past operations

Miguel Correia da Silva Guerra Nunes

Faculty of Aerospace Engineering

TU Delft

Reserve Fleet Optimisation

Assessing airline operational performance impact of capacity changes on past operations

by

Miguel Correia da Silva Guerra Nunes

to obtain the degree of Master of Science
at the Delft University of Technology,
to be defended publicly on Friday April 26, 2024.

Student number:	5523419	
Project duration:	February 2023 – April 2024	
Thesis committee:	Dr. A. Bombelli	TU Delft, Supervisor
	Ir. P.C. Roling	TU Delft, Chair
	Dr. F. Oliviero	TU Delft, Examiner
	M. Ansems	SWISS International Airlines
	Ir. D. Gougousis	SWISS International Airlines

Acknowledgements

This report constitutes the final deliverable for the Master of Science in Aerospace Engineering degree at the Delft University of Technology. I would like to use this section to acknowledge certain people to whom I am grateful for the support they have given me throughout the course of this graduation project.

I would like to start by thanking my supervisor Dr. Alessandro Bombelli, for the constructive and encouraging discussions throughout the project. Your guidance has played a pivotal role in shaping the outcome of this thesis.

At SWISS International Airlines I would like to thank Dimitrios Gougousis for always being available to answer my questions, discuss my doubts, and give me critical feedback. Additionally, I would like to thank Laura Braun, Harald Falkner, Mark Ansems, Oliver Ruegg and Lucas Weber for their valuable expertise and feedback at key stages of the project. Additionally, I would like to thank Dominique Heilmann for facilitating my progression from intern to MSc Thesis student and for supporting my choice of topic. Finally, I would like to thank all my former colleagues at SWISS for creating such a positive and stimulating work environment during my internship, and for sharing your knowledge with me. I am very grateful for having had the opportunity to confirm my passion for the aviation industry during both of these endeavors.

I would like to thank my friends for their cherished company throughout this project and the whole duration of my studies. Their laughter, support, valuable feedback, and comforting presence have been key in navigating the challenges and successes encountered along the way.

I'm immensely grateful to my girlfriend for her great encouragement and support. Thank you for being there for me and for being an essential part of my life, Joana.

Finally, I would like to thank my family. Though we are often far away from each other, you are always with me in spirit. To my parents, the culmination of my academic studies has been made possible by your continued mental and financial support. I am forever grateful for the incredible opportunities you have enabled me to have.

Miguel Correia da Silva Guerra Nunes
Évora, April 2024

Contents

List of Figures

List of Tables

List of Abbreviations

AOCC	Airline Operations Control Center
ARIMA	Autoregressive Integrated Moving Average
ATA	Actual Time of Arrival
ATD	Actual Time of Departure
ATFM	Air Traffic Flow Management
DOC	Direct Operating Cost
DQL	Double Q-Learning
FIFO	First In, First Out
FVNS	Scheduled Time of Departure
GBM	Gradient Boosting Regression Tree Method
KPI	Key Performance Indicator
LNS	Large Neighborhood Search
LSTM	Long Short-Term Memory
MAE	Mean Absolute Error
MAPE	Mean Absolute Percentage Error
MILP	Mixed Integer Linear Programming
MIPSSM	Mixed Integer Programming Simulation Scenario Model
MSC	Margin for Successful Connection
MSE	Mean Squared Error
NPS	NetPromoter Score
OR	Operating Room
QL	Q-Learning
RF	Random Forests
RMSE	Root Mean Squared Error
SCC	Shortest Cycle Cancellation
STA	Scheduled Time of Arrival
STD	Scheduled Time of Departure
SVM	Support Vector Machine
TSN	Time-Space Network

Introduction

During my master's degree at the TU Delft, I had a very rewarding internship at SWISS International Airlines´ Operations Research & Development department. Towards the end of that experience and as I was actively searching for a thesis topic, interest was expressed in a detailed study of the potential impact of making changes to the reserve fleet maintained during the airline´s Summer Schedule. This undertaking would provide a scientific basis for the yearly decision making process behind choosing the reserve fleet to be kept over the airline's entire Summer Schedule. Currently, this decision is made after a meeting between the Network Planning and Operations Departments, where the first advocates for extra planes to introduce in normal operations planning and scheduling and the second for extra aircraft to add to the reserve fleet to increase operational stability and customer satisfaction. However, as the basis for the decision over the past few years has been limited to the operational performance indicators achieved with the operated reserve fleet during past Summer Schedules, the decision has typically been to maintain the same fleet as the previous years´s.

Given these premises, the research objective that this work tries to reach can be formulated as follows: **To assess the airline operational performance impact of reserve capacity changes on past operations.**

This thesis report is organized as follows: In Part I, the scientific paper is presented, containing the main literature supporting this research, the problem statement, the solution approach used, and the results. Part II contains the Literature Study completed to initiate this project.

I

Scientific Paper

Reserve Fleet Optimisation: Assessing airline operational performance impact of capacity changes on past operations

Miguel Correia da Silva Guerra Nunes,*

Delft University of Technology, Delft, The Netherlands

Abstract

Airline reserve fleet capacity is a strategic resource that provides airlines the means to add robustness to their schedules. This paper focuses on leveraging past data on airlines' execution of planned schedules (resulting delays, cancellations and missed connections), along with the disruptions that prohibited the flawless execution of the planned schedules, in order to assess the impact of reserve fleet composition changes on past days of operations. The fleet schedule is modeled as a set of parallel time-space networks, and a Mixed Integer Linear Programming model is defined and solved with the objective of minimising passengers' disruption costs incurred during execution of the planned schedule. By adjusting the reserve fleet within the model, the resulting impact on disruption costs and operational performance indicators is quantified. The approach presented can be applied on heterogeneous fleets and to hub-and-spoke airlines. Its performance is discussed using a case study on the network of a Full Service Carrier operating in Europe. Simulated results of all reserve fleet scenarios considered are compared with recorded passenger disruption data from the same period, demonstrating the model's capacity to emulate historically observed operational performance. Notably, the model (considering the same reserve fleet as was operated) exhibits an average deviation of 7.9% in disruption costs when compared with recorded data. The analysis of resulting disruption costs stemming from reserve fleet composition changes reveals that adding an A220 aircraft to the airline's reserve fleet would result in a slightly higher reduction (-5.7%) of disruption costs compared to adding an A320 aircraft (-5.3%). Conversely, removing an A320 aircraft from the reserve fleet would lead to a significantly higher increase in disruption costs (+20.5%) compared to removing one A220 aircraft as reserve (+16.9%). Consequently, when considering changes to the reserve fleet operated, the airline should prioritise alterations on the A220 portion of their fleet.

1 Introduction

The airline industry operates within a complex framework that spans from long-term strategic planning to day-to-day operational execution. At the strategic level, airlines meticulously plan their fleet composition, route networks, and scheduling over extended time horizons, often spanning decades. However, the efficacy of these plans is continuously tested by the operational realities faced on a daily basis. Factors such as adverse weather conditions, congested airport hubs, and unexpected aircraft maintenance issues frequently disrupt the meticulously crafted flight schedules. These disruptions, ranging from flight cancellations to delays and missed connections, not only incur substantial expenses for airlines, but also disrupt passengers' travel plans, leading to dissatisfaction and potential financial liabilities.

In response to these challenges, Airline Operations Control Centers (AOCCs) play a pivotal role in mitigating disruptions by swiftly recalibrating flight, crew, and passenger itineraries to minimise operational costs and customer inconvenience. One potential strategy to mitigate the effects of disruptions is the utilisation of reserve fleet capacity. Although not typically factored into long-term fleet planning strategies, this reserve capacity serves as a crucial buffer, enabling airlines to reduce delays and swiftly replace inoperable aircraft, thereby maintaining operational continuity. However, the decision to maintain reserve capacity represents a significant investment, as it entails foregoing potential revenue from planned flights. Thus, striking a balance between long-term strategic planning and responsive operational management is imperative for airlines seeking to optimise efficiency while navigating the dynamic challenges of the industry.

In light of these considerations, there is a need for comprehensive research to evaluate the effectiveness and impact of reserve fleet composition changes on airlines' operational performance. By leveraging historical data on airlines' execution of planned schedules and the disruptions encountered during operations, this study aims to quantify the value proposition of varying levels of reserve fleet capacity on past operations. Through rigorous analytical modelling and simulation techniques, the research seeks to provide actionable insights for airlines seeking to optimise their fleet planning strategies and enhance operational resilience.

*MSc Student, Sustainable Air Transport, Faculty of Aerospace Engineering, Delft University of Technology

The remainder of the paper is structured as follows. First existing literature is reviewed in Section 2. Section 3 introduces the problem context. Section 4 describes the structure of the proposed framework, explaining each of its particularities in detail. Section 5 presents the case study that is used to demonstrate and validate the methodology as it could be applied in practice, along with the obtained results. Section 6 draws the conclusions of this work. Finally, Section 7 discusses the main limitations of the work, as well as relevant lines for further research to be conducted.

2 Literature Review

This research aims to obtain a comprehensive understanding of the impact that an airline's reserve fleet (and possible changes in its composition) has on its disruptions and operational Key Performance Indicators (KPIs). This section addresses modelling techniques frequently addressed in research for airline disruption management, as well as explores previous studies made resulting in the determination of the added value of reserve fleet capacity and in which way it can best be used.

2.1 Disruption Management

Airline disruption management has seen an increasing amount of attention since it was pioneered four decades ago by Teodorović and Guberinić (1984), due to its potential to ameliorate the service of airlines. The initial efforts focused on solving conflicts for a single resource at a time, as they are resolved in the AOCC, which is subdivided in different departments, each with a specific area of operations to monitor and to plan (e.g Network Operations Control, Crew Control, Passenger Care Center), where there are human specialists who focus on one resource and take action when necessary, coordinating with the other departments when needed (Castro et al., 2014). This is because it becomes too complex to make decisions regarding multiple resources at once, and so the industry standard is to solve the problem sequentially. While the sequential approach offers a quick resolution, it may not always yield the most optimal outcome due to the interdependence of resources. Although a schedule may be optimal for one resource, it could create conflicts for others. Consequently, recent research has concentrated on developing integrated decision-support tools that simultaneously recover aircraft, crew, and passenger itineraries or a combination of these three, optimising solution quality and reducing costs incurred by the airlines. However, current computer systems still demand significant computational time to solve these models, making it impractical to implement such systems for airlines in their AOCC, where a fleet-level solution is typically required within a tight timeframe of 2 minutes (Vink et al., 2020). Researchers and practitioners face a trade-off between the complexity of the models and the time required to find a solution, which is a key topic of discussion in the existing literature related to disruption management.

To address the trade-off between model complexity and solution time, modern research is focused on employing scope-limiting algorithms (frequently by means of machine learning techniques) that generate a subset of recovery options deemed likely to contain the optimal solution. While these algorithms can produce high-quality solutions, they are heuristic in nature and may not always guarantee the best outcome. However, as the purpose of this research is to study past days of operations and how the reserve fleet effected them and would have effected had there been a different composition of reserve fleet, computation time will not be a major limiting factor. The review will thus focus on the state-of-the-art modelling of the Airline Recovery Problem, without considering performance in terms of computing time.

Vos et al. (2015) were the first to propose incorporating parallel time-space networks (one per aircraft) to solve this problem. These networks represent the aircraft in a two-dimensional plane where one axis is time and the other are the airports considered, and allow for the modelling of cancellations, flights and delayed flights. The existence of one network per aircraft enables explicit modelling of individual aircraft and their maintenance constraints. Using Linear Programming the total disruption costs are minimised for one full day of operations. They further deviate from other researchers by introducing a dynamic framework for aircraft schedule recovery. Instead of assuming a static disruption scenario with full knowledge of all disruptions in advance, this framework handles disruptions as they occur and builds on the solutions of previous disruptions. The model is tested on data from an African hub-and-spoke airline with a 150 flight network. The author concludes that the static approach leads to an underestimation of disruption costs of 30.8%, when compared with the dynamic approach. This research assumes crew are always available and passenger connections are not considered.

Marla et al. (2017) incorporate cruise speed control in addition to traditional recovery actions by creating alternative flight plans corresponding to higher cruise speed with pre-determined cost. Similarly to Vos et al. (2015), their formulation incorporates parallel time-space networks, which enable explicit modelling of individual aircraft, but include passenger recovery, by explicitly assigning cost to delayed flights that carry connecting passengers. Departure time decisions are accounted for by introducing multiple flight arc copies representing different departure time alternatives. Similarly, cruise speed control alternatives are considered by generating additional flight arc copies for various speed options associated with each departure time alternative. This approach requires a discretisation of the cruise speed options and naturally enlarges the size of the generated network. The solution is also dynamic, i.e. disruptions are solved as they occur. The model is tested on data from a European hub-and-spoke airline with a 250 flight network. The proposed model is compared to a baseline sequential recovery model, different variants of an aircraft-centric model with flight planning, and different variants of a passenger-centric model without flight planning. To assess the performance of the models, the authors consider 60 delay scenarios (only this disruption type is considered) based on historical data from the airline. They conclude that their enhanced recovery models offer benefits in terms of reducing overall costs and minimising passenger-related delay costs for the airline, compared to existing approaches.

Vink et al. (2020) also adopt a similar approach to Vos et al. (2015) by employing parallel time-space networks and an aircraft selection algorithm, albeit with a slight modification. Aircraft maintenance schedules and passenger itineraries are modeled, and each flight's delay cost is precomputed considering the connections that will be lost if the arrival of inbound flights is delayed. Hassan (2019) extends their work by considering the option of the outbound flight also being delayed when determining whether a connection is lost. However, this consideration is only made for connecting flights of disrupted flights. As the primary focus of their research is to enable real-time operational use of the model, the tests focused on solution times. Both models are tested on 10 different disruption scenarios, each involving 2 to 6 disruptions (combinations of delays, cancellations and aircraft unavailability) occurring at various times of the day.

2.2 Airline Reserve Fleet Capacity

The effects of having reserve fleet capacity on the fleet plan are multiple. Since choices regarding fleet size have to be made years before schedule execution, a better understanding of the impact of reserve fleet capacity and its effectiveness can contribute towards airlines' better fleet plan creation and subsequent higher margins. However, research into this particular added capacity that some airlines choose to work with has been very limited. Reserve fleet capacity gives the airline an opportunity to quickly offer more flights in case of higher demand, but it also entails the risk of overcapacity in case of decreasing demand. It is an option that provides airlines the means to add robustness to their schedules, by mitigating delays, enabling passenger connections that would otherwise not be possible, and substituting inoperable aircraft in need of unexpected maintenance. All of these consequences have a positive impact in the Quality of Service Index of airlines, which values an airline's set of flights offered in an O-D market, relative to competitors, and is used to estimate the market share potential of new routes and incremental flights (Belobaba et al., 2015). In addition to this improvement of passenger experience, the potential decrease of secondary costs associated with cancellations, missed connections, and delays (such as delay compensation or rebooking of passengers) needs to be considered along with the increase of total fleet operation costs in case of additional capacity to ascertain the break even point financially for the airline, helping it to achieve higher margins.

Groeneveld (2021) studied the impact of reserve fleet capacity in delay mitigation, proposing three relevant case studies to be carried out to study different effects of the operation of reserve fleet. In the first, using Monte Carlo simulations on an existing flight schedule and a rule-based delay mitigation model, the effectiveness of reserve fleet capacity (in the form of standby unscheduled aircraft) is analysed. The break even point for having one standby reserve was found to be a fleet size of 48 aircraft, considering a trade-off between additional operational costs and potential saved delay costs. In the second, the impact of different delay durations on the effectiveness of reserve fleet capacity is studied. It was found that reserve fleet capacity becomes more effective in high delay scenarios with high outlier delays. In the third, a new schedule was used in which one rotation per day would be reallocated to the standby aircraft. The model was then rerun, which resulted in a lower performance of the reserve capacity in terms of efficiency. As delay predictions' accuracy improve this option of spreading reserve fleet capacity over the schedule is suggested to be revisited. An important aspect of this work is that all calculations are made considering a point to point network (part of a European hub-and-spoke carrier), and thus does not accurately represent the delay costs of a hub-and-spoke network, which in all likelihood would benefit more from on time performance (passenger connections are affected by delays).

Varenna (2023) worked on a relevant modular, stochastic discrete event simulation model of airline operations named ANEMOS (Airline Network and Maintenance Operations Simulation), with the objective of constructing a model capable of testing policies, plans, and scenarios involving both network and maintenance operations with the goal of understanding the performance of the system as a whole. The paper details the steps taken to simulate the network and maintenance operations of the inter-continental fleet of hub-and-spoke carriers. The case study presented in the paper, to validate the model and show its capabilities, was developed in collaboration with a major European carrier. It considered a fleet of 50 aircraft of four different models, and examined the effects introducing a second reserve aircraft to the fleet. The impact of the second reserve was quantified and compared to the situation with only one reserve aircraft available in terms of cancellation factor (the percentage of cancelled rotations with the second reserve aircraft decreased to less than half), delay reduction (probability of observing a delay longer than one hour decreased by 0.4% and a delay longer than two hours decreased by 0.1%), and economic value (savings achieved with the added reserve aircraft were estimated to be at most around

10% of the profit the same aircraft could produce for the airline in case of normal operations in the airline's fleet, meaning it would not be economically profitable for the airline to activate a second reserve aircraft). Simulated results of the baseline scenario are compared with historical data, and it is shown that the model closely resembles historically observed operational performance. Both Varenna (2023) and Groeneveld (2021) propose further research considering the flow of passengers in order to better measure airline performance in terms of missed connections.

While existing studies have made significant strides in developing models and methodologies for optimising operational performance and mitigating disruptions, several research gaps are evident. One notable gap lies in the limited exploration of the interaction between reserve fleet adjustments and their impact on operational outcomes within the context of executed operations. While some studies have examined the effectiveness of reserve fleet capacity in delay mitigation and schedule recovery, a noticeable gap persists in the literature regarding the assessment of how alterations to the reserve fleet might have impacted past operational performance. Additionally, the majority of existing research in the field of Disruption Management tends to adopt a forward-looking perspective, aiming to develop real-time decision-support tools for AOCC. However, there is a lack of studies that take a retrospective approach, leveraging insights from executed operations to inform strategic decisions regarding reserve fleet management. There is thus a clear research gap in understanding how adjustments to the reserve fleet composition affect operational performance metrics, such as disruption costs and passenger satisfaction. Addressing this gap would provide valuable insights for airlines seeking to optimise their reserve fleet strategy and enhance operational resilience.

3 Problem Statement

Initially, additional factors and critical considerations in the context of this research are introduced. Subsequently, a discussion on recovery actions available to airlines for mitigating passengers' delays, cancellations, and missed connections is provided. Finally, leveraging this enhanced understanding, the problem statement is introduced.

Unexpected events, such as bad weather conditions, mechanical failures or traffic congestion, are unavoidable for airlines and may result in aircraft unavailability and delays. Original schedules created for the aircraft and crew, and passenger itineraries may be disrupted as a consequence. Walker (2019) noted that over 22% of all flights in Europe were delayed (flights arriving over 15 minutes after Scheduled Time of Arrival) in 2019 as a result of disruptions, after considerable efforts and investments were made by airlines to improve 2018 record high delay levels, resulting in a decrease of almost 2%. In addition, operational cancellations decreased from 2.0% in the previous year to 1.7% in 2019. Despite this slight improvement, 2019 was still the third worst year in the last decade, with en-route Air Traffic Flow Management (ATFM) delays during the summer season remaining a problem for airlines, which indicates an upwards trend that will have to be properly dealt with by the airline industry.

In a paper commissioned by Amadeus, Gershkoff (2016) revealed the estimated cost of disruptions to airlines to be 8% of airline revenue, or 60 billion USD worldwide, which explains the need for effective disruption management strategies to be in place. The objectives of these strategies are: fulfill the customer promise by ensuring passengers and their luggage reach their intended destination promptly and as per the booked service level; minimise actual costs, which encompass excessive crew expenses, compensation costs, accommodation and lodging expenses for disrupted passengers and crew, as well as ticket expenses with alternative airlines; swiftly return to the original plan and resume operations as quickly as possible (Kohl et al., 2007).

Passengers whose itinerary is delayed (over 3 hours)/cancelled are entitled to compensations according to Regulation (EC) No 261/2004 (Cook and Tanner, 2015). The industry refers to these and operational costs as hard costs. In addition to these, however, another category of less measurable costs must be considered, originated by the opportunity cost of a passenger not choosing the same airline in the future due to an endured inconvenience. These costs, known as soft costs, are highly non-linear and constitute a big part of the total disruption costs. One metric used by companies to measure these costs across various sectors is the Net Promoter Score (NPS), a customer loyalty measure, due to its simplicity, ease of interpretation, low costs of calculation, and, overall, its confirmed positive and statistically significant relationship with profitability (Korneta, 2018).

Most large airlines operate AOCCs to perform on-the-day coordination of schedule execution. Their purpose is to monitor the progress of operations, to flag actual or potential problems, and to take corrective measures in response to unexpected events. In most airlines, representatives of key airline functions work together to ensure smooth schedule execution, in a process that is dynamic and cannot really be formulated explicitly, as described by Kohl et al., 2007. Whenever disruptions occur, a solution must be found to recover operations. In order to decide on which solution to adopt, the passenger, crew and aircraft perspectives have to be taken into account. From the passenger perspective it may, for example, make sense to delay an outbound flight (i.e. a flight departing an airline's hub) to ensure that passengers on a delayed inbound flight will be able to make their connection. This option must then be evaluated from the crew (e.g. they have strict legal limits of sequential hours of work) and aircraft (e.g. the propagation of delays may result in even higher costs later on in the day) perspectives. This process can go until an option that is legal and acceptable from all perspectives, has been found. Once a decision has been made, it is essential to execute it and maintain ongoing monitoring. This disruption management process differs from most control processes in complex systems involving humans, with its' particularly extensive range of available options and the intricate computational challenges associated with evaluating their effects.

To identify all possible schedule recovery options, the strategies that the AOCC can use include:

- **Delaying a flight**: the most commonly applied measure, that results in the propagation of the delay in the network.
- **Cancelling a flight or flights**: most often in the form of cancelling rotations (sequence of flight legs that start at a station and end at that same station) (Rosenberger et al., 2004).
- **Exchanging assigned aircraft**: swapping two rotations' assigned aircraft, when more buffer time can be achieved, and when it results in the reduction of total delay and/or cancellation of cheaper flights.
- **Activating a reserve aircraft**: assigning disrupted flight to a reserve aircraft (usually kept on standby at the hub). When previously assigned aircraft arrives to hub (in case the disruption is a delay) it becomes the new reserve.
- **Increasing/decreasing aircraft speed**: important tool for delay absorption, although its costs (additional fuel) must be taken into account (Marla et al., 2017).
- **Ferrying aircraft**: flying an aircraft without passengers to an airport to cover a flight that would otherwise have to be cancelled from that airport.

