Passenger-Centric Integrated Airline Schedule and Aircraft Recovery

Luis Cadarso
Aerospace Systems and Transport Research Group, European Institute for Aviation Training and Accreditation (EIATA), Rey Juan Carlos University, Fuenlabrada, Madrid, 28943, Spain
Email: luis.cadarso@urjc.es

Vikrant Vaze
Thayer School of Engineering, Dartmouth College
14 Engineering Drive, Hanover NH 03755, United States of America
Email: vikrant.s.vaze@dartmouth.edu

Abstract
Airlines are known to compete for passengers, and airline profitability heavily depends on the ability to estimate passenger demand, which in turn depends on flight schedules, fares and the number of seats available at each fare, across all airlines. Interestingly, such competitive interactions and passenger substitution effects may not be limited to the planning stages. Existing regulations in some countries and regions impose monetary compensations to passengers in case of disruptions, altering the way they perceive the utility of other travel alternatives after the disruption starts. These passenger rights regulations may act as catalysts of passengers' response to recovered schedules. Ignoring such passenger response behavior under operational disruptions may lead airlines to develop subpar recovery schedules. We develop a passenger response model and embed it into a novel integrated optimization approach that recovers airline schedules, aircraft, and passenger itineraries while endogenizing the impacts of airlines' decisions on passenger compensation and passenger response. We also develop an original solution approach, involving exact linearization of the non-linear passenger cost terms, combined with delayed constraint generation for ensuring aircraft maintenance feasibility and an acceleration technique that penalizes deviations from planned schedules. Computational results on real-world problem instances from two major European airlines are reported, for scenarios involving disruptions, such as delayed flights, airport closures and unexpected grounding of aircraft. Our approach is found to be tractable and scalable, producing solutions that are superior to airline's actual decisions, and highly robust in the face of passenger response uncertainty. Of particular relevance to the practitioners, our simulation results highlight that accounting for passengers' disruption response behaviors – even in a highly approximate manner – yields significant benefits to the airline, as compared to not accounting for them at all, which is the current state-of-the-art.

Keywords: airline, disruption, recovery, aircraft, tail, flight schedule, maintenance, passenger response.

1. Introduction

Airline schedules generated months before the day of operations are seldom

implemented exactly as planned, mainly due to disruptions (Barnhart and Vaze, 2015). Airlines often need to adjust the schedule for the time period of the disruption, and then carry out further recovery steps in order to get the operations back on track. Consequently, when the disruptions happen, airlines cannot guarantee that the original passenger itineraries will be preserved. Instead, disruptions usually require the airline to reschedule some passengers onto alternative itineraries to take them to their destinations as soon, and as inexpensively, as possible.

During the last two decades, with worsening passenger delays and disruptions, multiple government regulations have been proposed and implemented to protect air passengers' rights. These legislations typically include specific clauses providing monetary or in-kind compensations to the passengers when they face certain types of schedule disruptions. The specifics of these regulations differ in different countries and regions of the world. These include, for example, the European Union Regulation (EC) 261/2004, the United States' air "Passenger Bill of Rights", and the "Montreal Convention".

1.1 Passenger Rights Regulations

The Montreal Convention (IATA, 1999), signed in 1999 and effective since 2003, covers passengers on international flights between nations that honor the regulation. It is signed and recognized by more than 120 countries around the world, including the United States and the members of the European Union (EU). It provides rights and potential compensation for passengers suffering from several types of flight disruptions, including delays, flight cancelations, and denied boardings.

The United States also has its own laws to protect passengers who are denied boarding and/or are experiencing long tarmac delays. In 2008, the Department of Transportation instituted the law protecting passengers in case of denied boarding due to overbooking, clearly establishing compensation levels to disrupted passengers (USDOT, 2008). Similarly, in 2010, a rule was developed to enhance passenger protections in case of long tarmac delays, ensuring that passengers' essential needs are met, i.e., food and potable water are provided (USDOT, 2010).

The EU law also protects passengers and it is one of the most comprehensive such regulations in the world. It states that, in many instances, airlines are legally and financially responsible for flight issues. EU Regulation (EC) No 261/2004 (European Commission, 2004) repealed Regulation (EC) 295/91 and went into effect in 2005. Several interpretations of the Court of Justice of the European Union (CJEU), starting from 2009 (CJEU, 2009), have stated passenger rights strictly, prohibiting airlines from evasion of established obligations. Depending on the characteristics of the flight, disruption scenario, and the destination under consideration, this law establishes different reimbursement levels for denied boardings, flight cancelations, and long delays. In addition to monetary compensation, it includes other rights such as the obligation to inform passengers of their rights and care in terms of meals, hotel and communications.

Due to these different regulations, airlines incur additional costs from disruptions, and

hence stand to gain considerably by explicitly accounting for these costs when they make their operations recovery decisions. Yet, we are not aware of any study in the existing literature that attempts to explicitly include these impacts when solving the airline recovery problem.

1.2 Airline Disruption Management

There are several ways of dealing with disruptions: proactively, reactively, or both. We describe these ideas in Sections 1.2.1 and 1.2.2.

1.2.1 Proactive Approach: Robust Scheduling

Being proactive means building robustness into the schedules in an intelligent way, typically by placing buffers strategically to absorb disruption effects. The goal of robust planning is to make schedules less sensitive to disruptions. On the flipside, there is usually a price to be paid for adding robustness to schedules. This price may be in terms of the need to use more aircraft to fly the same schedule, or additional crew salary costs due to extra standby crews. Moreover, these robust schedules must be planned weeks before the day of operations, when there is still a tremendous amount of uncertainty about how the actual operations will manifest. In the current airline practice, the scheduling problem is often solved sequentially, despite the existence of some approaches for partial integration (Cadarso et al., 2017; Cacchiani and Salazar-González, 2016). Consequently, robustness may be introduced at any stage(s) of airline planning. There are several noteworthy studies on robust scheduling applications – see Sohoni et al. (2011) and Jiang and Barnhart (2013) for robust flight schedule optimization, Lan et al. (2006) and Yan and Kung (2018) for robust approaches to reduce delay propagation in aircraft routings, and Antunes et al. (2019) for robust crew pairing optimization. Ahmadbeygi et al. (2010) and Chiraphadhanakul and Barnhart (2013) studied slack reallocation in order to reduce schedule perturbations during operations. Cadarso and Marín (2013) and Cadarso and de Celis (2017) developed integrated robust approaches to optimize schedule design and fleet assignment, especially addressing the demand uncertainty. Gao et al. (2009) and Dunbar et al. (2012, 2014) developed integrated robust approaches for aircraft and crew optimization. Recently, Marla et al. (2018) performed a detailed comparison of the advantages and disadvantages of multiple robust scheduling paradigms.

1.2.2 Reactive Approach: Recovery from Disruptions

Even with robust plans, there is often a need to recover operations once there is a disruption. In the current airline practice, the recovery problem is often solved sequentially (Barnhart and Vaze, 2015) – the typical recovery sequence includes schedule and aircraft recovery, followed by crew recovery and then the passenger recovery. A main recovery decision is about timetabling, i.e., whether to maintain flight punctuality or to delay or cancel certain flights. Next, the aircraft recovery problem is solved to reassign individual aircraft to fly the repaired schedule, while matching passengers' transportation needs with available seating capacity and satisfying aircraft maintenance requirements. Also, crews may have to be rescheduled while complying with the government regulations and collective bargaining labor agreements. Finally, the disrupted passengers are re-accommodated on the available seats in the recovered network in order to take them to their destinations in an efficient manner. Clausen et

al. (2010) presented a review of airline disruption management, including aircraft, crew, passenger and some integrated recovery approaches, with a special emphasis on aircraft and crew recovery. During operations recovery, decisions must be made in a near real-time mode, and therefore, computational complexity should be reduced as much as possible. That combined with the inherent difficulties in implementing significant schedule changes, it is common practice to minimize the extent to which the schedules are readjusted (Cadarso et al., 2015).

Rosenberger et al. (2003) presented and heuristically solved an optimization model for aircraft rerouting and flight cancelation. Lettovský et al. (2000) developed a solution framework, which provides, in almost real time, a recovery plan for reassigning crews to restore a disrupted crew schedule. Yu et al. (2003) proposed a decision-support system for Continental Airlines to generate globally optimal, or near-optimal, crew-recovery solutions. Zhang et al. (2016) proposed a three-stage sequential math-heuristic to efficiently solve the schedule, aircraft and passenger recovery problems. However, the sequential approach has limitations (Petersen et al., 2012): computing a new timetable without accounting for aircraft and passenger considerations may produce a suboptimal, or even infeasible timetable, for aircraft assignment and passenger recovery purposes.

Recent advances in computing technology and algorithms have allowed partial integration of the various components of the recovery problem. Examples include Sinclair et al. (2014), Maher (2015a), and Marla et al. (2016). Sinclair et al. (2014) developed an iterative heuristic for integrated aircraft and passenger recovery. Maher (2015a) proposed an integrated airline recovery approach, combining the schedule, crew, and aircraft recovery stages, and used column-and-row generation to solve it. Marla et al. (2016) proposed to minimize the sum of multiple operating costs, including incremental fuel costs, swap costs, and passenger-related delay costs. They introduced a model that approximately captures the tradeoff between fuel burn and delays.

1.2.3 Passenger impacts of airline disruption recovery

Passenger demand for air transportation is defined for each combination of origin airport, destination airport, desired departure time and desired arrival time. While flights are also defined in a similar manner (i.e., as a combination of origin airport, destination airport, scheduled departure time and scheduled arrival time), many combinations of origin and destination airports are typically not served by non-stop flights at all. Therefore, passengers travel on itineraries, which are sequences of one or more non-stop flights traversed from the passenger's origin to the passenger's destination. On the day of operations, passenger level of service may be measured in terms of the amount of delay that passengers suffer when arriving at their destination as compared to their planned arrival time. Explicit modeling of these network effects and the itinerary-based passenger demand concepts in the recent studies has significantly enhanced the modeling and understanding of passenger demand and flow.

Passenger rights regulations, which mandate monetary compensations to passengers in cases of disruptions, alter the way the passengers perceive the utility of other alternatives once their planned itinerary gets disrupted. Compensation is typically

dependent on distance flown and the extent of itinerary disruption. Monetary compensation guarantees may empower the passengers to use the compensation money (even though it is often received at a future date) to book a flight on another airline to reach their destination sooner. These passengers typically do not use their meal/hotel vouchers; instead they reach their destinations sooner with a new booking, and then file for a monetary compensation. We call them *phantom* passengers, defined as the passengers with a confirmed reservation, who decide not to show up for their recovered itinerary in case of a disruption to their original schedule.

Due to these additional complexities (i.e., network effects, itinerary-based modeling and compensations), many studies have either solved approximate models or resorted to sequential solution approaches by breaking the model down into several stages, to handle the computational complexity of the large-scale instances. Note that, to the best of the authors' knowledge, no prior recovery study has focused on passenger compensation and response. Moreover, in the existing recovery literature, the inability to apply exact optimization methods has necessitated many researchers to rely on heuristic solution approaches, leading to poor and/or non-guaranteed solution quality in terms of optimality gaps.

Some recent studies have focused on understanding, quantifying, and minimizing the impacts of delays, disruptions and recovery on the airline passengers. In one of the first such studies, Bratu and Barnhart (2006) presented airline schedule recovery models and algorithms for simultaneous aircraft and passenger recovery, while identifying if reserve crews are needed, to find the optimal trade-off between airline operating costs and passenger delay costs. They explicitly modeled passenger disruptions, recovery options, and delay costs, capturing relevant hotel costs, and ticket costs for passengers recovered by other airlines. However, they found that the model could not be solved in real time for day-long decision windows and proposed an approximation to model delay costs.

Petersen et al. (2012) presented an optimization-based approach to solve the fully integrated recovery problem but limited the recovery options to solve it tractably. They decomposed passenger recovery problem itself into two sequential problems: the itinerary recovery problem and the passenger re-accommodation model. The former generates a set of eligible itineraries for each origin-destination pair that minimize the aggregate delay costs, while the latter allocates passengers to these itineraries to minimize the total assignment cost. The problem is solved using a Benders Decomposition-based heuristic. Maher (2015b) improved the model in Maher (2015a) to implicitly include passenger recovery options for better estimation of flight cancelation costs. He considered passengers in an integrated recovery problem through the cancelation variables, modeled as knapsack variables, to describe reallocation options for passengers. This modeling approach provided a simple and effective method to consider passengers in recovery. However, the approach assumed that all passengers book single-flight journeys, which is usually not the case in reality. Hu et al. (2016) proposed an integrated recovery model of aircraft and passengers to reschedule disrupted flights while minimizing passenger inconvenience by swapping aircraft, delaying and canceling flights, or reassigning passengers. Aircraft maintenance was not accurately captured in their model, which was solved using a GRASP-based heuristic.

Arıkan et al. (2017) integrated aircraft, crew and passenger recovery decisions over a simplified network. They proposed a flight network representation different from the commonly used time-space network representation and flight string representation. They placed a special emphasis on passenger recovery, modeling each passenger explicitly. However, for larger networks they used approximations of passenger delay costs due to the inability to solve the exact model. Another limitation pointed out in Arıkan et al. (2017) was the inability to assign attributes to paths, thus limiting the modeling of aircraft maintenance restrictions. In this paper, we present an approach to model aircraft maintenance in detail, while explicitly including airport maintenance capacity constraints.

Airlines may lose additional passengers if the recovery itineraries are not acceptable to the passengers in terms of their level of service. For example, the datasets provided by our airline partners showed that, on an average, over 13% of the passengers did not show up for their recovered itineraries. The inability to account for this important phenomenon may produce inefficient recovery schedules, for example by providing more seats than needed on certain recovery itineraries. These issues were reported and identified as significant operational concerns by our airline partners. All prior studies ignore such responses by passengers to disruptions and recovered schedules (Cadarso et al., 2013). In fact, existing literature implicitly or explicitly assumes that all rebooked passengers show up for their recovered flights. Passengers' recovery response may be significantly influenced by the operator's recovery actions, level of compensation and the provided level of service. Ours is the first study to explicitly account for these effects. Given the considerable pressure in the past few years by many EU governments on their respective national flag carriers to improve passenger rebooking processes in the event of disruptions, we believe that this is a particularly important and timely research direction.

