Dynamic Disruption Management in Airline Networks Under Airport Operating Uncertainty

Dynamic Disruption Management in Airline Networks under Airport Operating Uncertainty

Lavanya Marla
Department of Industrial and Enterprise Systems Engineering, University of Illinois at Urbana-Champaign, Urbana, IL 61801

Alexandre Jacquillat
Sloan School of Management, Massachusetts Institute of Technology, Cambridge MA 0214

Jane Lee
Department of Civil and Environmental Engineering, University of Illinois at Urbana-Champaign, Urbana, IL 61801

Air traffic disruptions result in flight delays, cancellations, passenger misconnections, and ultimately high costs to aviation stakeholders. This paper proposes a jointly reactive and proactive approach to airline disruption management, which optimizes recovery decisions in response to realized disruptions and in anticipation of future disruptions. The approach forecasts future disruptions partially and probabilistically by estimating systemic delays at hub airports (and the uncertainty thereof) and ignoring other contingent disruptions. It formulates a dynamic stochastic integer programming framework to minimize network-wide expected disruption recovery costs. Specifically, our Stochastic Reactive and Proactive Disruption Management (SRPDM) model combines a stochastic queuing model of airport congestion, a flight planning tool from Boeing/Jeppesen and an integer programming model of airline disruption recovery. We develop a solution procedure based on look-ahead approximation and sample average approximation, which enables the model's implementation in short computational times. Experimental results show that leveraging even partial and probabilistic estimates of future disruptions can reduce expected recovery costs by 1–2%, as compared to a myopic baseline approach based on realized disruptions alone. These benefits are mainly driven by the deliberate introduction of departure holds to reduce expected fuel costs, flight cancellations and aircraft swaps.

Key words: airline disruption management, stochastic optimization, integer programming, queuing model

1. Introduction

The formation and propagation of operating disruptions across spatial-temporal networks create missed revenue opportunities, resource wastage, employee overtime shifts and reduced customer satisfaction, leading to financial and welfare losses in industries such as supply chains, transportation, telecommunications, and medical services. As a prime example, flight delays and cancellations create significant costs across air traffic networks—estimated at over $30 billion in the United States in 2007 (Ball et al. 2010). Disruption management thus lies at the core of airline operations.

Airline disruption management interventions can be broadly classified into two categories: reactive and proactive interventions. Proactive interventions stem from the airline robust planning literature: they provide *a priori* operating plans (e.g., flight schedules, aircraft routings and crew

pairings) that can respond effectively to future disruptions. However, they do not adjust operating plans dynamically as operating disruptions are realized. Reactive interventions stem from the airline recovery literature: they provide *a posteriori* recovery plans in response to observed disruptions (e.g., whether, when and with which aircraft to operate each flight) to minimize the costs of bringing operations back to normal. However, they do not anticipate future disruptions that are likely to occur across the airline's network of flights—thus potentially resulting in sub-optimal decisions when future operations themselves depart from planned operations.

Instead, this paper proposes an original approach to disruption management that is jointly reactive and proactive—by simultaneously responding to past disruptions and anticipating future disruptions. A major challenge is that future disruptions can only be characterized *probabilistically* and *partially*. First, air traffic operations are subject to significant uncertainty, so disruptions cannot be anticipated in advance exactly and with certainty. Second, operating disruptions stem from systemic and contingent sources. Systemic disruptions arise from congestion resulting from more flights being scheduled than available capacity at busy airports. These disruptions can be estimated by means of stochastic queuing models, as shown by Pyrgiotis et al. (2013) and Jacquillat and Odoni (2015b). Contingent disruptions include other delay sources, such as aircraft maintenance, late crews, late passenger boarding, etc. In comparison, contingent disruptions are very difficult to anticipate. This paper integrates probabilistic forecasts of systemic disruptions across networks of operations into a dynamic and stochastic optimization framework for airline disruption recovery.

Specifically, this paper makes the following contributions:

It develops an original approach to network-wide disruption management that proactively leverages partial and probabilistic disruption forecasts into reactive disruption recovery. The approach relies on a dynamic and stochastic optimization model that optimizes recovery decisions, given observed disruptions and the probability distribution of future systemic disruptions. This paper thus integrates, for the first time to our knowledge, principles from the robust airline planning literature into the disruption recovery literatures. As compared to existing disruption recovery approaches, the proposed framework results in larger and more complex optimization problems, but can reduce expected disruption costs through more flexible and robust disruption management.

It formulates a Stochastic Reactive and Proactive Disruption Management (SRPDM) model to optimize network-wide airline disruption recovery under airport queuing stochasticity. SRPDM is formulated as a stochastic integer program using a probabilistic time-space network representation. It combines: (i) the stochastic and dynamic queuing model from Jacquillat and Odoni (2015a), which yields the probability distribution of delays over time at each hub airport; (ii) a flight planning tool from Boeing/Jeppesen, which provides routing, speed and altitude options for each flight, along with corresponding flying times and fuel costs; and (iii) the deterministic model of

recovery optimization from Marla et al. (2017). This provides the first model of network-wide air traffic optimization that applies a stochastic queuing model at multiple airports simultaneously.

It develops an efficient approximate algorithm that can solve the SRPDM within reasonable computational times, consistent with the airline disruption recovery literature and with practical implementation requirements. The size of SRPDM increases with the scale of the network, the time horizon of recovery, and the number of systemic disruption scenarios. In realistic instances, the model's sheer size makes direct implementation highly intractable. To solve it efficiently, this paper develops an approximate solution algorithm based on look-ahead approximation (by optimizing recovery decisions for a restricted time window) and sample average approximation (by leveraging a sampled set of disruption scenarios). The algorithm solves SRPDM iteratively over a rolling horizon (we use a one-hour rolling horizon in this paper, but the proposed algorithm can be applied more or less frequently in practice). Using real-world scheduling data from Delta Air Lines, we show that, at any point in time, the proposed algorithm can solve SRPDM in 3-5 minutes—which is consistent with earlier airline recovery models and with practical requirements. Ultimately, these computational results demonstrate the model's implementability in practice.

It shows that our jointly reactive and proactive approach to disruption management can significantly enhance recovery decisions, as compared to purely reactive approaches. Since our approximate algorithm does not yield solution quality guarantees, we compare the recovery solutions obtained with our modeling and computational framework to those obtained with a myopic baseline approach that does not leverage forecasts of future disruptions. For this comparison, we use disruption realizations derived from real-world data. Results suggest that our approach reduces expected recovery costs by 1–2%. Moreover, we find no disruption instance in which our approach increases recovery costs (it performs either as well as or better than the baseline). Stated differently, our stochastic optimization approach reduces expected operating costs without introducing additional risk in airline recovery. These benefits are mainly driven by deliberately introducing departure holds at key points in the network to reduce expected fuel costs, flight cancellations and aircraft swaps. This approach is particularly beneficial for airlines with concentrated operations at hub airports and with congested hubs. Ultimately, these case study results demonstrate the benefits of proactively leveraging even partial and probabilistic information on future disruptions and applying even an approximate stochastic optimization algorithm to enhance airline recovery decisions.

We review the literatures on robust airline planning and disruption recovery in Section 2. We describe our dynamic decision-making approach in Section 3. Section 4 formulates SRPDM and describes our solution algorithm. Our experimental setting is detailed in Section 5. Section 6 reports computational results, showing the benefits of the proposed modeling and computational approach. Section 7 summarizes our findings and outlines opportunities for future research.

2. Literature Review

The management of flight disruptions is one of the foremost objectives of air traffic management and airline operating systems. Major air traffic management interventions include the optimization of airport operations (Balakrishnan and Chandran 2010, Simaiakis et al. 2014, Jacquillat and Odoni 2017), air traffic flow management (Bertsimas et al. 2011, Vossen et al. 2012), and airport demand management (Zografos et al. 2012, Jacquillat and Odoni 2015a, Ribeiro et al. 2017). From an airline's perspective, minimizing disruptions comprises two main steps: (i) robust airline planning—to reduce its vulnerability to future disruptions (a proactive method) and (ii) disruption recovery—to re-allocate resources and minimize the impact of observed disruptions (a reactive approach). We review these two bodies of work in this section.

Robust airline planning involves optimizing planning decisions (such as flight schedules, fleet assignments, aircraft routings and crew schedules) to minimize the cost of operating disruptions, if and when they occur. This literature includes two main approaches: (i) those that minimize the *impact* of delays, and (ii) those that minimize the *occurrence* of (propagated) delays.

The first category designs strategic plans to respond effectively to future disruptions. Rosenberger et al. (2004) create fleet assignments with "short cycles" to minimize the ripple effects of cancellations. Smith and Johnson (2006) restrict the number of aircraft types at airports to create swapping opportunities. Sohoni et al. (2011) formulate probabilistic service level constraints, under block-time uncertainty. Arıkan et al. (2013) propose robust scheduling and network planning strategies, under delay propagation uncertainty. Pita et al. (2012) integrate airport congestion estimates into flight scheduling and fleet assignment. Froyland et al. (2013) optimize aircraft routing, given the uncertainty of future disruptions and resulting recovery. Other studies incorporate robustness into crew pairing (Schaefer et al. 2005, Yen and Birge 2006, Shebalov and Klabjan 2006).

The second category designs strategic plans to minimize delay propagation across flight networks. Lan et al. (2006) distinguish primary vs. propagated delays, and propose optimization models to prevent delay propagation through aircraft routing (by allocating schedule slack where it is most critical) and schedule retimings (by adjusting flights' departure and arrival times to reduce passenger misconnections). Ahmadbeygi et al. (2010) and Borndörfer et al. (2010) optimize aircraft routings and schedule re-timings to minimize propagated delays. Cadarso and Marín (2011) optimize flight scheduling and fleet assignment to avoid passenger misconnections, by allocating schedule slack accordingly. Dunbar et al. (2012) and Dunbar et al. (2014) optimize aircraft routings and crew pairings to minimize propagated delay—assuming deterministic and stochastic primary delays, respectively. Yan and Kung (2016) use robust optimization to capture the uncertainty on primary delays into the optimization of aircraft routings. Marla et al. (2018) compare chance programming, stochastic programming and robust optimization for aircraft routing.

This paper departs from this literature in two ways. First, robust airline planning focuses on strategic decisions, made prior to the day of operations in anticipation of future disruptions. In contrast, this paper optimizes airline recovery decisions, made during the day of operations in response to observed disruptions. Second, the vast majority of the robust airline planning literature focuses on propagated delays (due to insufficient buffers). In this work, we further break down non-propagated (primary) delays into "systemic" and "contingent" delays, and use a queuing model of airport congestion to capture systemic disruptions—in addition to propagated delays.

Airline disruption recovery involves optimizing operating decisions in response to observed disruptions during the day of operations, in order to minimize the costs of bringing operations back to normal. The main recovery levers include, from the least to most disruptive, aircraft and crew swaps (i.e., changes in aircraft-flight assignments and crew pairings), departure holds (i.e., voluntary introduction of flight departure delays), passenger re-accommodations, and flight cancellations (see Barnhart and Vaze 2015). Typical trade-offs involve operating flights close to their schedule to minimize delays vs. introducing departure holds to ensure connectivity.

Starting with initial aircraft recovery heuristics from Teodorović (1984), researchers have designed large-scale optimization algorithms to deal with realistic problem instances (see, e.g., Jarrah et al. 1993, Cao and Kanafani 1997, Clarke and Naryadi 1995, Yan and Yang 1996). Thengvall et al. (2000) extend basic aircraft recovery models to minimize changes in aircraft routings and to capture airlines' preferences. Rosenberger et al. (2003) jointly optimize departure holds (i.e., flight rescheduling) and aircraft reroutings. Eggenberg et al. (2010) add operational constraints to ensure that the airline can comply with aircraft maintenance, crew recovery, and passenger accommodation requirements. Other studies focused on crew recovery, following aircraft recovery decisions (Wei et al. 1997, Lettovský et al. 2000, Yu et al. 2003).

Subsequent studies integrate the problems of aircraft, crew and passenger recovery. Zhang et al. (2015) address aircraft and crew recovery in sequence. Jozefowiez et al. (2013) and Zhang et al. (2016) present three-step heuristics that sequentially solve schedule recovery, aircraft recovery and passenger recovery. Bratu and Barnhart (2006) combine aircraft and passenger recovery. Petersen et al. (2012) propose a fully integrated model of schedule, aircraft, passenger and crew recovery, solved with Benders decomposition, column generation and row generation. Follow-up studies have developed algorithms for large-scale integrated recovery problems, using large-scale neighborhood search (Sinclair et al. 2014), a reduced time-band representation (Hu et al. 2015), greedy randomized adaptive search (Hu et al. 2016), and row-and-column generation (Maher 2016). Marla et al. (2017) integrate flight planning (i.e., aircraft routing, flying altitude and speed) into aircraft and passenger recovery—showing that flight planning provides an additional recovery lever.

In contrast with this body of work, our paper leverages forecasts of future systemic disruptions—and the uncertainty thereof—into recovery optimization. To our knowledge, the literature on recovery optimization under uncertainty is limited. Abdelghany et al. (2004) propose a heuristic simulation to project flight delays and Abdelghany et al. (2008) integrate delay uncertainty into a myopic optimization model of disruption recovery. Jafari and Zegordi (2010) dynamically optimize aircraft recovery and passenger re-accommodation, on a relatively small-scale network and with two disruption scenarios. McCarty and Cohn (2018) propose a two-stage stochastic program to pre-emptively change passenger itineraries, before misconnections occur, under uncertainty on delay propagation. Our paper shares similarities with this approach but also exhibits two differences: (i) we focus on aircraft recovery, as opposed to passenger recovery, and (ii) we incorporate forecasts of propagated delays as well as systemic delays resulting from airport congestion.

Summary. This paper augments the prior literature in two major ways:

1. We propose the first jointly reactive and proactive approach to airline disruption management that optimizes aircraft recovery in response to observed disruptions, while anticipating future disruptions (partially and probabilistically). This approach differs from the airline disruption recovery literature by proactively leveraging forecasts of future disruptions. As such, it shares similarities with the robust airline planning literature, but it deals with a tactical disruption recovery problem—as opposed to a strategic planning problem.
2. We integrate probabilistic forecasts of systemic delays arising from demand-capacity imbalances at busy airports—in addition to propagated delays—into disruption management. This is achieved by embedding future disruption scenarios obtained from a stochastic queuing model of airport congestion into a prescriptive optimization framework of disruption recovery.

3. Modeling Framework

We now formulate our dynamic decision-making framework for aircraft recovery. The model optimizes recovery and flight planning decisions. Recovery decisions include departure times (i.e., when to operate each flight), aircraft-flight assignments (i.e., whether to "swap" aircraft or not), and flight cancellations. Flight planning decisions include aircraft routing, flying altitude and flying speed. Together, these two sets of decisions determine recovery costs (i.e., delay costs, swap costs and cancellation costs) and flight operating costs. Unlike existing approaches, our framework optimizes these decisions in response to observed disruptions as well as given forecasts of future disruptions—thus providing a jointly reactive and proactive approach to disruption management.

Specifically, disruptions observed at any point in time can be classified into three categories:

- *Propagated disruptions*: past disruptions spreading across spatial-temporal networks, due to insufficient buffers in the schedule to absorb upstream delays.

- *Systemic disruptions*: congestion at hub airports induced by demand-capacity imbalances.
- *Contingent disruptions*: other inefficiencies within airline and passenger operations (e.g., aircraft maintenance, late crews, late passenger boarding).

At any decision time, the decision-maker observes all operating disruptions. However, future disruptions are only known partially and probabilistically. First, our approach (like existing ones) captures propagated disruptions resulting from recovery decisions. Second, our approach (unlike existing ones) considers probabilistic forecasts of systemic disruptions—obtained from a stochastic and dynamic queuing model at hub airports. Third, our approach ignores future contingent disruptions in the optimization, but these still realize randomly at each decision point.