The availability of reserve aircraft is an option that provides airlines the means to add robustness to their schedules. Their effective use holds the potential to yield significant operational advantages, diminishing delays and facilitating continued operations in the event of sudden inoperability of assigned aircraft. This not only enhances customer satisfaction but also contributes to a reduction in operational costs. Despite these potential advantages, leveraging an airline's reserve capacity presents inherent challenges, notably due to the significant investment required to maintain this reserve capacity. Airlines' conventional revenue models rely on flight schedules as densely packed as possible, and the deviation from this norm by under utilising aircraft suggests a potential revenue loss stemming from the choice of maintaining some aircraft as reserves. As such, a detailed examination of the delicate balance between operational benefits and the financial implications of including reserve capacity in airlines' fleet planning strategies.

The availability of past data on airlines' execution of planned schedules (resulting delays, cancellations and missed connections), along with the disruptions that prohibited the flawless execution of the planned schedules provides an opportunity to assess the impact and effectiveness of reserve fleet composition changes on past days of operations. The aim of the research is to identify the added/reduced value of an increased/decreased reserve fleet capacity and to determine in which way it could have best been used on individual past days of operations. By defining costs associated with a given passenger's delay, cancellation or missed connection, the outcome of this research should give airlines the possibility of determining the resulting impact on disruption costs stemming from a change in their reserve fleet.

4 Methodology

This research aims at identifying the potential impact of reserve fleet composition changes on airline's past days of operations, based on planned and executed schedules and recorded disruptions. In order to achieve this, a disruption management model was created to simulate the AOCC decisions made throughout the day as a response to the disruptions occurring. This model makes it possible to compare airline performance indicators achieved with different reserve fleet compositions (based on the resulting execution of the schedule).

The methodology outlined in this study is tailored for hub-and-spoke network structures and is applicable to airlines operating a heterogeneous fleet. For each past day of operations, the input to the model comprises a flight schedule (planned and executed, including fleet assignment), a list of the recorded AOGs (Aircraft On Ground), the reserve fleet operated and the variations in this fleet to be considered. The resulting execution of the day's schedule (delays, cancellations and missed connections) are then calculated for all compositions of the reserve fleet requested. A schematic overview of this framework is presented in Figure 1.

The next subsection describes the parallel time-space network approach adopted in this work, followed by the novel distinction between the static and dynamic networks. Next, the method of including delays as they occurred on the day is explained in detail. The Mixed Integer Linear Programming model used to optimise schedule recovery is then presented, with its nomenclature and mathematical formulation. The modelling of connecting passengers is explained in the last subsection.

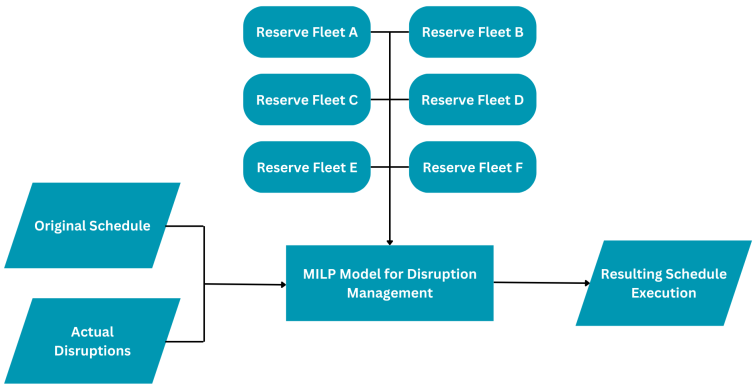

Figure 1: Flowchart of the proposed framework.

4.1 Parallel time-space networks

In this work, parallel time-space networks are utilised to model each individual aircraft, as previously proposed by Vos et al. (2015) and Vink et al. (2020). One separate network is created for each aircraft. Time is discretised in homogeneous time-steps, and the departure and arrival times are rounded to the nearest time-step. The advantage of individual aircraft networks lies in the ability to use specific tail numbers, enabling the solution of the scheduled recovery problem together with the aircraft routing recovery problem. In addition, aircraft specific constraints such as forced AOGs and the source and desired sink nodes (at the end of the day) of the aircraft, can be modeled.

A simplified example of one of the networks is shown in Figure 2, where each node represents an airport at a specific point in time. Nodes are created for all airports in the network and at every time step in the time window. Different arcs are used to model aircraft movements as time advances. Flight arcs represent flights from one airport to another. The origin node of the arc represents the origin airport at the departure time and the terminating node represents the destination airport and the arrival time. Delayed flight arcs are used to model delayed flights. These arcs are created for each possible delay (up to maximum considered delay). Ground arcs represent the time spent at an airport, where the aircraft stays at the same airport.

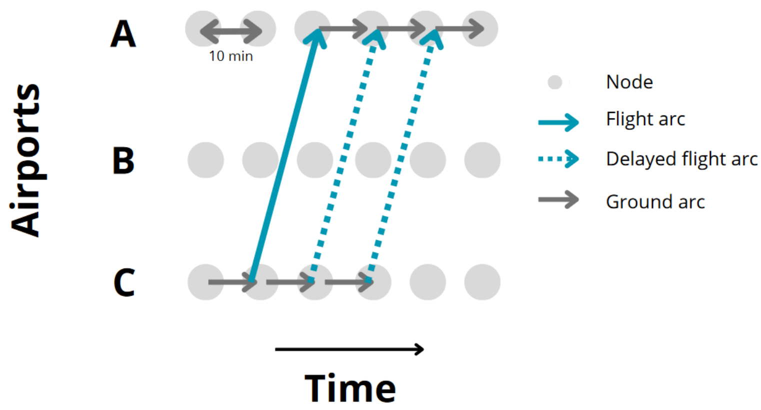

Figure 2: Ground, flight and delay arcs in time-space network (Hassan, 2019).

4.2 Distinction between dynamic and static network

This study makes a crucial distinction between the dynamic and static components of airline networks, which is key to assessing how adjustments to the reserve fleet might have impacted past operations. Typically, an airline's reserve fleet does not include aircraft of every type found in its full fleet. Airline networks contain some flights that cannot be operated by aircraft that the reserve fleet contains or could eventually contain according to long term company plans. For example, aircraft designated for short-haul flights within an airline's network are unsuitable for long-haul operations due to their limited range. Furthermore, within the short-haul network itself, there are flights that cannot be serviced by certain aircraft types in the fleet. This is often due to airline policies, such as crew certification restrictions, where pilots are typically qualified to operate only one specific family of aircraft.

Consequently, a subset of the fleet that will not be significantly affected by changes in the reserve capacity (due to the specificity of aircraft requirements for certain flights) is excluded from the time-space networks constructed for this study, and the flights operated by this subset constitute the static network of the study. These flights must not be completely disregarded in the study as they transport a significant portion of the airline's connecting passengers. In order to properly determine missed connections, for all flights in the static network, the only aspect considered in the model is the delay occurring in the airline's hub/hubs. For example, if a flight departing from a hub has a delay on departure of 1 hour and a delay on arrival of 30 minutes, the 1 hour delay on departure is what the model considers as the flight's Relevant Delay (RD). Due to this explained particularity, these delays also are not discretised, as they are in the case of flights in the dynamic network, modelled with the time-space networks.

4.3 Modelling delays and their propagation

Previous work by Vos et al. (2015) and Vink et al. (2020) in the field of disruption management has incorporated Turn Around Time (TAT), specific to each aircraft type, in the time-space networks by way of extending the flight and delayed flight arcs created, increasing the flight times by the relevant TAT. This ensures that aircraft do not have the possibility of departing on their next flight before turn around is complete. In this work, the executed schedule (Actual Times of Departure and Arrival, ATD and ATA respectively) on any particular past day of operations serves as the basis, enabling the more accurate representation of delays incurred in all flights throughout the day.

The following methodology is followed to determine the flight and delayed flight arcs length: (1) each aircraft's list of flown flights (on day in analysis) is ordered chronologically; (2) each flight f arc duration (flight and delayed flight) is calculated as the difference between the Actual time of Departure of flight immediately after flight f and the Actual time of Departure of flight f. The two exceptions to this are: if the flight is the last of the day or the difference between the calculated arc duration and scheduled flight time is higher than a predefined threshold, then the arc duration is the sum of actual flight time and turn around time specific to aircraft used. This method leverages the knowledge of executed operations, as it is considered that the AOCC's decisions to optimise schedule recovery were the best possible and considered factors not analysed in this study (e.g. crew considerations). This method allows the model to accurately consider all delays occurring throughout the dynamic network, most of which occur during aircraft turn-around. Furthermore, this allows for an accurate representation of flights' speed increases/decreases, without the introduction of additional flight and delayed flight arc copies as proposed by Marla et al. (2017) to consider speed alterations. It is worth noting that despite the fact this process depends on executed operations, each flight arc is set to depart at the Scheduled Time of Departure (STD) of its corresponding flight and delayed flight arcs are simply copies of the flight arcs, shifted by their specific delay at the departure and arrival nodes. Table 1 provides an example application of the proposed methodology, showing the STD, STA (Scheduled Time of Arrival), ATD and ATA of four consecutive flights, along with the resulting flight arc durations.

Table 1: Flight arc duration calculation for one aircraft's flights.

Flight #	**STD**	**STA**	**ATD**	**ATA**	**Arc length** (*mins*)
1	07:10	10:30	07:20	10:30	240
2	11:20	14:20	11:20	14:10	300
3	15:00	18:30	16:20	19:30	240
4	19:20	22:40	20:20	23:20	220

4.4 Model formulation

In this section, the Mixed Integer Linear Programming (MILP) model simulating the airline's disruption management is presented. Nomenclature is introduced first, after which the model formulation is presented. The following subsections explain some of the key features of the model in more detail.

The model's allowed recovery options include delaying or cancelling flights and changing the appointed aircraft to fly each flight (reserve aircraft can be assigned).

4.4.1 Nomenclature

Sets	
P	Aircraft in the dynamic fleet
F	Flights allocated to/flown by aircraft present in P in the operated schedule
F_C	Flights pairs transporting connecting passengers (both from the static and dynamic network)
F_p	Flights flown by/allocated to (in case of cancellations only) aircraft p
$F_{1^{st}}$	First flights of the day allocated to/flown by aircraft present in P in the operated schedule
$F_{\Delta>0}$	Flights that were flown with a positive departure delay Δ
F_m	Flight arcs corresponding to flights that would be in the air during forced grounding event m
D	Possible delays for each flight
D_f	Delays lower than actual departure delay of flight f
D_m	Delay arcs corresponding to delayed flights that would be in the air during forced grounding event m
G	Ground arcs associated with F
N	Intermediate nodes associated with F
E	Aircraft types included in P
A	Airports associated with F
T	Travel classes (Economy, Business and First class)
$O(n,p)$	Arcs originating at node n for aircraft p
$T(n,p)$	Arcs terminating at node n for aircraft p
$Z(a,e)$	Arcs terminating at airport a for aircraft type e
P_f	Aircraft in the (selected) fleet allowed to fly flight f
M_p	Forced groundings (AOGs) of aircraft p

Decision variables	
$\delta_{F_{p,f}}$	Binary variable equal to 1 if aircraft p is allocated to flight arc f, zero otherwise
$\delta_{D_{p,f,d}}$	Binary variable equal to 1 if aircraft p is allocated to flight f using delay arc d, zero otherwise
δ_{C_f}	Binary variable equal to 1 if flight f is cancelled, zero otherwise
$\delta_{MC_{a,b}}$	Binary variable equal to 1 if connection from flight arc a to flight arc b is made impossible, zero otherwise
$\delta_{G_{p,g}}$	Binary variable equal to 1 if aircraft p uses ground arc g, zero otherwise
$s_{a,e}$	Integer slack variable for aircraft type e balance at airport a at end of time window
s_f	Binary slack variable equal to 1 if aircraft used to fly flight f is changed when compared to executed operations

Parameters	
$DC_{f,d}$	Delay cost of flight f for delay d (per passenger)
CC_f	Cancellation cost of flight f (per passenger)
$MC_{a,b}$	Missed connection cost of connection from flight a to flight b (per passenger)
$MSC_{a,b}$	Margin for successful connection of connection from flight a to flight b
L_d	Length of delay d
α_f	Binary variable equal to 1 if flight f was not cancelled on day in analysis, zero otherwise
β_f	Binary variable equal to 1 if flight f is part of the dynamic network, zero otherwise
RD_f	Delay occurring at hub airport of static flight f
$H_{f,t}$	Number of booked passengers on flight f on travel class t
$HC_{a,b,t}$	Number of booked passengers connecting from flight a to flight b on travel class t
VF_t	Value factor of passengers travelling on travel class t
$M_{a,e}$	Cost of missing an aircraft of type e at airport a at the end of the time window
M_f	Cost of swapping aircraft used to fly flight f
ϵ	Maximum number of allowed tail swaps (compared to executed operations)
$B_{n,p}$	Flow Balance at node n for aircraft p. 0 at intermediate nodes, 1 at start node where aircraft p is supplied
$K_{a,e}$	Number of aircraft of type e demanded at airport a at the end of time window
M	Large constant

4.4.2 Formulation

Minimise:

$$\sum_{p\in P}\sum_{f\in F}\sum_{d\in D}\sum_{t\in T} DC_{f,d}\cdot H_{f,t}\cdot VF_t\cdot \delta_{D_{p,f,d}} + \sum_{f\in F}\sum_{t\in T} CC_f\cdot H_{f,t}\cdot VF_t\cdot \delta_{C_f} + \sum_{a\in F_C}\sum_{b\in F_C}\sum_{t\in T} MC_{a,b}\cdot HC_{a,b,t}\cdot VF_t\cdot \delta_{MC_{a,b}} + \sum_{a\in A}\sum_{e\in E} M_{a,e}\cdot s_{a,e} + \sum_{f\in F} M_f\cdot s_f \tag{1}$$

Subject to:

$$\sum_{p\in P}\left(\delta_{F_{p,f}} + \sum_{d\in D_f}\delta_{D_{p,f,d}}\right) = 0 \qquad \forall f\in F_{1^{st}}\cap F_{\Delta>0} \tag{2}$$

$$\sum_{p\in P}\left(\delta_{F_{p,f}} + \sum_{d\in D}\delta_{D_{p,f,d}}\right) + \delta_{C_f} = 1 \qquad \forall f\in F \tag{3}$$

$$\sum_{g\in G\cap O(n,p)}\delta_{G_{p,g}} - \sum_{g\in G\cap T(n,p)}\delta_{G_{p,g}} + \sum_{f\in F\cap O(n,p)}\delta_{F_{p,f}} - \sum_{f\in F\cap T(n,p)}\delta_{F_{p,f}} + \sum_{f\in F}\sum_{d\in D\cap O(n,p)}\delta_{D_{p,f,d}} - \sum_{f\in F}\sum_{d\in D\cap T(n,p)}\delta_{D_{p,f,d}} = B_{n,p} \qquad \forall n\in N, \forall p\in P \tag{4}$$

$$\sum_{g\in G\cap Z(a,e)}\delta_{G_{p,g}} + \sum_{f\in F\cap Z(a,e)}\delta_{F_{p,f}} + \sum_{f\in F}\sum_{d\in D\cap Z(a,e)}\delta_{D_{p,f,d}} = K_{a,e} - s_{a,e} \qquad \forall a\in A, \forall e\in E \tag{5}$$

$$\delta_{F_{p,f}} + \sum_{d\in D}\delta_{D_{p,f,d}} + \delta_{C_f} = 1 - s_f \qquad \forall p\in P, \forall f\in F_p \tag{6}$$

$$s_f = 0 \qquad \forall f \in F_{1^{st}} \tag{7}$$

$$s_f \leq \epsilon \qquad \forall f \in F \tag{8}$$

$$\begin{aligned}
M \cdot \delta_{MC_{a,b}} \geq\ & RD_a \cdot \alpha_a \cdot (1-\beta_a) - RD_b \cdot \alpha_b \cdot (1-\beta_b) + \\
& \left(\sum_{p\in P}\sum_{d\in D} L_d \cdot \delta_{D_{p,a,d}}\right) \cdot \beta_a - \left(\sum_{p\in P}\sum_{d\in D} L_d \cdot \delta_{D_{p,b,d}}\right) \cdot \beta_b \\
& + \frac{M}{2} \cdot (1-\alpha_a) \cdot (1-\beta_a) + \frac{M}{2} \cdot (1-\alpha_b) \cdot (1-\beta_b) \\
& + \frac{M}{2} \cdot \delta_{C_b} \cdot \beta_b + \frac{M}{2} \cdot \delta_{C_a} \cdot \beta_a - MSC_{a,b} \qquad \forall a, b \in F_C
\end{aligned} \tag{9}$$

$$\sum_{f\in F_m} \delta_{F_{p,f}} + \sum_{d\in D_m} \delta_{D_{p,f,d}} = 0 \qquad \forall m \in M_P \tag{10}$$

$$\sum_{p\in P\setminus P_f} \left(\delta_{F_{p,f}} + \sum_{d\in D} \delta_{D_{p,f,d}}\right) = 0 \qquad \forall f \in F \tag{11}$$

$$\delta_{F_{p,f}} \in \{0,1\} \qquad \forall p \in P, \forall f \in F \tag{12}$$

$$\delta_{D_{p,f,d}} \in \{0,1\} \qquad \forall p \in P, \forall f \in F, \forall d \in D \tag{13}$$

$$\delta_{G_{p,g}} \in \{0,1\} \qquad \forall p \in P, \forall g \in G \tag{14}$$

$$\delta_{C_f} \in \{0,1\} \qquad \forall f \in F \tag{15}$$

$$\delta_{MC_{a,b}} \in \{0,1\} \qquad \forall a, b \in F \tag{16}$$

$$s_f \in \{0,1\} \qquad \forall f \in F \tag{17}$$

$$s_{a,e} \in \mathbb{Z}_{\geq 0} \qquad \forall a \in A, \forall e \in E \tag{18}$$

A MILP model is proposed, based on the one presented by Vink et al. (2020), that minimises passengers' disruption costs. The objective function (1) includes disruption costs and penalties with the goal of limiting the schedule recovery options taken to a minimum. The first three terms define the disruption costs (delays, cancellations and missed connections respectively). These costs are defined by the airline at a passenger level (the cost is defined per passenger facing the particular disruption on the particular flight or flight pair in case of a missed connection). As such, the costs are multiplied by the number of passengers facing the disruption on each travel class and the Value Factor (defined by the airline) of each of these. The last two terms define the artificial penalties added to minimise changes made to the executed schedule.

The model is subject to constraints defined in Equations (2) to (18). To ensure that each day of operations starts with initial delays (as actually occurred), constraint (2) ensures that if an aircraft's first departure of the day was delayed (in executed operations), the flight is prohibited from departing either on schedule or with a delay lower than the occurred delay. All other delays occurring throughout the day are modelled through the Time-Space Networks created, as explained in the previous subsection. Constraint (3) ensures each flight is either executed as scheduled, delayed or cancelled. Flow balance is ensured by two constraints. For the intermediate and start nodes (4), balance is maintained for all nodes in each aircraft time-space network: if an aircraft is supplied to a certain node, it should also depart from that node. For nodes at the end of the time window, the demanded number of aircraft of a certain type is ensured with constraint (5). These constraints must be included to guarantee that aircraft are in the right position to resume scheduled operations on the next day of operations, as the model is only meant for a full day of operations and any aircraft not available on the next day of operations leads to additional disruption costs. However, this constraint includes a slack variable which has a penalty assigned in the objective function (1) for situations where the requested aircraft cannot be at that airport at the end of the time window.

Using Constraint (6), a penalty is assigned in the objective function (1) whenever the aircraft used to operate one flight is different from the executed schedule. In case the flight was cancelled (on the day in analysis), this penalty is assigned if the aircraft used to operate the flight is not the same as planned. scheduled routing of an aircraft is changed. If no aircraft swaps are made when compared to the execution of the schedule on the day in analysis, no penalty is assigned. Despite the static disruption approach taken in this work, using constraint (7), the model is prohibited from completing tail swaps on all first flights of the day. This measure decreases the disparity between this static approach and a dynamic disruption approach. By restricting tail swaps at the beginning of the day, the model simulates a scenario where no preemptive decisions are taken at the start of the day, thereby accentuating the impact of disruptions on operational continuity. The objective of this study makes it pertinent to add constraint (8), limiting the number of these swaps up to a maximum amount to be defined by the airline.

In order to ensure missed connections are flagged, constraint (9) checks if the combination of delays/cancellations incurred on 2 connecting flights renders the connection impossible, flagging those that are guaranteed missed connections. This inclusion of passenger connections is further detailed in Section 4.5.

Constraint (10) guarantees aircraft grounding during the periods of time the aircraft was inoperable (AOG). This ensures that all flight arcs and delay arcs corresponding to flights that are in the air at the aircraft's grounding period (AOG) are prohibited. This constraint enables the consideration of reduced reserve fleet scenarios. This can be achieved by means of adding AOGs to the operated reserve aircraft during the periods they served as reserves. On the other hand, considering augmented reserve fleets is possible by incrementing the set P of aircraft in the dynamic fleet. Constraint (11) ensures that flights are only performed by aircraft that are deemed suitable for the particular flight by the airline, due to conditions such as range and number of required seats. Constraints (12) to (18) define the decision variables as being binary or integer.

4.5 Connecting passengers

Particularly for hub-and-spoke airlines, connecting passengers amount to a significant portion of the airline's customers. Disruptions leading to missed connections often result in large compensation costs for the airline, and in unhappy customers who have to wait for an alternative connection in an airport. This research aims to include the effects of passengers who miss their connection in the aircraft recovery problem, despite the necessary increase in complexity of the MILP formulation previously analysed. To achieve this, each connection's Margin for a Successful Connection (MSC) is precomputed, representing the amount of time that each connection can shorten by such that it is still possible (considering a fixed minimum connecting time, specific to each airport where a connection occurs). To exemplify how this approach can determine whether any two flights' connection is missed, Figure 3 presents a connection between 2 flights, part of the dynamic network. If both flights are on time, the passengers' connecting time is 50 minutes. Assuming a minimum connecting time of 30 minutes, the MSC for this pair of flights is then set to 20 minutes, meaning that, for the connection to be made, the difference between the delay incurred on flight 1 and the delay incurred on flight 2 has to be lower or equal to the MSC of 20 minutes. In case of a cancellation on any of the 2 connecting flights, the model penalises the cancellation two-fold (by flagging the missed connection as well, as seen in constraint (9)). In case any or both of the connecting flights are part of the static network, the information regarding the delay at the hub and whether the flight was cancelled is used for flagging missed connections.