The aim of our research is to develop and solve an optimization model that integrates schedule, aircraft, and passenger recovery while avoiding many of these aforementioned approximations on the passenger modeling side. In our approach, passengers are individually traced through the entire networks, leading to exact modeling of their delays. However, passenger traceability and determination of itinerary delays inherently produce non-linearities when calculating the total delays and costs. These non-linearities are tackled in our study by introducing auxiliary variables and constraints, in turn providing a suitable framework for solving real-world large-scale case studies to within small optimality gaps in realistic runtime budgets. Also, a simulation approach for passenger behavior under disruption is presented to evaluate every recovery plan more accurately.

1.3 Contributions

We present a passenger-centric mixed-integer, non-linear programming model for the schedule, aircraft and passenger integrated recovery problem. For the first time, we explicitly account for passenger response to disruptions as a driver of airline costs. Our major contributions are threefold.

First, from a modeling standpoint, we develop a comprehensive approach which

captures regulated passenger rights and passengers' responses to recovery plans. Passenger response depends on recovered schedules, and specifically on their arrival time at destination, and may vary with the fare class since fare class is an (imperfect) indicator of willingness to pay. While determination of itinerary delays is usually straightforward, total passenger delay calculation is necessarily non-linear. Therefore, we model passenger flows comprehensively. A flight-based network structure is employed, including continuous flight re-timing. Despite the flight-based network structure, we are able to model aircraft maintenance in detail, including airport maintenance capacity constraints that are typically not modeled in the past studies. Also on the modeling side, we minimize schedule changes with respect to planned operations.

Second, from the solution process standpoint, we use exact methods to solve the model. Commercial solvers are usually not able to efficiently handle the combination of binary variables and non-linear expressions in such large-scale problem instances. To overcome this, auxiliary variables and constraints are introduced to linearize non-linear relationships in an exact, rather than approximate, manner. Our approach uses delayed constraint generation to comply with aircraft maintenance requirements including maximum flying hours and maximum number of take-offs and landings until the next maintenance.

Finally, we evaluate and validate our approach with realistic case studies featuring recovery periods of up to 3 days, over 1400 flights and up to 5 fleet types. We consistently obtain small optimality gaps in computational times that are small enough to be compatible with real-time operational needs. Optimization and simulation results and comparisons under real-world scenarios validate the effectiveness of the proposed approach. From a practical standpoint, our results show that capturing passengers' disruption response behaviors – even in a highly approximate manner – is significantly more profitable to the airline, as compared to not modeling them at all, which is the current state-of-the-art.

1.4 Paper Outline

The remainder of this paper is organized as follows. Section 2 describes the problem in detail. Sections 3 and 4, respectively, present our passenger response modeling approach and our mathematical formulation for an integrated recovery. Section 5 details computational experiments, and their results, for real-world problem instances and airline-provided datasets. Section 6 presents two prominent extensions and additional results demonstrating the generalizability and scalability of our approach. Section 7 summarizes the major findings.

2. Problem Description

This section describes in detail the airline recovery problem. The airline network is represented by the well-known time space network where nodes represent airports at various points in time and arcs represent flights (called flight arcs) as well as time spent on ground at an airport (called ground arcs). The airports are characterized by the number of airport operations that can be performed per unit time. Arrival and departure slot availability is assumed to be known for each airport.

2.1 Disruptions

There are different types of disruptions, some are large-scale and others small-scale. Small-scale disruptions usually feature delays to some of the flights, due to various factors such as weather conditions, crew shortages and airport congestion. They may or may not require any flight cancelations or major changes to the schedule. Nevertheless, they may have a considerable impact on operations and costs. Large-scale disruptions, on the other hand, usually lead to at least some flight cancelations, and their effects typically propagate throughout the airline's network. A common cause of a large-scale disruption is an airport closure due to extreme weather conditions.

2.2 Schedule and Aircraft Recovery

The objective of the schedule recovery problem is to minimize the costs associated with the new flight schedules, including those corresponding to delays and cancelations. In addition to disruption-imposed physical constraints, infrastructure availability heavily influences the ability to re-schedule flights. Airport slots at the schedule-coordinated airports, aircraft maintenance capacity available at maintenance stations, as well as airspace capacity, may be scarce resources. Therefore, a schedule recovery process needs to explicitly account for such infrastructure availability constraints. Additional complexity arises because of the impact of recovery actions on several different parts of the air transportation network, including aircraft and passengers.

The aircraft recovery problem reassigns individual aircraft to fly the repaired schedule, while ensuring aircraft maintenance-feasibility. Airworthiness is a measure of an aircraft's suitability for operating a safe flight. An airworthiness certificate is a document provided by a national or regional agency (e.g., the Federal Aviation Administration in the US and European Union Aviation Safety Agency in Europe) which grants authorization to use that aircraft to operate a flight. This authorization can be maintained only by performing the required maintenance actions. Because reassigning aircraft to flights changes the number of flight hours and the location of aircraft during the planning period, required maintenance actions may also be affected. If the airworthiness certificate is lost as a result, the aircraft may have to be grounded. Consequently, each aircraft's maintenance requirements and maintenance opportunities must be carefully identified and tracked while re-scheduling.

A maintenance opportunity (or a maintenance arc) is defined as a ground arc that meets two criteria: both its end nodes are at an airport which is a maintenance station for that airline and the ground arc spans at least as much time as the maintenance process needs. These ground arcs are confined to the time period between a flight arrival and the subsequent flight departure. Note that if two flights are assigned the same aircraft, it is easy to know whether a maintenance opportunity exists between the arrival of the first flight and the departure of the next. Once an aircraft traverses a maintenance arc (i.e., once it undergoes maintenance), its status is reset allowing it to continue flying. Moreover, maintenance opportunities are limited by the available number of workshops. At each maintenance station (i.e., airport) the number of tails simultaneously undergoing maintenance must be limited to the number of workshops. Every maintenance station features one or more workshops, and the sequence of tails

undergoing maintenance in the same workshop is called a maintenance sequence.

Note that the capacity of an airline network is characterized in terms of the number of flights on each nonstop segment and the seating capacity on each of them. Because it is common for airlines to own or lease multiple fleet types, reassigning an aircraft whose capacity is as close as possible to the capacity of the originally assigned aircraft for that flight is crucial to maintain the level of service to passengers. In particular, assigning a smaller aircraft to a flight, which originally had a larger one assigned to it, may cause additional disruption to passengers due to lack of seating capacity. A larger than originally planned aircraft size assigned to a flight, in turn, may result in the airline flying empty seats, while also likely reducing the seat availability elsewhere in the network due to such an aircraft swap.

2.3 Passenger Itinerary Recovery

The passenger recovery process aims to mitigate passenger inconvenience. Given the new recovered schedule in terms of flight times and fleet assignment, the objective is to serve all passengers while maintaining the passenger level of service. The greater the delays and disruptions, the poorer the offered level of service to passengers. If passengers cannot reach destinations as planned (due to itinerary cancelation, missed connections, etc.), airlines try to re-schedule passengers such that they arrive at their destinations as quickly as possible. If reassignment is not possible during the considered planning horizon, i.e., there is no feasible itinerary offered by the airline under consideration to take the passengers to their destinations, passengers are usually either assigned to another airline or to the next planning horizon, increasing reassignment cost and/or additionally increasing the total delay. Passenger reassignment is usually associated with two types of costs: re-accommodation cost and passenger inconvenience cost. The first one is the cost of the assistance provided to passengers, i.e., meals, hotels, and accommodation on another airline if needed, while the second one is usually a proxy, which measures the loss of goodwill due to the disruption and the subsequent recovery of the itinerary.

As mentioned before, passenger recovery is further complicated by existing passenger rights regulations and the associated monetary compensations, which impact passenger behavior. The EU regulation establishes a compensation of €250 to €600 depending on distance flown, magnitude of the disruption (e.g., whether the delay is over 3 hours), and the type of disruption (e.g., large delays, cancelations, or denied boardings). In case of delays, for a given distance, compensation is a step-function of the arrival delay in minutes. The monetary compensation may result in passengers not showing up for the flight on which they are re-accommodated, especially in cases where the re-accommodation corresponds to large delays. Therefore, airlines may lose passengers if the recovery is not good enough. Not accounting for this fact may produce inefficient schedules, for example, by providing more seats than needed on certain flights. Section 3 models the effects of compensation on passenger response to delays and recovered schedules.

2.4 Crew Recovery

The objective of the crew recovery process is to find a minimum cost reassignment of

crews to a disrupted flight schedule while satisfying all operational and legal constraints. During the planning phase, crew pairings – i.e., sequences of flights, as well as stop and rest periods – are designed to be flown by specific crews. Pairings span one or more days and are composed of duties. Each duty is a day-long work period consisting of one or more flights. Rosters are composed of pairings, and they usually cover a time period of one month. Duties, pairings and rosters must satisfy a myriad of constraints, including constraints on minimum rest between two consecutive duties, maximum time away from the crew's base, maximum flying time per duty, etc. The strictness and complexity of these crew regulations imply that small changes in schedules, e.g., flight retimings due to disruptions, may completely disrupt crew schedules – sometimes even making them infeasible to operate.

3. Passenger Response Model

Passenger related delay costs, namely re-accommodation, inconvenience and compensation costs, are non-decreasing functions of delays at destination. Moreover, passengers can react to recovered schedules by deciding not to fly, changing flights within the same airline – an option that is sometimes provided by the host airline, or even by changing the airline, while partly recovering or outright losing the original ticket fare with host airline. Ignoring passengers' response to the schedule recovery actions may not be realistic, especially if a sizeable fraction of passengers does not show up for the recovered schedules. According to the aforementioned passenger rights regulations, compensations to passengers depend on delays and distances to be flown. The greater the delays and distances, the greater the monetary compensation. Passenger rights regulations usually establish step functions to determine the amount of compensation; e.g., the EU regulation lists three delay levels for calculating compensation.

Additionally, our conversations with our industry partners as well as the datasets described in this section show that the higher the compensation the greater the probability of the passengers making a new reservation with another airline to reach their destination. These passengers are what we have defined earlier as *phantom* passengers: the passengers who have a confirmed reservation but do not show up for their rebooked itinerary in case of a disruption. This section presents a regression model to explicitly account for passenger response to recovered schedules. Its goal is to estimate the passenger no-show rate (i.e., phantom rates) once the recovered schedule is revealed, as a function of monetary compensations to delays and the delays themselves.

3.1 Available Data

Datasets from multiple realistic case studies featuring several origin-destination pairs and multiple delay levels were made available for this research by our airline partners. Because our industry partners are based in Europe, at least three delay levels should be considered as stated in the EU regulation: no-compensation level, first level of compensation and second level of compensation. We further divided the no-compensation delay range into three levels to result in a total of five delay levels, such that for each value of delay at destination, a delay level of 1, 2, 3, 4, or 5 is assigned. Delay levels from 1 to 4 span 60 minutes each (0-59, 60-119, 120-179 and 180-239 minutes), while delay level 5 corresponds to greater than or equal to 4 hours of delays.

We had 30 different disruption scenarios available for building the passenger response model. These scenarios feature information on fare-class, delay at destination in minutes, total number of passengers for each combination of fare class and delay, and the corresponding number of phantom passengers. Because of data limitations, the information on passengers on different origin-destination pairs, flying the same fare-class and suffering the same delay level is aggregated to produce stable phantom rate forecasts.

Once data aggregation is done, it is organized in a table featuring 379 rows, one for each combination of fare-class and delay level. Table 1 shows an example row in our dataset. A phantom rate can be calculated for each row. Here, phantom rate is defined as the fraction or probability of passengers, in each combination of fare-class and delay level, that do not show up for a recovered schedule. The aim of the passenger response model in the next subsection is to estimate these phantom rates for recovered schedules. Note that fare-class is defined by positive integers from 1 to 15 where a lower value indicates a higher fare-class.

Table 1: Example of the available information regarding phantom passengers

Fare Class	Delay Level	Total Number of Passengers	Number of Phantom Passengers
1	4	12	4

Because of lack of data, the responses of passengers to denied boardings and cancelations are assumed to be the same as those of passengers suffering a delay corresponding to the highest delay level.

3.2 Regression Model

We employ linear regression to model the relationship between phantom rates and explanatory variables. Our initial attempts at developing a detailed multivariate model to identify the dependence of phantom rates on compensations, distances and delay levels led to low statistical significance values. This is primarily due to the limitations of the available dataset combined with the expected high correlations among fares, distances, and compensations. Note that airlines typically have access to more fine-grained and richer datasets that should enable the development of more advanced statistical estimation models. If and when such models become available, they can be equally easily integrated into our integrated recovery formulation in Section 4. Also note that distances are static and cannot be controlled by airlines, as opposed to delays, which can be controlled by airlines to some extent. Therefore, it is essential that we capture the dependence of passengers' responses, first and foremost, on delay levels. However, the passengers may respond very differently to the same delay level depending on the fare class. This is so because the fare-class is often associated with the trip purpose and socio-economic characteristics of the passengers. More expensive fare-classes usually mean greater purchasing power, and therefore, a greater willingness to pay to arrive on-time (or closer to on-time).

Consequently, we specify a model to estimate phantom rate as a function of delay level

and fare-class: $\theta_{p,v}^{p',\zeta} = \alpha^{\zeta} + \beta^{\zeta} v + \varepsilon_{\zeta}$. Here, $\theta_{p,v}^{p',\zeta}$ is the phantom rate for passengers in fare class v displaced from itinerary p to itinerary p' experiencing an arrival delay of level ζ, α^{ζ} is an intercept specific to each delay level, β^{ζ} is the coefficient for fare-class specific to each delay level, and ε_{ζ} is the corresponding error term.

3.3 Estimation and Results

This regression model has two parameters, α^{ζ} and β^{ζ}, to be estimated for each delay level. As mentioned in Section 3.1, we consider five delay levels, following the idea behind the EU regulation, and dividing the no-compensation delay range into three delay levels. So the first and second compensations levels correspond to delay levels 4 and 5, respectively. Phantom rates are only estimated for delay levels with monetary compensation. We assume that without monetary compensation, the way the passengers perceive the utility of other alternatives is not altered and hence they do not decide to switch airlines. Moreover, the data provided by our airline partners included no phantom passengers for other delay levels, thus validating our approach. Delay levels 1, 2 and 3 incur an inconvenience cost as a proxy for the loss of passenger goodwill, while delay levels 4 and 5 incur both an inconvenience cost and a compensation cost.