Our modeling architecture is shown in Figure 1. It starts by applying the stochastic queuing model to generate probabilistic forecasts of systemic disruptions, using data on flight schedules and airport capacities (Section 3.1). It then optimizes recovery decisions dynamically. We divide the day into T periods, indexed by $t = 1, \cdots, T$. In each period, the state of the system is determined from past operations and observed (propagated, systemic and contingent) disruptions. Flight planning options are generated with an engineering tool provided by Boeing/Jeppesen called JetPlan (Section 3.2). We represent recovery and flight planning options in a probabilistic time-space network of operations (Section 3.3). We then optimize airline recovery decisions to minimize expected recovery costs. This is cast as a dynamic program (Section 3.4). However, the size of the problem makes it intractable, so we propose in Section 4 a solution procedure based on look-ahead approximation and sample average approximation. This procedure relies on the Stochastic Reactive and Proactive Disruption Management (SRPDM), which optimizes recovery decisions across a sampled set of disruption scenarios for a given look-ahead window. The recovery plan is used to define the state of the system in the next period. The process is repeated until the end of the horizon.

A few observations on our problem are noteworthy. First, the approach developed in this paper can be applied to any airline network but is likely to be most beneficial for hub-and-spoke airlines. Second, this paper focuses on schedule and aircraft recovery. We leave the integration of other recovery decisions, such as passenger and crew recovery, into our stochastic optimization framework for future research. In practice, passenger and crew recovery can be optimized subsequently, given the aircraft recovery plan. Last, we solve the aircraft recovery problem for each fleet type independently. This is consistent with the existing literature and current practice—as aircraft swaps typically occur within each fleet to minimize interference with passenger and crew itineraries.

3.1. Queuing Model of Systemic Disruptions at Hub Airports

The stochastic and dynamic queuing model is applied at each hub airport to forecast future systemic disruptions. This approach characterizes the airport as a queuing system, in which service is provided by the runway system and aircraft join the system when they are ready to take off or to

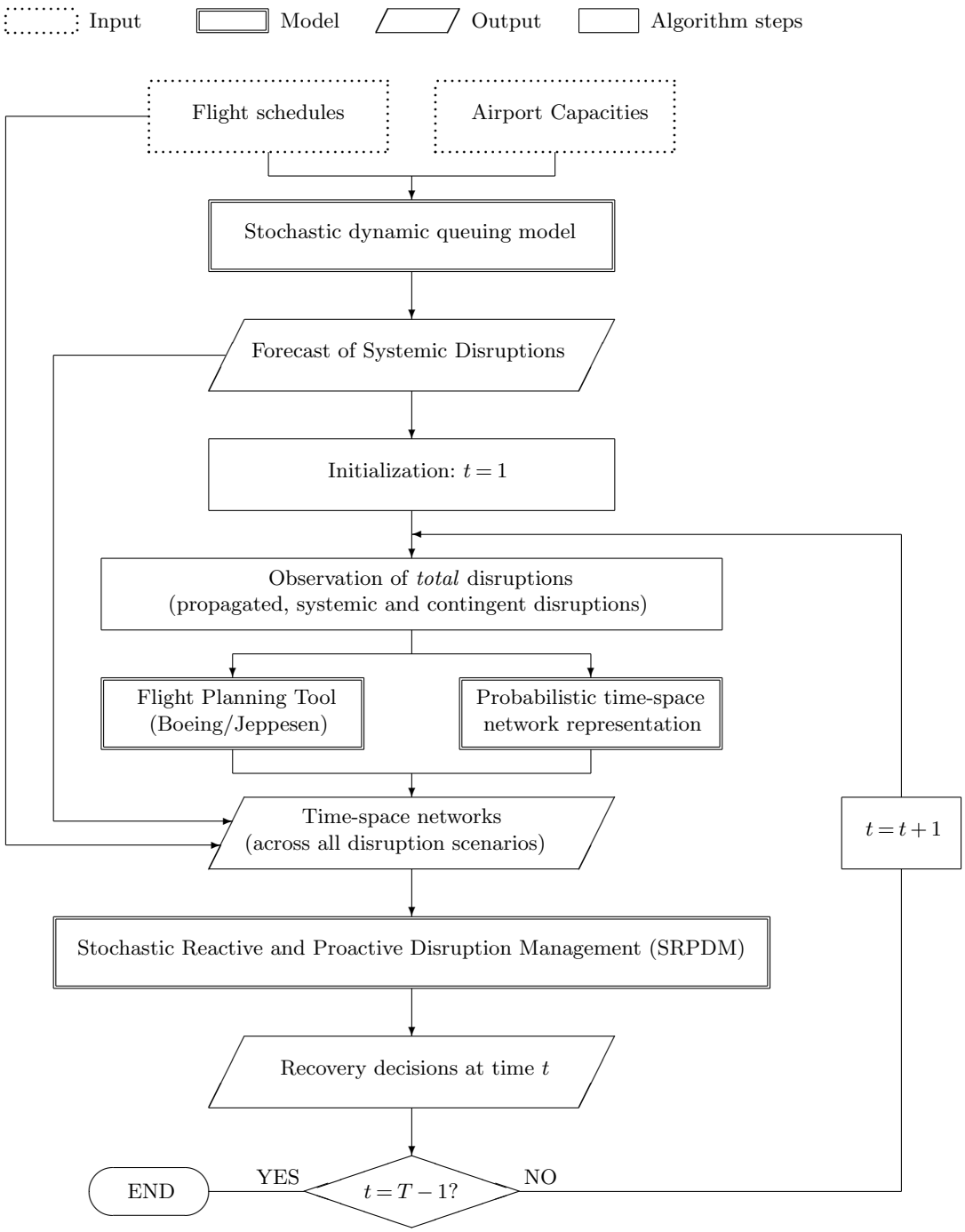


Figure 1 **Modeling architecture.**

land. The model takes as inputs the schedule of flights and the runway capacity at each airport. It returns the probability distribution of flight delays at each time of the day, which is then used to sample disruption scenarios and to construct our time-space networks (see Figure 1).

Specifically, we model arrival and departure delays at each airport by means of dynamic and stochastic $M(t)/E_3(t)/1$ queuing models. In other words, the arrival and departure demand pro-

cesses are both modeled as Poisson processes, and the arrival and departure service processes are modeled as Erlang processes of order 3. The model is non-stationary, i.e., demand and service rates are time-varying to reflect changes in flight schedules and airport capacities over the day. We divide the day of operations into periods of length $S = 15$ minutes. The demand rates (λ_s in period s) are determined by the number of flights scheduled. The service rates (μ_s in period s) are constrained by the airport's capacity. To reflect air traffic operating procedures, we integrate a dynamic programming model that optimizes arrival and departure service rates, under capacity constraints, by selecting runway configurations (i.e., the set of active runways) and balancing arrivals and departures (Jacquillat et al. 2016). This approach approximates the dynamics and magnitude of delays at busy airports with good accuracy (Jacquillat and Odoni 2015b).

The state-transition diagram of the $M(t)/E_3(t)/1$ queuing system is shown in Figure 2. It relies on the characterization of an Erlang process of order 3 and rate μ as the succession of 3 Markovian "stages of work", each completed at rate 3μ. The state i defines the number of remaining stages of work. Let u be a time index that varies continuously, and $P_i(u)$ be the probability of being in state i at time u. Equations (1)–(5) show the system of Chapman-Kolmogorov first-order differential equations describing the evolution of $P_i(u)$ in period s, with u varying between $(s-1)S$ and sS. The practical queue capacity is denoted by N. The system is empty at the beginning of the day. We solve Equations (1)–(5) using the built-in differential equation solver ode45 in MATLAB 8.1.

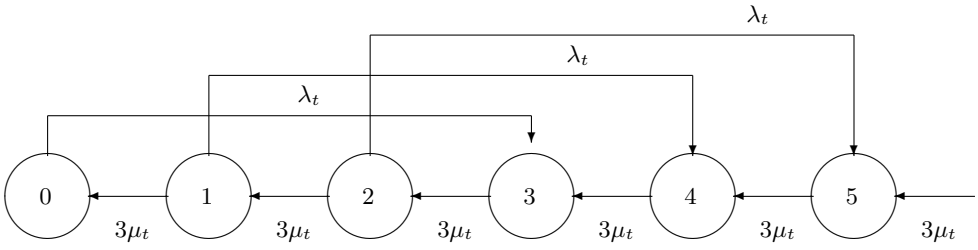


Figure 2 State-transition diagram of the $M(t)/E_3(t)/1$ queuing system.

$$\frac{dP_0(u)}{du} = -\lambda_t P_0(u) + k\mu_t P_1(u) \tag{1}$$

$$\frac{dP_i(u)}{du} = -(\lambda_t + k\mu_t)P_i(u) + k\mu_t P_{i+1}(u) \qquad \forall i \in \{1, \cdots, k\} \tag{2}$$

$$\frac{dP_i(u)}{du} = \lambda_t P_{i-k}(u) - (\lambda_t + k\mu_t)P_i(u) + k\mu_t P_{i+1}(u) \qquad \forall i \in \{k+1, \ldots, (N-1)k\} \tag{3}$$

$$\frac{dP_i(u)}{du} = \lambda_t P_{i-k}(u) - k\mu_t P_i(u) + k\mu_t P_{i+1}(u) \qquad \forall i \in \{(N-1)k+1, \ldots, kN-1\} \tag{4}$$

$$\frac{dP_{kN}(u)}{du} = \lambda_t P_{k(N-1)}(u) - k\mu_t P_{kN}(u) \tag{5}$$

We denote the sample space of systemic disruption profiles by $\tilde{\mathcal{Q}}$, indexed by $\tilde{q} = 1, \cdots, \tilde{Q}$, each occurring with a probability $\xi_{\tilde{q}}$. The set $\tilde{\mathcal{Q}}$ includes all queue length combinations in all time periods

and at all airports. Across T periods and in a network of K airports, the cardinality of $\tilde{\mathcal{Q}}$ is thus $(N+1)^{2TK}$ (since the arrival and departure queue lengths can each take any of the values $0, \cdots, N$ in each period and at each airport). Even for a short horizon and a small network, integrating the full range of airport congestion outcomes into recovery optimization is highly intractable. Therefore, we will proceed by Monte Carlo sampling to generate representative scenarios from this probability distribution. We denote by $\mathcal{Q}$ the set of such sampled scenarios, indexed by $q = 1, \cdots, Q$.

Two comments on this queuing model are noteworthy. First, the model is applied independently at each airport. This is motivated by the fact that airport operating stochasticity primarily stems from local factors (e.g., variations in flight operations, weather, aircraft mix, etc.). Second, look-ahead stochastic disruption scenarios are generated once for all periods $t = 1, \cdots, T$. In practice, delays occurring from period t onward obviously depend on realized congestion at period t. However, our approach aims to capture delay forecasts that are *available to an airline*. In the current environment, the level of collaboration between traffic managers and airlines is such that information on the exact number of queuing aircraft at each airport is not publicly shared in real time. Therefore, we adopt a conservative approach that only leverages the information that is available before the day of operations (e.g., the schedule of flights) or can be estimated from historical records of operations (e.g., airport capacity estimates). In future work, this assumption can be relaxed by integrating a dynamic queuing update mechanism into the framework shown in Figure 1—thus identifying the benefits resulting from real-time information sharing between operating entities.

3.2. Flight Planning

We leverage in this paper the JetPlan tool from Boeing/Jeppesen—a flight planning software used by many airlines to plan their flight trajectories prior to departure. JetPlan takes as inputs the flight's scheduled departure and arrival times, anticipated weather patterns, aircraft and engine configurations, and the aircraft's payload (including cargo, passengers, luggage, and fuel). It generates flight planning options (including aircraft routing across waypoints from origin to destination, flight speeds, and flying altitudes) and estimates the resulting fuel costs and travel times. In our framework, JetPlan is used to generate flight copies in the time-space networks (see Figure 1).

Specifically, the flight trajectories generated by JetPlan are expressed as a function of an engineering-based metric called the cost-index (CI). CI is defined as the ratio of the flight's time-related costs (determined by the flight's duration and aircraft, passenger and crew connectivities) divided by the fuel cost. CI can be interpreted as the amount of additional fuel worth burning (relative to the minimum fuel burn to operate the flight) to save one unit of time. The most fuel-efficient flight plan is referred to as CI0. The larger the CI value, the higher the fuel costs and the lower the flying times. Oftentimes, operating all flights at CI0 result in delays and lost connectivity,

especially under disruptions; therefore, flights are planned at a slightly higher CI, typically around CI30. In our experiments, we use as discrete inputs corresponding to CI0 (a conservative flight plan), CI30 (the baseline option), CI70, CI100, and CI700 (increasingly aggressive flight plans).

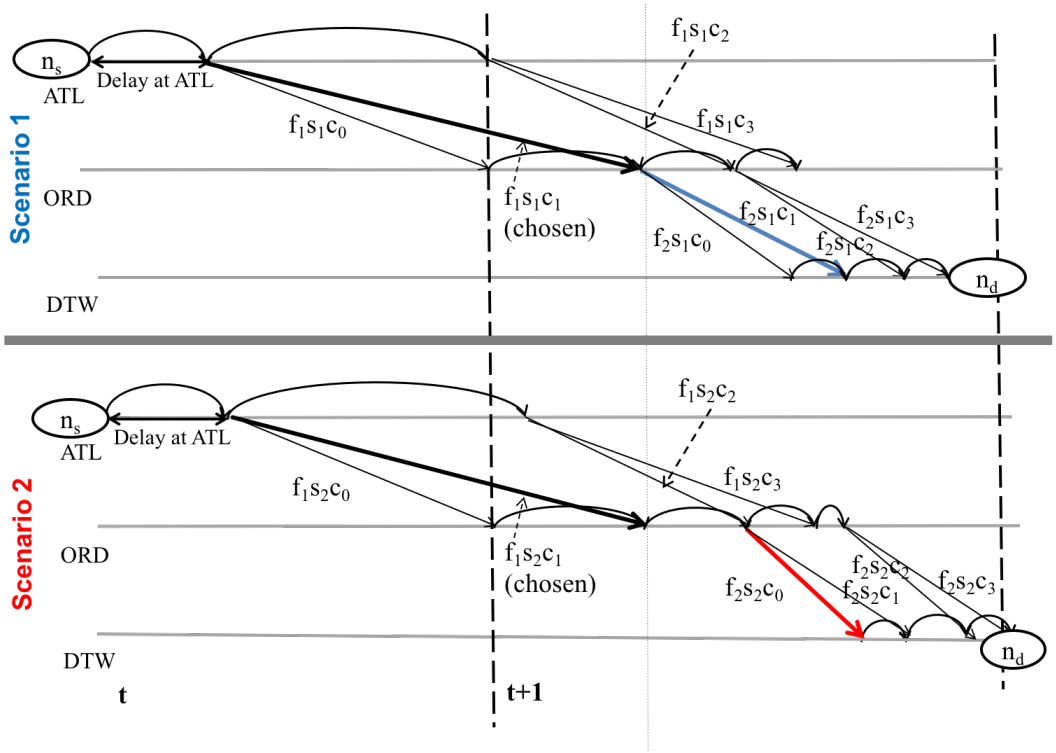
3.3. Time-space Network Representation

In each period, we construct a time-space network for each aircraft, comprising all flights that the aircraft can operate by the end of the horizon (period T). This representation starts with the aircraft's current location, observed disruptions, probabilistic forecasts of future systemic disruptions (Section 3.1), and possible flight plans (Section 3.2). For each flight, the network defines several *copies*, each associated with departure and arrival times and with a flight plan. It is then used to formulate a multi-commodity network flow model, with each aircraft treated as a commodity—enabling to optimize flight-aircraft pairings among a huge number of options (see Figure 1).