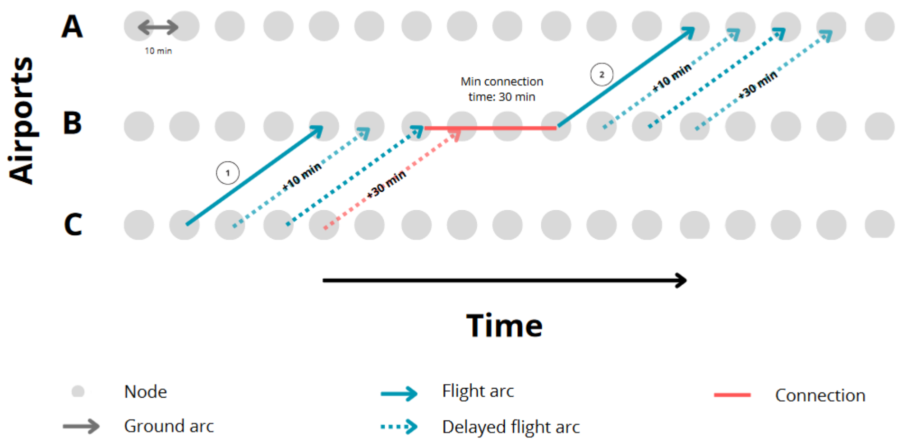

Figure 3: Connection example. Margin for Successful Connection (MSC)= 20 minutes.

5 Experimental Campaign

This section presents an implementation of the proposed framework that was developed in collaboration with a Full-Service Carrier operating in Europe and introduces an example case study used to validate the model and test its capabilities. The proposed case study investigates the effects of changes to the reserve fleet maintained to a partner airline's operational performance on past days of operations. A comprehensive sensitivity analysis of the parameter ϵ is conducted, followed by the presentation of final study outcomes, delineating the impact of these fleet adjustments. Four main indicators are used for the evaluation of the results of this case, as indicated by the airline: (1) the total disruption costs (along with a breakdown into delay, cancellation and missed connection costs); (2) the punctuality; (3) the heavy delay index; (4) the percentage of missed connections.

5.1 Description of case study

Several past days of operations were studied. In all of them, the airline operated a fleet of over 110 aircraft (over 70 as part of the dynamic fleet and over 40 as part of the static fleet), serving 120 airports on over 500 daily flights (less than 25% part of the static network), with over 3000 connections between them (each connection with a specific number of passengers). Flight schedule (planned and executed), fleet (including reserves) composition, disruption records, booked passengers per flight and connecting passengers were obtained from the airline.

The dynamic fleet is split over seven aircraft types of two different families:

- A220 family: A221; A223.
- A320 family: A319; A320; A32A; A321; A32B.

The airline operated a reserve fleet composed of aircraft of both families. Due to crew considerations and passenger capacity requirements, flights executed by aircraft in one family cannot be assigned to an aircraft of the other family by the model.

The recovery model was implemented in Python using Gurobi 11.0.0 to solve the MILP problem. Computational tests were conducted on a computer using a 64 bit Intel i7-7700 2.80 GHz processor and 16 GB RAM. For all considered reserve fleet scenarios, the MILP model encompassed between 1 and 1.2 million decision variables and between 460 and 540 thousand constraints. The models were solved to optimality and solution times spanned from 2 to 5 minutes.

5.2 Additional assumptions

To measure disruption costs, the airline provided reference values from an internal study, measuring the average passenger's Net Promoter Score (NPS) negative impact of the three main types of disruptions (cancellations, delays and missed connections). This metric measures individual customer loyalty with a score that spans from -100 to +100. All reserve fleet scenarios tested in the case study on past operational days were conducted using parameters determined by the airline, taking into account the constructed model. These parameters include:

- Time is discretised in 10 minute time steps.
- Flights can be delayed up to 3.5 hours.
- One passenger's delay on any flight leads to a maximum negative NPS impact of 30 (cost is directly proportional to the duration of the delay, as delay costs are assumed to be linear with time).
- One passenger's flight cancellation leads to a negative NPS impact of 59.
- One passenger's missed connection leads to a negative NPS impact of 85.
- One Business Class passenger is valued as 4 Economy Class passengers.
- One First Class passenger is valued as 8 Economy Class passengers.
- One tail swap (in relation to executed operations) costs the equivalent to 100 Economy Class passengers being delayed for one hour.
- One missing aircraft at the end of the time window at the designated airport costs the equivalent to 100 Economy Class passengers having their flights cancelled.
- Minimum connecting time at airline's hubs is 40 minutes.

5.3 Sensitivity analysis: ϵ parameter

The core of this research is the analysis of the operational impact that changes in the reserve fleet used by an airline can make. The alternative execution of the planned schedule involving minimal disruption costs is highly dependent on the parameter ϵ, which is the maximum number of allowed aircraft swaps (in comparison with the execution of the schedule on the day in analysis). This section will investigate the sensitivity of the model created to changes in this parameter. It is pertinent to not allow this number to be very big, so as to increase interpretability of the results and ensure (as much as possible) that the aircraft used to fly the majority of the day's flights remain unchanged and the swaps that are chosen are the most impactful. This is because any swap, although deemed possible by the model, could in reality be infeasible due to constraints not considered in this study (such as crew considerations or night bans in specific airports). On the other hand, this parameter cannot be set to 0 since the objective of this study is to change the composition of the reserve fleet available for disruption management, and so some swaps will always be necessary.

To determine the sensitivity of the solution to this parameter, one particular day of operations of the peak of the Summer Schedule (most passengers flying and most disruptions) was used as basis. No changes in the reserve fleet operated on this day were considered for this analysis. Resulting disruption costs obtained, sum of the first three terms in the objective function (1), are presented in Figure 4. The leftmost bar of the chart presents the observed disruption costs as comparison. It is important to note that the model's validation is possible when analysing the results obtained when setting ϵ to 0, essentially simulating the execution of the schedule as was observed. Any small deviations (from observed numbers) seen here can be explained by the existence of a time step of 10 minutes in the Time-Space Networks created to simulate the movement of the dynamic fleet, or by additional effects and discrepancies between reality and the MILP. This discretisation is responsible for the over/under estimation of delays, which can impact cancellations and missed connections as well.

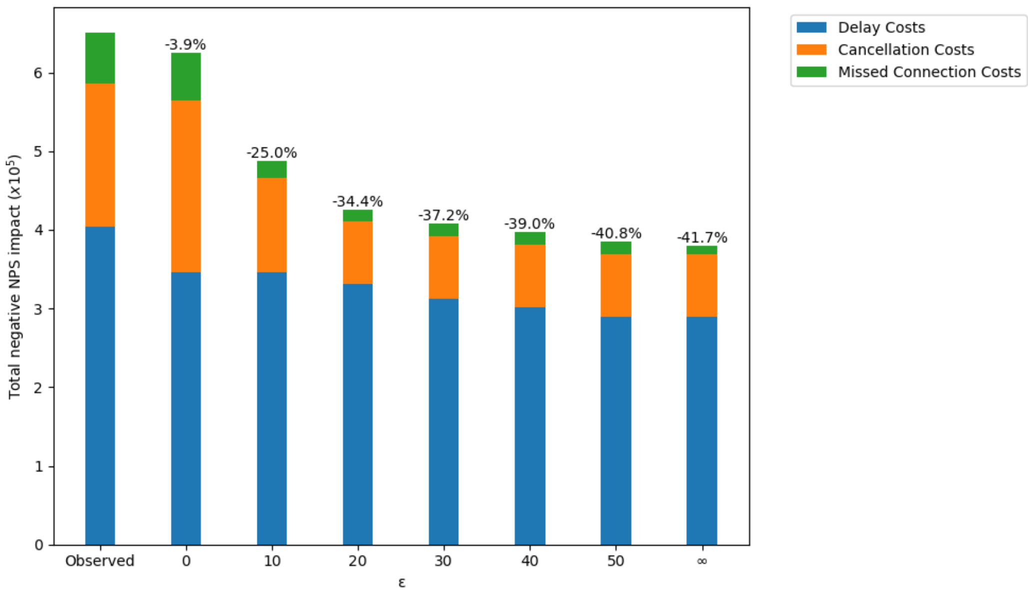

Figure 4: Sensitivity of Total Disruption Costs to changes in ϵ.

As expected, the disruption costs decrease as the model is given freedom to swap more aircraft used to operate specific flights. However, it is evident that as ϵ increases, the decrease in disruption costs slows down. When setting no limit to ϵ, removing constraint (8), the number of swaps deemed optimal is 52. It becomes apparent that the model optimises the execution of the planned schedule better than was actually possible on each day of operations. The biggest reason for this discrepancy is the knowledge of all disruptions (either AOGs or delays modelled through the Time-Space networks created) at the start of the day. Despite the fact that the model is not allowed to complete tail swaps on the first flights of the day, this constraint is not enough to As such, results obtained should be interpreted as the best possible outcome of the day.

This sensitivity analysis performed was used as a tool for deciding the value of ϵ to be set for the simulations to be run with different reserve fleet compositions. It was decided to set ϵ to 50 for the remaining analysis that will be discussed in the next subsections, to allow for a reasonable amount of swaps when introducing/removing aircraft from the reserve capacity.

5.4 Impact of Reserve Fleet changes

In order to determine the impact that changes in the composition of the reserve fleet kept could have made on past operations, one particular week of operations of the peak of the Summer Schedule (most passengers flying and most disruptions) was used as basis. Several changes to the reserve fleet operated on all days were considered for this analysis, proposed by the partner airline. These variations are presented in Table 2. The Baseline scenario consists of the reserve fleet actually maintained by the airline on each day of operations. The letters in the table will be used in the raminder ofthe article as codes for the corresponding scenarios.

Table 2: Reserve Fleet scenarios in analysis and corresponding letters.

Letter	Scenario: Reserve Fleet available
O	Observed: Actual execution of schedule
V	Baseline ($\epsilon = 0$)
B	Baseline ($\epsilon = 50$)
$\mathbf{B + A2}$	Baseline + one A220 ($\epsilon = 50$)
$\mathbf{B + A3}$	Baseline + one A320 ($\epsilon = 50$)
$\mathbf{B + A5}$	Baseline + one A220 + one A320 ($\epsilon = 50$)
$\mathbf{B - A2}$	Baseline - one A220 ($\epsilon = 50$)
$\mathbf{B - A3}$	Baseline - one A320 ($\epsilon = 50$)
$\mathbf{B - A5}$	Baseline - one A220 - one A320 ($\epsilon = 50$)

Among these scenarios that were tested, it is important to reiterate their relevance for this study. The observed execution of the schedule (O), obtained from the partner airline, allows for a simple extraction of all performance indicators that will be analysed and serves as a term of comparison for all obtained results. In order to ascertain whether the proposed framework is able to mimic that same execution of the schedule, the model is run with ϵ set to 0 (scenario V), disallowing any tail swaps (compared with executed operations). The model is also run with ϵ set to 50 (at most 50 tail swaps compared to executed operations), in order to have a term of comparison for the remaining scenarios introducing changes in the reserve capacity. Excluding this last scenario B from analysis would lead to an overestimation of the impact of the proposed changes to the reserve fleet on the airline's operations, as this exclusion would prompt a comparison between the results obtained with varying reserve fleets and the one obtained without allowing the model to complete tail swaps (thus optimising the execution of the schedule with the full knowledge of the disruptions of the day). The remaining scenarios refer to the model runs with changes to the reserve fleet operated.

A graphical representation of the total disruption costs calculated with the model for each of these reserve fleet adjustments, in each of the days of the week, is presented in Figure 5, along with the observed disruption costs for comparison. These costs, as previously stated, are expressed in terms of the negative Net Promoter Score (NPS) impact of all passenger's disruptions. As the focus is on the negative impact, this metric can range from 0 to 100 per passenger. Thus, the fact that results have an order of magnitude of 10^5 means that the number of dissatisfied customers has an order of magnitude of at least 10^3. These numbers include the passengers facing delays and cancellations on the airline's dynamic network, along with all passengers facing missed connections in the full network.

A daily analysis is pertinent for this analysis due to the fact that the the day-of-week cycle is the largest source of passenger demand variation in the airline industry (Swan, 2002), as is observed in Figure 5. Similarly to the previous subsection, each day's two leftmost bars serve as validation for the model, as the first bar presents observed disruption costs and the second bar represents the model's resulting disruption costs when ϵ is set to 0. A comparison between the two shows that the model successfully mimics the execution of the schedule during this week of operations. Furthermore, these results are in line with expectations in the sense that disruption costs are always decreased as the reserve fleet is incremented and always increased as it is reduced.

This daily breakdown of disruption costs per scenario allows for a detailed and insightful perspective. For example, on Tuesday and Wednesday (two least busy days in the airline industry) of this peak week of operations, results show that the removal of an aircraft of the A220 family from the reserve fleet maintained does not involve a significant increase in disruption costs. On the other hand, reinforcing the reserve fleet with this type of aircraft on both days provides almost no value to the airline. It might be thus beneficial for the airline to decide to operate a reserve fleet dependent on day of the week, where during these days, one of the aircraft of the A220 family normally maintained as reserve can operate extra scheduled flights, generating more revenue for the company.

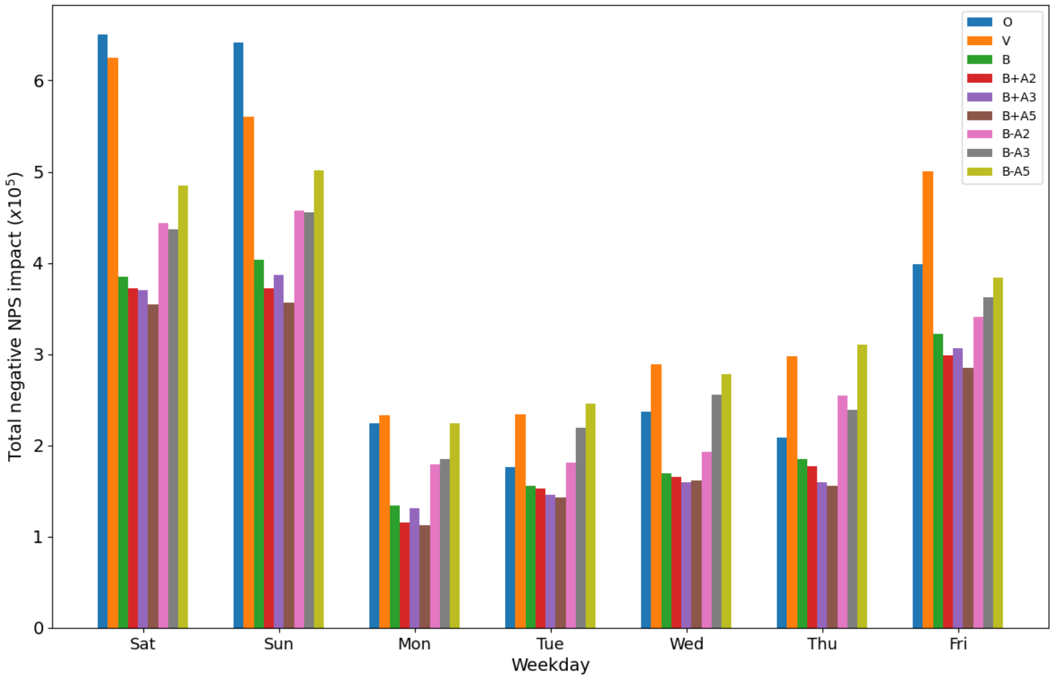

Figure 5: Daily Total Disruption Costs during peak week of operations.

To enable an analysis of the full week, a graphical representation of the average disruption costs calculated with the model for each of these reserve fleet adjustments is presented in Figure 6, along with the observed disruption costs for comparison. As can be seen, the difference between the observed disruption costs and the validating scenario V is only of 7.9%, which validates the model as discussed in the analysis of Figure 4. This representation of the results obtained allows for an informed choice regarding which aircraft should be prioritised when deciding to reinforce/reduce the size of the reserve fleet maintained. In this case, adding an A220 aircraft to the reserve fleet would result in a slightly higher reduction of disruption costs than the one achieved with an added A320 aircraft. On the other hand, removing an A320 aircraft from the reserve fleet would result in a significantly higher increase of disruption costs than the one achieved with one less A220 aircraft as reserve. Therefore, when considering changes to the reserve fleet operated, the airline should prioritise alterations on the A220 portion of the fleet. As expected, the most costly reserve fleet scenario is the one with the least amount of aircraft.

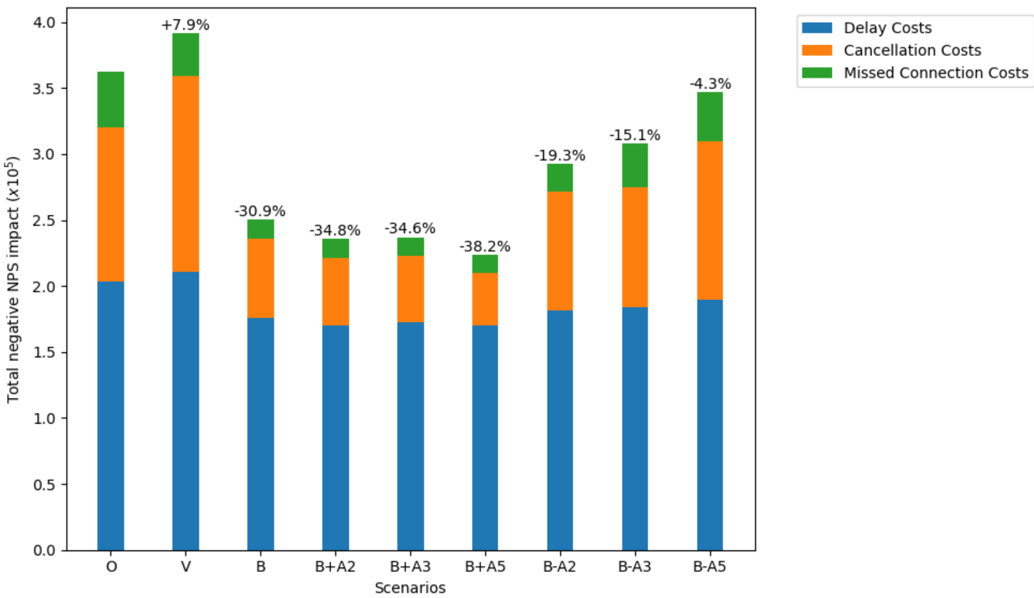

Figure 6: Average Total Disruption Costs for all reserve fleet scenarios during peak week of operations.

These results also provide the opportunity to determine the impact of considering a static approach to disruption management, where disruptions are known at the start of the day. As can be seen in Figure 6, not considering the dynamic nature of disruptions leads to an underestimation of disruption costs of on average 31%, in line with previous results obtained by Vos et al. (2015) and Vink et al. (2020).

For further analysis, Table 3 contains the same results for each scenario, along with the difference in percentage when compared with the observed disruption costs and when compared with the model results, considering the reserve fleet actually maintained by the airline during this week and ϵ set to 50. When considering increases to the size of the reserve fleet operated, these results allow for a particularly insightful analysis. Focusing on scenario B+A2, the results obtained can be interpreted in the following manner: the maximum impact that can be obtained from this change in the reserve fleet is a reduction of 34.8% (column ΔO) in passenger's disruption costs, since, as previously discussed, the model possesses knowledge of all disruptions of the day at its start, which is not the case in real operations; the minimum impact that can be obtained from this change in the reserve fleet is a reduction of 5.7% (column ΔB) in passengers' disruption costs, as the inclusion of the extra aircraft as reserve enabled the model to achieve this improved performance when compared with the same model with no extra reserve capacity. To elaborate on this last insight, it has been shown that the proposed framework underestimates disruption costs, since disruptions are known at the start of the day. Because of this, in all scenarios the reserve fleet's use is optimised throughout the day of operations to an extent that would be impossible to mimic in real operations. As such, the value proposition of an altered reserve fleet according to the proposed framework (when compared with scenario B) is underestimated.

Table 3: Disruption costs comparison for all reserve fleet scenarios.

Scenario	Total Disruption Costs ($\times 10^5$)	$\Delta O(\%)$	$\Delta B(\%)$
O	3.62	-	-
V	3.91	+7.9%	-
B	2.50	-30.9%	-
B + A2	2.36	-34.8%	-5.7%
B + A3	2.37	-34.6%	-5.3%
B + A5	2.24	-38.2%	-10.6%
B − A2	2.93	-19.3%	+16.9%
B − A3	3.08	-15.1%	+20.5%
B − A5	3.47	-4.3%	+38.5%

As the costs used for this study were measured in terms of negative Net Promoter Score (NPS) impact, it becomes imperative to contextualise these figures by comparing them to the estimated financial benefits provided by Network Planning. This comparison entails evaluating the potential profits generated if the airline were to incorporate additional aircraft into normal operations, thereby operating more flights with paying passengers, in contrast to maintaining the aircraft as reserve capacity. Such an analysis would enable the airline to make informed decisions regarding the financial justification of their reserve fleet investment, assessing the cost-effectiveness of different allocation strategies.

The impact of the proposed changes to the reserve fleet was also measured in terms of the airline's 3 key performance indicators, which are used to measure performance in the AOCC: punctuality, heavy delay index, percentage of missed connections. The first is defined as the percentage of flights in the dynamic network with a departure delay of less than 15 minutes. It should be noted that flights in the static network are not included, as their times of Departure and Arrival are obtained from the observed execution of the schedule, since possible changes in the reserve fleet are assumed to not impact their performance. The second is defined as the percentage of flights in the dynamic network with an arrival delay equal to or higher than 30 minutes. The third is defined as the percentage of connecting passengers in the full network (static + dynamic) who are able to make their connection.

The evident deviations between the observed performance indicators and the ones resulting from scenario V can be attributed to the discretisation of time in the Time-Space Networks created to simulate the movement of the dynamic fleet. This discretisation significantly impacts all three of these indicators, given their inherently discrete nature. They are all measured based on specific time thresholds (15-minute departure delay for punctuality, 30-minute arrival delay for heavy delay index and 40 minutes connecting time at hubs). It is essential to recognise that these performance indicators operate on a different scale compared to disruption costs, which may be continuous in nature. As a result, the impact of discreteness on these performance indicators can be more pronounced, leading to larger differences between simulated and observed outcomes. As such, the remaining analyses will focus on the differences observed between the varying reserve fleet scenarios (from B onwards).

The results for the punctuality metric for all scenarios are shown in Figure 7. These results are in line with expectations in the sense that punctuality is always increased as the reserve fleet is incremented and always decreased as it is reduced. While the addition of aircraft to the reserve fleet can make a small impact on punctuality (increase of 1% in case of an A220 and 0.5% in case of an A320), there is evidently almost no improved performance in this metric when comparing the scenario with two extra aircraft to the added A220 scenario. Bigger deviations are seen when considering reduced reserve capacity. Removing an A220 aircraft from the reserve fleet is preferable (when compared with an A320) since the first option results in a punctuality almost 2% higher than the latter. Finally, it can be seen that removing the two aircraft types from the reserve fleet would result in a punctuality decrease of over 3%.