Table 2 shows the regression results from using the standard method of least squares. To evaluate fit, R^2 value is calculated, i.e., the percentage of the response variable variation that is explained by the proposed linear model. Figures 1 and 3 show the fitted models for delay levels 4 and 5, respectively, and Figures 2 and 4 show the corresponding residual plots.

The results in Table 2 show that the greater the delay, the greater the probability of becoming phantom, as reflected in the intercept specific to each delay level, α^{ζ}, which is positive for both levels, and is bigger for the higher delay level. Also, lower the fare, lower the phantom rate for both delay levels, as demonstrated by the negative values of β^{ζ} coefficients. This indicates that lower fare-class passengers are more reluctant to pay for a rebooked itinerary on another airline, consistent with their generally lower willingness-to-pay, than higher fare-class passengers. Overall, the proposed model correctly describes a high percentage of the variation in phantom rates for each delay level, as suggested by the high R^2 values. In addition, the distribution and absolute values of the residuals (Figures 2 and 4) confirm that the proposed regression model explains the variation in phantom rates quite well.

Table 2: Linear regression results for phantom rate estimation

Delay Level	α^{ζ}	β^{ζ}	R^2
4	0.290	-0.016	0.853
5	0.473	-0.028	0.814

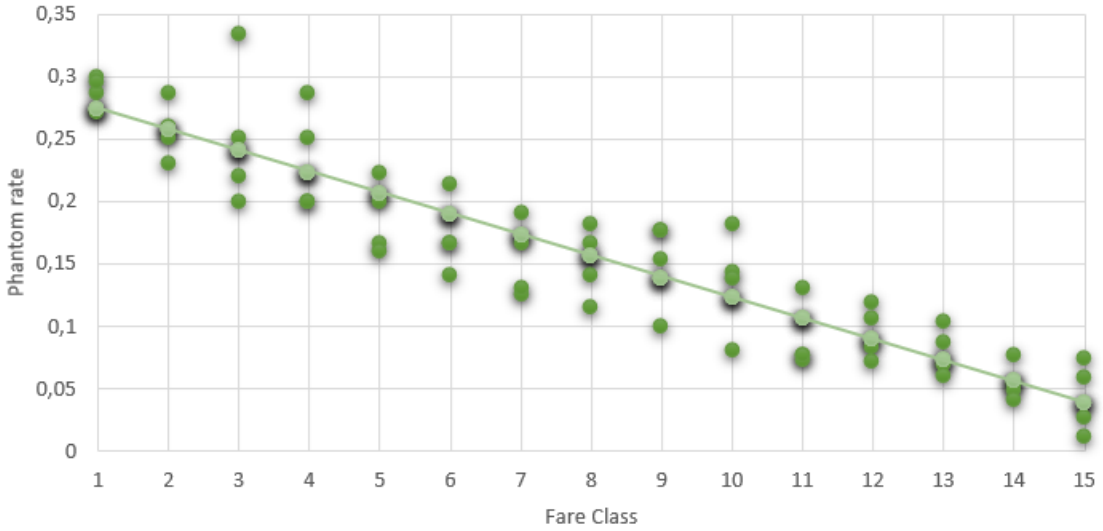

Figure 1: Linear regression for phantom rates at delay level 4 (180-239 min).

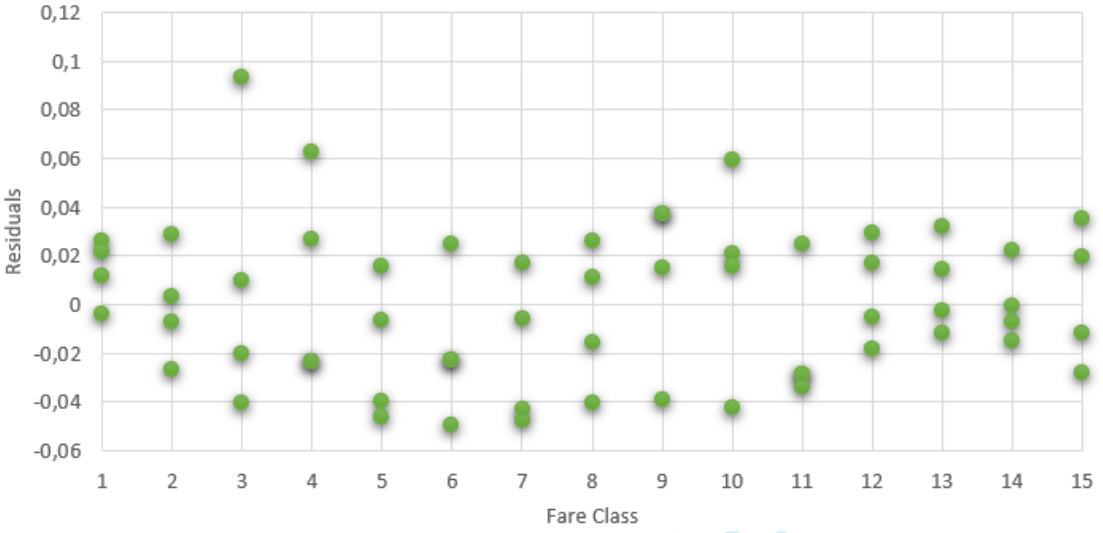

Figure 2: Residuals of the linear regression at delay level 4 (180-239 min).

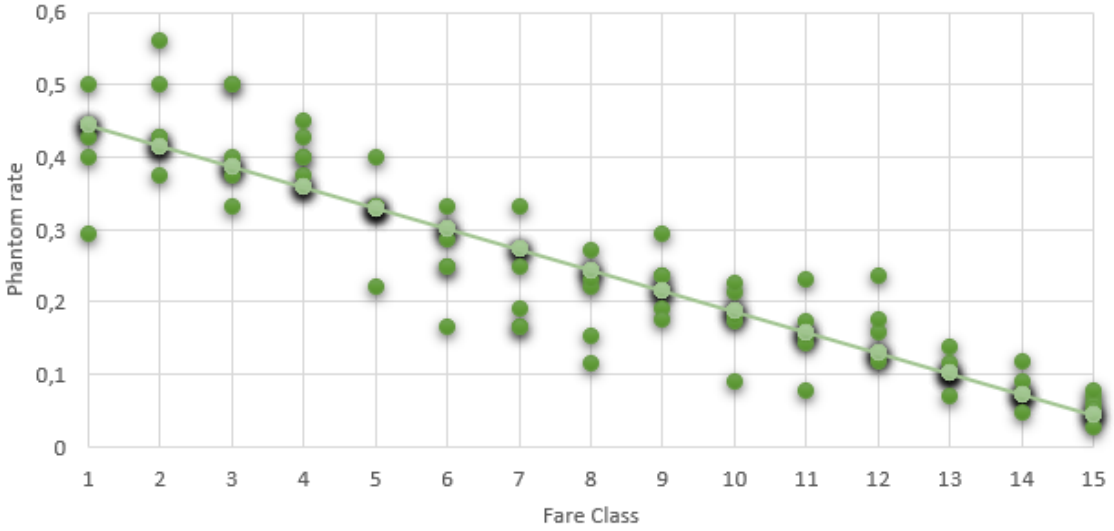

Figure 3: Linear regression for phantom rates at delay level 5 (≥ 240 min).

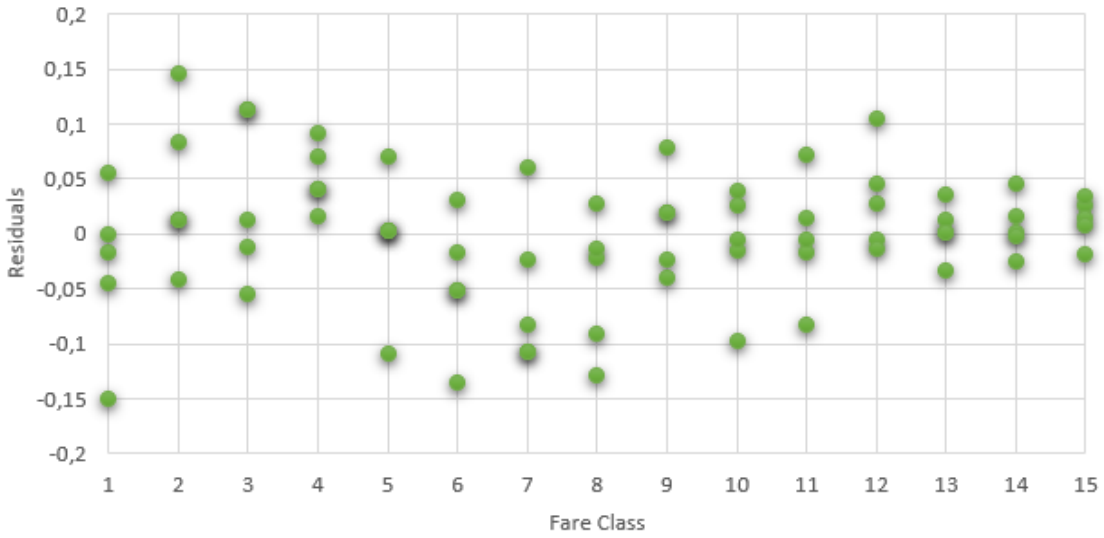

Figure 4: Residuals of the linear regression at delay level 5 (≥ 240 min).

4. Optimization Model Formulation

The aim of the Integrated Model for Airline Operations Recovery (IMAOR) is to determine the optimal set of schedule, aircraft and passenger recovery actions that will minimize the total recovery cost. The complexity of crew schedules, driven by the myriad constraints that must be satisfied, implies that changes in flight and/or aircraft schedules might potentially disrupt them and hence we must consider the way these changes may impact crew schedules. But, because of the need for real-time recovery decisions, the involved entities, namely aircraft, crew and passengers, are either modeled approximately or even neglected (Bratu and Barnhart, 2006; Maher, 2015a and 2015b; Arıkan et al., 2017) or the overall problem is solved heuristically (Petersen et al., 2012; Zhang et al., 2015). Because the emphasis of the model presented in this paper is on schedule, passenger and aircraft recovery, crew schedules are indirectly accounted for so that their consistency is traced and maintained as much as possible. This is consistent with integrated recovery studies in the literature (Bratu and Barnhart, 2006).

We use a flight-based model that captures aircraft maintenance constraints in a new way using delayed constraints generation. The model takes, as known input parameters, the following: the set of airports, the set of flights, slot availability (measured in terms of the number of arrival and departure operations) at each airport during each time period, aircraft availability, maintenance data (i.e., rules, stations and capacity), disruption information (i.e., nature, place, time, and duration), originally scheduled passenger flows (by fare-class and itinerary), passenger compensation rules and the estimated phantom rates model.

We refer to individual aircraft as individual *tails* since the unique identifying number for an aircraft is typically written on its tail. The notation in IMAOR is defined as follows:

4.1 Sets

- T: set of tails.

- Π: set of time instances evenly distributed throughout the recovery period. $|\Pi|$ determines the resolution of the temporal mesh.
- K: set of airports.
- F: set of flights.
- P: set of itineraries. Each itinerary $p \in P$ is an ordered set of flights. P also contains the null itinerary, defined as the one that the passengers are considered to be reassigned to when reassignment is not possible during the recovery period.
- Υ: set of fare-classes.
- Z: set of delay levels.
- C: set of feasible crew pairings. They can be thought of as ordered sets of flights.
- MO: set of maintenance opportunities, defined by pairs of flights (f, f') which if operated consecutively by the same tail could potentially allow the tail to undergo maintenance between them (depending on the exact flight departure and arrival times in the recovery solution). See Constraints (14)-(17) in Section 4.7 for more details.
- AA (DA): set of arrival (departure) slots, indexed by as (ds).
- $AA_f \subseteq AA$ ($DA_f \subseteq$ DA): set of arrival (departure) slots compatible with flight f. It is the set of all slots at the flight's destination (origin) airport, which are at or after the planned arrival (departure) time of the flight.
- $F_t \subseteq$ F: set of flights compatible with tail t.
- $F_\pi^\Pi \subseteq$ F: set of flights scheduled to finish at or before time instance π.
- $FA_k \subseteq$ F ($FD_k \subseteq$ F): set of flights that arrive in (depart from) airport k.
- $F_{as}^{\mathrm{AA}} \subseteq$ F ($F_{ds}^{\mathrm{DA}} \subseteq$ F): set of flights compatible with arrival slot as (departure slot ds).
- $T_f \subseteq$ T: set of tails compatible with flight f.
- $CO_p \subseteq$ P: subset of itineraries compatible with a reassignment from an original itinerary p. By definition, CO_p always includes the itinerary p itself, i.e., $p \in CO_p$.
- $CF_f \subseteq$ F: set of flights compatible with a connection from flight f. For a precise definition, please refer to Section 4.7.
- CF_p^P: set of all ordered flight pairs (f, f') of consecutive flights in itinerary p.
- $K^m \subseteq K$: subset of airports that have maintenance workshops.
- $T^m \subseteq T$: subset of tails that might need maintenance during the recovery horizon, depending on their extent of utilization.
- $\Pi^m \subseteq \Pi$: set of time instances at which maintenance requirements are checked by the model.
- CF_c^C: set of all ordered pairs (f, f') of consecutive flights in the same duty within crew pairing c. The time between the arrival of flight f and the departure of flight f' is known as the crew sit time.
- BF_c^C: set of all ordered pairs (f, f') of consecutive flights in crew pairing c, where f is the last flight in a duty and f' is the first flight in the next duty. The time between the arrival of flight f and the departure of flight f' is known as the crew rest time.

4.2 Parameters

- $oc_{t,f}$: cost of operating flight f with tail t.