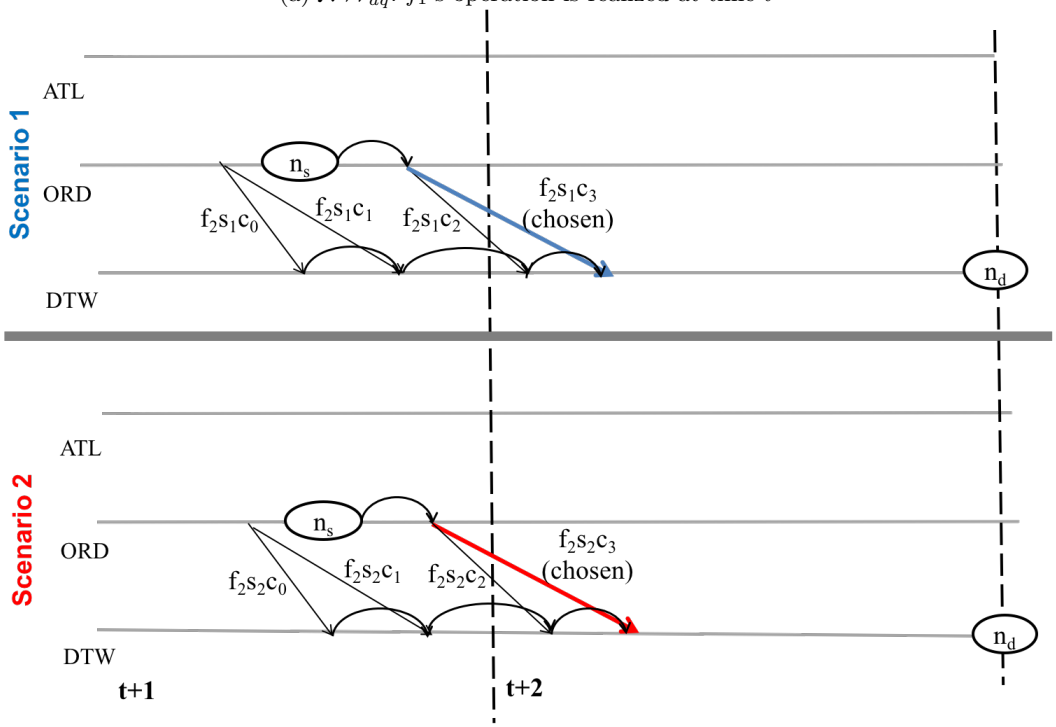
Figure 3 shows an example of such time-space network representation in periods t (Figure 3a) and $t+1$ (Figure 3b). Let us denote by $\widehat{\mathcal{NW}}_a^t$ the time-space network in period $t = 1, \cdots, T$ for aircraft $a \in \mathcal{A}$. Each node in $\widehat{\mathcal{NW}}_a^t$ represents a combination of time and location. Each arc represents a possible flight arc (straight line in Figure 3) or a ground arc, that is, the aircraft's turnaround from one arrival node to a subsequent departure node at the same airport (curved line). We add a supply node n_s denoting the location and time where the aircraft is currently available, and a demand node n_d representing the end of the aircraft's operation.

Flight copies are widely used in airline recovery, but we highlight here two particular features of our approach. First, flight copies differ not only by departure time (thus capturing recovery decisions, such as departure times, aircraft-flight assignments, and cancellations) but also by flight duration (thus capturing flight planning decisions, such as route, altitude, speed). This is consistent with the model from Marla et al. (2017). For instance, in Figure 3, flight f_1 has four copies ($f_1s_1c_0$, $f_1s_1c_1$, $f_1s_1c_2$ and $f_1s_1c_3$), where $f_1s_1c_2$ and $f_1s_1c_3$ involve larger departure delays than $f_1s_1c_0$ and $f_1s_1c_1$, and $f_1s_1c_1$ and $f_1s_1c_3$ have a longer en-route time than $f_1s_1c_0$ and $f_1s_1c_2$. Second, our time-space networks are subject to uncertainty regarding future disruptions. Thus, $\widehat{\mathcal{NW}}_a^t$ represents a probabilistic time-space network at this point. In Section 4, we shall develop scenario-based time-space network representations. This probabilistic time-space network representation lies at the core of the stochastic approach to airline recovery and flight planning developed in this paper.

The time-space network representation captures the evolution of the system over time by tracking each aircraft's movement over time and space and creating flight copies based on the latest (propagated, systemic and contingent) disruptions observations. To illustrate the dynamics of the system, the figure shows two scenarios, assumed to be equally probable. Flight f_1 is scheduled to operate in period t, before uncertainty is resolved. Decisions related to flight f_1 thus need to



(a) $\widehat{\mathcal{NW}}_{aq}^{t}$: f_1's operation is realized at time t



(b) $\widehat{\mathcal{NW}}_{aq}^{t+1}$: f_2's operation is realized at time $t+1$

Figure 3 **Example of probabilistic time-space network representation.**

be identical across all scenarios; in our example, copy c_1 is selected. Then, the model anticipates to operate copy c_1 of flight f_2 in scenario 1 (a slower option since departure delay is small) and copy c_0 in scenario 2 (a faster option since departure delay is larger). Figure 3b shows the state of the system at time $t+1$; the supply node n_s updates the aircraft's availability, depending on prior decisions and realized disruptions. The model is solved again to optimize recovery decisions in period $t+1$—at that point, the operating decision for flight f_2 (namely, operating copy c_3) is identical in scenario 1 and in scenario 2. The process is repeated until the end of the horizon.

3.4. Stochastic Optimization Approach

Finally, our stochastic optimization model determines, in each period $t = 1, \cdots, T$, the airline's recovery and flight planning decisions. It takes as inputs the probabilistic time-space networks (Section 3.3) of all aircraft (see Figure 1). For each flight, it selects at most one copy across all time-space networks—thus ensuring that each flight is either covered by one aircraft or cancelled.

We cast this problem as a finite-horizon dynamic program. We describe it in this section, and motivate our look-ahead and sample average approximations (detailed in Section 4).

Let $\mathcal{F}_t$ denote the set of flights scheduled to depart in period $t = 1, \cdots, T$. Let $\widehat{\mathcal{K}}^t_{fa}$ be the set of copies in the time-space network $\widehat{\mathcal{NW}}^t_a$ associated with flight $f \in \mathcal{F}_t$. Let ρ_k be the fuel cost associated with copy $k \in \widehat{\mathcal{K}}^t_{fa}$ (obtained from the flight plan). Let δ_k be its delay cost (obtained from its departure time). Let σ_k be its swap cost, incurred if aircraft a is different from the one that was originally planned to operate flight f. Let γ_f be the cost of cancelling flight $f \in \mathcal{F}_t$. Note that these cost parameters can capture non-linearities (e.g., non-linear costs of delays).

State variable: The state variable tracks the physical state of the airline's fleet and observed disruptions. The physical state can be represented by two vectors $\boldsymbol{\theta^t}$ and $\boldsymbol{l^t}$, each defined over $a \in \mathcal{A}$. For each aircraft $a \in \mathcal{A}$, θ^t_a and l^t_a denote, respectively, its latest arrival time and its arrival airport. Note that θ^t_a can either correspond to a past time stamp (if aircraft a is on the ground at time t) or a future one (if aircraft a is in the air at time t). Observed disruptions are represented by a vector $\boldsymbol{D^t}$ defined over $f \in \mathcal{F}_t$, where D^t_f denotes the departure delay of flight $f \in \mathcal{F}_t$ observed at time t. The state variable, denoted by $\boldsymbol{R^t}$, is thus given by:

$$\boldsymbol{R^t} = \left(\boldsymbol{\theta^t}, \boldsymbol{l^t}, \boldsymbol{D^t}\right). \tag{6}$$

The vector $\boldsymbol{R^t}$ is used to construct the time-space networks $\widehat{\mathcal{NW}}^t_a$ for all $a \in \mathcal{A}$.

Decision variables: All recovery and flight planning decisions are captured by the set of copies selected across all time-space networks $\widehat{\mathcal{NW}}^t_a$ for $a \in \mathcal{A}$. We capture them with two decision vectors $\boldsymbol{\hat{x}^t}$ and $\boldsymbol{\hat{z}^t}$, where $\boldsymbol{\hat{x}^t}$ is defined over $a \in \mathcal{A}$ and $k \in \cup_{f \in \mathcal{F}_t} \widehat{\mathcal{K}}^t_{fa}$ and $\boldsymbol{\hat{z}^t}$ is defined over $f \in \mathcal{F}_t$. Specifically,

$\hat{x}^t_{ka}$ is equal to 1 if copy k is selected and flown by aircraft a, and 0 otherwise; and $\hat{z}^t_f$ is equal to 1 if flight f is cancelled, and 0 otherwise. Our decision variable, denoted by $\boldsymbol{U^t}$ given by:

$$\boldsymbol{U^t} = \left(\hat{\boldsymbol{x}}^{\boldsymbol{t}}, \hat{\boldsymbol{z}}^{\boldsymbol{t}}\right). \tag{7}$$

Recovery and flight planning decisions are subject to a set of constraints (detailed in Section 4). We denote here the decision space by $\boldsymbol{\mathcal{U}^t}$.

Objective function. Our cost function, denoted by $C_t(\boldsymbol{R^t}, \boldsymbol{U^t})$, is defined as the total cost of recovery across all flights $f \in \mathcal{F}_t$, including fuel, delay, swap and cancellation costs. It is given by:

$$C_t(\boldsymbol{R^t}, \boldsymbol{U^t}) = \sum_{a\in\mathcal{A}} \sum_{f\in\mathcal{F}_t} \sum_{k\in\widehat{\mathcal{K}}^t_{fa}} \left(\rho_k + \delta_k + \sigma_k\right) \hat{x}^t_{ka} + \sum_{f\in\mathcal{F}_t} \gamma_f \hat{z}^t_f. \tag{8}$$

Transition function: The transition function describes the recovery process and the dynamic realization of disruptions between t and $t+1$. It can be represented by a function f_t as follows:

$$\boldsymbol{R^{t+1}} = f_t\left(\boldsymbol{R^t}, \boldsymbol{U^t}\right). \tag{9}$$

The recovery process updates the arrival airport and arrival time of each aircraft a. For example, if a flight is operated by aircraft a from airport K to airport L, then l^t_a is updated to airport L and θ^t_a is updated to its planned arrival time at airport L. Conversely, if an aircraft is not assigned to any departing flight at time t, then its availability remains unchanged. Specifically, we have:

$$(\theta^{t+1}_a, l^{t+1}_a) = \begin{cases} (\theta^t_a, l^t_a), & \text{if } \hat{x}^t_{ka} = 0, \text{ for all } k \in \cup_{f\in\mathcal{F}_t} \widehat{\mathcal{K}}^t_{fa}, \\ (\overline{\theta}_k, \overline{l}_k), & \text{if } \hat{x}^t_{ka} = 1, \text{ for some } k \in \cup_{f\in\mathcal{F}_t} \widehat{\mathcal{K}}^t_{fa}, \end{cases}$$

where $\overline{\theta}^k_f$ and $\overline{l}^k_f$ denote the time and location of arrival of flight copy k, respectively.

Realized disruptions are written as the sum of propagated, systemic and contingent disruptions, denoted respectively by $\overline{\boldsymbol{PD}}^{\boldsymbol{t+1}}$, $\overline{\boldsymbol{SD}}^{\boldsymbol{t+1}}$ and $\overline{\boldsymbol{CD}}^{\boldsymbol{t+1}}$. Specifically, we have:

$$\boldsymbol{D^{t+1}} = \overline{\boldsymbol{PD}}^{\boldsymbol{t+1}} + \overline{\boldsymbol{SD}}^{\boldsymbol{t+1}} + \overline{\boldsymbol{CD}}^{\boldsymbol{t+1}}, \text{ with: } \begin{cases} \overline{\boldsymbol{PD}}^{\boldsymbol{t+1}} \text{ realized from disruption decisions,} \\ \overline{\boldsymbol{SD}}^{\boldsymbol{t+1}} \text{ realized from } \xi \text{ (Section 3.1),} \\ \overline{\boldsymbol{CD}}^{\boldsymbol{t+1}} \text{ realized from an unknown distribution.} \end{cases} \tag{10}$$

Bellman Equation: Let $J_t(\boldsymbol{R^t})$ be the optimal cost-to-go in period t. The terminal cost in period $T+1$ is 0. Therefore, the Bellman equation is given as follows, where the expectation of the future cost-to-go is taken over the probability distribution ξ of future systemic disruptions (Section 3.1):

$$J_t(\boldsymbol{R^t}) = \min_{\boldsymbol{U^t}\in\boldsymbol{\mathcal{U}^t}} \left\{ \sum_{a\in\mathcal{A}} \sum_{f\in\mathcal{F}_t} \sum_{k\in\widehat{\mathcal{K}}^t_{fa}} \left(\rho_k + \delta_k + \sigma_k\right) \hat{x}^t_{ka} + \sum_{f\in\mathcal{F}_t} \gamma_f \hat{z}^t_f + E_\xi\left[J_{t+1}\left(f_t\left(\boldsymbol{R^t}, \boldsymbol{U^t}\right) | \boldsymbol{R^t}\right)\right] \right\}, \tag{11}$$

$$J_{T+1}(\boldsymbol{R^{T+1}}) = 0. \tag{12}$$

Unfortunately, this Bellman equation is too complex to be solved exactly by backward induction. Indeed, the state space grows exponentially as a function of the number of aircraft and the sample space of systemic disruptions (itself exponentially large), and the decision space grows exponentially as a function of the number of flights, the number of recovery options and the number of flight plans. The dynamic program can thus quickly become computationally intractable for real-sized instances—a well-known "curse of dimensionality" (Powell 2007).

4. Solution Approach and SRPDM Formulation

We propose a solution procedure based on *look-ahead approximation* and *sample average approximation* to solve the decision-making problem from Section 3. Look-ahead approximation involves estimating the cost-to-go function using a restricted time window, rather than the full horizon (Powell 2007, Bertsekas 2005, 2012). Sample average approximation involves estimating the cost-to-go function using a sampled set of scenarios of future systemic disruptions, rather than their full probability distribution (Kleywegt et al. 2002). At each decision point, we apply our Stochastic Reactive and Proactive Disruption Management (SRPDM) model, formulated as a stochastic integer program in Section 4.1. SRPDM optimizes recovery decisions for the look-ahead horizon based on the sampled disruption scenarios. Similar solution approaches have been applied to such problems as vehicle routing (Secomandi 2001) and job shop scheduling (Meloni et al. 2004).

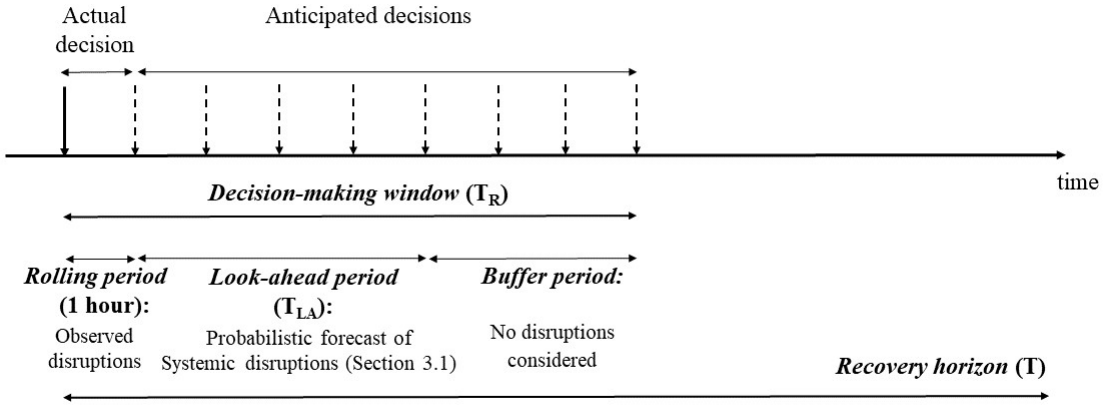


Figure 4 **Look-ahead framework for dynamic disruption management.**

Our look-ahead procedure is shown in Figure 4, with decision points indexed by $t = 1, \cdots, T$. The interval between periods t and $t+1$ is referred to as the *rolling period*. At each decision point, we observe realized disruptions for the rolling period and derive scenarios of systemic disruptions for a given *look-ahead period*. SRPDM is applied to derive disruption recovery and flight planning decisions for the rolling and look-ahead periods. We add a *buffer period* to handle cases in which

the disruptions are too large to accommodate all flights within the look-ahead period. The buffer period also ensures consistency of near-term operating decisions with flight schedules beyond the look-ahead window. We denote by T_R the length of the decision-making window and by $T_{LA}(<T_R)$ the length of the look-ahead period. The objective at time t is given by:

$$\min_{\substack{\boldsymbol{U^t}\in\boldsymbol{\mathcal{U}^t}\\ \cdots \\ \boldsymbol{U^{t+T_R}}\in\boldsymbol{\mathcal{U}^{t+T_R}}}} \sum_{\tau=t}^{t+T_R} E_\xi \left(\sum_{a\in\mathcal{A}} \sum_{f\in\mathcal{F}_\tau} \sum_{k\in\widehat{\mathcal{K}}^\tau_{fa}} (\rho_k+\delta_k+\sigma_k)\,\hat{x}^\tau_{ka} + \sum_{f\in\mathcal{F}_\tau} \gamma_f \hat{z}^\tau_f \right) \tag{13}$$