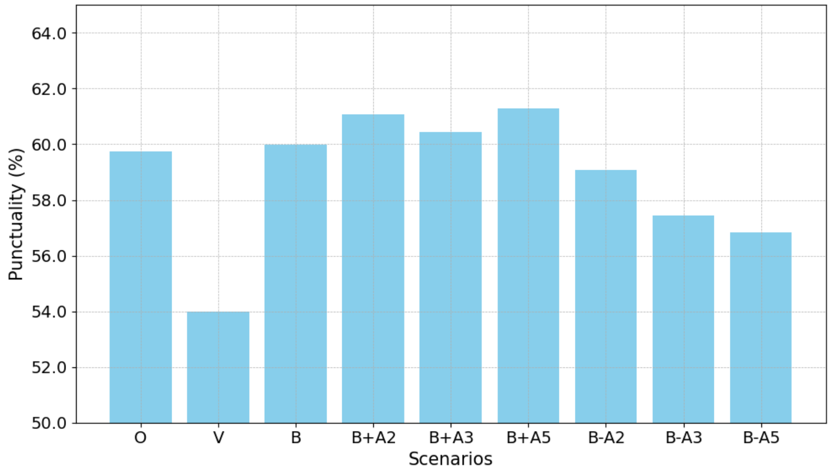

Figure 7: Punctuality for different Reserve Fleet scenarios.

The heavy delay index metric obtained for all scenarios is represented in Figure 8. These results are generally in line with expectations in the sense that that heavy delays increase as the reserve fleet is reduced and decrease as it is incremented. When comparing the two rightmost bars, however, it can concluded that removing two aircraft from the reserve fleet results in fewer heavy delays compared to the scenario where only one A320 is removed. While this deviation may seem unusual, it does not indicate an error. Rather, it underscores a key aspect of the mathematical formulation analysed in Section 4.4. The objective of this formulation is to minimise total disruption costs. Therefore, this discrepancy suggests that achieving this objective with a reserve fleet reduced by one A220 and one A320 aircraft may yield more favourable outcomes in terms of heavy delay reduction compared to simply removing one A320 aircraft. Upon further analysis, while the addition of aircraft to the reserve fleet can make a small impact on heavy delays (decrease of 0.5% in case of an A320 and slightly higher in case of an A220), there is evidently no improved performance in this metric when comparing the scenario with two extra aircraft to the added A220 scenario. On the other hand, reducing the A320 portion of the reserve fleet by one aircraft results in an increase in the amount of heavy delays of over 1%. This increase is slightly lower when considering the reserve fleet option with the least reserve aircraft, as previously discussed. Removing one A220 results in an almost identical heavy delay index to the one observed with the baseline reserve fleet.

The results for percentage of missed connections for all scenarios are shown in Figure 9. This indicator suffers no significant changes when considering increased reserve capacity. However, a slight increase in missed connections is noticeable (<0.2%) when introducing an extra A220 aircraft in the reserve fleet. This discrepancy can be attributed to similar factors as observed in the analysis of heavy delay index. However, reducing the available reserve fleet results in an increase in this percentage of almost 1% (in case of one less A220), over 2% (in case of one less A320) and almost 2.5% (in the minimum reserve fleet scenario considered). For a hub-and-spoke carrier such as the partner airline, with a significant portion of customers connecting on to additional flights, it is important to consider these numbers when deciding whether to reduce the reserve fleet operated.

All of the preceding analyses provide valuable insights for guiding future decisions regarding potential changes to the reserve fleet operated by the airline. Notably, the examination of the three analysed airline performance indicators reinforces the conclusions drawn from the analysis of disruption cost impacts associated with variations in reserve fleet composition. Specifically, it becomes evident that when contemplating adjustments, the airline should prioritise modifications to the A220 portion of its reserve fleet. This targeted approach would ensure that operational resilience and performance optimisation efforts are strategically aligned with the most impactful areas of improvement within the fleet.

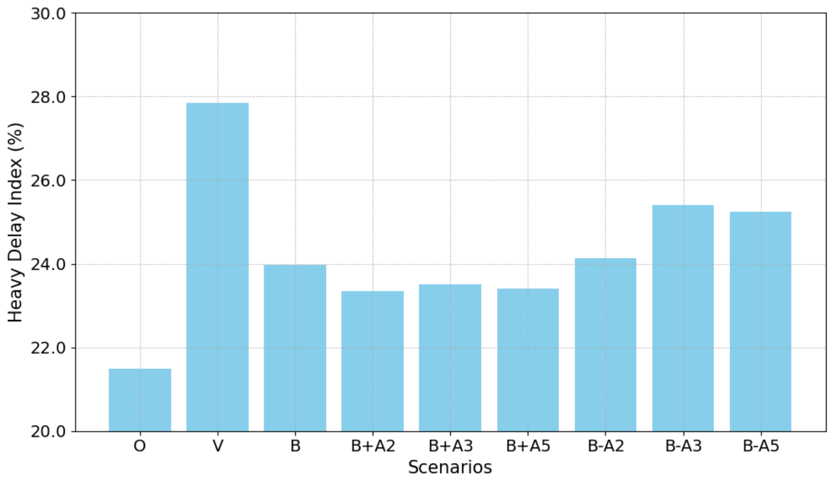

Figure 8: Heavy Delay Index for different Reserve Fleet scenarios.

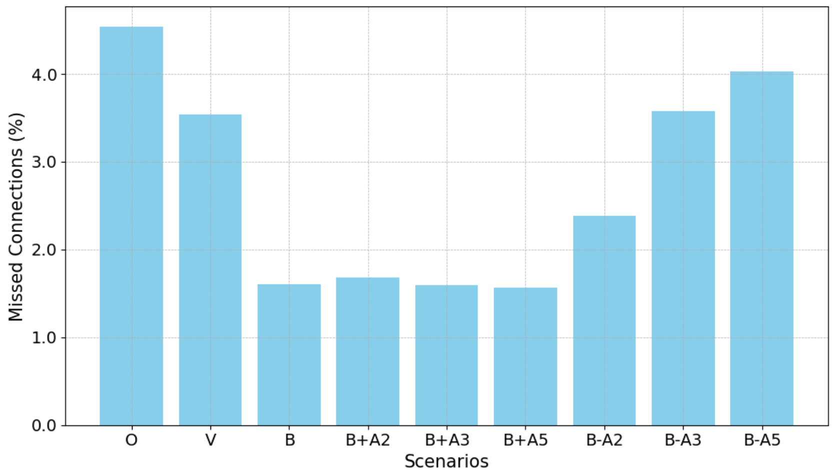

Figure 9: Percentage of missed connections for different Reserve Fleet scenarios.

6 Conclusions

In this paper, a novel post-operational framework is introduced to assess the repercussions of adjustments made to an airline's reserve fleet on previously executed operations. By incorporating data pertaining to the day's flight schedule, disruptions encountered, and passenger itineraries, this research offers a comprehensive analysis of the impacts of such modifications on incurred delays, cancellations and missed connections. This work represents a significant departure from traditional approaches by leveraging insights derived from executed operations, thus enabling a thorough examination of disruptions experienced on past operational days. Such an approach facilitates a comparison with the actual execution of the schedule, ensuring a deeper understanding of the implications of fleet adjustments on operational key performance indicators.

A case study developed in collaboration with a hub-and-spoke European carrier is presented to validate the framework and show its capabilities. The case study considers a fleet of over 110 aircraft (operating over 500 flights), and it investigates the effects of proposed variations to the reserve fleet on one week of the airline's past operations during the peak of the Summer Schedule. The impact of these changes is quantified in terms of disruption costs, punctuality, heavy delay index, and percentage of missed connections. Simulated results of all reserve fleet scenarios considered are compared with recorded passenger disruption data from the same period, demonstrating the model's capacity to emulate historically observed operational performance. Notably, the model exhibits an average deviation of 7.9% in disruption costs when compared with recorded data. When analysing a daily breakdown of the week under review, the obtained results suggest potential benefits for the airline in tailoring its reserve fleet operations based on the specific day of the week. By leveraging this insight, the airline could strategically deploy one of its reserve aircraft for additional scheduled flights during select days of the week, thereby augmenting revenue generation opportunities. Upon examining results for the average performance throughout the week per scenario, all performance indicators highlight a significant finding: the framework was able to identify the specific aircraft type that should be prioritised for adjustments within the airline's reserve fleet.

7 Discussion and Future Work

Three main limitations are identified in this research, by decreasing order of impact: static disruption approach, crew availability assumption, and static passenger itineraries. Firstly, the static disruption approach employed in this study assumes that disruptions affecting the airline's operations are known at the start of the day. However, in real-world scenarios, disruptions occur dynamically throughout the day, requiring airlines to adapt their strategies in real-time. Secondly, this research considered complete crew availability for any tail swaps and operations of reserve aircraft throughout the day. In reality, crew availability is subject to various factors such as duty hour regulations, rest requirements, and unforeseen absences. Thirdly, the static nature of passenger itineraries assumed in this research overlooks the dynamic nature of passenger-facing disruptions. In practice, passengers facing disruptions such as flight cancellations or delays are often rebooked or rerouted onto alternative flights to minimise inconvenience. Ignoring this aspect may lead to inaccurate assessments of disruption costs and operational performance.

This work has shown that not considering the dynamic nature of disruptions leads to an underestimation of disruption costs of on average 31%, in line with previous results obtained by Vos et al. (2015) and Vink et al. (2020). Despite the fact that this work developed a disruption management model as well, its ultimate goal was to compare performance achieved with varying levels of reserve fleet. To achieve this, it was vital to consider the full extent of disruptions faced on past days of operations, instead of considering several representative disruption scenarios to analyse the performance of the model created. The novel approach presented in Section 4.3 to consider all delays happening throughout the dynamic network was proposed in order to simulate exactly the full extent of disruptions in the network. However, this approach came at the cost of prohibiting the simulation of the dynamic nature of disruptions. As such, future research could try to identify the original sources of the delays occurring in the network (of the past days of operations analysed), in order to use them as basis for a dynamic introduction of disruptions as in Vink et al. (2020)'s work.

Expanding the model to incorporate considerations of crew availability would afford a deeper understanding of its consequential impact on operational performance. By integrating this additional constraint, the study's outcomes would attain greater realism, aligning more closely with the complexities inherent in real-world airline operations. Moreover, the establishment of algorithms or methodologies geared towards dynamically simulating passenger rebooking and rerouting processes would substantially enhance the model's fidelity and precision in assessing disruption costs and operational performance metrics. This future work would play an important role in enabling a more nuanced differentiation between disruptions that can be effectively mitigated by the airline and those that result in stranded passengers, who cannot reach their desired destinations on their intended date.

References

Belobaba, P., Odoni, A., Barnhart, C. (2015). *The global airline industry.* John Wiley & Sons.

Castro, A.J.M., Rocha, A.P., Oliveira, E. (2014). *A new approach for disruption management in airline operations control.* Vol. 562. Springer.

Cook, A.J., Tanner, G. (2015). "The cost of passenger delay to airlines in Europe-consultation document". In.

Gershkoff, I. (2016). "Shaping the future of airline disruption management (IROPS)". In: *Commissioned by Amadeus.*

Groeneveld, D. (2021). "Reserve fleet capacity". In: URL: https://repository.tudelft.nl/islandora/object/uuid:6fe80463-d6f3-470e-a360-9dfdbd92f521?collection=education.

Hassan, L. (2019). "Aircraft Disruption Management: Increasing Performance with Machine Learning Predictions". In: URL: https://repository.tudelft.nl/islandora/object/uuid:07efd193-1d6a-43ad-9461-ccc885dfcb9a?collection=education.

Kohl, N., Larsen, A., Larsen, J., Ross, A., Tiourine, S. (2007). "Airline disruption management—Perspectives, experiences and outlook". In: *Journal of Air Transport Management* 13.3, pp. 149–162. ISSN: 0969-6997. DOI: https://doi.org/10.1016/j.jairtraman.2007.01.001. URL: https://www.sciencedirect.com/science/article/pii/S0969699707000038.

Korneta, P. (2018). "Net promoter score, growth, and profitability of transportation companies". In: *International Journal of Management and Economics* 54.2, pp. 136–148. DOI: doi:10.2478/ijme-2018-0013. URL: https://doi.org/10.2478/ijme-2018-0013.

Marla, L., Vaaben, B., Barnhart, C. (2017). "Integrated Disruption Management and Flight Planning to Trade Off Delays and Fuel Burn". In: *Transportation Science* 51.1, pp. 88–111. DOI: 10.1287/trsc.2015.0609. eprint: https://doi.org/10.1287/trsc.2015.0609. URL: https://doi.org/10.1287/trsc.2015.0609.

Rosenberger, J.M., Johnson, E.L., Nemhauser, G.L. (2004). "A Robust Fleet-Assignment Model with Hub Isolation and Short Cycles". In: *Transportation Science* 38.3, pp. 357–368. DOI: 10.1287/trsc.1030.0038. eprint: https://doi.org/10.1287/trsc.1030.0038. URL: https://doi.org/10.1287/trsc.1030.0038.

Swan, W.M. (2002). "Airline demand distributions: passenger revenue management and spill". In: *Transportation Research Part E: Logistics and Transportation Review* 38.3, pp. 253–263. ISSN: 1366-5545. DOI: https://doi.org/10.1016/S1366-5545(02)00009-1. URL: https://www.sciencedirect.com/science/article/pii/S1366554502000091.

Teodorović, D., Guberinić, S. (1984). "Optimal dispatching strategy on an airline network after a schedule perturbation". In: *European Journal of Operational Research* 15.2, pp. 178–182. ISSN: 0377-2217. DOI: https://doi.org/10.1016/0377-2217(84)90207-8. URL: https://www.sciencedirect.com/science/article/pii/0377221784902078.

Varenna, S. (2023). "A Stochastic Discrete Event Simulation of Airline Network and Maintenance Operations". In: URL: https://repository.tudelft.nl/islandora/object/uuid%5C%3Ae2de25a7-971e-4575-b2d9-e38392ff25ee?collection=education.

Vink, J., Santos, B.F., Verhagen, W.J.C., Medeiros, I., Filho, R. (2020). "Dynamic aircraft recovery problem - An operational decision support framework". In: *Computers Operations Research* 117, p. 104892. ISSN: 0305-0548. DOI: https://doi.org/10.1016/j.cor.2020.104892. URL: https://www.sciencedirect.com/science/article/pii/S0305054820300095.

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* 10. 18th Euro Working Group on Transportation, EWGT 2015, 14-16 July 2015, Delft, The Netherlands, pp. 931–940. ISSN: 2352-1465. DOI: https://doi.org/10.1016/j.trpro.2015.09.047. URL: https://www.sciencedirect.com/science/article/pii/S2352146515002343.

Walker, C. (2019). "All-Causes Delay and Cancellations to Air Transport in Europe: Annual Report for 2019". In: *Eurocontrol: Brussels, Belgium.* URL: https://www.eurocontrol.int/archive_download/all/node/12132.

II

Literature Study
previously graded under AE4020

1

Introduction

The airline planning process consists of multiple steps that are performed sequentially over a period of decades. These steps can be classified based on the time frame between the decision-making point and the day of operations, which can span from long-term planning to short-term execution. Another perspective for categorisation is based on the business impact, with decisions having implications at strategic, tactical, and operational levels. All the steps that are involved in airline planning are depicted schematically in figure 1.1.

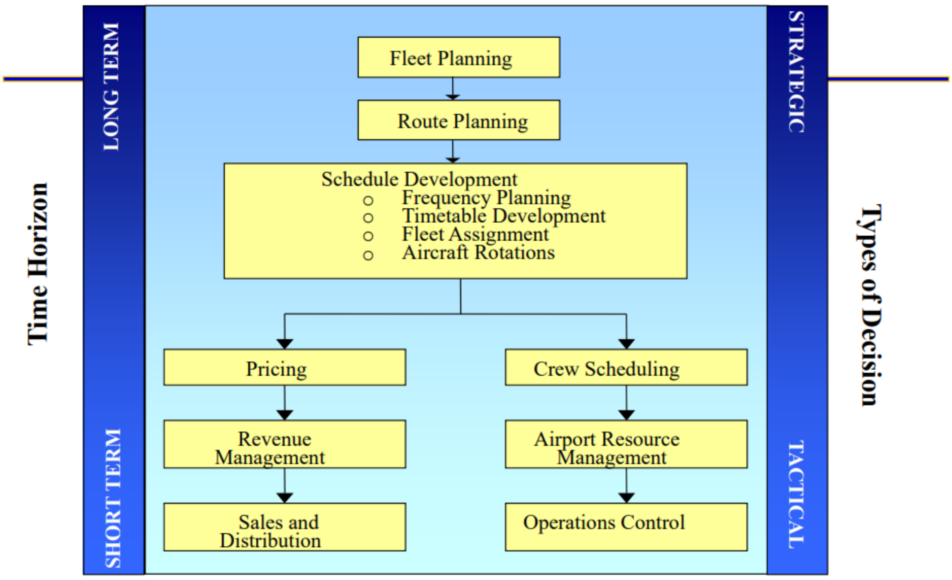

Figure 1.1: The Airline Planning Process (Belobaba et al., 2015).

On the operational level, daily operations in the airline industry are hindered by differences between scheduled and actual performance. Poor weather conditions, congestion at hubs, and aircraft mechanical problems are just a few of the causes that prevent airlines from operating their flight schedules as planned, which often result in flight cancellations, delays, passenger missed connections and even diversions. These disruptions create great expenses for airlines and passengers alike and disrupt the original schedule of the airline, preventing a regular day of operations. Much of the work carried out at the Airline Operations Control Centre (AOCC) involves dealing with these disruptions, repairing flight schedules, aircraft schedules, crew schedules and passenger itineraries, and making sure that the inevitable increase to the airline´s operational costs (e.g. additional crew overtime, increased fuel usage, passenger delay compensation or re-accommodation costs) are minimized as much as possible. One of the tools that are available for airlines to decrease the effect of these disruptions is their reserve fleet capacity, and while this capacity is not usually part of airlines' long term fleet planning strategy, its effective use can make a substantial difference by decreasing delays, enabling flights to take place when their assigned aircraft becomes suddenly inoperable, and thus increasing customer satisfaction and decreasing operational costs. However, this obviously represents a big investment, since airplanes' lack of utilization in the traditional sense of having planned flights involves a certain loss of possible profits that the airline could be making (Eurocontrol, 2019).

Due to the immensely complex nature of disruption management, most airlines have not yet automated this crucial aspect of daily operations to the same extent as crew and fleet scheduling. At the AOCC, humans must determine whether a particular event or sequence of events should be characterized as a disruption the airline must act upon. Typically this will be triggered by a message from computerized systems consolidating operational data, but the role of the AOCC operators is to determine whether events are sufficiently serious to require an action and also to define the scope of the disruption. Several apparently unrelated events may for example be relevant to be treated together as one disruption because the same resources are likely to be involved in its' resolution. The operators' decision in many cases will be partly dependent on information and judgments that will not exist in computer systems. The industry's current disruption management strategy is to only use computer systems to process the large amounts of operational data and present potentially critical events for a human, generating and evaluating possible solutions for a given disruption (Kohl et al., 2007).

As mentioned, most airlines still rely on human AOCC operators to handle daily disruptions. These past decisions, if properly documented, provide an opportunity to study the effectiveness of past uses of airlines' reserve fleet and to study the effect of different fleet compositions on past daily operations. The proposed research aims to make a significant contribution to this field by optimising reserve fleet in order to maximise airlines' financial profit and concluding on future years' various reserve fleet compositions' resulting KPIs. Moreover, the hereby presented literature study has allowed to get an overview of the current state of the art in topics related to reserve fleet planning, needed for the identification of key research gaps that could be further explored in the master's thesis.

1.1. Research Definition

To create a model that structurally incorporates reserve fleet capacity into the airline planning process, an overview of the existing literature was made. This literature review aims to provide a comprehensive analysis of airline fleet planning, disruption management and reserve planning in other domains so that valuable insights for further research in the field can be extracted. To guide the investigation, a series of research questions and sub-questions were set:

1. How is reserve fleet incorporated in the airline fleet planning problem?
    (a) How is the AOCC work carried out and which tools do they have available for solving disruptions?
    (b) Which techniques have been used to determine airlines' optimal fleet composition?
    (c) How have the effects of reserve fleet on airlines' results been studied?
2. What are the state-of-the-art disruption management tools?
    (a) How have the Aircraft Recovery and Integrated Recovery Problems been solved?
    (b) How have these tools modelled the propagation of delays in the network?
3. How is reserve capacity planning processed in other domains?
    (a) What are some commonly employed techniques in domains that share certain similarities with aircraft reserves?
    (b) How can these techniques be applied to reserve fleet planning in the context of airline operations?
4. What type of forecasting method is indicated to predict financial savings achieved in daily operations?

(a) What are some commonly employed forecasting methods and the distinctive characteristics that set them apart in terms of accuracy, applicability, and predictive power?

(b) How can the resulting predictions be validated?

1.2. Report Structure

The literature review aims to provide a comprehensive overview of airlines' reserve fleet planning, focusing on the various methods employed, their results, and key conclusions drawn. Chapter 2 delves into the airline fleet planning problem, highlighting the existent studies on reserve capacity's effect on airline performance. On the other hand, chapter 3 outlines the different solutions in airline disruption management, providing insight as well into how delay propagation has been previously modelled. Chapter 4 describes the methods that have been used for reserve planning in other domains. Finally, chapter 6 synthesizes the current deficiencies and opportunities for research in the field, from which research questions and objectives for the project will be derived.

A summary of the stated report outline is given in Table 1.1, including the research questions (see section 1.1) addressed in each chapter.

Table 1.1: Report outline including the research questions assessed in each chapter.

Chapter	**Content**	**Research Questions**
1. Introduction	Background and Research Questions	-
2. Airline Fleet Planning	Fleet Composition Problems and Reserve Fleet Capacity	1a & 1b & 1c
3. Airline Disruption Management	Airline Recovery Problem and Delay Propagation	2a & 2b
4. Reserve Capacity Planning in other domains	Reserve Planning in other industries and Reserve Crew Planning	3a & 3b
5. Forecasting	Machine Learning forecasting techniques and accuracy evaluation	4a & 4b
6. Conclusions	Current State-of-the-art summary highlighting research gap	-

2
Airline Fleet Planning

Airline fleet planning involves a complex process of determining the optimal size and composition of an airline's fleet to meet current and future demand. In this chapter, the current state of literature on airline fleet planning, and the extent to which reserve capacity planning is addressed, will be explored. By doing so, the first (***How is reserve fleet incorporated in the airline fleet planning problem?***) research question will be addressed.

