

Equitable and Efficient Coordination in Traffic Flow Management

Cynthia Barnhart

Department of Civil and Environmental Engineering, Massachusetts Institute of Technology,
Cambridge, Massachusetts 02139, cbarnhart@mit.edu

Dimitris Bertsimas

Sloan School of Management, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139,
dbertsim@mit.edu

Constantine Caramanis

Department of Electrical and Computer Engineering, The University of Texas at Austin, Austin, Texas 78712,
caramanis@mail.utexas.edu

Douglas Fearing

Operations Research Center, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139,
dfearing@mit.edu

When air traffic demand is projected to exceed capacity, the Federal Aviation Administration implements traffic flow management (TFM) programs. Independently, these programs maintain a first-scheduled, first-served invariant, which is the accepted standard of fairness within the industry. Coordinating conflicting programs requires a careful balance between equity and efficiency. In our work, we first develop a fairness metric to measure deviation from first-scheduled, first-served in the presence of conflicts. Next, we develop an integer programming formulation that attempts to directly minimize this metric. We further develop an exponential penalty approach and show that its computational performance is far superior and its tradeoff between delay and fairness compares favorably. In our results, we demonstrate the effectiveness of these models using historical and hypothetical scenarios. Additionally, we demonstrate that the exponential penalty approach exhibits exceptional computational performance, implying practical viability. Our results suggest that this approach could lead to system-wide savings on the order of \$25 to \$50 million per year.

Key words: traffic flow management; ground holding programs; equitable flight delay

History: Received: June 2009; revisions received: March 2010, March 2011; accepted: August 2011. Published online in *Articles in Advance* February 8, 2012.

1. Introduction

The Federal Aviation Administration (FAA) and the airline industry face tremendous challenges from unexpected weather-induced reductions in system capacity and resulting delays. The U.S. Congress Joint Economic Committee estimates that during calendar year 2007, 2.75 million hours of flight delays led to approximately \$25.7 billion in costs to the U.S. economy—\$12.2 billion in increased airline operating costs, \$7.4 billion in passenger time lost, and \$6.1 billion in costs to related industries (Joint Economic Committee 2008). To put this in perspective, the Air Transportation Association's 2008 Annual Report lists total profits for U.S. airlines of approximately \$3 billion for the 2006 operating year and \$5 billion for the 2007 operating year (Air Transportation Association 2008).

Delays over the national air space are projected to outpace increases in overall traffic. The FAA's 2004 Airport Capacity Benchmark Report demonstrates that many major U.S. airports regularly operate at or

near peak capacity (e.g., Hartsfield-Jackson Atlanta International, New York LaGuardia, and Chicago O'Hare; U.S. Department of Transportation 2004). As with any queueing system, there is a nonlinear relationship between delay and changes in demand when operating under these conditions (de Neufville and Odoni 2003). Increasing capacity by building additional runways and airports is logistically complex because of cost, space limitations, and environmental regulations. Additionally, projects of this type often take a decade or more to plan and complete. It is thus critical to consider tools to improve operational efficiency. In §1.4 we detail the contributions of this paper. But first, to put them in context, we provide a brief discussion of the existing traffic flow management tools and a nonexhaustive literature review.

1.1. Traffic Flow Management

Traffic flow management (TFM) refers to a set of strategic practices utilized by the FAA to ensure safe

operations while attempting to minimize costs associated with delay. TFM activities occur on the day of operations and generally affect a significant subset of airline traffic (e.g., all flights into a major airport). According to data publicly available from the U.S. Bureau of Transportation Statistics, we estimate that TFM activities account for approximately 30% of all air transportation delays (see calculations in §4.5 for details). Based on factors such as number of runways, runway configuration, scheduled personnel coverage, and weather forecasts, the FAA determines maximum capacities for resources in the U.S. air transportation system. These resources include arrival runways, departure runways, and air sectors in the National Airspace System (NAS). TFM programs are initiated only when significant imbalances between demand and capacity are expected, as in the midst of a severe storm. Minor to moderate inconsistencies between capacity and demand are resolved through localized air traffic control (ATC) techniques (e.g., speed adjustments, vectoring, or airborne queueing). Since the air traffic controllers' strike in 1981, the primary tool used for TFM has been the ground delay program (GDP). In a GDP, the FAA controls the arrival rate into a reduced-capacity airport by coordinating departure times for affected flights. The goal is to allow each aircraft to proceed safely to its destination with minimal airborne delay. Airspace flow programs (AFPs), first introduced in 2006, are operated much like GDPs. The FAA uses an AFP to control the arrival rate into a flow constrained area (FCA), for example, a reduced-capacity air segment of the NAS. The papers by de Neufville and Odoni (2003), and Ball et al. (2007) provide further details regarding the TFM problem and its extensions. To understand the prevalence of these programs, we list in Table 1 the number of days in 2007 that the corresponding numbers of GDPs and AFPs were enacted (e.g., there were only 16 days with no GDPs or AFPs). Thus, on approximately 40% of the days during 2007, at least one GDP and at least one AFP was in place. Although the number of GDPs varies significantly, the number of AFPs rarely exceeds two.