We now use our sample average approximation to approximate Equation (13). As mentioned in Section 3.1, we proceed by Monte Carlo sampling to approximate the expectation operator, using the set $\mathcal{Q}$ of disruption scenarios. As detailed in the next section, we define a time-space network for each aircraft $a\in\mathcal{A}$ in each scenario q, and denote by $\mathcal{K}^t_{faq}$ be the set of copies in the corresponding time-space network associated with flight $f\in\mathcal{F}_t$. We also introduce variables x^t_{kaq} and z^t_{fq} as the counterparts of $\hat{x}^t_{ka}$ and $\hat{z}^t_f$ in scenario q. These decision variables are scenario-dependent, thus capturing the flexibility of adapting future recovery and flight planning decisions as a function of the scenario realization. We also define non-anticipativity constraints in Section 4.1 to ensure the consistency of near-term decisions across all scenarios. The objective function becomes:

$$\min_{\substack{\boldsymbol{U^t}\in\boldsymbol{\mathcal{U}^t}\\ \cdots \\ \boldsymbol{U^{t+T_R}}\in\boldsymbol{\mathcal{U}^{t+T_R}}}} \sum_{\tau=t}^{t+T_R} \frac{1}{Q} \sum_{q\in\mathcal{Q}} \left(\sum_{a\in\mathcal{A}} \sum_{f\in\mathcal{F}_\tau} \sum_{k\in\mathcal{K}^\tau_{faq}} (\rho_k+\delta_k+\sigma_k)\,x^t_{kaq} + \sum_{f\in\mathcal{F}_\tau} \gamma_f z^t_{fq} \right) \tag{14}$$

We now use Equation (14) to develop the mathematical formulation of our SRPDM. We calibrate our look-ahead and sample average approximations in Section 5.4. Like any approximation scheme, our algorithm induces a trade-off between speed and solution quality. In our implementation, we strive to obtain solutions within a few minutes of computation—consistently with practical requirements and with earlier studies in airline recovery (see, e.g., Petersen et al. 2012, Maher 2016, Marla et al. 2017). Note, moreover, that our solution algorithm does not provide solution quality guarantees. We shall thus compare in Section 6 its solution to that of a myopic baseline (defined in Section 4.3)—underscoring the benefits of our modeling approach and approximate algorithm as compared to existing recovery approaches that do not leverage forecasts of future disruptions.

4.1. Stochastic Reactive and Proactive Disruption Management (SRPDM)

SRPDM builds upon the model from Marla et al. (2017), but extends it to capture partial and probabilistic forecasts of future disruptions. SRPDM optimizes disruption recovery decisions (i.e., departure holds, aircraft swaps, flight cancellations) and flight planning decisions (i.e., flying altitude, speed and route) to minimize *expected* recovery costs across all disruption scenarios (Equation (14)). It is formulated as a stochastic integer program, based on non-anticipativity constraints for first-stage decision variables and scenario-dependent constraints for subsequent periods.

Sets

$\mathcal{F}_\tau$: Set of flights scheduled to depart in period $\tau = t, \cdots, t+T_R$

$\mathcal{A}$: Set of available aircraft

$\mathcal{Q}$: Set of sampled disruption scenarios

$\mathcal{NW}^t_{aq}$: Time-space network corresponding to aircraft a in scenario q at time t

$\mathcal{K}^t_{faq}$: Set of copies of flight $f \in \mathcal{F}^R_t$ in network $\mathcal{NW}^t_{aq}$ from aircraft a in scenario q at time t

$\mathcal{G}^t_{aq}$: Set of ground arcs connecting pairs of nodes in $\mathcal{NW}^t_{aq}$

$\mathcal{N}^t_{aq}$: Set of nodes in $\mathcal{NW}^t_{aq}$

$\mathcal{I}^t_{naq}$: Set of incoming arcs to node $n \in \mathcal{N}^t_{aq}$ in $\mathcal{NW}^t_{aq}$

$\mathcal{O}^t_{naq}$: Set of outgoing arcs to node $n \in \mathcal{N}^t_{aq}$ in $\mathcal{NW}^t_{aq}$

We define here a time-space network $\mathcal{NW}^t_{aq}$ in each scenario $q \in \mathcal{Q}$ for each aircraft $a \in \mathcal{A}$. The earlier probabilistic network representation $\widehat{\mathcal{NW}}^t_a$ is equivalent to the collection $\left(\mathcal{NW}^t_{aq}\right)_{q\in\mathcal{Q}}$. The same observation applies to $\mathcal{K}^t_{faq}$. By construction, the time-space networks coincide for all flights in the rolling period across scenarios, i.e., $\mathcal{K}^t_{f,a,q_1} = \mathcal{K}^t_{f,a,q_2}$ for all $f \in \mathcal{F}_t$, $a \in \mathcal{A}$ and $q_1, q_2 \in \mathcal{Q}$. However, the networks may differ for the flights scheduled in the look-ahead and buffer periods to reflect the various operating conditions across disruption scenarios.

Parameters

δ_k: Delay cost associated with copy $k \in \mathcal{K}^t_{faq}$, over all $f \in \cup_{\tau=1,\cdots,T_R}\mathcal{F}_\tau, a \in \mathcal{A}, q \in \mathcal{Q}$

ρ_k: Fuel cost associated with copy $k \in \mathcal{K}^t_{faq}$, over all $f \in \cup_{\tau=1,\cdots,T_R}\mathcal{F}_\tau, a \in \mathcal{A}, q \in \mathcal{Q}$

σ_k: Aircraft swap cost associated with copy $k \in \mathcal{K}^t_{faq}$, over all $f \in \cup_{\tau=1,\cdots,T_R}\mathcal{F}_\tau, a \in \mathcal{A}, q \in \mathcal{Q}$

γ_f: Cost of cancellation of flight $f \in \cup_{\tau=1,\cdots,T_R}\mathcal{F}_\tau$

$$s^t_{naq} = \begin{cases} 1 & \text{if aircraft } a \in \mathcal{A} \text{ starts in } \mathcal{NW}^t_{aq} \text{ at node } n \\ -1 & \text{if aircraft } a \in \mathcal{A} \text{ ends in } \mathcal{NW}^t_{aq} \text{ at node } n \\ 0 & \text{otherwise} \end{cases}$$

The swap cost σ_k depends only on which aircraft is used to operate copy k. Specifically, we have $\sigma_{k_1} = \sigma_{k_2}$ for $k_1, k_2 \in \mathcal{NW}^t_{aq}$ for all a and q; and $\sigma_k = 0$ if $k \in \mathcal{NW}^t_{a_0q}$ for all $q \in \mathcal{Q}$, if the flight was originally planned to be operated by a_0. Moreover, $s^t_{naq} = 1$ (or -1) indicates that node n is the supply (or demand) node for aircraft a in $\mathcal{NW}^t_{aq}$, and $s^t_{naq} = 0$ means node n is an intermediate airport location. These parameters will be used to formulate the flow balance constraints.

Decision Variables

$$x^t_{kaq} = \begin{cases} 1 & \text{if copy } k \in \cup_{\tau=1,\cdots,T_R} \cup_{f\in\mathcal{F}_\tau} \mathcal{K}^t_{faq} \text{ is selected with aircraft } a \text{ in scenario } q \text{ at time } t \\ 0 & \text{otherwise} \end{cases}$$

$$y^t_{gaq} = \begin{cases} 1 & \text{if ground arc } g \text{ in } \mathcal{NW}^t_{aq} \text{ of aircraft } a \text{ is selected in scenario } q \text{ at time } t \\ 0 & \text{otherwise} \end{cases}$$

$$z^t_{fq} = \begin{cases} 1 & \text{if } f \in \cup_{\tau=1,\cdots,T_R}\mathcal{F}_\tau \text{ is cancelled in scenario } q \text{ at time } t \\ 0 & \text{otherwise} \end{cases}$$

The decision variables include first-stage variables, which are determined at time t before any scenario is realized, and scenario-dependent variables throughout the look-ahead and buffer periods.

This structure enables different decisions to be made across scenarios in periods $t+1,\cdots,t+T_R$. But decisions made for the rolling period (period t) will be subject to non-anticipativity constraints. Only those decisions in the rolling period are to be executed at time t.

Formulation

$$\min_{x,y,z} \quad \sum_{\tau=t}^{t+T_R} \frac{1}{|\mathcal{Q}|} \sum_{q\in\mathcal{Q}} \left(\sum_{a\in\mathcal{A}} \sum_{f\in\mathcal{F}_\tau} \sum_{k\in\mathcal{K}^t_{faq}} (\rho_k+\delta_k+\sigma_k) x^t_{kaq} + \sum_{f\in\mathcal{F}_\tau} \gamma_f z^t_{fq} \right) \tag{15}$$

$$\text{s.t.} \quad x^t_{k,a,q_1} = x^t_{k,a,q_2} \qquad \forall k\in\mathcal{K}^t_{f,a,q_1}, \forall f\in\mathcal{F}_t, \forall q_1,q_2\in\mathcal{Q}, \forall a\in\mathcal{A} \tag{16}$$

$$z^t_{f,q_1} = z^t_{f,q_2} \qquad \forall f\in\mathcal{F}_t, \forall q_1,q_2\in\mathcal{Q} \tag{17}$$

$$\sum_{a\in\mathcal{A}} \sum_{k\in\mathcal{K}^\tau_{faq}} x^t_{kaq} + z^t_{fq} = 1 \qquad \forall f\in\mathcal{F}_t\cup\cdots\cup\mathcal{F}_{t+T_R}, \forall q\in\mathcal{Q} \tag{18}$$

$$\sum_{g\in\mathcal{I}^t_{naq}\cap\mathcal{G}^t_{aq}} y^t_{gaq} + \sum_{k\in\mathcal{I}^t_{naq}\setminus\mathcal{G}^t_{aq}} x^t_{kaq} + s^\tau_{naq} = \sum_{g\in\mathcal{O}^t_{naq}\cap\mathcal{G}^t_{aq}} y^t_{gaq} + \sum_{k\in\mathcal{O}^t_{naq}\setminus\mathcal{G}^t_{aq}} x^t_{kaq} \qquad \forall n\in\mathcal{N}^t_{aq}, \forall a\in\mathcal{A}, \forall q\in\mathcal{Q} \tag{19}$$

$$x^t_{kaq}\in\{0,1\} \qquad \forall k\in\mathcal{K}^t_{faq}, \forall f\in\mathcal{F}_\tau, \forall a\in\mathcal{A}, \forall q\in\mathcal{Q}, \forall\tau=t,\cdots,t+T_R \tag{20}$$

$$y^t_{gaq}\in\{0,1\} \qquad \forall g\in\mathcal{G}^t_{aq}, \forall a\in\mathcal{A}, \forall q\in\mathcal{Q} \tag{21}$$

$$z^t_{fq}\in\{0,1\} \qquad \forall f\in\mathcal{F}_\tau, \forall q\in\mathcal{Q}, \forall\tau=t,\cdots,t+T_R \tag{22}$$

The objective function (15) minimizes expected recovery costs, averaged across all $|\mathcal{Q}|$ sampled scenarios. Constraints (16) and (17) are non-anticipativity constraints that ensure that first-stage decisions in the rolling period are identical across all scenarios. Constraints (18) ensure that a copy of each flight is selected or the flight is cancelled. Constraints (19) maintain flow conservation: if an aircraft is incoming to a node, it must also be outgoing from that node—except at the source and destination, which have an outgoing and an incoming aircraft, respectively. This formulation also ensures that each aircraft reaches its final destination by the end of the day—in high-disruption instances where this would lead to infeasibility, this formulation could be easily modified by imposing aggregate constraints ensuring, for instance, that a minimal number of aircraft would end their routes at each given airport. Constraints (20)–(22) define the domains of all variables.

4.2. Rolling Algorithm

We synthesize our dynamic solution procedure in Algorithm 1. The algorithm iterates over the recovery horizon $\{1,\cdots,T\}$. At each time period t, it generates the time-space networks $\mathcal{NW}^t_{aq}$, and solves SRPDM. From one period to the next, our state variable $\boldsymbol{R^t}=(\boldsymbol{\theta^t},\boldsymbol{l^t},\boldsymbol{D^t})$ is updated based on prior recovery decisions and revealed disruptions. We discuss below the creation of the time-space networks and other steps ensuring the feasibility and practicality of the solution.

Algorithm 1 Solution algorithm.

get R_1;
for each $t \in \{1, 2, \dots T\}$ **do**
  get $\mathcal{F}_t, \cdots, \mathcal{F}_{t+T_R}$
  maxWindow = initialHoldingWindow;
  holdingInterval = 10 minutes;
  set feasibleSolution = false;
  while feasibleSolution = false *and* maxWindow$\leq$ maxWindowLimit **do**
    for each $a \in \mathcal{A}$, each $q \in \mathcal{Q}$ **do**
      Generate $\mathcal{NW}^t_{aq}$; ▷ See Figure 5
    end for
    Solve *SRPDM*, (see Equations (15) to (22)) ▷ See Section 4.1
    if SRPDM feasible and less than Γ cancellations **then**
      feasibleSolution = true;
      for each $a \in \mathcal{A}$, each $f \in \mathcal{F}_t$, each $q \in \mathcal{Q}$, each $k \in \widehat{\mathcal{K}}^t_{faq}$ **do**
        if $x^t_{kaq} \forall q \in \mathcal{Q} \equiv x^t_{ka} = 1$ **then**
          Update $\boldsymbol{R^t}$ with new location and time of all aircraft $a \in \mathcal{A}$;
        end if
      end for
    else
      maxWindow=maxWindow+δ
    end if
  end while
  if feasibleSolution = false **then**
    $\tilde{A}$ = set of aircraft that cause infeasibility
    maxWindow = initialHoldingWindow;
    for each $a \in \mathcal{A} \setminus \tilde{A}$, each $q \in \mathcal{Q}$ **do**
      Generate $\mathcal{NW}^t_{aq}$; ▷ See Figure 5
    end for
    Solve *SRPDM* ▷ See Section 4.1
    for each $a \in \mathcal{A}$, each $f \in \mathcal{F}_t$, each $q \in \mathcal{Q}$, each $k \in \widehat{\mathcal{K}}^t_{faq}$ **do**
      if $x^t_{ka} = 1$ **then**
        Update $\boldsymbol{R^t}$ with new location and time of aircraft a;
      end if
    end for
  end if
end for

Time-space network generation. The process underlying the creation of the time-space networks $\mathcal{NW}^t_{aq}$ is shown in Figure 5. It starts by reading the flight schedule, the flight planning options, the latest availability of each aircraft (which depend on its past flight assignments, past disruptions,

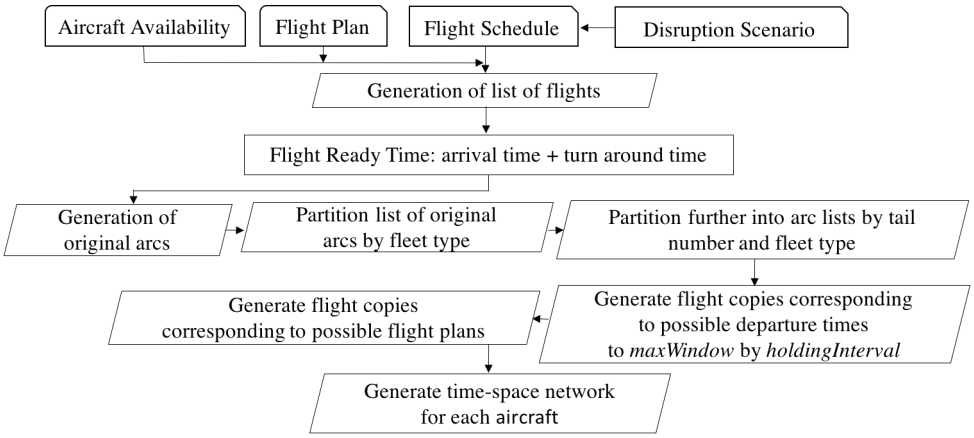


Figure 5 Process used to create time-space networks $\mathcal{NW}^t_{aq}$.

and turnaround times), and the disruption scenario considered. This information is used to generate a set of flight copies that can be operated by the aircraft in the decision-making window. For each flight, the first copy departs at the flight's scheduled departure time shifted by its delay observed at time t, or at the time when the aircraft becomes available—whichever comes later. The additional copies correspond to added departure delays and/or alternative flight plans.