- fc_f: extra fuel cost for delay absorption (through cruise speed increases) per minute for flight f.
- dc_f: delay cost per minute of arrival delay of flight f.
- $rc_p^{p'}$: re-accommodation cost for a passenger reassigned from itinerary p to itinerary p'. It includes items such as meal and hotel costs, if any.
- cc_c: cost of disrupting a crew pairing c due to flight cancelation, or insufficient crew sit or rest time between flights. It is a proxy to the cost of crew repositioning, or needing extra crews, for operating flights belonging to disrupted pairings.
- q_t: passenger seating capacity of the aircraft with tail t.
- $mtt_{f,t}^{f'}$: minimum turn time between flights f and f' for tail t.
- $mct_p^{f,f'}$: minimum connection time between flights f and f' in passenger itinerary p.
- $ct_f^{f'}$: planned connection time between flights f and f'. It equals scheduled departure time of flight f' minus the scheduled arrival time of flight f.
- $it_{as}^{AA}(ft_{as}^{AA})$: start (end) time of arrival slot as.
- $it_{ds}^{DA}(ft_{ds}^{DA})$: start (end) time of departure slot ds.
- $st_f^a(st_f^d)$: scheduled arrival (departure) time for flight f.
- $\theta_{p,v}^{p',\zeta}$: phantom rate for passengers in fare class v reassigned from a disrupted itinerary p to a new itinerary p' experiencing an arrival delay of level ζ.
- n_p^v: number of passengers in fare class v that are originally scheduled to take itinerary p.
- M: a big positive number.
- $lf_{f,p} \in \{0,1\}$: 1 if flight f is the last flight of itinerary p, and 0 otherwise.
- $\hat{x}_{t,f} \in \{0,1\}$: 1 if flight f was originally scheduled to be operated by tail t, and 0 otherwise.
- $tb_{t,k} \in \{0,1\}$: 1 if tail t is in airport k at the beginning of the recovery period, and 0 otherwise.
- ϑ_ζ: upper bound on the delay, expressed in minutes, corresponding to delay level ζ.
- κ: per-flight schedule change penalty for not operating the flight using the originally planned tail.
- abh_t: number of block hours available at the beginning of the recovery horizon for tail t before it requires maintenance.
- sbh_f: scheduled number of block hours, during the recovery period, for flight f.
- mbh_t: maximum allowable number of block hours between maintenances for tail t.
- mt_t: time required for performing the maintenance for tail t.
- $pc_p^{p',\zeta}$: sum of the cost of the loss of goodwill and the compensation cost (if any) for a passenger who was scheduled to take itinerary p and is reassigned to itinerary p', if the passenger's destination arrival delay via itinerary p' compared to the planned arrival time of itinerary p corresponds to delay level ζ.
- mst_c (mbt_c): required minimum crew sit (rest) time for crew pairing c.
- $\bar{\gamma}_f$: maximum possible delay absorption (through cruise speed increases) for

flight f, expressed in minutes.

- $\bar{\delta}_f$: maximum allowable delay for flight f.
- sb_f: schedule buffer time for flight f. Buffer times are usually added to flight times during the scheduling stage proactively to enhance recovery potential.
- $\underline{mtt}_f = min_{f' \in CF_f, t \in T_f}\{mtt_{f,t}^{f'}\}$: minimum of the minimum turn times for flight f.
- $sc_{as}^{AA}(sc_{ds}^{DA})$: capacity of the arrival slot *as* (departure slot *ds*).
- $maxct_f$: maximum connection time (for a tail) for flying a flight immediately after flight f, calculated as the maximum connection time to any other connecting flights that can be flown by the same tail.
- aw_k: number of available maintenance workshops in airport *k*.

4.3 Variables

- $x_{t,f} \in \{0,1\}$: 1 if tail t is assigned to flight f, and is 0 otherwise.
- $z_f \in \{0,1\}$: 1 if flight f is canceled, and is 0 otherwise.
- $y_{f,f'} \in \{0,1\}$: 1 if flights f and f' are consecutively flown by the same tail, and is 0 otherwise. This also allows the possibility of having the tail undergo a maintenance between flights f and f' and still be called consecutive.
- $\sigma_f \in \{0,1\}$: 1 if flight f is the last flight for a tail in the recovery period, and is 0 otherwise. Similarly, $\sigma_t^m \in \{0,1\}$ is 1 if tail t is the last one undergoing maintenance in the maintenance sequence of a workshop, and is 0 otherwise.
- $\rho_f \in \{0,1\}$: 1 if flight f is the first flight for a tail in the recovery period, and is 0 otherwise. Similarly, $\rho_t^m \in \{0,1\}$ is 1 if tail t is the first one undergoing maintenance in the maintenance sequence of a workshop, and is 0 otherwise.
- $\varphi_{t,f} \in \{0,1\}$: 1 if flight f is the first flight for tail t in the recovery period, and is 0 otherwise. Similarly, $\varphi_{t,f}^m \in \{0,1\}$ is 1 if tail t is the first one undergoing maintenance in a workshop in the recovery period and it does so immediately after flying flight f, and is 0 otherwise.
- $m_{t,t'} \in \{0,1\}$: 1 if tails t and t' consecutively undergo maintenance in the same workshop, and is 0 otherwise. Similarly, $m_{t,t',k}^m \in \{0,1\}$ is an auxiliary variable that is 1 if tails t and t' consecutively undergo maintenance in the same workshop at airport k, and is 0 otherwise.
- $v_{as,f}^{AA} \in \{0,1\}$ ($v_{ds,f}^{DA} \in \{0,1\}$): 1 if arrival slot as (departure slot ds) is assigned to flight f, and is 0 otherwise.
- $\alpha_p^{p',\zeta} \in \{0,1\}$: 1 if the arrival delay of itinerary p' with respect to the planned arrival time of itinerary p corresponds to delay level ζ, and is 0 otherwise.
- $\lambda_p \in \{0,1\}$: 1 if itinerary p is disrupted, and is 0 otherwise.
- $\omega_{t,f} \in \{0,1\}$: 1 if tail t undergoes maintenance immediately after operating flight f, and is 0 otherwise.
- $\mu_c \in \{0,1\}$: 1 if crew pairing c is disrupted, and is 0 otherwise.
- $\gamma_f \in \mathbb{R}^+$: amount of delay absorbed due to increased cruise speed of flight f.
- $imt_t \in \mathbb{R}^+ (fmt_t \in \mathbb{R}^+)$: starting (ending) time of maintenance operations for tail t.
- $h_p^{p',v} \in \mathbb{R}^+$: number of passengers in fare class v reassigned from itinerary p to

itinerary p'.

- $\tau_p^{p'} \in \mathbb{R}^+$: arrival delay of itinerary p' with respect to planned arrival time of itinerary p.
- $\delta_f^d \in \mathbb{R}^+$ ($\delta_f^a \in \mathbb{R}^+$): departure (arrival) delay of flight f.

The integrated airline recovery model is as follows.

4.4 Objective Function

$$\min z = \sum_{t\in T}\sum_{f\in F_t} oc_{t,f}x_{t,f} + \sum_{f\in F} fc_f\gamma_f + \sum_{f\in F} dc_f\delta_f^a + \sum_{v\in \mathrm{Y}}\sum_{p\in \mathrm{P}}\sum_{p'\in \mathrm{CO}_p} rc_p^{p'}\left(1-\sum_{\zeta\in \mathrm{Z}}\theta_{p,v}^{p',\zeta}\alpha_p^{p',\zeta}\right)h_p^{p',v} + \sum_{v\in \mathrm{Y}}\sum_{p\in \mathrm{P}}\sum_{p'\in \mathrm{CO}_p}\sum_{\zeta\in \mathrm{Z}} pc_p^{p',\zeta}\alpha_p^{p',\zeta}h_p^{p',v} + \sum_{c\in C} cc_c\mu_c + \kappa\sum_{t\in T}\sum_{f\in F_t}\left|x_{t,f}-\hat{x}_{t,f}\right| \quad (1)$$

The objective function in (1) is the sum of seven terms. The first six terms are as follows, and in the following order: flight operating cost, extra fuel consumption cost due to increased cruise speed as a first order approximation (Delgado and Prats, 2009), flight delay costs (including, extra crew, maintenance and fuel costs), passenger re-accommodation cost (e.g., meal cost, hotel cost, etc.), passenger delay cost (including compensation cost and cost of lost goodwill) and crew cost due to disrupted crew pairings, which approximately captures the cost of crew repositioning, or needing extra crews (Bratu and Barnhart, 2006).

The last term penalizes aircraft changes compared to the originally planned schedule $\hat{x}_{t,f}$. This term serves two purposes. First, it incentivizes the model to pick an aircraft schedule that is close to the original schedule, thus reducing operational deviations from the plan. This provides operational benefits to the airline since fewer changes are more desirable, all else being equal. Second, as will be demonstrated in Section 5, this term enables faster solution of the optimization model. Cadarso et al. (2015) introduced a more complex approach to control the number of schedule changes but it resulted in increased computational times. The set-partitioning based recovery model by Liang et al. (2018) provides an alternative way of incorporating schedule deviation penalty. To the best of the authors' knowledge, ours is the first attempt to include such a non-linear term to accomplish an operational goal (of low deviance from schedule) and a computational goal (of shorter solution time) simultaneously. Note that the fourth, fifth and seventh terms in the objective function are non-linear.

Next, we present the list of constraints. We categorize the constraints into nine main groups, with each group detailed in one of the next nine subsections.

4.5 Flight Schedule Constraints

$$\sum_{t\in T_f} x_{t,f} + z_f = 1, \quad \forall f \in F \quad (2)$$

Constraints (2) ensure that every flight must be either flown using exactly one aircraft or must be canceled.

4.6 Sequencing and Fleet Size Constraints

$$\sum_{f' \in CF_f} y_{f,f'} + \sigma_f = 1 - z_f, \quad \forall f \in F \tag{3}$$

$$\sum_{f \in F: f' \in CF_f} y_{f,f'} + \rho_{f'} = 1 - z_{f'}, \quad \forall f' \in F \tag{4}$$

$$1 + x_{t,f'} \geq x_{t,f} + y_{f,f'}, \quad \forall f \in F, f' \in CF_f, t \in T_f \cap T_{f'} \tag{5}$$

$$1 + x_{t,f} \geq x_{t,f'} + y_{f,f'}, \quad \forall f \in F, f' \in CF_f, t \in T_f \cap T_{f'} \tag{6}$$

$$\rho_f + x_{t,f} \leq 1 + \varphi_{t,f}, \quad \forall f \in F, t \in T_f \tag{7}$$

$$\sum_{f \in F_t \cap FD_k} \varphi_{t,f} \leq tb_{t,k}, \quad \forall t \in T, k \in K \tag{8}$$

Constraints (3) (respectively, Constraints (4)) are the flight sequencing constraints ensuring that each non-canceled flight has another flight after (respectively before) it operated by the same tail, unless it is the last (respectively first) flight in the recovery period operated by that tail. Constraints (5) and (6) together ensure that consecutive flights to be operated by the same tail are actually assigned the same tail. Constraints (7) state that if a flight is assigned a tail and it is the first flight in a sequence of flights, it must be the first flight of that tail. Constraints (8) state that a flight chosen to be a first flight in a sequence of flights can only be assigned to tails whose initial location matches the flight's departure airport. Note that these constraints allow for a tail to be not used at all during the recovery period.

4.7 Flight Delay Constraints

$$\delta_f^a \geq \delta_f^d - \gamma_f - sb_f, \quad \forall f \in F \tag{9}$$

$$\delta_{f'}^d \geq \delta_f^a + mtt_{f,t}^{f'} - ct_f^{f'} - M\left(3 - x_{t,f} - x_{t,f'} - y_{f,f'}\right), \quad \forall f \in F, f' \in CF_f, t \in T_f \cap T_{f'} \tag{10}$$

Constraints (9) relate the departure and arrival delays of each flight via delay absorption through increased cruise speed. Constraints (10) relate the arrival delay of one flight to the departure delay of the next flight operated by the same tail, by accounting for delay propagation. Note that the set of flights compatible with flight f, CF_f, is the set of flights that are scheduled to depart from the arrival airport of flight f, and fulfill the following two conditions: $st_{f'}^d + \bar{\gamma}_{f'} + \bar{\delta}_{f'} \geq st_f^a + \underline{mtt}_f$ and $st_{f'}^d \leq st_f^a + \bar{\delta}_f + maxct_f$.

4.8 Maintenance Constraints

We divide the set of available aircraft into two categories: those that might require maintenance (depending on the hours flown) during the recovery horizon, and those that will not. In order to check maintenance requirements, a discretized mesh of time instances is created. It is defined by a subset (Π^m) of time instances at which maintenance feasibility of all aircraft tails will be checked, and consists of time instances that are evenly distributed throughout the recovery horizon. Mesh sizes involve a clear tradeoff: a very fine mesh will ensure feasibility of the aircraft recovery plans, but will create additional computational burden from a large number of extra constraints. On the other hand, a very coarse mesh may reduce the computational burden, but could lead to maintenance-infeasible aircraft recovery plans. Therefore, we attempt to find the mesh that is fine enough to ensure feasibility while being coarse enough to ensure computational tractability. Note that the elements in Π^m are not known a priori. Delayed constraints generation is employed to determine the appropriate size of Π_m.

Maintenance check

$$abh_t \geq \sum_{f \in F_t \cap F_\pi^\Pi} sbh_f x_{t,f} - \sum_{f \in F_t \cap F_\pi^\Pi} mbh_t \omega_{t,f}, \quad \forall t \in \mathrm{T}^m, \pi \in \Pi^m \tag{11}$$
$$\omega_{t,f} \leq x_{t,f}, \quad \forall f \in F, t \in T_f \cap \mathrm{T}^m \tag{12}$$
$$\omega_{t,f} \leq \sum_{f' \in F:(f,f') \in MO, t \in T_{f'}} y_{f,f'}, \quad \forall f \in F, t \in T_f \cap \mathrm{T}^m \tag{13}$$

Constraints (11) ensure that the tails that might need maintenance during the recovery horizon do not exceed the available number of block hours until they undergo such maintenance. Constraints (12) require that if a tail undergoes maintenance after a flight f, it must have been assigned to flight f. Constraints (13) make sure that a tail t can undergo maintenance only if it is assigned consecutive flights f and f' that present a maintenance opportunity in the gap between them.