Table 1 Number of Days in 2007 and Corresponding Numbers of TFM Programs of Each Type

Number of AFPs	Number of GDPs												Total
	0	1	2	3	4	5	6	7	8	9	10	11	
0	16	24	31	44	26	28	15	13	8	2	2	1	210
1	5	10	13	7	5	11	5	6	3	0	0	1	66
2	2	15	13	14	7	16	5	9	3	2	1	0	87
3	0	0	0	0	0	1	0	0	0	0	0	0	1
4	0	0	0	0	0	0	0	0	1	0	0	0	1
Total	23	49	57	65	38	56	25	28	15	4	3	2	365

GDPs and AFPs are used in concert with a three-stage, collaborative approach to decision making. In the first stage, the FAA allocates arrival slots to airlines by applying the ration-by-schedule (RBS) method for each TFM program. Arrival slots are allocated according to the original schedule ordering, as described in detail in the following section. Although fairness is a subjective criterion, the RBS approach is generally considered fair within the airline industry because it maintains a first-scheduled, first-served invariant. In the second stage, airlines respond to the schedule disruption. Each airline is allowed to make changes to the schedule within the context of the slots allocated to it. For instance, an airline can swap arrival slots for two of its own flights as long as the swap does not cause either flight to depart prior to its posted departure time. Airlines can also choose to cancel flights in response to operational constraints on aircraft routing, crew assignments, and so forth. In the third stage, the FAA accepts the changes proposed by all airlines. These changes, when combined, constitute a capacity-feasible schedule because each airline is only allowed to make changes within the set of slots allocated to it. Subsequently, the FAA attempts to improve the schedule by filling any gaps created by cancellations or operator-announced delays. This procedure, known as compression, is described in detail in Vossen and Ball (2005). After compression, the new schedule proposal is sent to the airlines and the process is repeated as necessary.

1.2. Coordinating Multiple Programs

In RBS, arrival slots for a single resource, either an airport in a GDP or an FCA in an AFP, are allocated to flights according to the original schedule order. For FCAs, the scheduled arrival order is based on the estimated arrival times into the FCA (i.e., the scheduled departure time plus the estimated en route time to reach the FCA). Once the controlled arrival slots have been allocated for a resource, each affected flight receives a corresponding controlled time of departure (CTD) from its origin, converting the allocated arrival slot into a departure delay at the airport of origin.

When multiple TFM programs are implemented on the same day, applying RBS for each independently may lead to a single flight receiving conflicting CTDs (e.g., from a GDP and one or more AFPs). To resolve these conflicts, the FAA uses heuristics to determine a CTD for each flight based on the order in which the programs are initiated over the course of the day. When AFPs were created in 2007, there was long history of successful implementation of GDPs; thus, GDPs were given priority by default. Specifically, a flight already affected by a GDP at the time an AFP is initiated is exempted from the AFP (reducing the AFP capacity for other affected flights). But for a flight