Generating flight copies requires assumptions on the granularity and scope of the time-space network. Granularity refers to the interval between consecutive copies, named *holding interval*. The smaller the holding interval, the larger the decision space but the better the solution. We use a holding interval of 10 minutes. Scope refers to the largest allowed departure hold, named *maximum holding window*. For instance, a maximum holding window of 1 hour implies that at most 7 copies of each flight can be created for a given flight plan (associated with holds of $0, 10, \cdots, 60$ minutes). This information is used to generate the flight arcs and ground arcs across the network.

Ensuring global feasibility. Under large disruptions, the flow balance constraints in SRPDM may lead to infeasibility or result in inadequately large numbers of flight cancellations, due to the maximum holding window. This happens when no aircraft is available until the end of the allowed holding window for a given flight, thereby violating the flow balance constraint at the supply node. In practice, airlines need to balance objectives of minimizing the largest flight delays (captured by the maximum holding window) and of minimizing the number of cancellations. For this reason, we impose an upper bound Γ to the number of cancellations, and increase the maximum holding window iteratively until a solution with fewer than Γ cancellations is obtained. In our computations, we use a value Γ equal to 5% of the total number of flights in SRPDM at each time t.

We initialize the maximum holding window value and, if a feasible solution that cancels fewer than Γ flights is found, we update the state variable and proceed to the next period. If the problem

is infeasible or results in more than Γ cancellations, we increase the maximum holding window by increments of δ. As soon as a feasible solution that cancels fewer than Γ flights is found, we update the state variable and proceed to the next period. If no such solution is found after the maximum holding window reaches a pre-specified upper bound, we remove the restriction on the number of cancellations and re-solve the model. We report these parameters in Section 5.4.

4.3. Myopic Baseline

Before proceeding to the computational implementation of the modeling and computational framework developed in this paper, we outline the baseline approach used as a benchmark. Specifically, we consider a baseline that optimizes recovery decisions myopically, without considering future disruptions. This follows the approach from Marla et al. (2017). Under this approach, the decision-maker observes disruptions at time t, and optimizes recovery decisions over the full planning horizon. It does capture propagated disruptions, but ignores the creation of future primary (systemic or contingent) disruptions. Since the baseline approach is less computationally complex than our stochastic optimization approach, it can be solved as a single integer program from period t up to the terminal period T at each decision point. We still implement it on a rolling basis to capture the realizations of (systemic and contingent) disruptions—observed at each period.

This myopic baseline relies on a single time-space network for each aircraft, captured by sets $\tilde{\mathcal{K}}^t_{fa}$—analogous to $\mathcal{K}^t_{faq}$. Similarly, the decision variables are written as $\tilde{x}^t_{ka}$ and $\tilde{z}^t_f$, defined for all flights $f \in \mathcal{F}_t \cup \cdots \cup \mathcal{F}_T$ (or their copies). The variables $\tilde{x}$ and $\tilde{z}$ are subject to similar constraints as Equations (16)–(22). The objective function becomes:

$$\min_{x,z} \quad \sum_{\tau=t}^{T} \left(\sum_{a \in \mathcal{A}} \sum_{f \in \mathcal{F}_\tau} \sum_{k \in \tilde{\mathcal{K}}^\tau_{fa}} (\rho_k + \delta_k + \sigma_k) \tilde{x}^t_{ka} + \sum_{f \in \mathcal{F}_\tau} \gamma_f \tilde{z}^t_f \right). \tag{23}$$

5. Experimental Setup

We now implement our approach computationally to quantify the benefits of SRPDM, as compared to the myopic baseline. All models are implemented in the Java programming language interfaced with IBM ILOG CPLEX 12.6.1 on a workstation running at 1.8 GHz with 80 GB RAM.

5.1. Network Description

We consider the network of flights of Delta Air Lines, a major US hub-and-spoke airline. This choice is arbitrary, and does not reflect Delta's operating practices. While our model is expected to bring stronger benefits for hub-and-spoke carriers, it can be applied to any airline network. Delta Air Lines leverages six airports as hubs of operations (New York's LaGuardia (LGA) and John F. Kennedy (JFK), Atlanta (ALT), Detroit (DTW), Minneapolis-Saint Paul (MSP), and

Salt Lake City (SLC)). We obtain flight schedules and fleet assignments from the Aviation System Performance Metrics (ASPM) database maintained by the Federal Aviation Administration (FAA). We consider 3 fleet types: Airbus 319 (A319), Airbus 320 (A320), and Boeing 757-200 (B752).

Table 1 Summary statistics for each network.

	July 15			July 16			July 17			July 18		
Metric	A319	A320	B752	A319	A320	B752	A319	A320	B752	A319	A320	B752
# flights	256	260	306	266	247	312	279	254	311	276	262	314
percentile (2014)	33	80	40	80	30	60	100	50	60	95	80	65
# arr. – JFK	2	3	20	4	3	21	4	4	4	4	3	24
# arr. – LGA	19	9	1	20	9	2	23	10	4	23	8	2
# arr. – ATL	24	25	122	25	21	121	26	23	91	24	21	128
# arr. – MSP	26	38	27	27	37	27	28	38	24	28	34	29
# arr. – SLC	19	41	21	20	39	22	22	39	21	22	38	22
# arr. – DTW	36	25	16	38	22	18	37	22	16	37	22	16
Total	126	141	207	134	131	211	140	136	160	138	126	221
# dep. – JFK	2	3	20	4	3	24	4	4	7	4	3	24
# dep. – LGA	20	9	1	20	9	2	23	10	4	22	10	2
# dep. – ATL	24	25	117	26	21	123	26	22	89	24	22	124
# dep. – MSP	26	38	25	27	37	31	28	38	24	28	38	27
# dep. – SLC	19	40	22	21	39	21	22	38	20	22	38	22
# dep. – DTW	37	24	16	37	24	18	33	21	16	38	21	17
Total	128	139	204	135	133	219	136	137	160	137	132	216
less than 1 hour	9	2	0	10	0	0	9	0	0	9	0	0
1–2 hours	157	100	105	160	93	104	165	89	103	169	97	98
2–3 hours	54	72	75	57	64	74	63	65	71	57	71	78
3–4 hours	34	56	38	35	58	45	40	63	41	39	57	43
4–5 hours	2	25	42	4	27	42	2	28	48	2	29	49
5–6 hours	0	5	32	0	4	33	0	7	32	0	6	30
6+ hours	0	0	14	0	0	14	0	0	16	0	0	16
avg. flight time (min)	119	153	184	120	155	186	120	161	189	119	157	189

We consider the schedule of flights on four weekdays in July 2014. Table 1 reports characteristics of each fleet's network, and the corresponding percentile of the distribution of the number of daily flights from Delta Air Lines in 2014. Note, first, that our experimental setup captures the variability in Delta Air Lines' schedules: the four days under consideration range from the blue33$^{\text{rd}}$ percentile to the 100$^{\text{th}}$ percentile for the A319 fleet, from the 30$^{\text{th}}$ percentile to the 80$^{\text{th}}$ percentile for the A320 fleet, and from the 40$^{\text{th}}$ percentile to the 65$^{\text{th}}$ percentile for the B752 fleet. In terms of spatial concentration, the airline's network is tightly connected to the six hub airports—with around 50% of the arrivals and departures operated to or from a hub. Finally, the A319, A320 and B752 fleets cover increasingly long flights on average—thus offering different recovery opportunities.

5.2. Stochastic Model Inputs

We generate disruption scenarios at each of the six hubs (see Section 3.1). We use capacity data from Simaiakis (2012) and the Federal Aviation Administration (2004). We capture weather variations

by means of a two-state Markov chain with "Visual Meteorological Conditions" and "Instrumental Meteorological Conditions" states (used as proxies of "good" and "poor" weather, respectively). We estimate the transition probabilities from historical data.

Figure 6 shows the expected departure delays at each of the six hub airports, for each 15-minute period of the day. As expected, congestion levels exhibit significant variability from one airport to another, due to differences in underlying scheduling and capacity patterns. JFK and LGA are the most congested airports and SLC is the least congested one, with ATL, DTW and MSP lying in-between. Moreover, delay patterns also vary across airports. Some airports (e.g., DTW, SLC and MSP) operate a strongly "peaked" schedule, with alternating arrival and departure banks—resulting in a sequence of periods with high delays and periods with low delays (see Figure 8 in Appendix A). At other airports (e.g., LGA), the schedule of flights is relatively evenly distributed over the day, resulting in more steady congestion levels throughout the day. The other two airports (JFK and ATL) lie somewhat in-between: airport queues are less variable but still exhibit peaks and valleys. These different patterns underscore the potential value of capturing the dynamics of formation and propagation of delays by means of our queuing model for disruption recovery.

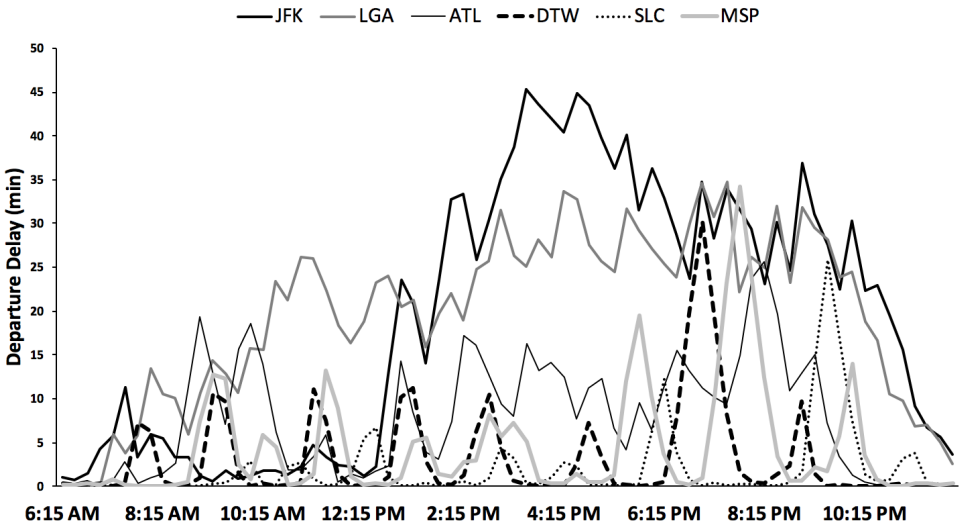


Figure 6 **Expected departure delay at the six hubs.**

5.3. Generation of Disruption Instances

In each period $t = 1, \cdots, T$, we generate disruption instances (i.e., the departure delay of each flight $i \in \mathcal{F}_t$). We aim here to replicate the dynamics of the system (captured by the variable $\boldsymbol{D^t}$) and assess the performance of our modeling and computational framework against the myopic baseline. For unbiased comparisons, we do *not* sample disruptions among the set of systemic disruption scenarios considered in our stochastic optimization framework, but instead we generate disruption

realizations from real-world data. This procedure captures *all* disruptions: we optimize recovery decisions based on delay propagation dynamics and probabilistic forecasts of systemic disruptions, but assess the resulting decisions against all (propagated, systemic *and* contingent) disruptions.

We use the departure delays from the Bureau of Transportation Statistics (BTS) database. But these delays result from combined propagated, systemic and contingent disruptions, so using them directly would result in double-counting propagated disruptions (which would carry over from previous time periods as well as appear in the newly generated disruptions). We thus need to infer the "new" (systemic and contingent) disruptions by subtracting propagated disruptions.

To this end, we use the inference method from Lan et al. (2006). We first sort the sequence of flight legs operated by each aircraft. We assume a minimum turnaround time of 30, 35 and 40 minutes for A319, A320 and B752 aircraft, respectively. For every pair of consecutive flights i and j operated by the same aircraft, we define the *slack* between i and j as the difference between the planned and minimum turnaround times between the two flights. The propagated delay of flight j is then computed as max(Arrival Delay of Flight i − Slack, 0). By subtracting this propagated delay from total delays, we obtain the newly created delays (from systemic and contingent disruptions).

Finally, for interpretability, we classify all disruption instances into "small", "medium" and "large" disruptions. The procedure is detailed in Appendix B.

5.4. Additional Settings and Parameters

In Algorithm 1, we use a rolling period of 1 hour and a decision-making window T_R of 7 hours. At each time period t, disruptions are observed for the one-hour rolling period. Systemic disruptions are forecasted for a look-ahead period T_{LA}, ranging from 0 to 4 hours. The remainder of the decision-making window is the buffer period, for which no new disruptions are considered.

Note that the hourly time discretization only plays a minor role computationally. As we shall see, the main driver of the problem's complexity is the look-ahead period T_{LA}, which impacts the set of flights $\mathcal{F}_t$ and the resulting number of scenario-dependent decision variables. The one-hour rolling period (which impacts the number of scenario-agnostic decision variables) restricts the number of iterations required to simulate the recovery dynamics over the full day of operations. In practice, however, an airline could implement the proposed model on a more frequent basis.

As described in Section 4.2, the departure holding window is set to an initial value and extended, in case of infeasibility, up to a maximum possible holding window. Table 2 reports the initial and maximum departure holding windows for each fleet type, used to balance computational runtimes and solution quality. This results in longer holding windows for B752, intermediate ones for A320, and shorter ones for A319—similar to the process applied at airlines' Operations Control Centers.

Aircraft swap costs (σ_k) are set to \$500. Cancellation costs (γ_f) are estimated as the cost of re-accommodating passengers on the next available flight—assuming 129, 155 and 291 seats for A319,

Table 2 Initial and maximum holding window by fleet type.

Fleet	Initial Holding Window	Maximum Holding Window
A319	30–40 minutes	90–360 minutes
A320	30–60 minutes	120–480 minutes
B752	50–110 minutes	400–550 minutes

A320 and B752 aircraft respectively, a load factor of 85%, and a cost of \$37.5 per hour of passenger delays (Cook and Tanner 2008b). Fuel costs (ρ_k) are set to \$0.53–\$0.73 per lb (International Air Transport Association 2010). We use a baseline value for the flight delay cost (δ_k) of \$10 per minute. This captures the direct costs of delays *to the airline*. We conduct sensitivity analyses by varying the parameter δ_k from its baseline value of \$10 per minute to a maximum value of \$77 per minute—obtained by fully internalizing the cost of passenger inconvenience, calculated as the product of the number of passengers on the flight and a value of time of \$37.5 per hour. This setup considers linear delay costs, consistent with the airline recovery literature (Maher 2016, Marla et al. 2017). However, delay costs may have increasing penalties (Cook and Tanner 2008a, Ball et al. 2010). We thus also consider a non-linear delay cost function in Section 6.4.2 to capture this.

6. Computational Results

We now evaluate the performance of SRPDM, as compared to the myopic baseline from Section 4.3. Unless otherwise specified, we implement the model with data from July 17, 2014, using 30 scenarios capturing delays at all six hubs for a look-ahead window of $T_{LA} = 4$ hours. First, we show that the proposed approach significantly reduces expected recovery costs (Section 6.1). We then demonstrate that, to be beneficial, the proposed approach needs to capture disruption forecasts over the full-scale network of operations and over an extended time horizon (Section 6.2). In Section 6.3, we discuss the computational performance of our approximate solution approach, showing that SRPDM is solved in reasonable computational times that are consistent with practical requirements. Section 6.4 then shows the robustness of the model's benefits to the schedule of flights and to the specification of the objective function. We synthesize the main insights in Section 6.5.

Before proceeding further, Table 3 shows outputs in three disruption instances to highlight the three main recovery mechanisms employed by SRPDM—and the various trade-offs. Instance 1 illustrates the *speed change* mechanism: if subsequent flights are likely to be delayed, SRPDM may deliberately slow down earlier flights to reduce fuel burn without impacting connectivity—albeit at a higher delay cost. Conversely, SRPDM can speed up some aircraft to maintain connectivity or facilitate swaps. Instance 2 illustrates the *departure hold* mechanism: SRPDM can deliberately delay departing flights (in conjunction with speed ups) to preserve connectivity. Instance 3 illustrates the *aircraft swap avoidance* mechanism. While aircraft swaps may be myopically less costly

than departure holds, they can also lead to cancellations at later times; the SRDPM can avoid this situation by anticipating future disruptions earlier, resulting in lower cancellation and swap costs.