2.1. Fleet composition problems

One of the keys to achieve airline success is bringing supply and demand as closely together as possible. In order to achieve this goal, an appropriate methodological approach for the fleet planning process needs to be followed. Fleet selection is an integral part of long-term planning process for airlines. The topic is addressed here to ascertain whether the existing literature considers aircraft to be used as reserves, and how the stochastic nature of passenger demand is dealt with, as that characteristic can be very similar to the stochastic nature of airline disruptions. Fleet selection is based on an iterative process of forecasting demand, considering aircraft types and creating a sustainable financial planning. The decisions made during this process lead up to the creation of a fleet plan. Airlines operate their fleet of aircraft over a relatively long time horizon during which the realized stochastic demand has the potential to profoundly impact the airlines' financial performance. This makes this plan a highly capital-intensive long-term commitment, associated with inherent risks.

Shube and John W Stroup, 1975 were the first to create a multi-year linear programming model to obtain optimal fleet compositions, and Schick and J. Stroup, 1981 reported their experience with the original version and its derivatives. The original model optimized the non-homogeneous fleet planning of a period of 10 years considering the expected demand. The suggested objective functions aim to minimize either DOC (direct operating cost) of all scheduled operations over the full 10 year period, the amount spent for new equipment, or the sum of DOC and cost of additional capital resulting from transactions implied by the fleet plan (i.e. debt), while satisfying a set of constraints regarding demand satisfaction, minimum and maximum flight frequencies per route, aircraft balance constraints and purchase constraints related to aircraft availability and maximum indebtedness. During that specific time period, the utilization of linear programming to solve mathematical models was relatively new. The primary focus of Schick and J. Stroup, 1981 centered around the applicability of a computerized fleet planning model based on linear programming to the airline industry. They noted that while the model generated realistic solutions by planning the fleet composition over multiple years and meeting demand, capacity, and frequency constraints, a key consideration was whether the airline industry, which traditionally approached fleet planning manually from a bottom-up perspective, was prepared to embrace this top-down computerized approach utilizing linear programming.

Santos, 2021 proposed a model for the fleet planning problem where the initial assumption is that the future demand per route is known (i.e., deterministic). The main objective of the model is to make decisions regarding aircraft acquisition and optimal assignment of aircraft to specific routes. This decision-making process takes into account various factors such as route demand, aircraft characteristics (e.g., ownership cost, cruise speed, seating capacity, block times), and the goal of maximizing profit.

The model formulates an objective function that thus seeks to maximize profit by considering both revenue and ownership costs. The leg-based mathematical model is based on the maximisation of profit and subject to six constraints: demand verification, capacity, continuity, aircraft productivity, range and runway. The following equations below describe the general model.

Objective function

The objective function is defined by equation 2.1.

$$\max(profit) =$$

$$\sum_{i\in N}\sum_{j\in N}\left[Yield * d_{ij} * (x_{ij} + w_{ij}) - \sum_{k\in K}(CASK^k * d_{ij} * s^k * z^k_{ij})\right] - \sum_{k\in K}(AC^k * C^k) \tag{2.1}$$

In equation 2.1, N forms a set of airports and K a set of aircraft types, with z^k_{ij} the number of flights from airport i to airport j completed with aircraft type k, x_{ij} the direct flow in number of passengers from airport i to airport j, w_{ij} the flow in number of passengers from airport i to airport j with transfer at the hub airport, and d_{ij} the distance between airports i and j.

The yield equals the revenue per RPK (revenue passenger kilometer) flown for aircraft type k. The $CASK^k$ is the unit operation cost per ASK (available seat kilometers) flown, s^k signifies the number of seats per aircraft, AC^k the number of aircraft type k in the fleet, and C^k the weekly lease cost.

Constraints

Demand The demand verification considers that the number of passengers should be less than the demand, as shown in equation 2.2, where q_{ij} is the travel demand between airports i and j.

$$x_{ij} + w_{ij} \leq q_{ij}, \text{ with } i, j \in N \tag{2.2}$$

Additionally, passengers only transfer at the hub airport if this hub is not their origin or destination, as presented in equation 2.3.

$$w_{ij} \leq q_{ij} * g_i * g_j, \text{ with } i, j \in N \tag{2.3}$$

In this equation g_i or g_j is activated if the hub is airport i or j. This is done in accordance to equation 2.4, with parameter g_a for each airport a.

$$g_a = \begin{cases} 0\ if\ a\ hub\ is\ located\ at\ airport\ a \\ 1\ otherwise \end{cases}, \text{ with } a \in N \tag{2.4}$$

Capacity Equation 2.5 models the constraint of capacity, for which the number of passengers on the aircraft should be lower than the available number of seats on the aircraft, which explains the inclusion of the Load Factor LF.

$$x_{ij} + \sum_{m\in N} w_{im} * (1 - g_j) + \sum_{m\in N} w_{mj} * (1 - g_i) \leq \sum_{k\in K} z^k_{ij} * s^k * LF, \text{ with } i, j \in N \tag{2.5}$$

Continuity The constraint of continuity signifies that the number of inbound aircraft should be equal to the number of outbound aircraft per airport, as shown in equation 2.6.

$$\sum_{j\in N} z^k_{ij} = \sum_{j\in N} z^k_{ji}, \text{ with } i \in N \text{ and } k \in K \tag{2.6}$$

Aircraft Productivity The aircraft productivity constraint, modelled in equation 2.7, considers that the number of hours of operation of an aircraft must be less than the average available hours of operation of that same aircraft. In this equation, sp^k resembles the speed of an aircraft type k, TAT^k the Turn-Around-Times of an aircraft type k, including landing and take-off times, and BT the average daily aircraft utilisation time in hours.

$$\sum_{i\in N}\sum_{j\in N}(\frac{d_{ij}}{sp^k} + TAT^k * z^k_{ij} \leq BT * 7 * AC^k, \text{ with } k \in K \tag{2.7}$$

Range Serving the purpose of constraining flight frequency to aircraft range limits, the range constraint is shown in equation 2.8, where R^k signifies the range of an aircraft type k.

$$z_{ij}^k \leq a_{ij}^k, \text{ with } a_{ij}^k = \begin{cases} 10000 \text{ if } d_{ij} \leq R^k \\ 0 \text{ otherwise} \end{cases}, \text{ with } i, j \in N \text{ and } k \in K \tag{2.8}$$

Runway In order to constrain flight frequency to airport runway limitations, the runway constraint is shown in equation 2.9, where rw^k signifies the required runway length of an aircraft type k, in order to safely take off and land, and rw_a is the runway length of airport a, both in meters.

$$z_{ij}^k \leq b_{ij}^k, \text{ with } b_{ij}^k = \begin{cases} 10000 \text{ if } rw^k \leq min(rw_i, rw_j) \\ 0 \text{ otherwise} \end{cases}, \text{ with } i, j \in N \text{ and } k \in K \tag{2.9}$$

Hsu et al., 2011 used dynamic programming to model a multistage stochastic airline fleet planning problem. In order to propose an optimal replacement schedule, a Grey forecasting model was used to develop a Markov-chain process to model the stochastic demand on a market level and a market share estimation that is only frequency dependent to calculate demand on an airline level. With the goal of optimising the proportion between purchased and leased aircraft in order to mitigate the effects of demand uncertainty, the model built on a given schedule of route flight frequencies and resulted in the conclusion that for markets with high fluctuations in demand, leasing should be prioritized over buying, which is consistent with previous work by Oum et al., 2000, even though Hsu et al., 2011 assume a given initial fleet as basis (Oum et al., 2000 size the fleet without considering an initial existing fleet). The objective function minimizes the cost of three cost terms per period over a multi-period planning horizon: operating costs; aircraft replacement costs; a penalty cost which increases with the inaccuracy level of the forecasted demand. While the utilization of an advanced demand forecasting method that considers cyclical demand represents a significant advancement, Hsu et al., 2011 point out that this approach does not incorporate the impact of random fluctuations in demand as result of, for example, terrorist attacks and aircraft accidents.

Teoh and Khoo, 2015 also note this drawback in earlier attempts in literature, and present a new method for modelling demand uncertainty. According to the authors, demand is more subjected to unexpected events and, not considering the possibility of disruptive events, the models do not resemble reality. In order to better achieve this, the formulate of a stochastic demand index (SDI) is proposed. The development of SDI involves several stages. Initially, potential unforeseen events like disease outbreaks and natural disasters are identified, and their probability distributions are determined based on historical data. Next, a Monte Carlo simulation is employed to model the likelihood of these uncertain events occurring. This information is then combined with a conventional demand forecast, which doesn't consider uncertain events, to calculate a single SDI for each operational period. The resulting SDI is subsequently utilized as an input for an optimisation model in fleet management.

2.2. Reserve fleet capacity

The effects of having reserve fleet capacity on the fleet plan are multiple. Since choices regarding fleet size have to be made made years before schedule execution, a better understanding of the impact of reserve fleet capacity and its effectiveness can contribute towards airlines´ better fleet plan creation and subsequent higher margins. However, research into this particular added capacity that some airlines choose to operate has been very limited. Reserve fleet capacity gives the airline an opportunity to quickly offer more flights in case of higher demand, but it also entails the risk of overcapacity in case of decreasing demand. It is a tool that provides airlines the means to add robustness to their schedules, by mitigating delays, enabling passenger connections that would otherwise not be possible, and substituting inoperable aircraft in need of unexpected maintenance. All of these consequences have a positive impact in the Quality of Service Index (QSI) of airlines, which values an airline's set of flights offered in an O-D market, relative to competitors, and is used to estimate the market share potential of new routes and incremental flights (Belobaba et al., 2015). In addition to this improvement of passenger experience, the potential decrease of secondary costs associated with cancellations, missed connections, and delays (such as delay compensation or rebooking of passengers) needs to be considered along with the increase of total fleet operation costs in case of additional capacity to ascertain the break even point financially for the airline, helping it to achieve higher margins.

Groeneveld, 2021 studied the impact of reserve fleet capacity in delay mitigation, proposing three relevant case studies to be carried out to study different effects of the operation of reserve fleet. In the first, using Monte Carlo simulations on an existing flight schedule and a rule-based delay mitigation model the effectiveness of reserve fleet capacity (in the form of standby unscheduled aircraft) is analysed- the break even point for having one standby reserve was found to be a fleet size of 48 aircraft, considering a trade-off between additional operational costs and potential saved delay costs (see figure 2.1). In the second, the impact of different delay durations on the effectiveness of reserve fleet capacity is studied. It was found that reserve fleet capacity becomes more effective in high delay scenarios with high outlier delays, which is expected as the operation of the additional capacity allows the airline to prevent the biggest delays in this case, as opposed to choosing one of the similar delays that it can be used to prevent. In the third, a new schedule was used in which one rotation per day would be reallocated to the standby aircraft. The model was then rerun, which resulted in a lower performance of the reserve capacity in terms of efficiency. This result was expected, as with an aircraft available at all times to tackle the biggest operational issue at hand at a given time airlines will in principle save the most delays, but it was important to see how much the difference was since the decrease of operational costs adherent to scheduling their standby aircraft at the position in their schedule where they have predicted to need this capacity should also be considered (bringing the standby unexpectedly into rotation from a far away parking spot is costly). As delay predictions´ accuracy improve this option of spreading reserve fleet capacity over the schedule is suggested to be revisited. An important aspect of this work is that all calculations are made considering a point to point network, and thus does not accurately represent the delay costs of a hub and spoke network, which in all likelihood would benefit more from on time performance (passenger's connections are affected by delays).

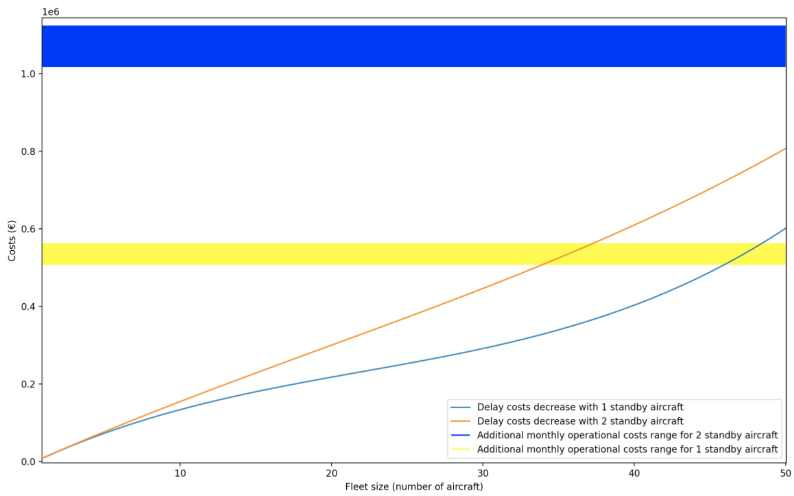

Figure 2.1: Resulting break even points of adding reserve capacity based on total system delay costs (Groeneveld, 2021).

Varenna, 2023 worked on a relevant modular, stochastic discrete event simulation model of airline operations named ANEMOS (Airline Network and Maintenance Operations Simulation), with the objective of constructing a model capable of testing policies, plans, and scenarios involving both network and maintenance operations with the goal of understanding the performance of the system as a whole. The paper details the steps taken to simulate the network and maintenance operations of the intercontinental fleet of hub-and-spoke carriers. As examples of use cases, the author mentions testing the performance of a flight schedule in combination with a maintenance schedule, investigating the effects of anticipating maintenance task execution, evaluating how different recovery policies influence the outcome of disruptions, or evaluating the effects of some external factors such as increased hub congestion on the airline's performance. The case study presented in the paper, to validate the model and show its capabilities, was developed in collaboration with a major European carrier. It considered a fleet of fifty aircraft of four different models, and the effects of adding a second reserve aircraft to the fleet were examined. The impact of the second reserve was quantified and compared to the situation with only one reserve aircraft available in terms of cancellation factor (the percentage of cancelled rotations with the second reserve aircraft decreased to less than half), delay reduction (probability of observing a delay longer than one hour decreased by 0.4% and a delay longer than two hours decreased by 0.1%), and economic value (savings achieved with the added reserve aircraft were estimated to be at most around 10% of the profit the same aircraft could produce for the airline in case of normal operations in the airline's fleet, meaning it would not be economically profitable for the airline to activate a second reserve aircraft), as seen in figure 2.2. Simulated results of the baseline scenario are compared with historical data, and it is shown that the model closely resembles historically observed operational performance.

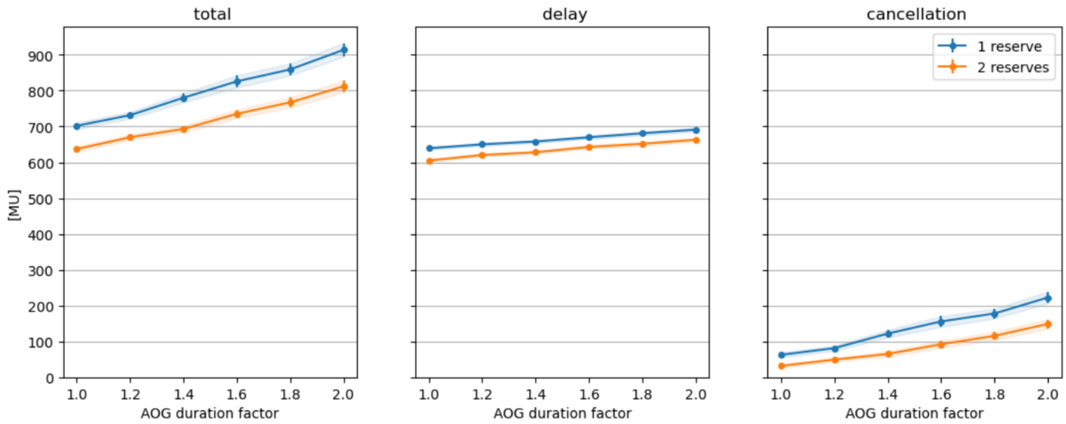

Figure 2.2: Delay, cancellation, and total costs of disruptions for different AOG duration factors and reserve aircraft scenarios (Varenna, 2023).

2.3. Discussion

This chapter provides a brief overview of the different airline fleet planning modelling perspectives. As there are many sources of uncertainty adherent to the problem, a literature review on some of them is presented. However, a research gap is identified, as prior studies have not considered the need for reserve aircraft when planning airline fleets for the future. This, however, is expected as, when planning for the long-term future, airlines do not know the level of disruptions they will face in such a long-term future and thus cannot accurately predict the need for these standby aircraft.

Finally, an overview of reserve fleet capacity studies is provided, with a detailed analysis of Groeneveld, 2021 and Varenna, 2023's work. Both rely on simulated daily operations to investigate the effects of adding an additional reserve aircraft to the airline fleet, which may be a source of possible inaccuracies, since not all resources involved are considered (none of them consider crew constraints). Possible future work in the field could consider past daily operation flight information and records of AOCC actions taken as a baseline scenario that naturally fulfils all constraints (since the events actually happened). Overall, this chapter provides valuable insights into the airline fleet planning problem, later focusing on the specific planning of reserve fleet.

3

Airline Disruption Management

Unexpected events, such as bad weather conditions, mechanical failures or traffic congestion, are unavoidable for airlines and may result in aircraft unavailability and delays. Original schedules created for the aircraft and crew, and passenger itineraries may be disrupted as a consequence. Walker, 2019 noted that over 22% of all flights in Europe were delayed (flights arriving over 15 minutes after Scheduled Time of Arrival) in 2019 as a result of disruptions, after considerable efforts and investments were made by airlines to improve 2018 record high delay levels, resulting in a decrease of almost 2%. In addition, operational cancellations decreased from 2.0% in the previous year to 1.7% in 2019. Despite this slight improvement, 2019 was still the third worst year in the last 10 years, with en-route Air Traffic Flow Management (ATFM) delays during the summer season remaining a problem for airlines, which indicates an upwards trend that will have to be properly dealt with by the airline industry.

In a paper commissioned by Amadeus in 2016, Gershkoff, 2016 revealed the estimated cost of disruptions to airlines to be 8% of airline revenue, or 60 billion USD worldwide, which explains the need for effective disruption management strategies to be in place. The objectives of these strategies are: fulfill the customer promise by ensuring passengers and their luggage reach their intended destination promptly and as per the booked service level; minimize actual costs, which encompass excessive crew expenses, compensation costs, accommodation and lodging expenses for disrupted passengers and crew, as well as ticket expenses with alternative airlines; swiftly return to the original plan and resume operations as quickly as possible (Kohl et al., 2007).

Passengers whose itinerary is delayed (over 3 hours)/cancelled are entitled to compensations according to Regulation (EC) No 261/2004 (A. Cook and Tanner, 2015). The industry refers to these and operational costs as hard costs. In addition to these, however, another category of less measurable costs must be considered, originated by the opportunity cost of a passenger not choosing the same airline in the future due to an endured inconvenience. These costs, known as soft costs, are highly non-linear and constitute a big part of the total disruption costs. Given the complexity of computing disruption costs, these are usually estimated in the literature. Reference values for these estimations are provided by Andrew J Cook and Graham Tanner, 2015.

Most large airlines operate AOCCs to perform on-the-day coordination of schedule execution. Their purpose is to monitor the progress of operations, to flag actual or potential problems, and to take corrective measures in response to unexpected events. In most airlines, representatives of key airline functions work together to ensure smooth schedule execution, in a process that is dynamic and cannot really be formulated explicitly, as shown in figure 3.1. Whenever disruptions occur, a solution must be found to recover operations. In order to decide on which solution to adopt, the passenger, crew and aircraft perspectives have to be taken into account. From the passenger perspective it may, for example, make sense to delay an outbound flight (i.e. a flight departing an airline's hub) to ensure that passengers on a delayed inbound flight will be able to make their connection. This option must then be evaluated from the crew (e.g. they have strict legal limits of sequential hours of work) and aircraft (e.g. the propagation of delays may result in even higher costs later on in the day) perspectives. This process can go until an option that is legal and acceptable from all perspectives, has been found. Once a decision has been made, it is essential to execute it and maintain ongoing monitoring. This disruption management process differs from most control processes in complex systems involving humans, with its' particularly extensive range of available options and the intricate computational challenges associated with evaluating their effects.

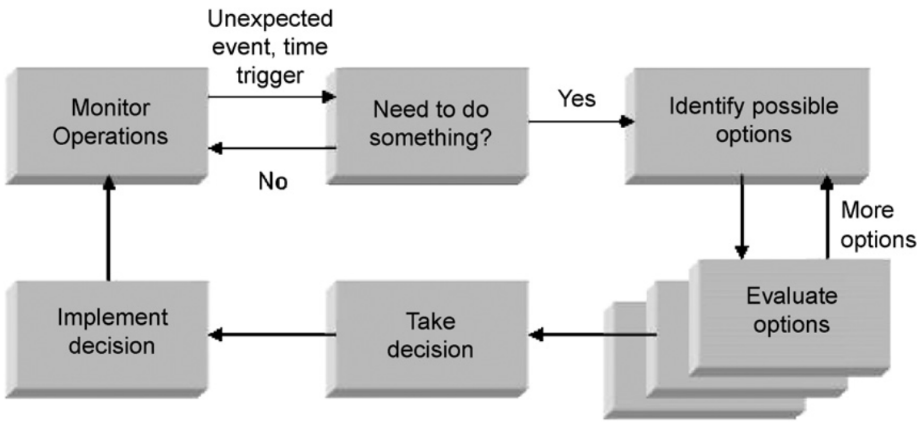

Figure 3.1: High-level view of airlines' Disruption Management process (Kohl et al., 2007)

To identify all possible schedule recovery options, the tools and strategies that the AOCC can use include:

- **Delaying a flight**: the most commonly applied measure, that results in the propagation of the delay in the network.
- **Cancelling a flight or flights**: most often in the form of cancelling rotations (sequence of flight legs that start at a station and end at that same station) (Rosenberger, Johnson, et al., 2004).
- **Exchanging assigned aircraft**: swapping two rotations' assigned aircraft, when more buffer time can be achieved, and when it results in the reduction of total delay and/or cancellation of cheaper flights.
- **Activating a reserve aircraft**: assigning disrupted flight to a reserve aircraft (usually kept on standby at the hub). When previously assigned aircraft arrives to hub (in case the disruption is a delay) it becomes the new reserve.
- **Increasing/decreasing aircraft speed**: important tool for delay absorption, although its costs (additional fuel) must be taken into account (Marla et al., 2017).
- **Ferrying aircraft**: flying an aircraft without passengers to an airport to cover a flight that would otherwise have to be cancelled from that airport.