Maintenance schedule

$$imt_t \geq st_f^a + \delta_f^a - M(1 - \omega_{t,f}), \quad \forall f \in F, t \in T_f \cap \mathrm{T}^m \tag{14}$$
$$imt_t \leq M \sum_{f \in F_t} \omega_{t,f}, \quad \forall t \in T_m \tag{15}$$
$$fmt_t = mt_t \sum_{f \in F_t} \omega_{t,f} + imt_t, \quad \forall t \in T_m \tag{16}$$
$$fmt_t \leq st_{f'}^d + \delta_{f'}^d + M(1 - y_{f,f'}), \quad \forall (f,f') \in MO, t \in T_f \cap T_{f'}, \cap \mathrm{T}^m \tag{17}$$

Constraints (14) state that if a tail undergoes maintenance after a flight, the starting time of maintenance is after the actual arrival time of the flight. Constraints (15) state that if a tail does not undergo maintenance then its starting time is set to zero. Constraints (16) determine the ending time of maintenance of a tail, which is the sum of the maintenance starting time and the required time to perform maintenance operations. Constraints (17) limit the ending time of maintenance operations to earlier than or equal to the actual departure time of the next flight operated by that tail.

Workshop schedule

$$\sum_{t' \in \mathrm{T}^m} m_{t,t'} + \sigma_t^m = \sum_{f \in F_t} \omega_{t,f}, \quad \forall t \in \mathrm{T}^m \tag{18}$$
$$\sum_{t \in \mathrm{T}^m} m_{t,t'} + \rho_{t'}^m = \sum_{f \in F_{t'}} \omega_{t',f}, \quad \forall t' \in \mathrm{T}^m \tag{19}$$
$$imt_{t'} \geq fmt_t - M(1 - m_{t,t'}), \quad \forall t, t' \in \mathrm{T}^m \tag{20}$$
$$2m_{t,t',k}^m \leq \sum_{f \in F_t \cap FA_k} \sum_{f' \in F_{t'} \cap FA_k} (\omega_{t,f} + \omega_{t',f'}), \quad \forall k \in K^m, t, t' \in \mathrm{T}^m \tag{21}$$
$$m_{t,t'} = \sum_{k \in K^m} m_{t,t',k}^m, \quad \forall\, t, t' \in \mathrm{T}^m \tag{22}$$
$$\rho_t^m + \omega_{t,f} \leq 1 + \varphi_{t,f}^m, \quad \forall f \in F, t \in T_f \cap \mathrm{T}^m \tag{23}$$
$$\sum_{t \in \mathrm{T}^m} \sum_{f \in F_t \cap FA_k} \varphi_{t,f}^m \leq aw_k, \quad \forall k \in K \tag{24}$$

Constraints (18)-(19) are logically similar to constraints (3)-(4). They determine the sequencing of maintenance operations of tails in workshops. Constraints (20) state that if two tails undergo maintenance consecutively in the same workshop, their maintenance starting and ending times must be compatible. Constraints (21) and (22) establish that if two tails consecutively undergo maintenance in the same workshop, they must do so at the same airport. Constraints (24) are workshop capacity constraints – the number of maintenance sequences should be at most equal to the number of available workshops for each airport, while constraints (23) help in modeling this workshop capacity constraint by correctly defining auxiliary variables $\varphi_{t,f}^m$.

Note that this modeling approach is general enough to account for different types of maintenance needs. In our computational experiments, the maintenance requirements are assumed to be dependent on the number of flying hours. However, if they depend

on other measures, e.g., on the number of flights, the approach could be easily adapted with the parameters abh_t, sbh_f, and mbh_t representing the number of flights instead of the number of flying hours. Similarly, if multiple different types of maintenance operations are to be simultaneously included in the model, our approach can easily accommodate those as well.

4.9 Airport Slot Constraints

$$it_{as}^{AA} \leq st_f^a + \delta_f^a + M\left(1 - v_{as,f}^{AA}\right), \quad \forall f \in F, as \in AA_f \tag{25}$$

$$ft_{as}^{AA} \geq st_f^a + \delta_f^a - M\left(1 - v_{as,f}^{AA}\right), \quad \forall f \in F, as \in AA_f \tag{26}$$

$$\sum_{as \in AA_f} v_{as,f}^{AA} = 1 - z_f, \quad \forall f \in F \tag{27}$$

$$\sum_{f \in F_{as}^{\mathrm{AA}}} v_{as,f}^{AA} \leq sc_{as}^{AA}, \quad \forall as \in AA \tag{28}$$

$$it_{ds}^{DA} \leq st_f^d + \delta_f^d + M\left(1 - v_{ds,f}^{DA}\right), \quad \forall f \in F, ds \in DA_f \tag{29}$$

$$ft_{ds}^{DA} \geq st_f^d + \delta_f^d - M\left(1 - v_{ds,f}^{DA}\right), \quad \forall f \in F, ds \in DA_f \tag{30}$$

$$\sum_{ds \in DA_f} v_{ds,f}^{DA} = 1 - z_f, \quad \forall f \in F \tag{31}$$

$$\sum_{f \in F_{ds}^{\mathrm{DA}}} v_{ds,f}^{DA} \leq sc_{ds}^{DA}, \quad \forall ds \in DA \tag{32}$$

Constraints (25) and (26) are the constraints on the start and end times, respectively, of arrival slots. Constraints (27) ensure that each non-canceled flight is assigned to exactly one arrival slot. Constraints (28) ensure that arrival slot capacity limits are not violated. Constraints (29) to (32) are the constraints analogous to constraints (25) to (28) respectively, but for departure slots.

4.10 Passenger Flow Constraints

$$\sum_{p' \in CO_p} h_p^{p',v} = n_p^v, \quad \forall p \in P, v \in \Upsilon \tag{33}$$

$$h_p^{p',v} \leq \left(1 - \lambda_{p'}\right) n_p^v, \quad \forall p \in P, v \in \Upsilon, \mathrm{p}' \in CO_p \tag{34}$$

$$\sum_{t \in T_f} q_t x_{t,f} \geq \sum_{v \in \Upsilon} \sum_{p \in P} \sum_{\mathrm{p}' \in CO_p : p' \ni f} \left(1 - \sum_{\zeta \in Z} \theta_{p,v}^{p',\zeta} \alpha_p^{p',\zeta}\right) h_p^{p',v}, \quad f \in F \tag{35}$$

Constraints (33) are passenger flow constraints; they ensure that all passengers are reassigned to some itinerary, which might be the same or different from their originally scheduled itinerary. Constraints (34) state that passengers can be reassigned only to the non-disrupted itineraries. Constraints (35) are passenger seating capacity constraints for each flight. The right-hand side represents the total number of passengers that do show up for their reassigned itineraries, summed across fare-classes, originally scheduled itineraries and reassigned itineraries. Note that the right-hand side is a non-linear function of the model's decision variables.

4.11 Itinerary Feasibility Constraints

$$\lambda_p \geq z_f, \quad \forall f \in F, p \in P : p \ni f \tag{36}$$

$$st_{f'}^d + \delta_{f'}^d - st_f^a - \delta_f^a \geq mct_p^{f,f'} - M\lambda_p, \quad \forall p \in P, (f, f') \in CF_p^P \tag{37}$$

Constraints (36) and (37) determine when an itinerary gets disrupted due to flight cancelations and due to flight retiming decisions, respectively.

4.12 Itinerary Delay Constraints

$$\tau_p^{p'} = \sum_{f' \in F} lf_{f',p'} \left(\delta_{f'}^a + st_{f'}^a\right) - \sum_{f \in F} lf_{f,p} st_f^a, \quad \forall p \in P, p' \in CO_p \tag{38}$$

$$\tau_p^{p'} \leq \sum_{\zeta \in Z} \vartheta_\zeta \alpha_p^{p',\zeta}, \quad \forall p \in P, p' \in CO_p \tag{39}$$

$$\sum_{\zeta \in Z} \alpha_p^{p',\zeta} = 1, \quad \forall p \in P, p' \in CO_p \tag{40}$$

Constraints (38) calculate the passenger arrival delay after the itinerary reassignment. Constraints (39) and (40) determine the passenger delay level for each pair of compatible itineraries, depending on the actual delay value in minutes.

4.13 Crew Constraints

$$\mu_c \geq z_f, \quad \forall f \in F, c \in C: c \ni f \tag{41}$$

$$st_{f'}^d + \delta_{f'}^d - st_f^a - \delta_f^a \geq mst_c - M\mu_c, \quad \forall c \in C, (f,f') \in CF_c^C \tag{42}$$

$$st_{f'}^d + \delta_{f'}^d - st_f^a - \delta_f^a \geq mbt_c - M\mu_c, \quad \forall c \in C, (f,f') \in BF_c^C \tag{43}$$

Constraints (41), (42) and (43), respectively, determine whether a crew pairing is disrupted due to flight cancelation, or insufficient crew sit or rest time between flights due to flight retiming.

4.14 Linearization of Non-linear Terms

The mathematical model given by (1)-(43) is a non-linear mixed-integer programming model. The non-linearities are in the objective function (1) and in constraints (35).

To improve the computational tractability of the model, we linearize it as follows. We introduce an additional variable $\beta_{p,v}^{p',\zeta}$, which equals the number of passengers in fare class v with originally scheduled itinerary p, reassigned to itinerary p', corresponding to an arrival delay level ζ. Therefore, $\beta_{p,v}^{p',\zeta}$ is defined as follows.

$$\beta_{p,v}^{p',\zeta} = \alpha_p^{p',\zeta} h_p^{p',v}, \quad \forall p \in P, p' \in CO_p, v \in \Upsilon, \zeta \in Z \tag{44}$$

Constraints (44) are themselves still non-linear, but they can be linearized as follows.

$$\beta_{p,v}^{p',\zeta} \leq h_p^{p',v} + M\left(1 - \alpha_p^{p',\zeta}\right), \quad \forall p \in P, p' \in CO_p, v \in \Upsilon, \zeta \in Z \tag{45}$$

$$\beta_{p,v}^{p',\zeta} \geq h_p^{p',v} - M\left(1 - \alpha_p^{p',\zeta}\right), \quad \forall p \in P, p' \in CO_p, v \in \Upsilon, \zeta \in Z \tag{46}$$

$$\beta_{p,v}^{p',\zeta} \leq M\alpha_p^{p',\zeta}, \quad \forall p \in P, p' \in CO_p, v \in \Upsilon, \zeta \in Z \tag{47}$$

Note that the set of constraints (45)-(47) imply that if $\alpha_p^{p',\zeta}$ is 0, then $\beta_{p,v}^{p',\zeta}$ must be zero (enforced by constraints (47)), and if $\alpha_p^{p',\zeta}$ is 1, then $\beta_{p,v}^{p',\zeta}$ must be equal to $h_p^{p',v}$ (enforced by constraints (45) and (46) together).

Therefore, constraints (35) may be replaced with constraints (45)-(48).

$$\sum_{t \in T_f} q_t x_{t,f} \geq \sum_{v \in \Upsilon} \sum_{p \in P} \sum_{p' \in CO_p : p' \ni f} \left(h_p^{p',v} - \sum_{\zeta \in Z} \theta_{p,v}^{p',\zeta} \beta_{p,v}^{p',\zeta}\right), \quad f \in F \tag{48}$$

Next, we linearize the objective function (1) in which the fourth, fifth and seventh terms are non-linear. The fourth and fifth terms are non-linear due to the bilinear term, $\alpha_p^{p',\zeta} h_p^{p',v}$, in them. This bilinear term is linearized by simply replacing it with the additional variables $\beta_{p,v}^{p',\zeta}$. Finally, the last term in the objective function, can be transformed as follows: $\left|x_{t,f} - \hat{x}_{t,f}\right| = \left(x_{t,f} - \hat{x}_{t,f}\right)^2 = {x_{t,f}}^2 - 2\hat{x}_{t,f}x_{t,f} + {\hat{x}_{t,f}}^2 = x_{t,f} - 2\hat{x}_{t,f}x_{t,f} + \hat{x}_{t,f}$. Note that this transformation works only because

$x_{t,f}, \hat{x}_{t,f}, |x_{t,f} - \hat{x}_{t,f}| \in \{0,1\}$ which ensures that ${x_{t,f}}^2 = x_{t,f}$, ${\hat{x}_{t,f}}^2 = \hat{x}_{t,f}$, and $(x_{t,f} - \hat{x}_{t,f})^2 = |x_{t,f} - \hat{x}_{t,f}|$. Therefore, the linearized version of the objective function (1) is given as (49).

$$\min z = \sum_{t\in T}\sum_{f\in F_t} oc_{t,f}x_{t,f} + \sum_{f\in F} fc_f\gamma_f + \sum_{f\in F} dc_f\delta_f^a + \sum_{v\in \Upsilon}\sum_{p\in \mathrm{P}}\sum_{p\prime\in \mathrm{CO}_p} rc_p^{p'}\left(h_p^{p',v} - \sum_{\zeta\in \mathrm{Z}}\theta_{p,v}^{p',\zeta}\beta_{p,v}^{p',\zeta}\right) + \sum_{v\in \Upsilon}\sum_{p\in \mathrm{P}}\sum_{p\prime\in \mathrm{CO}_p}\sum_{\zeta\in \mathrm{Z}} pc_p^{p',\zeta}\beta_{p,v}^{p',\zeta} + \sum_{c\in C} cc_c\mu_c + \kappa\sum_{t\in T}\sum_{f\in F_t}(x_{t,f} - 2\hat{x}_{t,f}x_{t,f} + \hat{x}_{t,f}) \quad (49)$$

Thus, objective function (49) combined with constraints (2)-(34), (36)-(43) and (45)-(48) provide a mixed-integer linear optimization version of our Integrated Model for Airline Operations Recovery (IMAOR). This is the version that we use for all computational experiments. This comprehensive model, containing 45 types of constraints, has been developed keeping in mind the needs of our two airline partners. Practitioners working on other airline networks may not find some of these constraints to be core to their business needs. However, the solution approach presented in this section is general enough to work well even if the users decide to eliminate some of the constraints.

5. Computational Experiments and Results

Computational experiments in this section are conducted using real-world problem instances based on IBERIA airline's network. It is a pure hub-and-spoke network, with 48 airports and 164 OD (origin-destination) airport pairs, with the only hub located in Madrid (MAD), which is also the only maintenance station featuring 4 workshops. This network was chosen because of the availability of detailed datasets (including data on *phantom passenger* rates, among others) provided to us by IBERIA airline for this research project.