Table 3 Experimental results of SRPDM vs. baseline in three disruption instances.

Experiments	Model	Total # Cancellations	Total Fuel Burn (lb.)	Total # of Swaps	Total Dep. Delay (min)	Total Arr. Delay (min)	Cost Savings Per Day (%)
Instance 1	Baseline	2	**1,511,000**	0	**556**	1,418	–
	SRPDM	2	**1,508,025**	0	**586**	1,476	0.16%
Instance 2	Baseline	**14**	1,457,276	6	**714**	1,313	–
	SRPDM	**10**	1,457,276	6	**950**	1,433	6.16%
Instance 3	Baseline	**13**	1,424,447	**4**	1,312	1,320	–
	SRPDM	**11**	1,446,258	**2**	1,318	1,335	3.60%

Throughout this section, we report the total costs obtained with the baseline model vs. the SRPDM over the full day of operations. Unless otherwise specified, we consider stochastic disruptions at the six hubs for a look-ahead period T_{LA} of 4 hours, and consider 30 disruption scenarios.

6.1. SRPDM Benefits

Table 4 reports the results of SRPDM and the myopic baseline across all disruption instances and fleet types. Note that SRPDM yields significant improvements over the myopic baseline: SRPDM reduces expected recovery costs by 1.5%, 1.8% and 1.9% for the A319, A320 and B752 fleets, respectively (calculated as the relative difference between the baseline and SRPDM costs over all disruptions). These expected savings can result in large financial gains for major airlines, underscoring the benefits of anticipating future disruptions into disruption recovery—even with partial and probabilistic forecasts of future disruptions and even with an approximate solution algorithm.

Under small disruptions, the strongest benefits are derived for the B752 fleet and, to a smaller extent, for the A320 fleet. In contrast, SRPDM does not lower costs of recovery for the A319 network. This mainly stems from the fact that the A319 network primarily consists of short-haul flights, which limits flexibility in terms of flight speed changes. Moreover, the A319 network is significantly smaller than the A320 and B752 ones, and thus less sensitive to disruptions.

When it comes to medium and large disruptions, SRPDM reduces expected recovery costs for all fleet types—from 1% to 2.5%. Note, again, the variability across disruption instances: SRPDM yields the same costs as the myopic baseline in some instances but large cost reductions in other instances. Such variability underscores the importance of capturing stochastic airport delays in airline recovery in instances where future disruptions at hub airports can be significant.

These results also highlight that the performance improvements from SRPDM mainly stem from an increased number and magnitude of departure holds—reflected by larger departure delays than

Table 4 Comparison of SRPDM over myopic baseline, for each fleet type and each disruption category.

Disruption Category	Fleet	Model	Total Cost ($)	Dep. Delay (min)	Arr. Delay Delay (min)	Fuel Burn ($)	Num. Speed Change	Num. Cancel.	Num. Swaps	Savings
Small	A319	Baseline (avg.)	310,587	951	543	257,608	15	11	0.25	–
		Baseline (min.)	255,425	604	410	257,261	15	6	0	–
		Baseline (max.)	352,312	1,245	688	258,647	15	15	1	–
		SRPDM (avg.)	310,587	951	543	257,608	15	11	0.25	0%
		SRPDM (min.)	255,425	604	410	257,261	15	6	0	0%
		SRPDM (max.)	352,312	1,245	688	258,647	15	15	1	0%
	A320	Baseline (avg.)	1,169,750	1,015	1,540	1,502,723	72	3	0.6	–
		Baseline (min.)	1,117,555	574	1,314	1,457,276	70	0	0	–
		Baseline (max.)	1,279,491	1,783	1,925	1,523,659	74	10	6	–
		SRPDM (avg.)	1,149,427	1,073	1,578	1,503,788	72	2	0.6	1.7%
		SRPDM (min.)	1,117,555	574	1,314	1,457,276	70	0	0	0%
		SRPDM (max.)	1,279,491	1,783	1,925	1,523,659	74	10	6	6.2%
	B752	Baseline (avg.)	982,914	1,517	1,334	1,658,401	50	2.4	0.25	–
		Baseline (min.)	928,083	579	1,009	1,657,556	49	1	0	–
		Baseline (max.)	1,114,213	2,616	1,910	1,662,632	50	6	2	–
		SRPDM (avg.)	943,657	1,575	1,367	1,658,401	50	1.3	0.08	3.9%
		SRPDM (min.)	928,083	579	1,043	1,657,556	49	1	0	0%
		SRPDM (max.)	1,008,525	2,750	1,972	1,662,632	50	3	1	19.7%
Medium	A319	Baseline (avg.)	500,281	1,929	1,284	458,580	23.8	7.2	1.4	-
		Baseline (min.)	314,116	727	585	281,761	16	0	0	-
		Baseline (max.)	1,174,571	3,900	3,134	1,502,174	73	14	4	-
		SRPDM (avg.)	490,556	1,978	1,331	457,260	23.7	6.7	1.3	2.4%
		SRPDM (min.)	314,116	727	602	296,284	17	0	0	0%
		SRPDM (max.)	1,174,571	3,900	3,134	1,502,174	73	14	4	27.1%
	A320	Baseline (avg.)	1,211,508	1,880	1,909	1,492,283	72	5	0.8	-
		Baseline (min.)	1,124,625	850	1,404	1,256,011	60	0	0	-
		Baseline (max.)	1,499,232	3,998	2,944	1,523,437	74	28	5	-
		SRPDM (avg.)	1,185,928	1,942	1,939	1,495,583	72	3.7	0.47	2.2%
		SRPDM (min.)	1,124,625	850	1,404	1,256,011	60	0	0	0%
		SRPDM (max.)	1,499,232	3,998	2,944	1,523,437	74	28	5	16.3%
	B752	Baseline (avg.)	1,033,625	2,452	1,694	1,658,712	50	3.53	1	-
		Baseline (min.)	930,693	1,219	1,061	1,657,556	48	1	0	-
		Baseline (max.)	1,605,263	4,042	3,080	1,665,240	50	19	9	-
		SRPDM (avg.)	1,020,740	2,464	1,702	1,658,613	50	3.2	1	1.1%
		SRPDM (min.)	930,693	1,239	1,122	1,657,556	48	1	0	0%
		SRPDM (max.)	1,495,119	4,042	3,080	1,665,240	50	16	8	7.4%
Large	A319	Baseline (avg.)	536,473	2,718	1,165	366,123	19.8	11.9	2.2	-
		Baseline (min.)	338,946	1,796	691	306,153	17	2	0	-
		Baseline (max.)	859,662	4,526	2,155	377,477	21	30	7	-
		SRPDM (avg.)	527,945	2,771	1,179	366,123	19.8	11.4	2.1	2.0%
		SRPDM (min.)	338,946	1,796	691	306,153	17	2	0	0%
		SRPDM (max.)	859,662	4,526	2,155	306,153	21	30	7	9.5%
	A320	Baseline (avg.)	1,241,417	2,679	2,273	1,483,771	72	6.4	1.7	-
		Baseline (min.)	1,143,749	1,208	1,320	1,424,447	69	2	0	-
		Baseline (max.)	1,339,418	4,710	3,236	1,524,743	74	14	4	-
		SRPDM (avg.)	1,223,880	2,744	2,307	1,486,887	72	5.4	1.4	1.4%
		SRPDM (min.)	1,143,749	1,182	1,335	1,446,258	70	2	0	0%
		SRPDM (max.)	1,293,446	4,710	3,236	1,524,743	74	11	4	6.4%
	B752	Baseline (avg.)	1,203,155	5,026	2,590	1,597,024	48	8.3	4	-
		Baseline (min.)	946,623	1,610	942	1,491,546	45	1	0	-
		Baseline (max.)	1,688,686	1,610	4,966	1,657,556	50	20	13	-
		SRPDM (avg.)	1,191,148	5,278	2,708	1,590,140	47	8	4.4	1.0%
		SRPDM (min.)	946,623	1,556	931	1,491,546	45	1	0	0%
		SRPDM (max.)	1,659,628	13,419	5,818	1,657,556	50	19	15	5.6%

with the myopic baseline. These departure holds can stem from "new" disruptions in the current period and/or the propagation of earlier events (e.g., earlier disruptions and earlier departure holds). Ultimately, the larger departure delays result in fewer cancellations and swaps (especially for the medium- and long-haul flights in the A320 and B752 networks, which offer stronger flight planning flexibility). Moreover, as long as disruptions remain relatively small, the airline can maintain network connectivity through departure holds. Ultimately, SRPDM increases low-impact recovery measures (departure holds) and decreases higher-impact recovery measures (cancellations and swaps)—thus mitigating expected recovery costs through more flexible and robust recovery.

Table 5 details the distribution of the SRPDM benefits across disruptions instances, for each fleet and disruption category. Except for the three small-disruption instances with the A319 fleet, the SRPDM reduces recovery costs in 8% to 43% of disruption instances but never increases them. This can be explained as follows. First, the SRPDM adds robustness into airline recovery through departure holds and slower flight plans—which are only applied if they markedly reduce the future likelihood of flight cancellations or aircraft swaps. In other words, the SRPDM plans for scenarios in the lower tail of the delay distribution. Second, in any time period, realized disruptions (which account for propagated, systemic and contingent disruptions) are unlikely to be much smaller than forecasted ones (which ignore contingent disruptions). Third, realized disruptions are even less likely to be much smaller than forecasted ones throughout the recovery horizon. So even if in a certain time period the SRPDM costs are higher than baseline costs, the added robustness resulting from SRPDM can be exploited at later decision points. From a practical standpoint, this result shows that our approach not only reduces expected recovery costs without increasing worst-case recovery costs—thus enhancing the average recovery efficiency without introducing additional risk.

Table 5 SRPDM results statistics.

	Small disruptions			Medium disruptions			Large disruptions		
Fleet	Num. Worse	Num. Better	Total instances	Num. Worse	Num. Better	Total instances	Num. Worse	Num. Better	Total instances
A319	0	0	3	0	1	12	0	2	10
	(0%)	(0%)	–	(0%)	(8.3%)	–	(0%)	(20%)	–
A320	0	7	16	0	3	11	0	2	7
	(0%)	(44%)	–	(0%)	(27%)	–	(0%)	(29%)	–
B752	0	3	12	0	5	17	0	2	8
	(0%)	(25%)	–	(0%)	(28%)	–	(0%)	(25%)	–

6.2. Impact of Spatial and Temporal Scale

We now show the importance of capturing systemic disruptions at scale in space (i.e., across all hubs in the networks of the airline's operations versus a subset of hubs) and time (i.e., for sufficiently long

look-ahead windows). Table 6 reports SRPDM results from the A320 fleet when future disruptions are captured (i) at a subset of three hubs (at JFK, ATL and MSP, but not at LGA, DTW and SLC); (ii) for a 2-hour look-ahead window (rather than a 4-hour look-ahead window); and (iii) for the full set of six hubs and a 4-hour look-ahead window (referred to as "full size").

Table 6 Performance of SRPDM, for different spatial and temporal scopes of disruptions predictions.

Model	Total Cost ($)	Dep. Delay (min)	Arr. Delay (min)	Fuel Burn (lb.)	# Speed Change	# Cancel.	# Swaps	Savings
Baseline (avg.)	1,212,395	1,975	1,960	1,491,618	72	5	1	–
Baseline (min.)	1,117,555	574	1,314	1,256,011	60	0	0	–
Baseline (max.)	1,499,232	4,710	3,236	1,524,743	74	28	6	–
SRPDM, 3 hubs (avg.)	1,191,783	2,679	2,402	1,508,036	73	11	4	0.05%
SRPDM, 3 hubs (min.)	1,117,555	1,015	1,284	458,580	24	2	0	0.00%
SRPDM, 3 hubs (max.)	1,499,232	4,710	3,236	1,524,743	74	13	4	1.54%
SRPDM, $T_{LA} = 2$ hrs (avg.)	1,191,783	2,342	2,230	1,498,243	72	7	2	0.46%
SRPDM, $T_{LA} = 2$ hrs (min.)	1,117,555	574	1,281	458,580	24	0	0	0.00%
SRPDM, $T_{LA} = 2$ hrs (max.)	1,319,100	4,710	3,236	1,524,743	74	13	6	6.17%
SRPDM, full size (avg.)	1,191,764	2,037	1,994	1,494,200	72	4	1	1.8%
SRPDM, full size (min.)	1,117,555	574	1,314	1,256,011	60	0	0	0.0%
SRPDM, full size (max.)	1,499,232	4,710	3,236	1,524,743	74	28	6	16.3%

First, the benefits of SRPDM increase significantly as stochastic disruptions forecasts are developed at 6 hub airports, as compared to a subset of 3 hubs. Around 25% of flights in the A320 network depart from or arrive at the subset of three hubs (JFK, ATL and MSP), while about 50% of flights depart from or arrive at the full set of six hubs. The results show that SRPDM provides minimal improvements over the myopic baseline when disruptions are only forecasted at the subset of three hubs—with average cost reductions of 0.05%. In contrast, SRPDM reduces expected costs by 1.8% on average when delays are predicted at the full set of six hubs, with reductions of up to 16% in some instances. These results demonstrate the value of capturing *network-wide disruptions*, and their stochasticity, by applying the queuing model at several hub airports simultaneously.

Recall, also, that each look-ahead disruption scenario relied on independent disruption forecasts across airports—ignoring potential network-wide correlations. But this restriction is conservative, as any benefits obtained with independent scenarios (as compared to the myopic baseline) can also be achieved with correlated scenarios. This is because we test the model on real-world disruption instances, which contain any correlations observed in practice. Ultimately, our results provide a lower bound of the benefits that could be obtained with correlated disruptions.

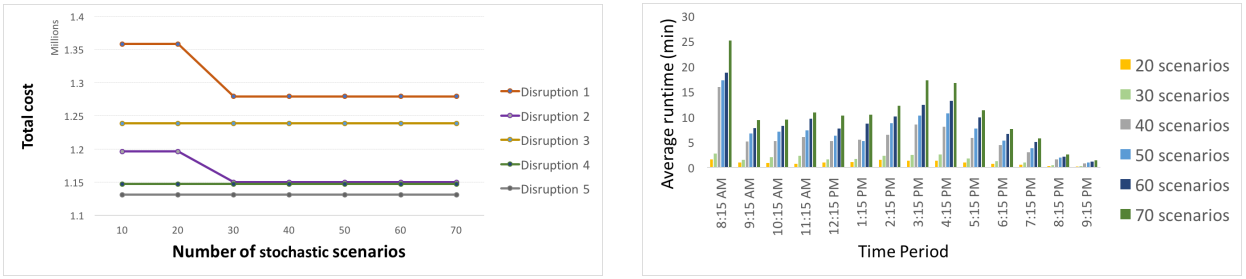
Finally, developing scenarios over extended time periods can further reduce recovery costs. Indeed, the benefits of SRPDM are larger with a longer look-ahead window of 4 hours than with a smaller look-ahead window of 2 hours (1.8% vs. 1.54% on average). This difference mainly stems

from the much lower recovery costs in the largest disruption instances—with a maximum cost reduction of 16.3% over the baseline with a 4-hour look-ahead, as opposed to 6.17% with a 2-hour look-ahead. By anticipating broader ranges of future disruptions, longer look-ahead windows T_{LA} mitigate expected recovery costs through added flexibility and robustness in decision-making.

6.3. Computational Performance

The main determinants of the model's size—and hence, of its computational performance—are the scope of the flight plans (determined by the maximum holding window and the holding interval, as described in Section 4.2) and the number of scenarios. Note that the size of SRPDM remains unchanged as disruptions are forecasted at more airports and/or over longer look-ahead windows.

Figure 7 shows the sensitivity of the solution quality and the runtimes as a function of the number of scenarios, over five randomly-generated disruption instances. The runtimes are given here for each iteration of the algorithm—thus reflecting its relevant computational requirements for the airline at any decision point. Figure 7a indicates that solution quality improves with 30 vs. 10 scenarios, but remains unchanged with 30 vs. 70 scenarios. At the same time, computational requirements increase non-linearly at each iteration with the number of scenarios (Figure 7b)—from 3–5 minutes with 30 scenarios to over 15 minutes with more scenarios. This indicates a "sweet spot" in the model's implementation with 30 scenarios, which yields a sample space that is large enough to ensure high solution qualities and small enough to derive solutions in reasonable runtimes.