Airline disruption management has seen an increasing amount of attention since it was pioneered three decades ago by Teodorović and Guberinić, 1984, due to its potential to ameliorate the service of airlines. The initial efforts focused on solving conflicts for a single resource at a time, as they are resolved in the Airline Operations Control Center (AOCC), which is subdivided in different departments, each with a specific area of operations to monitor and to plan (e.g Network Operations Control, Crew Control, Passenger Care Center), where there are human specialists who focus on one resource and take action when necessary, coordinating with the other departments when needed (Castro et al., 2014). This is because it becomes too complex to make decisions regarding multiple resources at once, and so the industry standard is to solve the problem sequentially. While the sequential approach offers a quick resolution, it may not always yield the most optimal outcome due to the interdependence of resources. Although a schedule may be optimal for one resource, it could create conflicts for others. Consequently, recent research has concentrated on developing integrated decision-support tools that simultaneously recover aircraft, crew, and passenger itineraries or a combination of these three, optimising solution quality and reducing costs incurred by the airlines. However, current computer systems still demand significant computational time to solve these models, making it impractical to implement such systems for airlines in their AOCC, where a fleet-level solution is typically required within a tight timeframe of 2 minutes (Vink et al., 2020). The recent progress in machine learning technology holds significant promise for addressing large and intricate problems within reduced timeframes. By harnessing the extensive volume of data, machine learning can provide valuable insights and predictions that expedite the decision-making process. When integrated with optimisation problems, it presents a promising avenue for developing faster and more practical models to recover airline schedules during challenging situations. This combination of machine learning and optimisation represents a compelling direction for advancing the field and finding efficient solutions in the context of airline schedule recovery.

For a more extensive literature review into airline disruption management, the reader is referred to Clausen et al., 2010 and Hassan et al., 2021, which include literature from the period of 1984-2017.

This section aims to provide a comprehensive review of the current knowledge on the airline recovery problem (considering only aircraft and integrating this resource with passengers and/or crew) and delay propagation modelling, highlighting the key challenges, opportunities, and directions for future research. By doing so, the second (***What are the state-of-the-art disruption management solvers?***) research question is addressed.

3.1. Aircraft Recovery

Before 2009, the majority of the literature relating to the airline recovery problem focuses on aircraft and schedule recovery. This focus simplifies the recovery problem, enables a sequential approach, and ensures that the most constraining and expensive resources' use is optimized (Hassan et al., 2021). Nevertheless, aircraft recovery remains a topic of ongoing research, with a particular emphasis on enhancing the complexity of models to accurately reflect real-world networks while concurrently reducing computation time. The objective is to develop more sophisticated approaches that can effectively handle the intricacies of aircraft recovery while still providing timely solutions. Researchers are continuously striving to strike a balance between modeling accuracy and computational efficiency in order to meet the demands of the dynamic aviation industry. In the following subsections the most relevant literature on aircraft recovery will be discussed and classified.

3.1.1. Initial Efforts

Teodorović and Guberinić, 1984 were the first to attempt to restore airline schedules after a disruption with the goal of minimizing total passenger delay. They used branch and bound on a 3-aircraft fleet (assuming only one aircraft type) with 8 flights and ignored maintenance constraints. Passengers are explicitly modeled, but they assume that all itineraries contain only a single flight leg. Teodorović and Stojković, 1990 extended this work by considering operating hours of the airports in the network and tested it on a network of 14 aircraft and 80 flights. A dynamic programming based approach is used where the goal was to minimize the total number of cancelled flights and, in case of an equal number of these between solutions, the total passenger delay. In 1995, the same authors included crew constraints in their work, presenting a heuristic based on the First In, First Out (FIFO) principle and a dynamic programming based sequential approach. The model uses the same objective function as the previous work of the same authors mentioned. The model was tested on over 1000 different randomly generated scenarios with varying fleet and flight numbers and varying disturbance types. All three of these papers were largely based on connection networks, as defined by Clausen et al., 2010.

Since the genesis of the efforts to solve the airline recovery problem, the range of recovery options, disturbances, and network sizes have increased greatly and will continue increasing. While exact optimisation algorithms have been developed, recent literature has increasingly incorporated the utilization of metaheuristics or scope-limiting algorithms to fulfil the tight time constraints of the AOCC.

3.1.2. Exact Optimisation Methods

Rosenberger, Johnson, et al., 2003 present an approach to address the aircraft recovery problem by formulating it as a set-partitioning problem that incorporates airport slot constraints. Their model incorporates an aircraft selection heuristic, which aids in selecting a subset of aircraft to be considered for recovery. This selected subset is then used to generate potential new routes, considering flight cancellation and delay as possible recovery options. The optimisation process is performed over the set-partitioning problem with the objective of minimizing delay, cancellations, and estimated costs associated with crew and passenger recovery. The paper introduces several variants of the model, each focusing on different objectives. Their method is validated using SimAir, a simulation of airline operations, described in Rosenberger, Schaefer, et al., 2002, on the daily flight schedule of fleets from a major domestic airline carrier. Results are presented for a recovery solution spanning 500 days of simulated airline operations, with solution times reported to be up to 9 minutes, and show reduced cancellations and passenger disruptions, while improving on-time performance, compared to the shortest cycle cancellation policy (SCC) from Rosenberger, Schaefer, et al., 2002. However, it's important to note that the model assumes the new flight schedule will operate as planned, after a disruption is solved, not considering future adverse conditions.

Eggenberg et al., 2010 extended the work of Bard et al., 2001 and presented a column generation algorithm where a time-band network model is used to show the aircraft routes. Each unit (the paper presents results for the Aircraft Recovery Problem and informs about the adaptations needed to use the same modelling framework for recovering passengers and crew) is associated with a specific recovery network and the model considers unit-specific constraints. The column-generation algorithm employed in this context guarantees global feasibility by adhering to the structural constraints of the problem. The conventional multi-commodity approach followed by Bard et al., 2001 encounters difficulties in incorporating unit-specific constraints, but this approach effectively addresses unit-specific constraints, thereby enhancing the overall feasibility of the solution. While the presented results seem promising and the majority of instances solve within 100 seconds, the authors report that for the most computationally expensive case (7 days of recovery period and 16 aircraft concerned by the disruptions) the run time exceeds 1 hour. The authors note, however, that the case instances are only tested with a homogeneous fleet (allowing for unconstrained swapping) and suggest as possible future work to test the model on fleets with multiple aircraft types.

Aktürk et al., 2014 were the first to consider aircraft cruise speed control as a recovery option to deal with the Aircraft Recovery Problem. In most literature, all delays propagate throughout the flight network since it is considered that en-route flight times are fixed (refer to section 3.3 for a detailed review on this topic). By increasing cruise speed, some delays can be absorbed at the cost of additional fuel. Increasing cruise speed for early arrival, for example, generates additional tail swap options which can drive down the cost of the recovery operation and compensate the additional fuel cost. Due to the non-linearity of fuel burn in cruise speed, the authors use a conic quadratic optimisation approach to solve the problem with minimization of recovery related costs like tail swap cost (a formulation for a heterogeneous fleet is also provided), increased fuel consumption and passenger delay cost. The model is tested on two data sets extracted from the U.S. Department of Transport database. These two schedules contain 114 and 207 flights, and 31 and 60 aircraft respectively. The solution time of the most computationally expensive case does not exceed 250 seconds. The paper concludes that significant cost savings can be achieved with cruise speed control, making it a suitable recovery approach to include in aircraft recovery studies.

Vos et al., 2015 deviate from other researchers by introducing a dynamic framework called the Disruption Set Solver (DSS) for aircraft schedule recovery. Instead of assuming a static disruption scenario with full knowledge of all disruptions in advance, the DSS framework handles disruptions as they occur and builds on the solutions of previous disruptions. Their formulation incorporates parallel time-space networks, which enable explicit modeling of individual aircraft. To limit the scope of the recovery process, the framework employs a candidate aircraft selection algorithm. The model is tested on four different cases using fleet data from Kenya Airways. In 93.3% of the cases, the DSS model provides a solution in under 10 minutes. The paper also compares the solutions obtained from the DSS framework with those of a static model that assumes all disruptions for the day are known upfront. The findings of the study suggest that using static disruption scenarios leads to an underestimation of costs and an overestimation of feasible schedule recovery. The dynamic approach offered by the DSS framework provides a more realistic assessment of the recovery process in the face of evolving disruptions.

3.1.3. Metaheuristics

Guimarans et al., 2015 introduce a large neighborhood search (LNS) algorithm that is based on a constraint programming formulation. To assess the robustness of the discovered solution, a simulation-based approach is employed, which takes into account the stochastic nature of real-life disruptions. This approach involves subjecting the solution to various simulated scenarios, and if it is found to be easily disrupted in many instances, it is rejected. This approach offers a significant advantage over deterministic recovery solutions, which do not consider the future potential for disruption. The model is tested on three real flight networks of varying sizes, ranging from 9 to 40 aircraft and 49 to 163 flights. In the largest case, the solution time is approximately 5 minutes, highlighting the efficiency of the algorithm.

Liang et al., 2018 introduce a column-generation based heuristic approach to address the aircraft recovery problem while considering constraints related to airport capacity and maintenance. In their model, they incorporate maintenance flexibility as an additional recovery option alongside flight delays, flight cancellations, and aircraft swaps. A notable novelty of their approach is the modeling of continuous delay options instead of discrete ones. The proposed model is tested on eight different scenarios involving fleets with varying fleet sizes (16 to 85 aircraft) and varying flight schedules (59 to 638 flights). A comparison is made between the performance of discrete (a flight is copied every 30 minutes until it reaches the maximum delay allowed (Maher, 2016)) and continuous delay models. The computational times remain below 6 minutes for all cases. The paper concludes that the continuous delay modeling offers two significant benefits. Firstly, the run time is at least as fast as the discrete version, with large-scale cases demonstrating up to 260 times faster computation. Secondly, the costs associated with recovery are reduced. Additionally, the implementation of flexible maintenance strategies is found to potentially reduce recovery costs by 20-60%.

Lin and Z. Wang, 2018 propose a two-stage fast variable neighborhood search (FVNS) algorithm designed for the aircraft recovery problem in the context of airport closures. In the first stage, a feasible flight plan is restored by introducing delays to all flights. In the second stage, the algorithm explores the local solution space to search for improved solutions. A distinctive feature of FVNS is that, after updating the optimal solution, it continues searching within the same neighborhood. Furthermore, FVNS excludes potentially sub optimal solutions, thereby reducing the search space. The model is tested on two flight networks obtained from a Taiwanese airline, one with 7 aircraft, the other with 12, and both consisting of 70 flights. The solution time is computed with a run time in the order of milliseconds. However, it should be noted that the disruption scenarios considered in the study are limited (one hour closure of 2 airports only), and a global optimum solution is not provided as a reference.

Lee et al., 2022 were the first to solve the Aircraft Recovery Problem efficiently utilizing a Reinforcement Learning (RL) approach. More specifically they implement Q-Learning and Double Q-Learning algorithms with the objective of recovering disrupted flight schedules caused by temporary closures of airports which lead to delay propagation. The paper highlights two important distinctive features that differentiate it from previous studies: re-timing certain flights and swapping aircraft can be implemented by the RL approach without relying on copies of flight arcs; by just modifying reward functions, the RL agent can flexibly adapt to various objectives. The authors demonstrate the effectiveness and flexibility of the proposed approach through computational experiments and a real-world case study with a domestic flight schedule from a South Korean airline, achieving maximum computation times of around 2 minutes and showing that the DQL algorithm outperformed the strategy that did not allow swapping aircraft chosen as comparison.

Table 3.1 shows an overview of all discussed aircraft recovery literature. Note that only the largest case dimensions and solution times are presented.

Table 3.1: Overview and classification for literature focusing on Aircraft Recovery

Reference	Type	Approach	Disruption Types				Recovery Actions					Data	Case Dimensions			CPU
			Delays	CNL	AC U/A	Airport Disruption	Delays	CNL	Swaps	Reserve AC	Speed Control		AC	Fleets	Flights	(s)
Rosenberger, Johnson, et al., 2003	EO	Set-Partitioning										G	96	3	469	16
Eggenberg et al., 2010	MH	Dynamic Programming with Col Generation										RL	100	1	760	63
Aktürk et al., 2014	MH	Conic Quadratic Mixed Integer Programming										RL	60	6	207	202
Vos et al., 2015	MH	Aircraft Selection Heuristic with MILP										RL	43	1		900
Guimarans et al., 2015	O	Constraint Programming with LNS and simulation										RL	40	1	163	<226
Liang et al., 2018	MH	Col Generation										G	44	1	638	356
Lin and Z. Wang, 2018	MH	Fast Variable Neighborhood Search										RL	12		70	<0.01
Lee et al., 2022	MH	Q-learning and Double Q-learning										RL	20	3	94	<180
Zhao et al., 2023	MH	Single and Multi-commodity network model										RL	73		207	<1

Table 3.2: Overview and classification for literature focusing on Integrated Recovery

Reference	Elements			Type	Approach	Disruption Types				Recovery Actions					Data	Case Dimensions			CPU
	AC	Pax	Crew			Delays	CNL	AC U/A	Airport Disruption	Delays	CNL	Swaps	Reserve AC	Speed Control		AC	Fleets	Flights	(s)
Bisaillon et al., 2011				MH	Rolling horizon time framework with MILP										RL	256	1	1423	<600
Petersen et al., 2012				MH	Benders' Decomposition										RL		3	800	1080
Sinclair et al., 2014				MH	Large Neighbourhood Search										RL	256	1	1423	<600
Sinclair et al., 2016				MH	Large Neighbourhood Search with Col Generation										RL	618	1	2178	<420
U. Arıkan et al., 2017				EO	Conic Quadratic MILP										RL	402	555	1254	480
Marla et al., 2017				EO	Rolling horizon time framework with MILP										RL			250	<120
Vink et al., 2020				EO	MILP combined with a fleet selection method										RL	100	2	600	<44

Abbreviations used in tables: U/A: Unavailable, AC: Aircraft, CNL: Cancellation, Pax: Passengers, CPU: Computation time in seconds, EO: Exact Optimisation, MH: Metaheuristics, O: Others, ‘ ’: Not mentioned or not relevant, G: Generated Data, RL: Real-life Data. Symbols used in tables: () considered; () not considered

3.2. Integrated Recovery

3.2.1. Initial Efforts

The integration of two or more recovery resources (aircraft, crew and passengers) is a difficult task, from a mathematical perspective, as well as a computational one. To the best of the author's knowledge, the first proposal of a truly integrated approach is the PhD Thesis of Lettovsky, 1997, where the author formulated the 'Airline Integrated Recovery' problem integrating the full scope of disrupted resources (aircraft, crew and passengers). In the thesis, a linear mixed-integer mathematical problem is introduced, aiming to maximize the total profit of the airline while considering the availability of all resources. A decomposition scheme is proposed, involving a master problem called the "Schedule Recovery Model" and three sub-problems known as the "Aircraft Recovery Model," "Crew Recovery Model," and "Passenger Flow Model." The Benders' decomposition technique is applied to obtain the solution. However, it is worth noting that an important limitation of the model is that it only considers flight crew, not including cabin crew in the optimisation process.

3.2.2. Exact Optimisation Methods

U. Arıkan et al., 2017 proposed a network flow model that accounts for the flow of each entity (aircraft, crew and passengers). Similarly to Aktürk et al., 2014, the authors implemented aircraft cruise speed control and proposed a conic quadratic mixed integer programming formulation. The flight network proposed by the authors is designed to maintain a manageable problem size, enabling real-time solutions, since it avoids the need for discretizing departure times and cruise speed decisions, and only partial networks are constructed, containing flight arcs that are likely to be used in the final solution. The model is tested on a flight network of a major U.S. airline operating 1254 daily flights and 402 aircraft, solving within 8 minutes. The considered disruptions were varying times of delay for a given flight, as well as varying closure times of a hub. The authors note that, especially in large networks, cruise speed control can be used as a valuable recovery option. They also manage to achieve solutions with a maximum optimality gap of 1.5% for networks having 288 and 473 flights, while average solution times are reduced from 6.7 to 1.7 minutes.

Marla et al., 2017 incorporate cruise speed control in addition to traditional recovery actions by creating alternative flight plans corresponding to higher cruise speed with pre-determined cost. Similarly to Vos et al., 2015, their formulation incorporates parallel time-space networks, which enable explicit modelling of individual aircraft, but include passenger recovery, by explicitly assigning cost to delayed flights that carry connecting passengers. Departure time decisions are accounted for by introducing multiple flight arc copies representing different departure time alternatives. Similarly, cruise speed control alternatives are considered by generating additional flight arc copies for various speed options associated with each departure time alternative. This approach requires a discretization of the cruise speed options and naturally enlarges the size of the generated network. The solution is also dynamic, i.e. disruptions are solved as they occur. The model is tested on data from a European hub-and-spoke airline with a 250 flight network. To comply with AOCC-required performance, the runtime is limited to 2 minutes. The proposed model is compared to a baseline sequential recovery model, different variants of an aircraft-centric model with flight planning, and different variants of a passenger-centric model without flight planning. To assess the performance of the models, the authors consider 60 delay scenarios based on historical data from the airline. They conclude that their enhanced recovery models offer benefits in terms of reducing overall costs and minimizing passenger-related delay costs for the airline, compared to existing approaches.

Vink et al., 2020 also adopt a similar approach to Vos et al., 2015 by employing parallel time-space networks and an aircraft selection algorithm, albeit with a slight modification. Instead of selecting aircraft based on ground time, the authors base their selection on aircraft location. Aircraft maintenance schedules and passenger itineraries are modeled. As the primary focus of their research is to enable real-time operational use of the model, it finds an initial trivial solution solution within seconds, which is promptly presented to the user. It then continues to refine the recovery problem by adding a single candidate aircraft per iteration, updating the best solution as each iteration is solved. This approach ensures that the model can be used in real-time by the AOCC, as it always provides a feasible solution for immediate action, even if the runtime becomes too high. The model is tested on 10 different disruption scenarios, each involving 2 to 6 disruptions (combinations of delays, cancellations and aircraft unavailability) occurring at various times of the day. The aircraft selection algorithm successfully identifies the best solution in 50 seconds or less, and in 90% of cases, the final solution matches the dynamic global optimum achieved by optimising with all aircraft. On average, the presented solutions are obtained in 22.1 seconds and are only 6% costlier than the dynamic global optimum, demonstrating the efficiency and effectiveness of the proposed model.

3.2.3. Metaheuristics

In 2009, the French Operational Research and Decision Support Society (ROADEF) organized an Operations Research challenge dealing with disruption management for commercial aviation. This challenge resulted in several publications and Bisaillon et al., 2011 won the first prize, by formulating an LNS heuristic that combined fleet assignment, aircraft routing and passenger assignment. It consisted of three phases: construction, repair and improvement. These phases destroy and repair parts of the solution in an iterative manner. The model constructs aircraft routes and passenger itineraries for the recovery period with the goal of minimizing operating cost and impact on passenger. The aim of the first two phases is to produce an initial solution that satisfies a set of operational and functional constraints. The third phase considers large schedule changes and tries to further optimise the solution while maintaining feasibility. Sinclair et al., 2014 built on this framework by making changes in each of the three phases, in order to further optimise the obtained solutions. In the construction phase, the heuristic prioritized the cancellation of aircraft that caused the highest cost. During the repair phase, the focus shifted towards re-booking disrupted passengers and utilizing spare aircraft to cover the cancelled flights identified in the construction phase. In the improvement phase, the authors aimed to accommodate disrupted passengers by introducing flight delays. The enhanced model was evaluated using the challenge's dataset, where the algorithm achieved 17 best solutions within five minutes and 21 best solutions within 10 minutes for 22 instances. In a subsequent study by Sinclair et al., 2016, they further extended the previous work by introducing a post-optimisation column generation heuristic that reduces the model size to improve solutions within reasonable run-times. By defining dual variables after solving the LP relaxation, the reduced cost of the variables is calculated and the negative values are considered when re-solving the LP problem. The model was tested on the challenge's dataset and found best known solutions to all scenarios. The authors suggest future research should focus on implementing a rolling-time horizon with the column-generating algorithm.

Petersen et al., 2012 presented an integrated optimisation approach that resembled the one used by Lettovsky, 1997, where they distinguish between four recovery sub-problems: schedule recovery, aircraft rotations, crew assignment and passenger assignment. A column generation algorithm is employed along with a candidate aircraft selection algorithm to reduce problem size. The authors tested the model with data from a regional US carrier that operates a hub-and-spoke network with 800 daily flights and 2 aircraft types. Three different disruption scenarios are considered: 50%, 75% and 100% reduction in hub flow rate (arrivals and departures). The results of the proposed integrated model are compared to the results of a sequential approach. The results show that the costs of the integrated approach are always equal or lower than the cost of the sequential approach. Although, for the largest case, the solution time is in excess of 30 minutes, the integrated approach resulted in a feasible solution in 100% of the cases, while the sequential approach only managed to do so for 75% of the cases.

Zhao et al., 2023 propose a two-stage approach to minimize the impact of daily disruptions on an airline's flight schedule and compare it to a rolling horizon approach that closely mimics current procedures used by several airlines. The authors address uncertainties in disruption duration and the timing of information availability, much like Vos et al., 2015. First, they solve a single-commodity network flow problem with side constraints to eliminate aircraft and flights that are not likely to be part of the optimal solution, and then a multi-commodity network flow problem to derive an optimal recovery plan. The objective is to minimize passenger delay costs, cancellation costs, curfew violations, and deviation from the original schedule. By considering a range of likely scenarios, the authors provide guidance and robust solutions for flight directors, which offer anywhere from a slight improvement over the rolling horizon approach to as much as a 35% improvement for the 150 instances investigated and are effective in finding (near-)optimal solutions quickly (<1 second). The research offers valuable insights and lays the groundwork for further investigations in dealing with uncertainty.

Table 3.2 shows an overview of all discussed aircraft recovery literature. Note that only the largest case dimensions and solution times are presented.

3.3. Delay Propagation

In order to properly compute total costs incurred by airlines with disruptions, it is vital to consider the propagation of delays and disruptions in the network, stemming from the constrained availability of resources. Optimising aircraft utilization requires airlines have tight turn-around times between flights, but this can increase the likelihood of delays in subsequent flights, and the study of this phenomenon can lead to an increased schedule robustness.

Wu, 2005 explores the inherent delay present in each aircraft rotation. To simulate operations and analyse this delay, a two-module model is employed in combination with a Monte Carlo simulation. The first module, known as the turnaround module, captures the dynamics of aircraft turnaround processes by means of a combination of two parallel Markov chains and discrete-event driven simulations. The en-route module models block time as a random variable drawn from a probability distribution. The model's validity is assessed using data from a point-to-point carrier, revealing that real-world delays are significantly higher than those simulated by the model. However, the results also demonstrate the model's capability to track delay propagation trends within a rotation. Additionally, the model is utilized to evaluate various scenarios involving the introduction of root delays in operations, and the schedule's performance is measured based on its ability to accommodate such delays.

Wong and Tsai, 2012 propose a survival model of flight delay propagation. In this model they consider the two types of delay, arrival and departure delays. They consider as survival time the time a delay will last, given that it has started or it has lasted for a certain amount of time. Using flight data from a Taiwanese airline, the authors utilize the Kaplan-Meier estimator to estimate the survival functions of arrival and departure delays. They employ the Cox proportional hazard model to assess the impact of different factors on delays. The results indicate that aircraft type, flight operations, and cargo mail handling significantly affect departure delays, while arrival delays are primarily influenced by block buffer time and weather conditions. The authors propose as future work the examination of the propagation of delays throughout the network by recursively applying the departure and arrival delay models. However, the implementation of this expanded model is not carried out in the study.