The network is operated using five different fleet types: A319, A320, A321, A340-200 and A340-600 featuring 141, 171, 200, 254 and 342 seats, respectively. A recovery horizon of three days is considered. During this time horizon, 1074 flights are scheduled to be operated, and 19 aircraft will potentially need to undergo (non-delayable) maintenance. The average duration of these 19 maintenance tasks is 12.4 hours. Crew recovery is out of scope of the experiments presented in this section because no crew data was made available by this airline. This eliminates constraints (41)-(43) and removes the sixth term in the objective function. The rest of the model remains intact. As will be seen later, this data limitation is overcome by the subsequent experiments in Section 6.2.

We coded the IMAOR in GAMS, using CPLEX 12.1 as the optimization solver on a computer with 32 GB RAM. Note that a computational time limit of 600 seconds was imposed on all experiments, even though in many cases an optimal or near-optimal solution was obtained much sooner than that. Our discussions with our airline partners revealed that this duration of 600 seconds (i.e., 10 minutes) is often perceived as an upper limit on the amount of time for which recovery models can be run on the day of operations, though much faster solutions are often strongly preferred.

5.1 Case studies

In this section, we evaluate our model's performance with problem instances focusing on two different types of disruptions: a small-scale and a large-scale disruption.

The small-scale disruption consists of a delayed flight departing from JFK (John F. Kennedy airport in New York City) and arriving at MAD. The flight suffers a departure delay of 320 minutes. Despite the relatively small-scale nature of this disruption, it must be noted that this flight is an important feeder in the network and therefore, the consequences of this disruption may be serious and important for both passengers and aircraft. Specifically, despite its small scale, it resulted in 32 directly disrupted itineraries corresponding to 1204 disrupted passengers. Out of these, 45% were nonstop passengers while the remaining 55% were connecting passengers. The main actions for passenger recovery included delaying connecting flights from MAD to accommodate passengers arriving on the delayed JFK-MAD flight, and reaccommodating those that misconnect on later outgoing flights from MAD. Also note that this is an actual disruption event in the historical data provided by the airline, and hence it enables generating practically relevant insights based on our approach. It is primarily used to compare the performance of the proposed IMAOR-based approach, against actual recovery operations performed by the airline and also against simplified recovery approaches.

The second problem instance involves a large-scale disruption, caused by a three-hour airport closure, at Barcelona (BCN) airport, which is the second most important airport in terms of IBERIA airline's operations. The closure is active from 16:30 to 19:30 on a weekday resulting in 53 directly disrupted itineraries corresponding to 1917 disrupted passengers. Out of these, 74% were nonstop passengers while the remaining 26% were connecting passengers. The main actions for passenger recovery included delaying or canceling flights arriving into or departing from BCN airport, as well as reaccommodating disrupted passengers on alternative itineraries. Due to the severe nature of the disruption, cancelations are unavoidable. This too is based on an actual, large-scale disruption that happened to the airline's schedule in the past. However, data on actual airline recovery operations under this disruption was not made available for this research. So, we cannot use this disruption instance to compare our solution with that used by the airline. But, this instance serves a different, important purpose – it allows us to demonstrate the tractability and scalability of our approach, by evaluating its effectiveness in handling such large-scale disruptions, which heavily constrain the network operations, in a reasonable amount of time with low optimality gaps. Note that, since the airport closures are typically out of control of an individual airline, Regulation (EC) No 261/2004 does not apply and passengers are not required to be compensated. However, they must be provided with re-accommodation. Since the large-scale disruption does not invoke passenger compensation and has only about 59% higher number of directly affected passengers compared to the smaller-scale disruption instance, the overall recovery costs have a similar order of magnitude across the two problem instances.

5.2 Choice of Mesh Granularity for Maintenance Feasibility Checks

An important parameter of the IMAOR model is the number of time instances at which maintenance requirements are checked, that is, the size of set Π^m. Another important

parameter is κ, which is the per-flight schedule change penalty for not operating the flight using the originally planned aircraft. Note that, unlike the remaining parameters in the model, the values of these two parameters are not provided exogenously to us by the airline. Therefore, in this and the next subsection, we perform computational experiments using our IMAOR-based approach to identify the tradeoffs associated with using different combinations of values of these two critical parameters.

In our first set of experiments, we identify the most appropriate value of $|\Pi^m|$. To isolate the effects of $|\Pi^m|$, we fix $\kappa = 2000$ for this particular set of experiments. For a given value of $|\Pi^m|$, the set Π^m gets automatically decided because we assume, as mentioned in Section 4.8, that the time instances are evenly distributed across the recovery horizon. Larger $|\Pi^m|$ increases the computational burden, while smaller $|\Pi^m|$ could lead to infeasible recovery plans. Table 3 demonstrates this tradeoff by presenting the number of aircraft exceeding available flying hours, and the solution time in seconds, for $|\Pi^m| = 0,3,6,9,12,15,18$, for both small-scale and large-scale disruptions. As expected, the computational burden increases and the number of infeasibilities decreases, with increasing values of $|\Pi^m|$. Note that, whether an aircraft recovery plan generated by IMAOR is actually maintenance feasible or not needs to be evaluated after solving the model to "optimality" first. Depending on the outcome, more constraints are added in the form of additional time instances for feasibility checks and the optimization run is continued. Thus, this process amounts to a delayed constraint generation approach. As shown in Table 3, the generated aircraft recovery plan becomes maintenance feasible when $|\Pi^m| = 18$ in this particular set of experiments, and thus terminates the delayed constraint generation process.

Using this insight, for all subsequent experiments presented in this paper, we started off with $|\Pi^m| = 18$, and considered adding more constraints if the solution is found to be infeasible. However, we found that the solution was always maintenance feasible in all our subsequent computational experiments when using $|\Pi^m| = 18$. Interestingly, these observations seem to suggest that the adequate mesh size is relatively stable for a given network and does not vary significantly from one problem instance to another and from one disruption to another. This could potentially reduce or eliminate the need for re-running the solution process with increasingly finer mesh sizes on the day of operations and hence further enhance the computational performance of our approach. Keep in mind, however, that these results still leave open the possibility that in some of the experiments in this section, a $|\Pi^m|$ value smaller than 18 could also have sufficed.

Table 3: Changes in the number of infeasibilities and solution times with mesh granularity

Mesh size (Number of time instances)	**Number of aircraft exceeding maximum flying hours**		**Solution time (s)**	
	Small-scale	Large-scale	Small-scale	Large-scale
0	4	6	81	245
3	4	4	103	337
6	3	4	114	478
9	2	3	129	589

12	1	2	156	600 (time limit)
15	1	1	169	600 (time limit)
18	0	0	197	600 (time limit)

5.3 Choice of Schedule Deviation Penalty Parameter

The choice of schedule deviation penalty parameter κ also involves an important tradeoff. Small κ can lead to a recovery plan too different from schedule and/or may cause longer solution times. Both these can adversely affect the usability of a recovery plan in practice. On the other hand, large κ can lead to larger recovery costs. Therefore, it is important to choose a value that balances these conflicting goals. In order to evaluate this tradeoff, we performed our second set of computational experiments by fixing $|\Pi^m| = 18$ while varying κ from 0 to 5000 in steps of 1000.

Table 4 describes model performance and solution quality for different values of κ. The first row shows the value of κ; the second main row shows the objective function value for each case study; and the third main row shows the objective function values without considering its last term, i.e., subtracting the κ-weighted penalty term. Thus, the third row provides the actual recovery cost value. The fourth main row shows the solution time in seconds. The fifth main row shows the optimality gap as provided by the mixed-integer linear programming (MILP) solver. With $\kappa = 0$, the model is unsolvable within 10 minutes and does not even yield a feasible solution. All it can provide is a lower bound (LB) on the optimal value of this minimization problem, which cannot be implemented in practice. However, it allows us to bound from below the lowest possible recovery cost (where recovery cost is defined as the sum of all terms in the objective function (49) excluding the last term). The sixth main row calculates the optimality gap as the difference between this lower bound value and the value in the third row of the table, divided by this lower bound value, represented as a percentage. The last (seventh) main row shows the number of aircraft schedule changes with respect to the original plan.

The model is found to be solvable to reasonable optimality gaps within 10 minutes of runtime for all nonzero κ values attempted in these experiments. The objective function shows a decreasing trend with decreasing κ, which is reasonable because operational deviations are decreasingly penalized. If we ignore the penalties, and instead focus on the actual recovery costs (as listed in the third main row), we see a similar but non-monotonic trend. As the value of κ decreases, costs generally decrease because the resulting solutions are more flexible to react to disruptions. At larger κ values, there is less flexibility to re-design plans and operations have to remain similar to the planned ones. For the rest of the computational experiments, we set the value of κ to 2000. This choice is based on the observation that this value of κ allows a fast solution (in ~3 minutes) to a recovery cost value that is just 0.05% worse than the best one, ensuring a faster solution (with 26% lower runtime) and 13% fewer schedule changes than those for $\kappa = 1000$. For the large-scale case study, $\kappa = 2000$ produces a solution with a recovery cost that is only 0.72% worse than the best one and 19% fewer schedule changes than those for $\kappa = 1000$ in the same computational time.

Table 4: Changes in model performance with schedule deviation penalty parameter κ

Item\ κ		5000	4000	3000	2000	1000	0
Objective function	Small-scale	13,795,810	13,787,760	13,753,050	13,671,030	13,642,410	No Solution (LB = 13,467,110)
	Large-scale	16,314,310	16,061,920	15,994,010	16,225,760	16,012,990	No Solution (LB = 15,319,540)
Objective function without penalty	Small-scale	13,715,810	13,723,760	13,699,050	13,619,030	13,612,410	-
	Large-scale	15,934,310	15,733,920	15,712,010	15,825,420	15,884,990	-
Solution time (s)	Small-scale	227	183	129	197	268	600 (time limit)
	Large-scale	600 (time limit)	600 (time limit)	600 (time limit)	600 (time limit)	600 (time limit)	600 (time limit)
Optimality gap (%) as provided by the MILP solver	Small-scale	0	0	0	0	0	-
	Large-scale	1.76	0.92	0.99	1.63	2.45	-
Optimality gap (%) w.r.t best known lower bound	Small-scale	1.84	1.90	1.72	1.13	1.08	-
	Large-scale	4.01	2.70	2.56	3.30	3.69	-
Number of schedule changes	Small-scale	8	8	9	13	15	-
	Large-scale	38	41	47	52	64	-

5.4 Benefits of Capturing the Phantom Passengers Phenomenon

An important novelty and contribution of this research is the explicit capture of the phantom passengers phenomenon. We now run our third set of computational experiments in this subsection with the goal of investigating the extent of benefits of capturing this phenomenon and identifying the optimal tradeoff between passenger compensation cost, flight operating cost and passenger re-accommodation costs.

Table 5 has five columns. The first one lists the row headings and the second, third and fifth list solutions as provided by the model for three different scenarios, while setting $|\Pi^m| = 18$ and $\kappa = 2000$. The second and third columns correspond to the small-scale disruption with and without accounting for phantom passengers. Column 5 corresponds to the large-scale case study. Since passengers are not required to be compensated for a disruption due to airport closure, no results are provided regarding phantom passengers for the large-scale disruption. Results in Column 5 demonstrate the tractability and scalability of our approach, by evaluating its effectiveness in solving such large-scale instances. Column 2 describes the recovery solution obtained by solving the IMAOR as presented in Section 4 using the phantom rate values as estimated in Section

3. On the other hand, Column 3 describes the recovery solution obtained by solving the IMAOR while setting all phantom passenger rates $\theta_{p,v}^{p',\zeta} = 0$. Column 4 lists the values in Column 3 minus the values in Column 2, divided by the values in Column 2.

For the small-scale disruption instance, the "Without Phantom Passengers" approach, as compared to the "With Phantom Passengers" approach, has higher values of all the passenger-related costs as provided by the optimization model. While the operating costs and the extra-fuel costs are lower, the overall objective function without penalty also increases slightly, meaning that overall, the "With Phantom Passengers" approach provides a slightly cheaper solution. It is important to note that the "With Phantom Passengers" approach gives a better solution in terms of service quality, which may be measured by re-accommodation cost, compensation cost and expected number of phantom passengers. Note that when passengers cannot be reassigned during the considered planning horizon, they are reassigned to the null itinerary, i.e., they are to be re-accommodated later and thus are "rejected" from the considered recovery period. Rejection costs capture the cost of inability to reassign passengers during the considered planning horizon, and are included in the re-accommodation cost term in the objective function (49) of the mathematical model presented in Section 4. When presenting the results in Table 5, we decided to split the overall re-accommodation costs (given by the fourth term in the objective function) into rejection costs (Row 8) and the other re-accommodation costs (Row 6) to highlight this distinction.

Under the large-scale disruption, the recommended recovery policy by the mathematical model includes canceling flights, changing fleet assignments to flights and re-accommodating passengers. Note that this corresponds to 15 cancelations and 52 schedule changes, reflecting the larger impacts of this disruption on airline operations, as compared to the previously discussed small-scale disruption.

Table 5: Comparison of optimization model solutions with and without capturing the phantom passengers phenomenon

Item	With Phantom Passengers (Small-scale)	Without Phantom Passengers (Small-scale)	Change (%) (Small-scale)	Without Phantom Passengers (Large-scale)
Objective function without penalty	13,619,030	13,630,425	+0.08	15,825,420
Operating cost	13,299,500	13,281,300	-0.13	13,276,470
Extra-fuel cost	18,523	11,054	-40.32	38,270
Delay cost	25,727	33,921	+31.84	47,875
Re-accommodation cost	14,100	25,070	+77.80	1,678,805
Compensation cost	261,160	279,080	+6.86	n/a
Rejection cost	0	0	0	576,000
# of canceled flights	0	0	0	15
# of schedule	13	12	-7.69	52

changes				
Expected # of phantom passengers	14.15	62.38	+340.84	n/a

5.5 Simulation-based Evaluation under Passenger No-show Uncertainty

Table 5 showed that when evaluated with a deterministic objective function, the IMAOR solutions that account for phantom passengers slightly outperform those which do not account for phantom passengers, as expected. However, to provide a more objective evaluation of the benefits of accounting for phantom passengers, we now perform a simulation-based evaluation under passenger no-show uncertainty by leveraging our access to historical data. Table 6 follows a similar structure as Table 5 and presents the results from simulation-based evaluation of each optimal solution. This evaluation approach overcomes some important limitations of the more direct optimization-based evaluation approach. Recall that the optimization model simply uses the phantom passenger rate to calculate the expected number of passengers that will actually show-up for each reassigned itinerary. This calculation ignores the fact that passengers' phantom behaviors (i.e., their decisions to show up or not) involve high degree of uncertainty from an airline operator's perspective. In reality, the phantom rates estimated in Section 3 are the estimated probabilities of each individual passenger not showing up for the reassigned itinerary and must be treated as such to allow a more precise evaluation of any recovery plan. Accounting for this uncertainty requires development of a simulator.