(a) SRPDM performance (b) Average computation times per time period

Figure 7 SRPDM performance vs. computation times over increasing numbers of scenarios.

Ultimately, SRPDM can be implemented in short computational times—consistent with earlier models of disruption recovery and with airline requirements. This strong computational performance would enable the implementation of SRPDM in practice. Our model could also provide close-to-optimal solution in shorter runtimes, should the airline need to (e.g., by considering a smaller scenario set, or by imposing a maximum runtime).

6.4. Robustness

At this point, we have shown that the proposed modeling and computational framework can mitigate expected recovery costs by 1–2%, as compared to the myopic baseline. We now establish the robustness of these findings. We first vary the schedule of flights by considering inputs from different days in July 2014. We then vary the delay cost parameter, thus changing the weights attributed to the different components of the objective function.

6.4.1. Impact of Flight Schedule. First, we consider flight schedules for four weekdays of July 2014 (July 15–18)—which are representative of the distribution of flight schedules over the entire year of operations (see Table 1). For each day, we generate the disruption instances from historical delays (Section 5.3) and future scenarios at each of the six hubs (Section 5.2).

Table 7 Results for the A320 fleet and a 4-hour look-ahead for multiple weekdays in July.

Day in July	Model	Total Cost ($)	Dep. Delay (min)	Arr. Delay (min)	Fuel Burn ($)	# Swaps	# Cancel.	# Speed Changes	Savings
15th	Baseline (avg.)	908,701	467	719	815,755	1	3	42	–
	Baseline (min)	864,396	185	582	800,320	0	2	38	–
	Baseline (max)	971,778	865	944	832,796	3	4	54	–
	SRPDM (avg.)	896,457	511	745	821,671	1	2	42	1.4%
	SRPDM (min)	858,852	223	582	799,903	0	0	38	0.0%
	SRPDM (max)	971,778	865	962	832,796	3	4	54	4.6%
16th	Baseline (avg.)	901,364	774	725	735,552	3	5	35	–
	Baseline (min)	814,419	140	514	695,086	0	2	32	–
	Baseline (max)	994,649	1,747	1,095	760,847	5	8	36	–
	SRPDM (avg.)	890,434	811	719	735,306	3	5	35	1.3%
	SRPDM (min)	787,796	278	539	695,086	0	0	32	0.0%
	SRPDM (max)	994,649	1,747	1,120	760,461	5	8	37	3.4%
18th	Baseline (avg.)	1,378,449	1,531	1,335	1,069,890	3	10	56	–
	Baseline (min)	1,333,659	830	1,109	1,038,907	1	7	52	–
	Baseline (max)	1,429,588	2,421	1,523	1,094,065	4	14	68	–
	SRPDM (avg.)	1,358,686	1,566	1,351	1,069,263	3	9	57	1.5%
	SRPDM (min)	1,314,221	918	1,133	1,038,907	1	7	53	0.0%
	SRPDM (max)	1,410,578	2,431	1,523	1,093,679	4	13	68	3.0%

Table 7 reports the outputs of SRPDM and the myopic baseline for July 15, 16 and 18 (the corresponding results for July 17 are shown in Table 4). The results confirm that SRPDM reduces expected recovery costs by 1–2%, as compared to the myopic baseline. As earlier, SRPDM increases departure delay to reduce the number of aircraft swaps, the number of cancellations and, in some cases, fuel burn. The largest recovery cost reduction, over all disruption instances, is lower than the corresponding one for July 17. This stems from the difference in realized disruptions: the benefits of SRPDM tend to be higher when realized disruptions occur in peak periods, resulting in higher delays at hub airports and higher downstream impacts throughout the network. In such instances,

the myopic baseline tends to increase aircraft speeds to ensure connectivity, while SRPDM leverages information on future disruptions to strategically introduce departure holds. As it turns out, July 17 had a higher incidence of such large peak-hour disruptions than other days. Nonetheless, the average recovery cost savings are consistent across all days—thus highlighting that our approach does not solely provide benefits during the busiest or least busy days of the year.

6.4.2. Impact of Delay Costs. Next, we establish the robustness of the benefits of SRPDM with respect to the objective function. Recall that SRPDM and the myopic baseline are formulated as multi-objective optimization problems that trade off fuel, delay, aircraft swap, and flight cancellation costs (Equation (15) and (23)). We keep all cost parameters unchanged, but vary the unit delay cost from \$10 to \$77 per minute (as described in Section 5.4). Table 8 reports the results of these experiments, for a random subset of "small" disruption instances.

We also consider non-linear (convex) delay costs: following Cook and Tanner (2008a), we set, for each copy $k \in \mathcal{K}^t_{faq}$, $\delta_k = C \times D_k \times \ln(D_k)$, where C is a constant and D_k is the delay (in minutes) of copy k. We choose C such that the average delay cost is equal to \$10/minute when applied to the solution obtained with a linear delay cost of \$10/minute—so that we change the distribution of delay costs but not their overall magnitude. These results are also presented in Table 8.

These results show that the relative reduction in recovery costs achieved by SRPDM, as compared to the baseline, is remarkably consistent across all (linear or non-linear) delay cost functions—ranging from 2.3% to 3.1% on average. As delay costs increase, both models reduce the incidence of departure holds—reducing average delays but increasing the number of flight cancellations and aircraft swaps. But regardless of the delay cost function, SRPDM results in higher departure and arrival delays than the myopic baseline but in fewer cancellations and (with one exception) fewer aircraft swaps. Ultimately, these results confirm that SRPDM reduces recovery costs by introducing strategic departure holds, over the full range of delay cost functions under consideration.

6.5. Summary

Our results suggest that the SRPDM can enhance airline disruption recovery decisions—by reducing recovery costs by 1–2% on average, as compared to a myopic baseline that does not anticipate future disruptions. These results are driven by reductions in flight cancellations and aircraft swaps, partially offset by increases in flight delays. In other words, SRPDM leverages information on a range of disruption scenarios by strategically introducing departure holds and adjusting flight plans, to avoid resorting to flight cancellations and aircraft swaps in subsequent time periods.

Moreover, the benefits of SRPDM increase with the scope of the disruptions under consideration: capturing systemic disruptions—and their stochasticity—from more hub airports and over longer look-ahead windows increases the value of SRPDM. This underscores the benefits of the framework

Table 8 Experimental results for different delay cost parameters.

Delay Cost ($/min)	Model	Total Cost ($)	Dep. Delay (min)	Arr. Delay (min)	Fuel Burn (lb.)	# Swaps	# Cancel.	# Speed Changes	Savings
$10	Baseline (min)	1,133,878	556	1,313	1,449,264	0	2	69	-
	Baseline (max)	1,358,403	1,202	1,599	1,523,785	6	14	73	-
	Baseline (avg.)	**1,205,434**	**895**	**1,463**	**1,495,607**	**1.1**	**5.1**	**72**	**-**
	SRPDM (min)	1,121,481	586	1,433	1,457,276	0	0	70	0.03%
	SRPDM (max)	1,279,491	1,222	1,750	1,522,769	6	10	74	6.2%
	SRPDM (avg.)	**1,167,328**	**1,002**	**1,535**	**1,497,605**	**1**	**3**	**72**	**2.8%**
$20	Baseline (min)	1,146,048	536	1,277	1,443,010	0	2	68	-
	Baseline (max)	1,366,260	1,192	1,566	1,524,265	6	14	72	-
	Baseline (avg.)	**1,214,304**	**864**	**1,412**	**1,489,390**	**1.1**	**5.4**	**70**	**-**
	SRPDM (min)	1,131,292	566	1,315	1,443,010	0	0	68	0.02%
	SRPDM (max)	1,289,708	1,212	1,717	1,523,511	6	10	73	5.9%
	SRPDM (avg.)	**1,177,269**	**971**	**1,484**	**1,491,388**	**1.1**	**3.4**	**71**	**3.1%**
$30	Baseline (min)	1,158,625	496	1,180	1,443,770	0	2	64	-
	Baseline (max)	1,373,611	1,162	1,451	1,526,435	6	14	68	-
	Baseline (avg.)	**1,223,163**	**820**	**1,318**	**1,484,522**	**1.1**	**5.6**	**67**	**-**
	SRPDM (min)	1,140,885	526	1,287	1,443,770	0	0	64	0.0%
	SRPDM (max)	1,299,419	1,162	1,520	1,525,214	6	10	68	5.7%
	SRPDM (avg.)	**1,187,198**	**923**	**1,376**	**1,486,723**	**1.1**	**3.6**	**67**	**3.0%**
$40	Baseline (min)	1,164,992	436	1,028	1,399,814	0	2	58	-
	Baseline (max)	1,363,486	1,102	1,261	1,529,358	5	14	63	-
	Baseline (avg.)	**1,229,222**	**758**	**1,168**	**1,478,142**	**1**	**5.9**	**61**	**-**
	SRPDM (min)	1,149,774	456	1,173	1,421,097	0	0	59	0.0%
	SRPDM (max)	1,272,351	1,112	1,402	1,528,129	7	9	64	7.2%
	SRPDM (avg.)	**1,191,866**	**867**	**1,246**	**1,479,936**	**1.3**	**3.8**	**62**	**3.1%**
$50	Baseline (min)	1,170,381	386	889	1,404,649	0	2	51	-
	Baseline (max)	1,370,553	1,052	1,120	1,534,742	5	14	57	-
	Baseline (avg.)	**1,238,187**	**710**	**1,005**	**1,483,210**	**1.1**	**5.9**	**54**	**-**
	SRPDM (min)	1,160,776	406	1,026	1,405,701	0	0	52	0.0%
	SRPDM (max)	1,301,195	1,062	1,206	1,534,214	5	11	58	5.3%
	SRPDM (avg.)	**1,207,499**	**815**	**1,073**	**1,482,628**	**1.3**	**4.1**	**55**	**2.5%**
$60	Baseline (min)	1,174,241	386	879	1,384,418	0	2	51	-
	Baseline (max)	1,395,891	1,032	1,099	1,535,943	6	16	57	-
	Baseline (avg.)	**1,253,596**	**698**	**988**	**1,481,443**	**1.5**	**6.4**	**55**	**-**
	SRPDM (min)	1,169,464	406	1,024	1,405,701	0	0	52	0.0%
	SRPDM (max)	1,328,993	1,042	1,185	1,535,415	6	11	58	5.8%
	SRPDM (avg.)	**1,218,281**	**806**	**1,062**	**1,483,390**	**1.4**	**4.3**	**56**	**2.8%**
$77	Baseline (min)	1,180,803	386	879	1,384,418	0	2	51	-
	Baseline (max)	1,404,425	1,032	1,099	1,536,319	6	16	57	-
	Baseline (avg.)	**1,265,439**	**697**	**987**	**1,481,550**	**1.5**	**6.4**	**55**	**-**
	SRPDM (min)	1,180,485	406	1,024	1,405,701	0	0	52	0.0%
	SRPDM (max)	1,341,709	1,042	1,185	1,535,791	6	11	58	5.5%
	SRPDM (avg.)	**1,231,967**	**805**	**1,060**	**1,483,496**	**1.4**	**4.3**	**56**	**2.6%**
non-linear cost (equivalent to $10)	Baseline (min)	1,142,595	556	1,418	1,447,980	0	2	71	-
	Baseline (max)	1,193,108	1,372	1,772	1,523,785	7	4	73	-
	Baseline (avg.)	**1,157,004**	**959**	**1,529**	**1,503,547**	**1**	**2.9**	**72**	**-**
	SRPDM (min)	1,111,532	586	1,471	1,447,980	0	0	71	0.0%
	SRPDM (max)	1,152,517	1,372	1,772	1,522,892	7	4	74	3.8%
	SRPDM (avg.)	**1,131,204**	**1,034**	**1,582**	**1,502,920**	**1**	**1.6**	**72**	**2.3%**

developed in this paper, which proposes a multi-stage decision-making under uncertainty approach combining, for the first time, a predictive queuing model—applied at several airports of the network simultaneously—into a prescriptive combinatorial optimization model of airline disruption recovery.

7. Conclusion

This paper proposes a jointly reactive and proactive approach to airline disruption management. This approach optimizes disruption recovery decisions while leveraging a partial and probabilistic forecast of future disruptions—by characterizing probabilistically future systemic disruptions (i.e., congestion at hub airports) but ignoring other contingent disruption forecasts (e.g., aircraft maintenance, late crews, late passenger boarding). We formulate a Stochastic Reactive and Proactive Disruption Management (SRPDM) that combines a stochastic queuing model of congestion (applied at several airports within a network), a flight planning tool from Boeing/Jeppesen, and an optimization model of airline disruption recovery. We design an efficient solution procedure based on look-ahead approximation and sample average approximation, which enables the model's implementation at any decision point in reasonable computational times—consistent with earlier recovery models and with practical airline requirements. Results suggest that leveraging even partial and probabilistic information on future disruptions and an approximate algorithm can enhance recovery decisions: SRPDM consistently performs as well as or better than a myopic baseline, ultimately reducing expected disruption costs without creating additional risk in airline recovery.

The implications of these results are threefold. First, airline recovery can be improved through more flexible and robust decision-making—by deliberately introducing departure holds and speed changes to mitigate the incidence of flight cancellations and aircraft swaps at later points in time. Second, airline operations can benefit from the elicitation of systemic disruption scenarios, especially in instances where flight networks are concentrated at hub airports and where hub airports are highly congested. Such scenarios can be constructed from information available offline, including flight schedules, historical records of airport operations, and weather forecasts. Last, further cost savings could potentially be achieved through online sharing of operating information between airline operators, airport operators, and air traffic managers. Most notably, continuous alignment on operating conditions, real-time congestion and delay forecasts could reduce system-wide uncertainty on future operations, thus permitting more effective recovery.

These results motivate future work on airline recovery optimization under uncertainty. First, this paper has relied on a simple prediction of future disruptions by applying the queuing model independently at each hub airport; future research could generate disruption scenarios that capture cross-airport correlations. Moreover, further research could investigate how to incorporate dynamic updates of delay predictions into recovery optimization—in line with the real-time information sharing paradigm mentioned above. Second, this paper has focused on aircraft recovery. An

important extension would involve developing a jointly reactive and proactive approach to the integrated problem of aircraft, passenger and crew recovery. Third, the approximate solution algorithm considered in this paper could be augmented with exact algorithms for multi-stage recovery optimization under uncertainty. The framework, model and algorithm proposed in this paper provide the foundations to explore these questions—toward more efficient, reliable and robust recovery.

References

Abdelghany K, Abdelghany A, Ekollu G (2008) An integrated decision support tool for airlines schedule recovery during irregular operations. *European Journal of Operational Research* 185(2):825–848.

Abdelghany K, Shah S, Raina S, Abdelghany A (2004) A model for projecting flight delays during irregular operation conditions. *Journal of Air Transport Management* 10(6):385–394.

Ahmadbeygi S, Cohn A, Lapp M (2010) Decreasing airline delay propagation by re-allocating scheduled slack. *IIE transactions* 42(7):478–489.

Arıkan M, Deshpande V, Sohoni M (2013) Building reliable air-travel infrastructure using empirical data and stochastic models of airline networks. *Operations Research* 61(1):45–64.

Balakrishnan H, Chandran B (2010) Algorithms for Scheduling Runway Operations Under Constrained Position Shifting. *Operations Research* 58(6):1650–1665.

Ball M, Barnhart C, Dresner M, Hansen M, Neels K, Odoni A, Peterson E, Sherry L, Trani A, Zou B (2010) Total Delay Impact Study. Technical report, National Center of Excellence for Aviation Operations Research, College Park, MD.

Barnhart C, Vaze V (2015) Irregular operations: Schedule recovery and robustness. *The Global Airline Industry* 253–274.

Bertsekas D (2005) *Dynamic Programming and Optimal Control*, volume I (Athena Scientific), 3rd edition.

Bertsekas D (2012) *Dynamic Programming and Optimal Control*, volume II (Athena Scientific), 4th edition.

Bertsimas D, Lulli G, Odoni A (2011) An integer optimization approach to large-scale air traffic flow management. *Operations Research* 59(1):211–227.