M. Arıkan et al., 2013 developed an analytical model for investigating delay propagation at a network level, based on publicly available data on airline on-time performance for all domestic flights within the United States. In their model, they consider stochastic block time, which accounts for both root delay and propagated delay. They propose the use of a log-Laplace distribution to capture the characteristics of a random variable that considers both block time and propagated delay, and evaluate the network using analytical methods. Their analysis reveals that deterministic approaches for delay propagation tend to overestimate the extent of delay propagation. By applying their models, they identified bottlenecks in airport capacity and critical flights within an airline's schedule.

Kafle and Zou, 2016 used econometric analysis to provide estimates on how much propagated delay is generated out of each minute of newly formed delay and to quantify the effects of different factors on the initiation and progression of propagated delay. They begin with developing an analytical model providing three different perspectives to quantify how flight delay propagates from any upstream node to any downstream node of the same aircraft. This analytical approach is applied to the US air transportation system using publicly available data. Then, the authors develop a joint discrete-continuous econometric model to investigate the factors influencing the initiation and progression of propagated flight delays, employing Heckman's two-step procedure. The findings are found to be robust across different scenarios, considering different buffer values and the absorption of delays by buffers. Their work enables the assessment of the significance of newly formed delays in the context of delay propagation, specifically at individual airports and for different airlines.

Wu and Law, 2019 develop a Bayesian network in a delay propagation tree framework to investigate delay propagation and delay causes, accounting for the 3 major airline resources: aircraft; crew (cabin crew and pilots); passengers. Historical data is used to train the model, but the model is capable of testing hypothetical scenarios for which historical data is not available, predicting delay propagation in the network, and diagnosing the primary cause of a delay. The authors demonstrate the value of their model in conducting network-wide 'what-if' scenario studies, making use of the inference ability of their Bayesian Network to predict posterior probabilities of flight delays.

3.4. Discussion

This chapter explores disruption management within the context of airline operations. It is vital to understand the growing occurrence of disruptions stemming from the rapid growth trend of air travel and the importance of the AOCC's work in solving these and thus minimising incurred costs. The various tools at the disposal of the operators were analysed and their way of work at the majority of airlines was described. The dissected literature offers insight into the many different methods used to solve both the airline recovery problem, and the integrated recovery problem (considering aircraft, passengers and/or crew), with the aim of being able to substitute human AOCC operators in the future. In addition, to be able to determine the effect of decisions different than the ones that were actually made in past daily operations, some delay propagation studies were reviewed.

With an increase in computing power and knowledge gained through research, it is likely that larger more complex problems will be developed in future. These models may incorporate expanded flight schedules, larger fleets, and longer time windows. Additionally, researchers are incorporating additional factors such as maintenance, airport operations, and recovery options to capture the intricacies of airline operations in their disruption management systems, and integrating more resources, such as crew and passengers. Moreover, recently, stochastic elements are included to capture the uncertainty of decision making related to the future state of the network.

Despite the progress made in modeling reality and developing efficient software solutions for AOCCs, there are still limitations to consider. The inclusion of various constraints and recovery options encountered in practice requires substantial computational power that is not currently commercially available. For example, considering passenger reallocation from competing airlines would necessitate encompassing the entire network of flights from all airlines, which is computationally intensive. Researchers and practitioners face a trade-off between the complexity of the models and the time required to find a solution (AOCCs have time requirements), which is a key topic of discussion in the existing literature related to disruption management.

To address the trade-off between model complexity and solution time, modern research often employs scope-limiting algorithms that generate a subset of recovery options deemed likely to contain the optimal solution. While these algorithms can produce high-quality solutions, they are heuristic in nature and may not always guarantee the best outcome. However, this is the only viable procedure when the goal is to satisfy the AOCC's required timeframe of 2 minutes in producing fleet-level solutions (Vink et al., 2020). Another approach involves providing the AOCC with at least one feasible solution within a specified time limit, which can be iteratively improved as better solutions are discovered. This approach ensures that the AOCC can utilize the software for every disruption, even though the solutions obtained within the time limit may not always be optimal. However, as the purpose of this research is to study past days of operations and how the reserve fleet effected them and would have effected had there been a different composition of reserve fleet, computation time will not be a major limiting factor.

4

Reserve Capacity Planning in other domains

As was made evident in chapter 2, and although most airlines have some aircraft ready to operate as reserves, this topic is not usually discussed in the planning stages of the airline fleet in previous literature. As such, the author extended the literature review in order to obtain an overview of the topic of reserve capacity planning in other domains. The goal of this chapter is therefore to determine how optimal reserve capacity is calculated across the different fields of study and how its' use is optimised, addressing the third (***How is reserve capacity planning processed in other domains?***) research question.

4.1. Reserve Planning in other industries

Public transportation networks are in many ways comparable to airline networks. Cats and Jenelius, 2015 consider the metro network of Stockholm and propose a way of using given reserve capacity to lower overall "welfare loss" in case of a disruption. The proposed model is divided into two parts, the first part identifies the most central links available for capacity enhancement based on the expected fraction of all passengers who travel during a given time interval that traverse the link and the second part simulates disruptions in order to determine the links that are the most saturated during a disruption and, of those, the ones whose saturation levels are sufficiently higher than the scenario with no disruptions. The method integrates stochastic supply and demand and dynamic route choice and is tested on the Swedish capital's metro network using BusMezzo, an agent-based public transport simulation model, leading to a choice of increasing service frequency by 50% on 2 segments of the network. Simulation of this change results in the calculation of the savings achieved by using the Swedish value-of-time, resulting in a sum of around 300,000 USD. This dynamic agent-based modelling enables the authors to capture the adaptive redistribution of passenger flows as well as cascading network effects.

Tang and Y. Wang, 2015 focus on allocating limited operating room (OR) capacity among subspecialties in hospitals, considering both elective and emergency surgeries. The authors address the challenge of uncertainty in surgery demand by developing a robust optimisation model that minimizes the worst-case revenue loss resulting from a shortage of OR resources. The model provides hospital administrators with valuable guidance on how to allocate OR capacity to different subspecialties and reserve capacity for emergency cases. By incorporating an adjustable robust optimisation framework, the proposed approach allows for flexibility in controlling the level of conservatism in the solutions. The computational results demonstrate the effectiveness of the robust optimisation method, showing that it closely approximates the expected objective value obtained from a scenario-based stochastic optimisation (difference $< 1\%$). Furthermore, the robust approach offers the advantage of limiting the worst-case outcome of the surgery capacity allocation problem (difference $>15\%$). Overall, this minimisation of revenue loss in the worst-case scenario where there is a shortage of OR resources can be linked to robust scheduling in the airline industry.

4.2. Airline Reserve Crew Planning

An extensive research topic into reserve capacity is that of airline reserve crew planning.

Sohoni et al., 2006 propose a model to integrate the crew planning and the rostering problem while optimizing for reserve crew increasing reserve availability and utilization metrics. Their approach consists of two steps. Firstly, they estimated reserve demand. This demand comes from two sources that were separately estimated- bidding-invoked conflicts and irregular daily operations. To estimate the demand from this second source, they used SimAir, a stochastic tool that simulates airline operations (Rosenberger, Schaefer, et al., 2002). The second step of the approach is also two-fold, as they begin with generating a minimum set of reserve duty periods and go on to generate schedules covering the reserve duty requirements. The models developed for both steps are standard set-covering and set-partitioning models. Empirical test results demonstrate the algorithm effectively increases reserve availability and minimizes the manpower needed.

Bayliss, Maere, et al., 2012 consider the reserve crew problem in reverse. It approaches the reserve crew problem by considering the probability of crew unavailability. Through an investigation of possible objective functions for the model it was found that minimizing the sum of these probabilities lead to reserve crew schedules with the most desirable properties (lowest cancellation rate and highest reserve crew utilisation rate). A multiple objective optimisation approach was considered where the number of reserve crew to assign is a variable along with the corresponding reserve crew schedule itself. A pruned dynamic programming based algorithm obtained optimal solutions, although this may not always be the case. As such, the authors proposed a backwards heuristic, which obtained good solutions very rapidly and was then used in the pruned dynamic programming approach as a lower bound estimation heuristic in order to improve the reliability of the method. However, the authors mention the inclusion of delayed flights as a future improvement since the reserve crew might become unavailable for covering a flight if it is delayed for too long, as well as including other recovery actions such as aircraft swaps and crew swaps (cheaper recovery actions).

Homaie-Shandizi et al., 2016 highlight the impact of pilot shortages on airlines' performance, and thus worked on predicting the number of reserve pilots required each month, based on data from previous months' reserve usage. The authors propose a supervised learning method based on decision trees to predict the total monthly pilot reserve hours. Their approach utilizes characteristics of the monthly schedule as factors and employs an iterative algorithm to make predictions. Because the flight schedule is used as an input in the model, the approach is sensitive to the dynamic aspects of schedule changes. The model is tested using real data and demonstrates a 15% improvement compared to the current predictions used in the airline. The proposed method accounts for the dynamic nature of the airline industry by considering changes in the flight schedule and employs a time-dependent classifier to handle longitudinal data. The authors use a learning process that consists of two phases: tree growing and tree pruning. This second phase is done by simulating the prediction process from the past and finding the tree that minimizes the errors obtained by means of a loop. Although the authors acknowledge some limitations of the approach, such as potential inaccuracies in the presence of outliers, they propose as future improvements model selection methods for choosing the best set of trees each month and focusing on predicting outliers to reduce prediction errors.

Bayliss, De Maere, et al., 2017 explore a scenario-based approach to represent and address the uncertainty associated with reserve crew demand. The authors' work focuses on airline reserve crew scheduling under crew absence uncertainty and delays in a single hub and spoke network. They propose a mixed integer programming approach called the mixed integer programming simulation scenario model (MIPSSM) to schedule the airline's reserve crew. The MIPSSM formulation incorporates disruption scenarios derived from simulations of the airline's operations without reserve crew. The objective is to find a reserve crew schedule that minimizes the overall level of disruption over the set of input scenarios. The author also explores modifications of the MIPSSM and introduces a heuristic solution approach. The results demonstrate that the MIPSSM-based approaches outperform alternative methods. The study highlights the importance of selecting appropriate disruption scenarios and discusses the practical and methodological contributions of the proposed approach. The work offers valuable insights into increasing airline schedule robustness and presents a novel scenario-based approach for reserve crew scheduling, addressing a challenging real-world problem in the airline industry.

4.3. Discussion

This chapter provides an overview of reserve planning problems in relevant domains, where similarities can be found with the airline reserve fleet planning problem aiming to be solved. The study of the optimal use of reserve capacity in a public transport network and the optimal assignment of OR rooms to scheduled and urgent surgeries were reviewed for this reason.

The Airline Reserve Crew Planning problem provided an added insight into crew considerations, which are important when dealing with disruptions. In addition, the approaches chosen to plan their schedules show promising results to be applied to the reserve fleet planning problem.

One reviewed paper that stood out for its relevance to the proposed research was Homaie-Shandizi et al., 2016's work. The authors achieved an added accuracy in predicting pilots' monthly reserve hours, based on data from previous months' reserve usage. Because the flight schedule is used as an input in their model, the approach is sensitive to the dynamic aspects of schedule changes. When predicting future reserve fleet effects on airline performance this approach should be considered. Overall, this chapter provides valuable insights into some literature on reserve planning, focusing on problems with similarities to the airline reserve capacity optimisation.

5

Forecasting

As the goal of the research is to provide insight into future optimal reserve fleets, it will be vital to determine how to use results achieved with the study of past days of operations to forecast the future. The topic of forecasting is extensively discussed in previous literature, applied to several industries. In this chapter a number of machine learning forecasting techniques will be discussed, along with accuracy evaluation metrics to determine their performance, addressing the fourth (***What type of forecasting method is indicated to predict financial savings achieved in daily operations?***) research question.

5.1. Accuracy Evaluation Metrics

Nearly all publications that address forecasting techniques also cover the assessment of forecasting precision. Establishing appropriate precision metrics is highly significant for several reasons. Quantifying performance is crucial for refining existing methods and determining the effectiveness of new approaches. Furthermore, performance metrics are essential for choosing and configuring the most suitable forecasting method for a particular dataset. Unfortunately, the array of metrics available for selection is extensive, making it impractical to identify a single universal benchmark. Most authors do not hinge on one metric, but use multiple metrics in parallel, each one with its advantages and limitations.

Within this segment, the most frequently employed evaluation metrics are presented and analysed, considering both their merits and drawbacks. Equations 5.1, 5.2, 5.3, and 5.4 illustrate the mathematical formulations of these evaluation metrics. Here, n signifies the prediction horizon, $\hat{y}_i$ denotes the predicted value, and y_i represents the test value.

Mean Absolute Error (MAE) quantifies the mean absolute difference between the predicted values derived from a trained model and the actual test values, as denoted by equation 5.1. The computed MAE aligns with the scale of both the predicted and test data. However, a notable drawback is encountered when attempting to compare MAE across diverse datasets due to the inherent challenge of reconciling variations arising from differing scales. This limitation hinders straightforward cross-dataset comparisons, meaning it should only be used in specific contexts where scale congruence is assured.

$$MAE = \frac{1}{n} \times \sum_{i=1}^{n} |\hat{y}_i - y_i| \tag{5.1}$$

Mean Absolute Percentage Error (MAPE) quantifies the mean absolute difference between the predicted values generated by a trained model and the corresponding test values, divided by the test value itself, as denoted by equation 5.2. Its expression as a percentage facilitates straightforward interpretation and comparison. However, a notable drawback with this metric is that if the test value is zero or near-zero, it leads to resulting MAPE values that are unbounded or divisions by zero, requiring careful consideration in the interpretation of MAPE outcomes.

$$MAPE = \frac{100}{n} \times \sum_{i=1}^{n} \left| \frac{\hat{y}_i - y_i}{y_i} \right| \tag{5.2}$$

Mean Squared Error (MSE) quantifies the average squared disparity between the predicted values generated by a trained model and the corresponding test values, as denoted by equation 5.3. This metric incorporates both the variance and bias of the estimator, providing a comprehensive measure of the model's prediction quality. This squared approach in penalizing larger differences imparts a heightened sensitivity to the metric, where substantial deviations are weighted more significantly. MSE serves as a valuable indicator of prediction accuracy, reflecting the model's performance by considering the magnitude of prediction errors.

$$MSE = \frac{1}{n} \times \sum_{i=1}^{n} (\hat{y}_i - y_i)^2 \tag{5.3}$$

Root Mean Squared Error (RMSE) is defined as the square root of the MSE, as illustrated by equation 5.4. RMSE exhibits reduced sensitivity to larger differences between the predicted and test values compared to MSE. The key advantage of RMSE lies in its preservation of the scale of the original data, facilitating a more straightforward interpretation of the resulting RMSE value.

$$RMSE = \sqrt{\frac{1}{n} \times \sum_{i=1}^{n} (\hat{y}_i - y_i)^2} \tag{5.4}$$

5.2. Machine Learning Methods

Forecasting disruption costs for the Summer Schedule (most profitable period of the year) is a crucial task for many airlines around the world, particularly for SWISS International Airlines. This task is challenging because the passenger demand can vary greatly depending on the day of the week, and the weather conditions and random maintenance issues that cause aircraft to be grounded are not straightforward. Accurate and precise disruption forecasting (considering various compositions of reserve fleet) enables airlines to better manage their resources, resulting in an improved customer experience and enhanced profitability.

In response to these forecasting needs, numerous statistical methods have been proposed in the past. However, these models face limitations in effectively capturing the intricate patterns of passenger demand and adapting to changes in demand dynamics. In recent years, advanced machine learning techniques have emerged as potent tools for forecasting. Distinguished by their capability to grasp long-term dependencies in data and their adeptness at learning from intricate patterns, these techniques represent a promising avenue for overcoming the shortcomings of traditional statistical models in the realm of forecasting.

The contemporary landscape of forecasting methods, especially within machine learning literature, is predominantly influenced by deep-learning-based forecasting models. In contrast, tree-based forecasting methods have not garnered as much academic attention (Januschowski et al., 2022). However, the results from recent forecasting competitions, including the M4 and M5 competitions (Makridakis et al., 2022), Global Energy Forecasting (GEFCom) competitions, and various Kaggle competitions (Bojer and Meldgaard, 2021), present a different narrative. In these competitions, tree-based approaches prominently secure top positions, which underscores the pragmatic efficacy and competitiveness of tree-based methods in real-world forecasting scenarios. For this reason and the fact that these methods are known to gracefully handle missing data (as opposed to other methods such as neural networks), it was decided by the author to focus on decision tree based methods for the forecasting of operational disruption costs that is intended for this research.

5.2.1. Random Forest

Random Forests (RF) are a type of ensemble machine learning algorithm widely applied in diverse tasks, including forecasting. This class of algorithms enhances prediction accuracy and stability by combining the outputs of multiple individual models. In the context of forecasting, RF operates by training numerous decision trees on distinct data subsets and subsequently averaging their predictions to generate a consolidated forecast. Its randomness stems from the fact that each decision tree is trained on a randomly selected subset of the data, and the ultimate prediction results from averaging the predictions of all the trees.

The utilization of RF for forecasting offers several advantages. Notably, RF demonstrates proficiency in handling high-dimensional data and accommodating nonlinear relationships between features and the target variable. Additionally, RF exhibits resilience to overfitting, a concern encountered with certain machine learning algorithms when applied to extensive datasets. The fine-tuning of RF is achievable through adjustments to hyperparameters, such as the number of trees in the forest or the maximum depth of the trees, enhancing the model's adaptability to specific forecasting tasks (Goehry et al., 2021).

Toqué et al., 2017 explores the forecasting of passenger flows in multimodal transport (trains, buses and trams). Two forecasting horizons are considered: long-term predictions aim to forecast passengers entering stations and boarding vehicles one year ahead; short-term predictions focus on the next 15-30 minutes. Two machine learning models (RF and Neural Networks) are employed on a real 2-year smart card dataset of a well-known major business district in Paris. Results validate the efficacy of these forecasting approaches and show that the best results are provided by the RF models for both forecasting horizons (RMSE and MAE were the accuracy metrics used). The paper underscores the value of such models in optimizing transportation operations and enhancing trip planning for citizens.

5.2.2. Gradient Boosting

Gradient Boosting is another decision-tree technique similar to RF. In this approach, numerous small decision trees are employed in order to make predictions. The methodology revolves around merging multiple weak predictive models (trees) into an ensemble, thereby enhancing the accuracy of predictions. Initially, given a set of input features and their corresponding output values, the algorithm trains a "weak" decision tree (the base learner) to predict output values from input features. Subsequently, the algorithm computes errors (differences between the predicted and actual values). The next phase involves training a second model, also referred to as the base learner, to predict these errors rather than the output values. This second model employs the errors as output values and the same input features as the first model. The algorithm combines predictions from both base learners iteratively, producing a more accurate ensemble model. This iterative process is known as boosting, and the gradient component is relative to the optimization of the loss function by minimizing its gradient or slope (Bentéjac et al., 2021). In essence, the algorithm aims to identify the direction in which the loss function decreases most rapidly and updates the model accordingly. The gradient boosting process iterates, with each new model trained to reduce the residual error of the preceding ensemble.

Zhang and Haghani, 2015 explore the application of the Gradient Boosting Regression Tree Method (GBM) in predicting freeway travel time. The study extensively discusses the model's distinctive features, and makes use of the model's suitability for capturing complex nonlinear relationships. A comparison is performed (MAPE the accuracy metric used) with the classical statistical model Autoregressive Integrated Moving Average (ARIMA) and RF, underscoring the GBM's superior performance, demonstrating its efficacy in accurate travel time prediction. As the prediction horizon (5 to 30 minutes ahead predictions are analysed) increases, all three models' performances drop, but the GBM method is less sensitive to prediction horizon and maintains a good prediction performance. The paper acknowledges challenges related to parameter optimization and computational time, providing valuable insights into considerations when implementing the GBM model, such as how performance varies with different choices of parameters (number of trees, learning rate and variable interactions). This research makes a noteworthy contribution to the literature dedicated to this decision tree based method, specifically showcasing the effectiveness of GBM in the context of travel time prediction.

5.2.3. Extreme Gradient Boosting

Extreme Gradient Boosting (XGBoost) stands as a scalable machine learning technique for tree boosting, developed and introduced by Chen and Guestrin, 2016. In numerous machine learning competitions, XGBoost has established itself as a frequently employed and successful algorithm, notably gaining popularity within the Kaggle community. Distinguished by its scalability in various scenarios, XGBoost exhibits a computational speed ten times faster than comparable methods and demands minimal memory from the operating system. This efficiency is particularly attributed to its handling of sparse data through a quantile sketch method, facilitating instance weights in tree learning. The combination of rapid computation, low memory usage, and adeptness in managing substantial data volumes positions

XGBoost as an attractive tool for data scientists, enabling the processing of millions of examples even on standard desktop configurations.

Xiong et al., 2022 demonstrated XGBoost's superior performance in predicting wind power compared to Support Vector Machine (SVM) and Long Short-Term Memory (LSTM) models (MAE and RMSE the accuracy metrics used). In addition, the authors use Bayesian hyper-parameter optimization to determine the optimal XGBoost parameters. Nielsen, 2016's work aimed to answer the question "Why Does XGBoost Win "Every" Machine Learning Competition?", having concluded that, when compared with other tree boosting algorithms, this method uses "a higher-order approximation at each iteration, and can thus be expected to learn better tree structures. Moreover, it provides clever penalization of individual trees."

5.3. Discussion

This chapter provides a brief overview of some forecasting methods, along with accuracy evaluation metrics to assess them. As there are many (namely statistical and machine learning) methods for forecasting extensively covered in the literature, a selection had to be made by the author. It was decided to focus this review on decision tree based machine learning models for their pragmatic efficacy and competitiveness in real-world forecasting scenarios.

All methods presented have been successfully applied in several use cases, and are excellent candidates to forecast future disruption costs associated with several compositions of reserve fleet, based on past days of operations' results.

6
Conclusion

This section states the proposed master thesis' research objective derived from the literature review performed and the research questions that must be answered to achieve the objective.

Overall, the literature on both airline fleet planning and disruption management is growing and diverse, although there is a considerable gap related to reserve capacity planning. The current state-of-the-art research in the topic determines the impact of a variable number of reserve aircraft based on completely simulated operations (Groeneveld, 2021; Varenna, 2023). To the best knowledge of the author, no studies have considered past operational data (most accurate representation of AOCC decisions available) on reserve aircraft usage as a baseline scenario to calculate resulting delays, cancellations and savings achieved from the use of the reserve aircraft available. The proposed research aims to make a significant contribution to this field by optimising reserve fleet in order to maximise airlines' financial profit and concluding on future years' various reserve fleet compositions' resulting KPIs.