Consequently, a Monte Carlo simulation of passengers' phantom behaviors is performed once the airline recovery plan is determined with the IMAOR optimization approach. Note that, ideally, we should use historical data directly to simulate no-shows, the times when passengers show up (if they show up), as well as the business rules around seat assignments. However, since we did not have access to such detailed historical data from the airline, we came up with the following approximations to best leverage the data available at our disposal. For this simulation, the no-show probability for each reassigned passenger for its reassigned itinerary is assumed to be equal to its phantom rate. For passengers that do show up, passenger arrival time for their reassigned itineraries is assumed to be random, distributed uniformly between 210 and 50 minutes before the reassigned itinerary departure time. All passengers that do show up, are assumed to block seats in a First-Come-First-Served manner. If more passengers show up than the available number of seats on their reassigned itinerary (which can happen due to the inherent stochasticity being modeled in the simulation approach), we assume that they are reassigned to the next available itinerary to their destination. Note that itineraries are ordered in terms of arrival times, from the one with the earliest arrival time to the one with the latest arrival time, followed by the null itinerary. For the large-scale disruption instance, no simulation is needed since the phantom passengers phenomenon is almost non-existent in the absence of the monetary compensations.

Table 6: Simulation-based evaluation of solutions with and without capturing the phantom passengers phenomenon

Item	With Phantom Passengers (Small-scale)	Without Phantom Passengers (Small-scale)	Change (%) (Small-scale)
Objective function without penalty	13,622,390	13,764,730	+1.04
Operating cost	13,299,500	13,281,300	-0.13
Extra-fuel cost	18,523	11,054	-40.32
Delay cost	25,727	33,921	+31.84
Re-accommodation cost	15,800	36,809	+132.96
Compensation cost	262,827	305,646	+16.29
Rejection cost	0	96,000	Nan
# of canceled flight	0	0	0
# of schedule change	13	12	-7.69
Expected # of phantom passengers	18.72	87.24	+366.02

The simulation-based evaluation brings out the benefits of capturing the phantom passengers phenomenon much more clearly and strongly. The simulation results show the same general trends the optimization results show, but are able to capture the stochasticity in passenger behavior thus providing a much more comprehensive and accurate comparison of the different solutions. In particular, the simulation results indicate that explicitly accounting for the phantom passengers phenomenon can save the airlines considerable costs, as much as 142,340 Euros over a three day recovery period. Thus, our approach has a significant potential for positive impact on the airlines' bottom-lines. Ignoring the phantom passengers phenomenon results in 133% higher re-accommodation cost and 16% higher compensation cost. It also leads to 96,000 Euros worth of rejection costs, all of which can be eliminated by accounting for the phantom passengers explicitly.

As shown in Section 4, phantom rates have been estimated according to the limited amount of historical data that was made available by our partner airline for this research. Because these rates play an important role in the presented approach, a sensitivity analysis is performed on them, which consists of solving the IMAOR optimization problem for the nominal phantom rates estimated by the regression model and then performing the simulation-based evaluation for different scenarios with varying actual phantom rates. Table 7 shows five different such scenarios one in each column, with $|\Pi^m| = 18$ and $\kappa = 2000$. Note that since all columns assume the same nominal phantom rate values, they amount to the exact same recovery solution and the exact same operating cost and extra-fuel cost. However, the columns differ in terms of the actual phantom rates. Columns 2, 3, 4, 5, and 6 show the results for actual phantom rates equal to the nominal phantom rates plus -2, -1, 0, +1 and +2 standard deviations respectively. Numbers in parentheses are percentage changes with respect to the scenario where the nominal rates are the actual rates, i.e., Column 4.

From Table 7, we observe that the expected number of phantom passengers increases as phantom rates increase. An increase in phantom passengers means lower re-accommodation cost as confirmed in the obtained solutions. Similarly, compensation cost also decreases as phantom rate increases. Overall, the objective function decreases as phantom rate increases. These results highlight a very important finding – even when the phantom rates are estimated very poorly (e.g., off by one or two standard deviations), the changes to the total recovery costs are much smaller (of the order of at most 0.05%) compared to the case where the phantom passengers phenomenon is completely ignored (causing over 1% cost increase as shown in Table 6). This confirms that neglecting phantom passengers may produce inefficient recovery schedules, while accounting for them in even an approximate manner with relatively poor datasets and/or simplified statistical estimation approaches can still yield significant benefits. Recent airline schedule design studies (e.g., Cadarso et al., 2017; Wei et al., 2020; Yan et al., 2022) have highlighted the value of incorporating passenger itinerary choice decisions into the strategic scheduling process. Our findings in this research, for the first time, highlight similar benefits of modeling passenger response behaviors when designing the recovery schedules as well.

Table 7: Sensitivity analysis on phantom rates (numbers in parentheses are percentage changes with respect to the middle column values)

Estimation Error	-2σ	$-\sigma$	-	$+\sigma$	$+2\sigma$
Objective function without penalty	13,628,630 (0.05%)	13,624,290 (0.01%)	13,622,390 (-)	13,618,570 (-0.03%)	13,616,470 (-0.04%)
Re-accommodation cost	19,680 (24.56%)	17,210 (8.92%)	15,800 (-)	13,760 (-12.91%)	13,210 (-16.39%)
Compensation cost	265,190 (0.90%)	263,321 (0.19%)	262,827 (-)	261,050 (-0.68%)	259,500 (-1.27%)
Rejection cost	0 (-)	0 (-)	0 (-)	0 (-)	0 (-)
Expected # of phantom passengers	2.47 (-86.81%)	10.32 (-44.87%)	18.72 (-)	24.61 (31.46%)	39.85 (112.87%)

5.6 Comparison with the Airline's Solution

Finally, and most importantly, it is insightful to compare the solution provided by our approach and the solution implemented by the airline. Table 8 shows this comparison for the small-scale disruption instance. The first main column corresponds to the recovery solution provided by the airline, the second one corresponds to that from our IMAOR-based optimization approach accounting explicitly for the phantom passengers, and the last one reports how different the IMAOR solution is compared to the Airline solution, as a percentage of the value for the Airline solution. Each of these main three columns feature two sub-columns, optimization and simulation, indicating how the solution evaluation was conducted.

As shown clearly in Table 8, the IMAOR solution substantially outperforms the solution implemented by the airline in terms of the overall objective function without the penalty term. In fact, the IMAOR solution is better than the Airline solution in all but two aspects: extra-fuel cost and number of schedule changes. We find that the mathematical model

uses these two variables, in particular, to better adapt the schedule to the disruption and therefore minimizes the rest of the costs. Simulation results confirm the trend showed by the optimization model in terms of passenger figures – in fact, the simulation results are even more impressive. Most importantly, these results highlight the value of our overall approach to the airline profit and cost. The simulation-based evaluation results in Table 8 show that over a three-day recovery period, the airline can save a total of 446,158 Euros (including a saving of 225,130 Euros in operating cost and a saving of 192,000 Euros in passenger rejection cost) by using our IMAOR-based approach as compared to the currently used solution by the airline.

Table 8: Comparison of our IMAOR-based recovery solution and the airline's real operations

	Airline		**IMAOR**		**Change (%)**	
	Opt.	Sim.	Opt.	Sim.	Opt.	Sim.
Objective function without penalty	13,865,381	14,068,535	13,619,010	13,622,377	-1.78	-3.17
Operating cost	13,524,630		13,299,500		-1.66	
Extra-fuel cost	6,248.14		18,523.21		+196.45	
Delay cost	35,217.35		25,727.00		-29.94	
Re-accommodation cost	17,800	21,300	14,100	15,800	-20.78	-25.82
Compensation cost	281,486	289,140	261,160	262,827	-7.22	-9.10
Rejection cost	0	192,000	0	0	0.00	-100.00
# of schedule changes	8		13		+62.50	
Expected # of phantom passengers	48.39	63.56	14.15	18.72	-70.75	-70.54

6. Extensions and Generalizations

Our computational results so far have demonstrated the consistent superiority of our IMAOR-based approach over various baselines. In this section, we will further evaluate the robustness, scalability and generalizability of our approach under various extensions – both on the modeling side as well as practical side.

6.1 Relaxing the Fuel Cost Approximation

The mathematical model given by the objective function (49) combined with constraints (2)-(34), (36)-(43) and (45)-(48) considers cruise speed optimization in an implicit manner. Specifically, the objective function includes extra fuel consumption due to increased cruise speed, but the model assumes that the extra fuel cost increases linearly with the minutes of delay absorbed by speed increases. Prior literature has demonstrated that the corresponding fuel costs vary nonlinearly, and have been previously modeled using conic quadratic mixed-integer programming (Arıkan et al., 2017). Because our linear approximation may introduce an error in the recovery cost calculation, we test two additional, more accurate, approaches for extra fuel cost calculation, such that we end up with three different models.

The first approximation is the one presented in the mathematical model in Section 4, namely, *Linear*. The second one (named *Piecewise-Linear*) is a generalization of the linear one. It models extra fuel cost as a piecewise linear convex function composed of

two pieces. The last one (named *Conic*) is based on conic optimization similar to Arıkan et al. (2017), and can be considered the most realistic approach available in the literature.

Table 9 shows the results from solving the three models for both disruption scenarios described in Section 5. Each row corresponds to one of the three models. The columns correspond to the objective function value as evaluated by the respective model, the optimality gap provided by the solver, the solution time, as well as another column called "Conic evaluation". This last column lists the objective function value evaluated as per the *Conic* model's objective function expression for the recovery solution obtained by each model. It evaluates the solution from each model based on the most realistic evaluation metric from the literature and allows us to impartially measure the accuracy of the approximate models (namely *Linear* and *Piecewise-Linear*). The values in parentheses in the "Conic evaluation" column indicate the cost increase caused by each approximate model compared with the *Conic* model.

The key findings from both disrupted scenarios are the same. For both small- and large-scale disruptions and for both *Linear* and *Piecewise-Linear* approximations, the cost increases caused by these simplified models are in the range of 0.003% to 0.007% – in other words, negligible. On the other hand, simplified models yield these high-quality solutions in remarkably shorter runtimes – between 42 times and 54 times shorter than using the *Conic* model. This table demonstrates that our simplified model produces almost equally high-quality solutions with much shorter runtimes, compared to the state-of-the-art cruise speed optimization model, thus validating our modeling choice and demonstrating its robustness to relaxing the cruise speed modeling assumption.

Table 9: Comparison of different models for capturing extra fuel consumption costs

Model	Small-scale				Large-scale			
	Objective function	**Optimality gap (%)**	**Solution time (s)**	**Conic evaluation**	**Objective function**	**Optimality gap (%)**	**Solution time (s)**	**Conic evaluation**
Linear	13,671,030	0	**197**	13,674,928 **(0.007%)**	16,225,760	1.63	**600** (time limit)	16,233,802 **(0.005%)**
Piecewise -Linear	13,675,180	0	**253**	13,674,518 **(0.004%)**	16,234,050	2.21	**600** (time limit)	16,233,421 **(0.003%)**
Conic	13,673,921	0	**10,651**	13,673,921	16,232,915	0	**31,354**	16,232,915

6.2 Scalability and Generalizability

All computational experiments in Section 5 were performed using the datasets from one particular hub-and-spoke network carrier; they were performed under two specific types of disruption instances (large delay to a flight and airport closure); and they both lacked crew schedule data and hence did not allow us to evaluate the crew recovery aspects of our integrated model. Next, we overcome each of these three limitations. To demonstrate the generalizability and scalability of our IMAOR-based approach, we conduct another large-scale case study, using a new dataset from a different type of airline – a point-to-point carrier, using a new type of disruption scenario – namely, aircraft on ground, and by fully incorporating the crew recovery aspects of our model in

the resulting computational experiments. Specifically, we now use real-world problem instances based on an airline flying a point-to-point medium range network, featuring 173 airports, and a fleet of approximately 120 aircraft of three fleet types: A319, A320 and A321. This case study spans a recovery horizon of two days, during which more than 1,400 flights are scheduled to be operated. The disruption scenario is driven by Aircraft On Ground (AOG), i.e., unexpectedly grounded aircraft due to causes such as malfunction and crew availability. Specifically, in this case, four aircraft are grounded each for an entire day due to technical issues. From a modeling standpoint, AOG amounts to a fleet size reduction. From a computational standpoint, this leads to a much larger model size, featuring 706,608 variables (including 699,871 binary), more than 40 million constraints and over 124 million non-zero elements. This corresponds to more airports, more aircraft, and more flights compared to either of the case studies in Section 5, and also additionally includes crew recovery.

For tackling this problem of formidable size, Table 10 shows the results from three different implementations of our IMAOR-based approach. A naïve direct implementation (last row of Table 10) directly applies the IMAOR-based approach to this large-scale instance. The naïve approach runs for just over 7 hours and is unable to find an optimal solution (as it runs out of memory); therefore, we report the obtained lower bound instead. Clearly, this is not useful from a recovery practitioner's standpoint, but it still serves as an important benchmark to evaluate the other two implementations. For solving large-scale, long-horizon recovery problems such as this one, a rolling horizon approach (RHA) is typically employed in practice. RHA is a heuristic which entails solving a sequence of optimization subproblems where each subproblem partitions the variables into three subsets. Variables in the first subset are fixed to values obtained in previous subproblems; the variables in the second subset are kept free; and the variables in the third subset are fixed to 0. RHA solves the problem in a step-by-step manner by dividing the recovery horizon into several periods, whose length impacts the solution quality and computational time. RHA serves two purposes. First, it accelerates the solution process. Second, it only requires the knowledge of disruptions in the short-term horizon corresponding to one period. By not needing access to information on disruptions in future periods, it reduces the input data requirements. This double reduction, in data needs and in the computational runtime, is very convenient in a dynamic, real-time setting where data availability and accuracy improve over time, but may lead to suboptimality. We implement RHA by dividing the recovery horizon into 4, 5, 6, and 7 periods and list the results in Columns 3, 4, and 5 of Table 10.