Borndörfer R, Dovica I, Nowak I, Schickinger T (2010) Robust tail assignment .

Bratu S, Barnhart C (2006) Flight operations recovery: New approaches considering passenger recovery. *Journal of Scheduling* 9(3):279–298.

Cadarso L, Marín Á (2011) Integrated robust airline schedule development. *Procedia-Social and Behavioral Sciences* 20:1041–1050.

Cao J, Kanafani A (1997) Real-time decision support for integration of airline flight cancellations and delays Part II: algorithm and computational experiments. *Transportation Planning and Technology* 20(3):201–217.

Clarke M, Naryadi Y (1995) The Airline Operation Control Centre : an overview of Garuda's Operation Control (EM) at Cengkereng Jakarta, Indonesia : final report to PT Garuda Indonesia. Technical report, Massachusetts Institute of Technology, Flight Transportation Laboratory.

Cook A, Tanner G (2008a) Dynamic Cost Indexing: Airline costs of delayed passengers and how to estimate full network delay costs. Technical report, Innovative Cooperative Actions of R&D in EUROCONTROL Programme CARE INO III.

Cook J, Tanner G (2008b) Innovative Cooperative Actions of Research & Development in EUROCONTROL Programme CARE INO III: Dynamic Cost Indexing: Airline costs of delayed passengers and how to estimate full network delay costs. Technical report, Transport Studies Group, University of Westminster, London.

Dunbar M, Froyland G, Wu CL (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 CL (2014) An integrated scenario-based approach for robust aircraft routing, crew pairing and re-timing. *Computers & Operations Research* 45:68–86.

Eggenberg N, Salani M, Bierlaire M (2010) Constraint-specific recovery network for solving airline recovery problems. *Computers & Operations Research* 37(6):1014–1026.

Federal Aviation Administration (2004) Airport Capacity Benchmark Report. Technical report.

Froyland G, Maher SJ, Wu CL (2013) The recoverable robust tail assignment problem. *Transportation Science* 48(3):351–372.

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, URL `https://doi.org/10.1016/j.tre.2016.01.002`.

Hu Y, Xu B, Bard JF, Chi H, Gao M (2015) Optimization of multi-fleet aircraft routing considering passenger transiting under airline disruption. *Computers & Industrial Engineering* 80:132 – 144, URL `http://dx.doi.org/https://doi.org/10.1016/j.cie.2014.11.026`.

International Air Transport Association (2010) Jet Fuel Price Development. Technical report.

Jacquillat A, Odoni A (2015a) An Integrated Scheduling and Operations Approach to Airport Congestion Mitigation. *Operations Research* 63(6):1390–1410.

Jacquillat A, Odoni A (2015b) Endogenous Control of Arrival and Departure Service Rates in Dynamic and Stochastic Queuing Models with Application at JFK and EWR. *Transportation Research Part E: Logistics and Transportation Review* 73(1):133–151.

Jacquillat A, Odoni A (2017) A New Airport Demand Management Approach Based on Targeted Scheduling Interventions. *Journal of Transport Economics and Policy* 51(2):115–138.

Jacquillat A, Odoni AR, Webster MD (2016) Dynamic control of runway configurations and of arrival and departure service rates at JFK airport under stochastic queue conditions. *Transportation Science* 51(1):155–176.

Jafari N, Zegordi SH (2010) The airline perturbation problem: considering disrupted passengers. *Transportation Planning and Technology* 33(2):203–220.

Jarrah A, Yu G, Krishnamurthy N, Rakshit A (1993) A Decision Support Framework for Airline Flight Cancellations and Delays. *Transportation Science* 27(3):266–280.

Jozefowiez N, Mancel C, Mora-Camino F (2013) A heuristic approach based on shortest path problems for integrated flight, aircraft, and passenger rescheduling under disruptions. *Journal of the Operational Research Society* 64:384–395.

Kleywegt AJ, Shapiro A, Homem-de Mello T (2002) The sample average approximation method for stochastic discrete optimization. *SIAM Journal on Optimization* 12(2):479–502.

Lan S, Clarke JP, 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, Nemhauser G (2000) Airline Crew Recovery. *Transportation Science* 34(4):337–348.

Maher SJ (2016) Solving the integrated airline recovery problem using column-and-row generation. *Transportation Science* 50(1):216–239.

Marla L, Vaaben B, Barnhart C (2017) 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 & Operations Research* 98:165–184.

McCarty LA, Cohn AEM (2018) Preemptive rerouting of airline passengers under uncertain delays. *Comput. Oper. Res.* 90:1–11, URL `http://dx.doi.org/10.1016/j.cor.2017.09.001`.

Meloni C, Pacciarelli D, Pranzo M (2004) A Rollout Metaheuristic for Job Shop Scheduling Problems. *Annals of Operations Research* 131:215–235.

Petersen J, Solveling G, Johnson E, Clarke J, Shebalov S (2012) An optimization approach to airline integrated recovery. *Transportation Science* 46(4):482–500.

Pita J, Barnhart C, Antunes A (2012) Integrated Flight Scheduling and Fleet Assignment under Airport Congestion. *Transportation Science* 47(4):477–492.

Powell WB (2007) *Approximate Dynamic Programming: Solving the Curses of Dimensionality* (Wiley-Interscience), 2nd edition.

Pyrgiotis N, Malone K, Odoni A (2013) Modelling Delay Propagation within an Airport Network. *Transportation Research Part C: Emerging Technologies* 27:60–75.

Ribeiro N, Antunes A, Jacquillat A, Odoni A, Pita J (2017) An optimization approach for airport slot allocation under IATA guidelines. *Transportation Research Part B: Methodological* 112:132–156.

Rosenberger JM, Johnson EL, Nemhauser GL (2003) Rerouting aircraft for airline recovery. *Transportation Science* 37(4):408–421.

Rosenberger JM, Johnson EL, Nemhauser GL (2004) A robust fleet-assignment model with hub isolation and short cycles. *Transportation science* 38(3):357–368.

Schaefer AJ, Johnson EL, Kleywegt AJ, Nemhauser GL (2005) Airline crew scheduling under uncertainty. *Transportation science* 39(3):340–348.

Secomandi N (2001) A Rollout Policy for the Vehicle Routing Problem with Stochastic Demands. *Operations Research* 49(5):796–802.

Shebalov S, Klabjan D (2006) Robust airline crew pairing: Move-up crews. *Transportation science* 40(3):300–312.

Simaiakis I (2012) *Analysis, Modeling and Control of the Airport Departure Process*. Ph.D. thesis, Massachusetts Institute of Technology.

Simaiakis I, Khadilkar H, Balakrishnan H, Reynolds T, Hansman J (2014) Demonstration of reduced airport congestion through pushback rate control. *Transportation Research Part A: Policy and Practice* 66:251–267.

Sinclair K, Cordeau JF, 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.

Smith BC, Johnson EL (2006) Robust airline fleet assignment: Imposing station purity using station decomposition. *Transportation Science* 40(4):497–516.

Sohoni M, Lee YC, Klabjan D (2011) Robust airline scheduling under block-time uncertainty. *Transportation Science* 45(4):451–464.

Teodorović D (1984) Optimal dispatching strategy on an airline network after a schedule perturbation. *European Journal of Operational Research* 15(2):178–182.

Thengvall BG, Bard JF, Yu G (2000) Balancing user preferences for aircraft schedule recovery during irregular operations. *IIE Transactions* 32(3):181–193.

Vossen T, Hoffman R, Mukherjee A (2012) Air Traffic Flow Management. *Quantitative Problem Solving Methods in the Airline Industry*, volume 169 of *International Series in Operations Research & Management Science*, 385–453 (Springer US).

Wei G, Yu G, Song M (1997) Optimization Model and Algorithm for Crew Management During Airline Irregular Operations. *Journal of Combinatorial Optimization* 1(3):305–321.

Yan C, Kung J (2016) Robust aircraft routing. *Transportation Science* 52(1):118–133.

Yan S, Yang DH (1996) A decision support framework for handling schedule perturbation. *Transportation Research Part B: Methodological* 30(6):405–419.

Yen JW, Birge JR (2006) A stochastic programming approach to the airline crew scheduling problem. *Transportation Science* 40(1):3–14.

Yu G, Argüello M, Song G, McCowan S, White A (2003) A New Era for Crew Recovery at Continental Airlines. *Interfaces* 33(1):5–22.

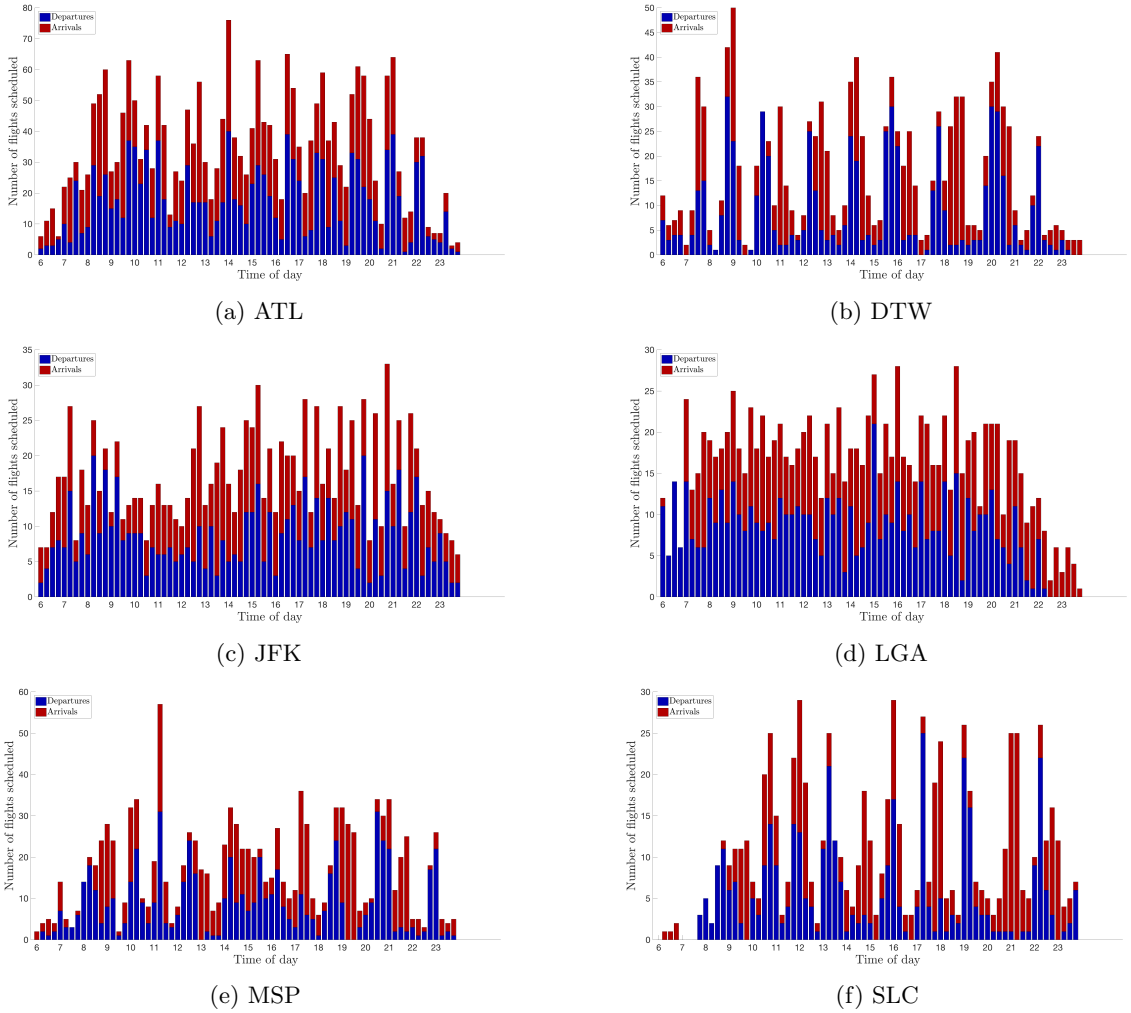
Zhang D, Lau HH, Yu C (2015) A two stage heuristic algorithm for the integrated aircraft and crew schedule recovery problems. *Computers & Industrial Engineering* 87:436–453.

Zhang D, Yu C, Desai J, Lau HH (2016) A math-heuristic algorithm for the integrated air service recovery. *Transportation Research Part B: Methodological* 84:211–236.

Zografos K, Salouras Y, Madas M (2012) Dealing with the Efficient Allocation of Scarce Resources at Congested Airports. *Transportation Research Part C: Emerging Technologies* 21(1):244–256.

Appendix A: Flight Schedules at the Hub Airports

Figure 8 plots the number of flights scheduled at each of the six hub airports under consideration on July 17, 2014 per 15-minute period of the day. Note that, at some airports, the schedule of flights is relatively evenly distributed over the day, whereas other airports operate a strongly "peaked" schedule. At one extreme, New York's LaGuardia (LGA) airport faces strong local demand and schedule limits (or "flight caps"), resulting in high scheduling levels and limited schedule variability throughout the day. At the other extreme, Detroit (DTW) operates arrival "banks" immediately followed by departure "banks" to enable passenger connections, resulting in a sequence of peaks and valleys. Similar scheduling patterns are observed at Salt Lake City (SLC) and, to a lesser extent, at Minneapolis Saint Paul (MSP). The remaining two airports (New York's John F. Kennedy (JFK) airport and Atlanta's (ATL) airport) fall in-between: the schedule exhibits peaks and valleys but milder variations than at DTW, SLC and MSP. These scheduling patterns are the primary determinants of the delay patterns shown in Figure 6.



(a) ATL (b) DTW

(c) JFK (d) LGA

(e) MSP (f) SLC

Figure 8 Schedule of flights at each airport on July 17, 2014.

Appendix B: Disruption Classification

We classify each disruption instance (see Section 5.3) into "small", "medium" and "large" disruption categories. We use a multinomial logistic regression with two independent variables: total delay and maximum delay occurring in the network. Our model is trained with selected scenarios that are *a priori* classified into the three categories, as a training step. We then perform a validation step, to predict the groups for all remaining scenarios. The proportional log-odds of a disruption scenario belonging to the small and medium category, versus the large category, is defined as the logarithm of the ratio of the two probabilities. The logistic regression model is provided in Equations (24) and (25), where P_s, P_m and P_l denote the probabilities of belonging to the large, medium and small delay categories, respectively.

$$\log\left(\frac{P_s}{P_l}\right) = \beta_0 + \beta_1^s \text{ total delay} + \beta_2^s \text{ max delay} \tag{24}$$

$$\log\left(\frac{P_m}{P_l}\right) = \beta_0 + \beta_1^m \text{ total delay} + \beta_2^m \text{ max delay} \tag{25}$$

The model's coefficients for each fleet type are reported in Table 9. First, all coefficients are negative: the larger the total delays and/or the maximum delays, the less likely the scenario under consideration is to be classified in the small or medium category, as compared to the large category. Second, all coefficients are smaller for the small delay category than for the medium delay category: all else being equal, the larger the total delays and/or the maximum delays, the less likely the scenario is to be classified in the small category, as opposed to the medium category. Third, all but one p-values are lower than 0.05, thus suggesting that the classification model is statistically significant. Last, we perform a cross-validation on a test set, and confirm zero misclassification. The number of scenarios into each category is reported in Table 10.

Table 9 Parameter estimates for multinomial logistic regression.

		small (vs. large)			medium (vs. large)		
Fleet	Predictor	Coef.	Std. err.	p-value	Coef.	Std. err.	p-value
A319	constant	16.729	2.488	0.000	8.485	2.068	0.000
	total delay	-0.007	0.001	0.000	-0.002	0.001	0.001
	max delay	-0.016	0.005	0.000	-0.013	0.004	0.001
A320	constant	22.054	3.138	0.000	10.361	2.406	0.000
	total delay	-0.008	0.001	0.000	-0.002	0.001	0.002
	max delay	-0.036	0.008	0.000	-0.016	0.005	0.001
A320	constant	21.968	2.902	0.000	11.464	2.293	0.000
	total delay	-0.002	0.001	0.006	0.000	0.000	0.130
	max delay	-0.078	0.013	0.000	-0.026	0.006	0.000

Table 10 Number of disruption instances in each category.

Fleet	A319	A320	B752
Small disruptions	3	16	12
Medium disruptions	12	11	17
Large disruptions	10	7	8