6.1. Research Objective

The research aims to investigate the effect of previous year's reserve aircraft usage data and use it to predict future financial savings with a variable composition of reserve fleet, so as to determine the optimal reserve fleet composition.

6.2. Research Questions

The literature review performed allowed to identify a gap on the airline reserve fleet planning problem, the optimisation of an airline's fleet in terms of the balance between reserve aircraft and aircraft with scheduled flights. Few have studied the impact of reserve aircraft on daily operations, and no work has considered past operational data on reserve aircraft usage to perform a prediction of the optimal reserve fleet for future years. Thus, a tool that uses airlines' operational past data to decide on their fleet of reserve aircraft before the creation of the next year's future schedules has never been assessed and published as for the author's knowledge.

For that reason, the current report proposes a research which aims to answer:

How to determine an airline's financially optimal reserve fleet based on past years' operational data?

This question is relevant as it would provide a scientific basis for the yearly decision making process behind choosing the reserve fleet that will be kept over SWISS International Airline's entire Summer Schedule. Currently, this decision is made after a meeting between the Network Planning and Operations Departments, where the first advocates for extra planes to introduce in normal operations planning and scheduling and the second for extra planes to add to the reserve fleet to increase operational stability and customer satisfaction. Therefore, a study that can predict the results that will be achieved with different compositions of the reserve fleet would be of great relevance for both academia and industry.

The proposed question will be assessed by answering a set of sub-questions:

- **How to compute savings achieved using flight and reserve aircraft usage data available for past years' Summer Schedules?**

  - How to organize the data coming from various sources?
  - How to simulate operations had no reserve aircraft been used, in order to determine the prevented delays, cancellations and missed connections?

  The various data coming from different sources will have to be integrated and combined in a way that optimises the future simulations that will have to be run. In order to ascertain the savings achieved, an accurate simulation of operations will have to be implemented, using scheduled and actual times of Departure and Arrival of daily flights, and finding an accurate delay propagation model.

- **How to compute savings that would have been achieved had the airline had a different composition of their reserve fleet?**
  - How to simulate past operations with a different number of reserve aircraft?
  - How to determine delay propagation resulting from a disruption that stems from the simulation (not coincident with actual past operations)?

  In order to calculate the benefits of the reserve aircraft, one must take into account a baseline scenario where no reserve aircraft were available as a disruption mitigation measure. To simulate this, the other main mitigation options available to AOCC operator of delaying flights (letting the delay propagate in the network), cancelling flights (most often cancelling rotations (Rosenberger, Johnson, et al., 2004)) and swapping aircraft (assigning rotations to a different aircraft) must be considered for each time the reserve aircraft were actually used and the effects of these changes must be modelled in order to achieve a full daily schedule of flights (with new Actual Times of Departures and Arrival and perhaps new cancellations). This simulation must be run for all fleet compositions that should be considered, taking into account as many real reserve aircraft usage decisions and their effects as possible.

- **How can the resulting savings of previous years' application of various reserve fleets be used to predict future years' results?**
  - Given the problem formulation, what form of machine learning algorithm and approach best suits the targeted solution?
  - Which metrics of the schedule should be used as input when applying the chosen machine learning algorithm?
  - How to validate the resulting predictions?

  After assessing a number of supervised learning techniques, the objective of the project is to provide insight into future optimal reserve fleets, meaning this step will be vital to finalise an answer to the research question. Comparing these monetary amounts with the Network Planning estimates of potential profit of additional aircraft with scheduled flights is of particular interest to the company.

- **What are the potential operational improvements that may arise from a detailed analysis of past years' reserve aircraft usage, and how can these be addressed?**
  - Are there any weekly/monthly trends that arise?
  - What are the most common causes of reserve aircraft utilization?

  Such a detailed analysis of past operations (on a daily level) will provide valuable insight to the company. One possible result from this analysis could be a recommendation for the operation of a variable reserve fleet dependent on the weekday of operations, as it is known that the day-of-week cycle is the largest source of passenger demand variation in the airline industry (Swan, 2002).

Bibliography

Aktürk, M. Selim, Alper Atamtürk, and Sinan Gürel (2014). "Aircraft Rescheduling with Cruise Speed Control". In: Operations Research 62.4, pp. 829–845. doi: 10.1287/opre.2014.1279. eprint: https://doi.org/10.1287/opre.2014.1279. url: https://doi.org/10.1287/opre.2014.1279.

Arıkan, Mazhar, Vinayak Deshpande, and Milind Sohoni (2013). "Building Reliable Air-Travel Infrastructure Using Empirical Data and Stochastic Models of Airline Networks". In: Operations Research 61.1, pp. 45–64. doi: 10.1287/opre.1120.1146. eprint: https://doi.org/10.1287/opre.1120.1146. url: https://doi.org/10.1287/opre.1120.1146.

Arıkan, Uğur, Sinan Gürel, and M. Selim Aktürk (2017). "Flight Network-Based Approach for Integrated Airline Recovery with Cruise Speed Control". In: Transportation Science 51.4, pp. 1259–1287. doi: 10.1287/trsc.2016.0716. eprint: https://doi.org/10.1287/trsc.2016.0716. url: https://doi.org/10.1287/trsc.2016.0716.

Bard, Jonathan F., Gang Yu, and Michael F. Arguello (2001). "Optimizing aircraft routings in response to groundings and delays". In: IIE Transactions 33.10, pp. 931–947. doi: 10.1080/07408170108936885. eprint: https://doi.org/10.1080/07408170108936885. url: https://doi.org/10.1080/07408170108936885.

Bayliss, Christopher, Geert De Maere, et al. (2017). "A simulation scenario based mixed integer programming approach to airline reserve crew scheduling under uncertainty". In: Annals of Operations Research 252, pp. 335–363.

Bayliss, Christopher, Geert De Maere, et al. (2012). "Probabilistic Airline Reserve Crew Scheduling Model". In: 12th Workshop on Algorithmic Approaches for Transportation Modelling, Optimization, and Systems. Ed. by Daniel Delling and Leo Liberti. Vol. 25. OpenAccess Series in Informatics (OASIcs). Dagstuhl, Germany: Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik, pp. 132–143. isbn: 978-3-939897-45-3. doi: 10.4230/OASIcs.ATMOS.2012.132. url: http://drops.dagstuhl.de/opus/volltexte/2012/3709.

Belobaba, Peter, Amedeo Odoni, and Cynthia Barnhart (2015). The global airline industry. John Wiley & Sons.

Bentéjac, Candice, Anna Csörgő, and Gonzalo Martínez-Muñoz (2021). "A comparative analysis of gradient boosting algorithms". In: Artificial Intelligence Review 54, pp. 1937–1967.

Bisaillon, Serge et al. (2011). "A large neighbourhood search heuristic for the aircraft and passenger recovery problem". In: 4OR 9, pp. 139–157.

Bojer, Casper Solheim and Jens Peder Meldgaard (2021). "Kaggle forecasting competitions: An overlooked learning opportunity". In: International Journal of Forecasting 37.2, pp. 587–603. issn: 0169-2070. doi: https://doi.org/10.1016/j.ijforecast.2020.07.007. url: https://www.sciencedirect.com/science/article/pii/S0169207020301114.

Castro, António JM, Ana Paula Rocha, and Eugénio Oliveira (2014). A new approach for disruption management in airline operations control. Vol. 562. Springer.

Cats, Oded and Erik Jenelius (2015). "Planning for the unexpected: The value of reserve capacity for public transport network robustness". In: Transportation Research Part A: Policy and Practice 81. Resilience of Networks, pp. 47–61. issn: 0965-8564. doi: https://doi.org/10.1016/j.tra.2015.02.013. url: https://www.sciencedirect.com/science/article/pii/S0965856415000300.

Chen, Tianqi and Carlos Guestrin (2016). "XGBoost: A Scalable Tree Boosting System". In: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. KDD '16. San Francisco, California, USA: Association for Computing Machinery, pp. 785–794. isbn: 9781450342322. doi: 10.1145/2939672.2939785. url: https://doi.org/10.1145/2939672.2939785.

Clausen, Jens et al. (2010). "Disruption management in the airline industry—Concepts, models and methods". In: Computers Operations Research 37.5. Disruption Management, pp. 809–821. issn: 0305-0548. doi: https://doi.org/10.1016/j.cor.2009.03.027. url: https://www.sciencedirect.com/science/article/pii/S0305054809000914.

Cook, AJ and G Tanner (2015). “The cost of passenger delay to airlines in Europe-consultation document”. In.

Cook, Andrew J and Graham Tanner (2015). “European airline delay cost reference values - updated and extended values (Version 4.1)”. In.

Eggenberg, Niklaus, Matteo Salani, and Michel Bierlaire (2010). “Constraint-specific recovery network for solving airline recovery problems”. In: Computers Operations Research 37.6, pp. 1014–1026. issn: 0305-0548. doi: https://doi.org/10.1016/j.cor.2009.08.006. url: https://www.sciencedirect.com/science/article/pii/S030505480900210X.

Eurocontrol (2019). “EUROCONTROL Think Paper 5 - Effects on the network of extra standby aircraft and Boeing 737 MAX grounding”. In: Eurocontrol: Brussels, Belgium. url: https://www.eurocontrol.int/sites/default/files/2020-01/eurocontrol-think-paper-5-737.pdf.

Gershkoff, Ira (2016). “Shaping the future of airline disruption management (IROPS)”. In: Commissioned by Amadeus.

Goehry, Benjamin et al. (2021). “Random forests for time series”. In.

Groeneveld, Diederick (2021). “Reserve fleet capacity”. In: url: https://repository.tudelft.nl/islandora/object/uuid:6fe80463-d6f3-470e-a360-9dfdbd92f521?collection=education.

Guimarans, Daniel, Pol Arias, and Miguel Mujica Mota (2015). “Large Neighbourhood Search and Simulation for Disruption Management in the Airline Industry”. In: Applied Simulation and Optimization: In Logistics, Industrial and Aeronautical Practice. Ed. by Miguel Mujica Mota, Idalia Flores De La Mota, and Daniel Guimarans Serrano. Cham: Springer International Publishing, pp. 169–201. isbn: 978-3-319-15033-8. doi: 10.1007/978-3-319-15033-8_6. url: https://doi.org/10.1007/978-3-319-15033-8_6.

Hassan, LK, Bruno Filipe Santos, and Jeroen Vink (2021). “Airline disruption management: A literature review and practical challenges”. In: Computers & Operations Research 127, p. 105137.

Homaie-Shandizi, Amir-Hosein et al. (2016). “Flight deck crew reserve: From data to forecasting”. In: Engineering Applications of Artificial Intelligence 50, pp. 106–114. issn: 0952-1976. doi: https://doi.org/10.1016/j.engappai.2016.01.028. url: https://www.sciencedirect.com/science/article/pii/S0952197616000324.

Hsu, Chaug-Ing et al. (2011). “Aircraft replacement scheduling: A dynamic programming approach”. In: Transportation Research Part E: Logistics and Transportation Review 47.1, pp. 41–60. issn: 1366-5545. doi: https://doi.org/10.1016/j.tre.2010.07.006. url: https://www.sciencedirect.com/science/article/pii/S1366554510000773.

Januschowski, Tim et al. (2022). “Forecasting with trees”. In: International Journal of Forecasting 38.4. Special Issue: M5 competition, pp. 1473–1481. issn: 0169-2070. doi: https://doi.org/10.1016/j.ijforecast.2021.10.004. url: https://www.sciencedirect.com/science/article/pii/S0169207021001679.

Kafle, Nabin and Bo Zou (2016). “Modeling flight delay propagation: A new analytical-econometric approach”. In: Transportation Research Part B: Methodological 93, pp. 520–542. issn: 0191-2615. doi: https://doi.org/10.1016/j.trb.2016.08.012. url: https://www.sciencedirect.com/science/article/pii/S0191261515302010.

Kohl, Niklas et al. (2007). “Airline disruption management—Perspectives, experiences and outlook”. In: Journal of Air Transport Management 13.3, pp. 149–162. issn: 0969-6997. doi: https://doi.org/10.1016/j.jairtraman.2007.01.001. url: https://www.sciencedirect.com/science/article/pii/S0969699707000038.

Lee, Junhyeok, Kyungsik Lee, and Ilkyeong Moon (2022). “A reinforcement learning approach for multi-fleet aircraft recovery under airline disruption”. In: Applied Soft Computing 129, p. 109556. issn: 1568-4946. doi: https://doi.org/10.1016/j.asoc.2022.109556. url: https://www.sciencedirect.com/science/article/pii/S1568494622006226.

Lettovsky, Ladislav (1997). Airline operations recovery: an optimization approach. Georgia Institute of Technology.

Liang, Zhe et al. (2018). “A column generation-based heuristic for aircraft recovery problem with airport capacity constraints and maintenance flexibility”. In: Transportation Research Part B: Methodological 113, pp. 70–90. issn: 0191-2615. doi: https://doi.org/10.1016/j.trb.2018.05.007. url: https://www.sciencedirect.com/science/article/pii/S0191261517310421.

Lin, Hongtao and Zhong Wang (2018). “Fast Variable Neighborhood Search for Flight Rescheduling After Airport Closure”. In: IEEE Access 6, pp. 50901–50909. issn: 2169-3536. doi: 10.1109/ACCESS.2018.2869842.

Maher, Stephen J. (2016). “Solving the Integrated Airline Recovery Problem Using Column-and-Row Generation”. In: Transportation Science 50.1, pp. 216–239. doi: 10.1287/trsc.2014.0552. eprint: https://doi.org/10.1287/trsc.2014.0552. url: https://doi.org/10.1287/trsc.2014.0552.

Makridakis, Spyros, Evangelos Spiliotis, and Vassilios Assimakopoulos (2022). “The M5 competition: Background, organization, and implementation”. In: International Journal of Forecasting 38.4. Special Issue: M5 competition, pp. 1325–1336. issn: 0169-2070. doi: https://doi.org/10.1016/j.ijforecast.2021.07.007. url: https://www.sciencedirect.com/science/article/pii/S0169207021001187.

Marla, Lavanya, Bo Vaaben, and Cynthia Barnhart (2017). “Integrated Disruption Management and Flight Planning to Trade Off Delays and Fuel Burn”. In: Transportation Science 51.1, pp. 88–111. doi: 10.1287/trsc.2015.0609. eprint: https://doi.org/10.1287/trsc.2015.0609. url: https://doi.org/10.1287/trsc.2015.0609.

Nielsen, Didrik (2016). “Tree boosting with xgboost-why does xgboost win" every" machine learning competition?” MA thesis. NTNU.

Oum, Tae Hoon, Anming Zhang, and Yimin Zhang (2000). “Optimal demand for operating lease of aircraft”. In: Transportation Research Part B: Methodological 34.1, pp. 17–29. issn: 0191-2615. doi: https://doi.org/10.1016/S0191-2615(99)00010-7. url: https://www.sciencedirect.com/science/article/pii/S0191261599000107.

Petersen, Jon D. et al. (2012). “An Optimization Approach to Airline Integrated Recovery”. In: Transportation Science 46.4, pp. 482–500. doi: 10.1287/trsc.1120.0414. eprint: https://doi.org/10.1287/trsc.1120.0414. url: https://doi.org/10.1287/trsc.1120.0414.

Rosenberger, Jay M., Ellis L. Johnson, and George L. Nemhauser (2003). “Rerouting Aircraft for Airline Recovery”. In: Transportation Science 37.4, pp. 408–421. doi: 10.1287/trsc.37.4.408.23271. eprint: https://doi.org/10.1287/trsc.37.4.408.23271. url: https://doi.org/10.1287/trsc.37.4.408.23271.

— (2004). “A Robust Fleet-Assignment Model with Hub Isolation and Short Cycles”. In: Transportation Science 38.3, pp. 357–368. doi: 10.1287/trsc.1030.0038. eprint: https://doi.org/10.1287/trsc.1030.0038. url: https://doi.org/10.1287/trsc.1030.0038.

Rosenberger, Jay M., Andrew J. Schaefer, et al. (2002). “A Stochastic Model of Airline Operations”. In: Transportation Science 36.4, pp. 357–377. doi: 10.1287/trsc.36.4.357.551. eprint: https://doi.org/10.1287/trsc.36.4.357.551. url: https://doi.org/10.1287/trsc.36.4.357.551.

Santos, Bruno F. (2021). AE4423Lct3.4- H&S Fleet Models. https://brightspace.tudelft.nl/d2l/le/content/398016/viewContent/2296805/View. Accessed: 2023–06-09.

Schick, GJ and JW Stroup (1981). “Experience with a multi-year fleet planning model”. In: Omega 9.4, pp. 389–396. issn: 0305-0483. doi: https://doi.org/10.1016/0305-0483(81)90083-9. url: https://www.sciencedirect.com/science/article/pii/0305048381900839.

Shube, David P and John W Stroup (1975). Fleet Planning Model. Tech. rep. Institute of Electrical and Electronics Engineers (IEEE).

Sinclair, Karine, Jean-François Cordeau, and Gilbert Laporte (2014). “Improvements to a large neighborhood search heuristic for an integrated aircraft and passenger recovery problem”. In: European Journal of Operational Research 233.1, pp. 234–245. issn: 0377-2217. doi: https://doi.org/10.1016/j.ejor.2013.08.034. url: https://www.sciencedirect.com/science/article/pii/S0377221713007182.

— (2016). “A column generation post-optimization heuristic for the integrated aircraft and passenger recovery problem”. In: Computers Operations Research 65, pp. 42–52. issn: 0305-0548. doi: https://doi.org/10.1016/j.cor.2015.06.014. url: https://www.sciencedirect.com/science/article/pii/S0305054815001653.

Sohoni, Milind G, Ellis L Johnson, and T Glenn Bailey (2006). “Operational airline reserve crew planning”. In: Journal of Scheduling 9.3, pp. 203–221.

Swan, William M. (2002). “Airline demand distributions: passenger revenue management and spill”. In: Transportation Research Part E: Logistics and Transportation Review 38.3, pp. 253–263. issn: 1366-5545. doi: https://doi.org/10.1016/S1366-5545(02)00009-1. url: https://www.sciencedirect.com/science/article/pii/S1366554502000091.

Tang, Jiafu and Yu Wang (2015). “An adjustable robust optimisation method for elective and emergency surgery capacity allocation with demand uncertainty”. In: International Journal of Production Research 53.24, pp. 7317–7328. doi: 10.1080/00207543.2015.1056318. eprint: https://doi.org/10.1080/00207543.2015.1056318. url: https://doi.org/10.1080/00207543.2015.1056318.

Teodorović, Dušan and Slobodan Guberinić (1984). “Optimal dispatching strategy on an airline network after a schedule perturbation”. In: European Journal of Operational Research 15.2, pp. 178–182. issn: 0377-2217. doi: https://doi.org/10.1016/0377-2217(84)90207-8. url: https://www.sciencedirect.com/science/article/pii/0377221784902078.

Teodorović, Dušan and Goran Stojković (1990). “Model for operational daily airline scheduling”. In: Transportation Planning and Technology 14.4, pp. 273–285. doi: 10.1080/03081069008717431. eprint: https://doi.org/10.1080/03081069008717431. url: https://doi.org/10.1080/03081069008717431.

Teoh, Lay Eng and Hooi Ling Khoo (2015). “Airline Strategic Fleet Planning Framework”. In: Journal of the Eastern Asia Society for Transportation Studies 11, pp. 2258–2276. doi: 10.11175/easts.11.2258.

Toqué, Florian et al. (2017). “Short long term forecasting of multimodal transport passenger flows with machine learning methods”. In: 2017 IEEE 20th International Conference on Intelligent Transportation Systems (ITSC), pp. 560–566. doi: 10.1109/ITSC.2017.8317939.

Varenna, Sara (2023). “A Stochastic Discrete Event Simulation of Airline Network and Maintenance Operations”. In: url: https://repository.tudelft.nl/islandora/object/uuid%5C%3Ae2de25a7-971e-4575-b2d9-e38392ff25ee?collection=education.

Vink, J. et al. (2020). “Dynamic aircraft recovery problem - An operational decision support framework”. In: Computers Operations Research 117, p. 104892. issn: 0305-0548. doi: https://doi.org/10.1016/j.cor.2020.104892. url: https://www.sciencedirect.com/science/article/pii/S0305054820300095.

Vos, Hans-Wieger M., Bruno F. Santos, and Thomas Omondi (2015). “Aircraft Schedule Recovery Problem – A Dynamic Modeling Framework for Daily Operations”. In: Transportation Research Procedia 10. 18th Euro Working Group on Transportation, EWGT 2015, 14-16 July 2015, Delft, The Netherlands, pp. 931–940. issn: 2352-1465. doi: https://doi.org/10.1016/j.trpro.2015.09.047. url: https://www.sciencedirect.com/science/article/pii/S2352146515002343.

Walker, C (2019). “All-Causes Delay and Cancellations to Air Transport in Europe: Annual Report for 2019”. In: Eurocontrol: Brussels, Belgium. url: https://www.eurocontrol.int/archive_download/all/node/12132.

Wong, Jinn-Tsai and Shy-Chang Tsai (2012). “A survival model for flight delay propagation”. In: Journal of Air Transport Management 23, pp. 5–11. issn: 0969-6997. doi: https://doi.org/10.1016/j.jairtraman.2012.01.016. url: https://www.sciencedirect.com/science/article/pii/S0969699712000178.

Wu, Cheng-Lung (2005). “Inherent delays and operational reliability of airline schedules”. In: Journal of Air Transport Management 11.4, pp. 273–282. issn: 0969-6997. doi: https://doi.org/10.1016/j.jairtraman.2005.01.005. url: https://www.sciencedirect.com/science/article/pii/S0969699705000141.

Wu, Cheng-Lung and Kristie Law (2019). “Modelling the delay propagation effects of multiple resource connections in an airline network using a Bayesian network model”. In: Transportation Research Part E: Logistics and Transportation Review 122, pp. 62–77. issn: 1366-5545. doi: https://doi.org/10.1016/j.tre.2018.11.004. url: https://www.sciencedirect.com/science/article/pii/S1366554517309857.

Xiong, Xiong et al. (2022). “A short-term wind power forecast method via xgboost hyper-parameters optimization”. In: Frontiers in energy research 10, p. 905155.

Zhang, Yanru and Ali Haghani (2015). “A gradient boosting method to improve travel time prediction”. In: Transportation Research Part C: Emerging Technologies 58. Big Data in Transportation and Traffic Engineering, pp. 308–324. issn: 0968-090X. doi: https://doi.org/10.1016/j.trc.2015.02.019. url: https://www.sciencedirect.com/science/article/pii/S0968090X15000741.

Zhao, Ai, Jonathan F. Bard, and J. Eric Bickel (2023). “A two-stage approach to aircraft recovery under uncertainty”. In: Journal of Air Transport Management 111, p. 102421. issn: 0969-6997. doi: https://doi.org/10.1016/j.jairtraman.2023.102421. url: https://www.sciencedirect.com/science/article/pii/S0969699723000649.