We also consider a second RHA implementation, named "Improved RHA". The first step in Improved RHA is the same as that in RHA – optimize recovery over the first time period. But after that, Improved RHA proceeds by solving for the entire remaining recovery horizon (after Period 1) while fixing the variables corresponding to the first period and keeping the rest of them free. This second step is allowed to run at most until the end of the first time period (while leaving enough slack – of 60 minutes – for the practitioners to implement changes), because that is the earliest time when some of the subsequent decisions from the model need to be implemented. Columns 6, 7, and 8 list the results for Improved RHA. Note that the optimality gaps in Columns 4 and 7 are calculated relative to the lower bound obtained by the Naïve Direct approach. Also, note

that the solution times in Columns 5 and 8 are the runtimes of the first period subproblem in RHA and Improved RHA followed, in square brackets, by the total runtime summed across all subproblems in RHA and Improved RHA. The runtime for the first subproblem is the most critical one since that determines the delay till implementation of the recovery solutions. But we also provide the total runtime across all subproblems for the sake of completeness.

As expected, Improved RHA performs better than RHA, but requires access to disruption information for the entire recovery horizon. More importantly, we observe that RHA and Improved RHA significantly reduce the required runtimes while producing high-quality solutions for such a large-scale problem instance. For example, even with an initial runtime of only 397 seconds the 5-period solution produces optimality gaps of 2.2-2.6% (with respect to the best known lower bound to the true global optimal) for both RHA and Improved RHA demonstrating the effectiveness of the rolling horizon implementations of our IMAOR-based approach. These additional computational experiments in this subsection show that our approach scales well to large-scale problems and generalizes successfully to different airline network types, to different disruption types, and to the integration with a limited capture of crew recovery.

Table 10: Comparison of the solutions from naïve direct, RHA, and Improved RHA implementations for different number of periods

	# periods	RHA			Improved RHA		
		Objective function	Optimality gap (%)	Solution time (s)	Objective function	Optimality gap (%)	Solution time (s)
Heuristic	7	14,290,900	7.12	67 [135]	13,812,445	3.53	67 [21,060]
	6	14,150,630	6.06	105 [241]	13,759,872	3.14	105 [25,200]
	5	13,683,090	2.56	397 [874]	13,641,289	2.25	397 [30,960]
	4	13,580,400	1.79	1,432 [2,269]	13,556,054	1.61	1,432 [39,600]
Naïve Direct		Lower bound			Solution time (s)		
		13,340,935			25,368		

7. Summary

The airline recovery process involves solving multiple problems such as schedule, aircraft, crew, and passenger itinerary recovery, with the ultimate goal of transporting passengers from their respective origins to destinations. Consequently, in the event of disruptions, the integration of these problems is needed to obtain the global optimum to the problem, in a passenger-centric manner. Traditionally, the airline recovery process has sidestepped customer behavior modeling – instead assuming that all passengers will be willing to fly on recovered itineraries, irrespective of their service quality or the associated compensations for the reduced level of service. However, existing passenger rights regulations around the world alter the way passengers perceive the utility of the different alternatives – including the recovered itinerary and the itineraries from competing airlines – enabling some passengers to not show up for the recovered itineraries and instead buy another ticket on a competing airline to reach

destination early. We call these passengers as phantom passengers.

We develop a comprehensive approach which captures regulated passenger rights and passengers' responses to recovery plans. At the core of our modeling approach is an integrated optimization model that includes schedule, aircraft, and passenger itinerary recovery. It also allows maintaining, to the extent possible, crew schedules' consistency and comprehensively captures aircraft maintenance considerations in detail, including airport maintenance capacity constraints that are typically not modeled in the literature. The optimization model also limits schedule deviations with respect to planned operations to facilitate operations to be pulled back to schedule expeditiously. We also develop and estimate a linear regression for passenger response modeling and integrate it into the overall optimization model using datasets provided by our airline partners. The passenger response model replicates the relationship between phantom rates and explanatory variables. Even with a limited amount of available data for the estimation process, the results show that accounting for passenger response in even an approximate manner can still yield significant benefits. We develop an original exact solution approach to solve our model to near-optimality leveraging three key ideas – exact linearization of the passenger-related non-linear constraints and objective function terms, careful penalization of schedule deviations to enable faster solutions, and delayed maintenance check constraint generation to strike the right balance between checking too frequently (causing model intractability) and too infrequently (causing solution infeasibility).

We evaluate disruption scenarios, such as large delay, airport closures and unexpected grounding of aircraft. In addition, we develop a simulation environment to capture the stochasticity in passenger no-show behavior thus providing a better comparison of the different solutions. We find that explicitly accounting for phantom passengers can save the airlines considerable costs. Moreover, these benefits are robust in the face of phantom rate uncertainty –neglecting phantom passengers produces inefficient recovery schedules, while accounting for them in even an approximate manner with relatively poor datasets and/or simplified statistical estimation approaches can still yield significant benefits. Consequently, the proposed modeling framework is attractive from the perspective of the airline operators.

Future research may focus on leveraging more detailed datasets on passenger behaviors under disruptions, developing more sophisticated statistical models to capture such behaviors, and more elaborate evaluation using an even richer set of historical scenarios. We obtained our datasets from a European hub-and-spoke carrier and another European point-to-point carrier for our computational experiments. Using these datasets, we have demonstrated that our approach scales well and demonstrates robustness to a critical modeling assumption. Future research may also focus on developing and evaluating enhanced optimization models and algorithms using other, even larger airline networks and under a wider variety of disruption scenarios.

8. Acknowledgments

This research was supported by Project Grants PID2020-112967GB-C33 by the Agencia

Estatal de Investigación, TRA2016-76914-C3-3-P by the Ministerio de Economía y Competitividad, Spain, and CAS19/00036 by the Ministerio de Ciencia, Innovación y Universidades, Spain.

9. References

- Ahmadbeygi, S., Cohn, A., & Lapp, M. (2010). Decreasing airline delay propagation by re-allocating scheduled slack. IIE transactions, 42(7), 478-489.
- Antunes, D., V. Vaze, and A. Antunes (2019). A Robust Pairing Model for Airline Crew Scheduling. Transportation Science. Transportation Science, 53(6), 1751-1771.
- Arıkan, U., Gürel, S., & Aktürk, M. S. (2017). Flight network-based approach for integrated airline recovery with cruise speed control. Transportation Science, 51(4), 1259-1287.
- Barnhart, C., & Vaze, V. (2015). Irregular operations: Schedule recovery and robustness. The Global Airline Industry, Second Edition, 2653-28674.
- Bratu, S., & Barnhart, C. (2006). Flight operations recovery: New approaches considering passenger recovery. Journal of Scheduling, 9(3), 279-298.
- Cacchiani, V., & Salazar-González, J. J. (2016). Optimal solutions to a real-world integrated airline scheduling problem. Transportation Science, 51(1), 250-268.
- Cadarso, L., & Marín, Á. (2013). Robust passenger-oriented timetable and fleet assignment integration in airline planning. Journal of Air Transport Management, 26, 44-49.
- Cadarso, L., Maróti, G., & Marín, Á. (2015). Smooth and Controlled Recovery Planning of Disruptions in Rapid Transit Networks. IEEE Trans. Intelligent Transportation Systems, 16(4), 2192-2202.
- Cadarso, L., & de Celis, R. (2017). Integrated airline planning: Robust update of scheduling and fleet balancing under demand uncertainty. Transportation Research Part C: Emerging Technologies, 81, 227-245.
- Cadarso, L., Vaze, V., Barnhart, C., & Marín, Á. (2017). Integrated airline scheduling: considering competition effects and the entry of the high speed rail. Transportation Science, 51(1), 132-154.
- Chiraphadhanakul, V., & Barnhart, C. (2013). Robust flight schedules through slack re-allocation. EURO Journal on Transportation and Logistics, 2(4), 277-306.
- Clausen, J., Larsen, A., Larsen, J., & Rezanova, N. J. (2010). Disruption management in the airline industry—Concepts, models and methods. Computers & Operations Research, 37(5), 809-821.
- Court of Justice of the European Union. (2009). Judgment of the Court (Fourth Chamber) of 19 November 2009. https://publications.europa.eu/en/publication-detail/-/publication/15b45be3-f094-4ac3-a6c1-1430a699cf50/language-en
- Delgado, L., Prats, X. (2009). Fuel consumption assessment for speed variation concepts during the cruise phase. ATM Economics Conference. Belgrade, Serbia (German Aviation Research Society and University of Belgrade), 1–12.
- Department of Transportation, United States (USDOT). (2008). PART 250—Oversales. https://www.govinfo.gov/content/pkg/CFR-2018-title14-

vol4/pdf/CFR-2018-title14-vol4-part250.pdf

- Department of Transportation, United States (USDOT). (2010). PART 259—Enhanced Protections for Airline Passengers. https://www.gpo.gov/fdsys/pkg/CFR-2012-title14-vol4/pdf/CFR-2012-title14-vol4-part259.pdf
- Dunbar, M., Froyland, G., & Wu, C. L. (2012). Robust airline schedule planning: Minimizing propagated delay in an integrated routing and crewing framework. Transportation Science, 46(2), 204-216.
- Dunbar, M., Froyland, G., & Wu, C. L. (2014). An integrated scenario-based approach for robust aircraft routing, crew pairing and re-timing. Computers & Operations Research, 45, 68-86.
- European Parliament and the Council of the European Union (2004). Regulation (EC) no 261/2004 of the European Parliament and of the Council of 11 February 2004 establishing common rules on compensation and assistance to passengers in the event of denied boarding and of cancelation or long delay of flights, and repealing Regulation (EEC) No 295/91. Official Journal of the European Union, https://eur-lex.europa.eu/resource.html?uri=cellar:439cd3a7-fd3c-4da7-8bf4-b0f60600c1d6.0004.02/DOC_1&format=PDF
- Gao, C., Johnson, E., & Smith, B. (2009). Integrated airline fleet and crew robust planning. Transportation Science, 43(1), 2-16.
- Hu, Y., Song, Y., Zhao, K., & Xu, B. (2016). Integrated recovery of aircraft and passengers after airline operation disruption based on a GRASP algorithm. Transportation research part E: logistics and transportation review, 87, 97-112.
- International Air Transport Association (IATA). (1999). Montreal Convention 1999. https://www.iata.org/policy/Documents/MC99_en.pdf
- Jiang, H., & Barnhart, C. (2013). Robust airline schedule design in a dynamic scheduling environment. Computers & Operations Research, 40(3), 831-840.
- Lan, S., Clarke, J. P., & Barnhart, C. (2006). Planning for robust airline operations: Optimizing aircraft routings and flight departure times to minimize passenger disruptions. Transportation science, 40(1), 15-28.
- Lettovský, L., Johnson, E. L., & Nemhauser, G. L. (2000). Airline crew recovery. Transportation Science, 34(4), 337-348.
- Liang, Z., Xiao, F., Qian, X., Zhou, L., Jin, X., Lu, X., & Karichery, S. (2018). A column generation-based heuristic for aircraft recovery problem with airport capacity constraints and maintenance flexibility. Transportation Research Part B: Methodological, 113, 70-90.
- Maher, S. J. (2015a). Solving the integrated airline recovery problem using column-and-row generation. Transportation Science, 50(1), 216-239.
- Maher, S. J. (2015b). A novel passenger recovery approach for the integrated airline recovery problem. Computers & Operations Research, 57, 123-137.
- Marla, L., Vaaben, B., & Barnhart, C. (2016). Integrated disruption management and flight planning to trade off delays and fuel burn. Transportation Science, 51(1), 88-111.
- Marla, L., Vaze, V., & Barnhart, C. (2018). Robust optimization: lessons learned from aircraft routing. Computers and Operations Research, 98, 165-184.
- Petersen, J. D., Sölveling, G., Clarke, J. P., Johnson, E. L., & Shebalov, S. (2012). An optimization approach to airline integrated recovery. Transportation Science,

46(4), 482-500.

- Rosenberger, J. M., Johnson, E. L., & Nemhauser, G. L. (2003). Rerouting aircraft for airline recovery. Transportation Science, 37(4), 408-421.
- Sinclair, K., Cordeau, J. F., & Laporte, G. (2014). Improvements to a large neighborhood search heuristic for an integrated aircraft and passenger recovery problem. European Journal of Operational Research, 233(1), 234-245.
- Sohoni, M., Lee, Y. C., & Klabjan, D. (2011) Robust airline scheduling under block-time uncertainty. Transportation Science, 45(4), 451-464.
- Wei, K., Vaze, V., & Jacquillat, A. (2020). Airline timetable development and fleet assignment incorporating passenger choice. Transportation Science, 54(1), 139-163.
- Yan, C., Barnhart, C., & Vaze, V. (2021). Choice-Based Airline Schedule Design and Fleet Assignment: A Decomposition Approach. Available at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3513164.
- Yan, C., & Kung, J. (2018). Robust aircraft routing. Transportation Science, 52(1), 118-133.
- Yu, G., Argüello, M., Song, G., McCowan, S. M., & White, A. (2003). A new era for crew recovery at continental airlines. Interfaces, 33(1), 5-22.
- Zhang, D., Lau, H. H., & Yu, C. (2015). A two-stage heuristic algorithm for the integrated aircraft and crew schedule recovery problems. Computers & Industrial Engineering, 87, 436-453.
- Zhang, D., Yu, C., Desai, J., & Lau, H. H. (2016). A math-heuristic algorithm for the integrated air service recovery. Transportation Research Part B: Methodological, 84, 211-236.