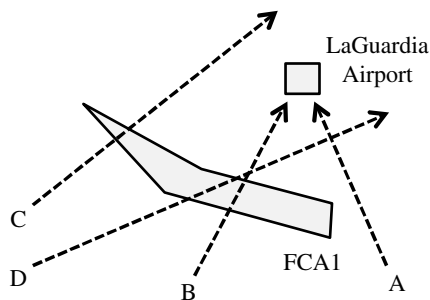


Figure 1 Visual Representation of Flight Routes for Flights A, B, C, and D

already affected by an AFP at the time a GDP is initiated, the CTD is modified to correspond to the GDP (Federal Aviation Administration 2005). This conflict-resolution heuristic, referred to as precedence RBS, is the default behavior in Flight Schedule Monitor, the application developed by Metron Aviation that the FAA uses to manage GDPs and AFPs (Metron Aviation 2009a). In March 2009, functionality was enabled in Flight Schedule Monitor that allows TFM managers to exempt AFP-affected flights from subsequent GDPs (Metron Aviation 2009b). Thus, the other conflict resolution approach we consider is a strict exemption-based heuristic in which a flight is given a CTD from the first GDP or AFP that affects its schedule and exempted from all future GDPs or AFPs. We refer to this conflict resolution heuristic as exemption RBS.

Consider the following example based on the four flight routes displayed in Figure 1 with planned schedule details listed in Table 2. At 17:00, an AFP is initiated for FCA1 with a controlled arrival rate of one flight every five minutes from 18:40 until 19:00. Subsequently, at 17:05, a GDP is initiated for New York's LaGuardia Airport (LGA), with arrivals into LGA restricted to one flight every 10 minutes from 18:55 until 19:15. Note that the time a TFM program is initiated determines which flights are affected, because flights already in the air at the time of initiation are exempted from the program. Performing RBS for each resource independently leads to the CTDs listed in Tables 3 and 4. In this case, flight B receives conflicting CTDs (17:15 from the AFP at FCA1 and 17:24 from the GDP at LGA). Thus, according to precedence RBS, flight B will be given a CTD of 17:24 (because the GDP at LGA takes precedence). This leads to the controlled schedule listed in Table 5. In Table 6, we provide

Table 2 Planned Departure and Arrival Times

Flight	Departure	FCA1	LGA
A	17:45	—	18:55
B	17:15	18:40	18:56
C	18:00	18:45	—
D	18:15	18:46	—

Table 3 RBS CTDs for FCA1

Flight	Slot	CTD
B	18:40	17:15
C	18:45	18:00
D	18:50	18:19

Table 4 RBS CTDs for LGA

Flight	Slot	CTD
A	18:55	17:45
B	19:05	17:24

the controlled schedule according to the exemption RBS (under the modified assumption that the GDP is implemented first).

It is important to note that, in this example, it is impossible to satisfy first-scheduled, first-served for each resource *and* simultaneously minimize system delay. That is, the controlled schedule must either deviate from first-scheduled, first-served or incur excess delay. This simple example illustrates the general tradeoff that exists between fairness and efficiency in the multiresource or network setting. We will consider these examples further in §2.1. For additional examples of this type, the reader is encouraged to review Lulli and Odoni (2007).

1.3. Literature Review

The first thorough review of the TFM problem is provided by Odoni (1987). Subsequently, two leading research paths have emerged and diverged. The first has focused primarily on collaboration and equity in the context of single-resource approaches. These approaches are applicable for a single TFM program or multiple, nonconflicting TFM programs. The second research path has focused primarily on computational efficiency in the context of network-wide,

Table 5 Controlled Departure and Arrival Times According to Precedence RBS

Flight	CTD	FCA1	LGA
A	17:45	—	18:55
B	17:24	18:49	19:05
C	18:00	18:45	—
D	18:19	18:50	—

Table 6 Controlled Departure and Arrival Times According to Exemption RBS

Flight	CTD	FCA1	LGA
A	17:45	—	18:55
B	17:24	18:49	19:05
C	18:05	18:50	—
D	18:24	18:55	—

or multiresource, approaches to TFM, in which typically all airports and air sectors are placed under the FAA's control. Research into single resource approaches has gained more traction within the industry primarily because of (i) the inclusion of collaboration and equity considerations and (ii) computational tractability (i.e., computations involving a full day of flights for a single resource run quickly). Arguably, it is because of failures in these areas that research into network approaches has gained less traction. Indeed, it is only recently that network formulations have begun to incorporate equity or collaboration considerations.

For the single-resource TFM problem, deterministic, static-stochastic, and dynamic-stochastic versions of the problem were first formulated in the early 1990s (see Richetta and Odoni 1993, 1994; Terrab and Odoni 1993). More recent research has extended these models to incorporate collaboration and equity (see Ball et al. 2003; Vossen and Ball 2005; Kotnyek and Richetta 2006). Vossen et al. (2003) define a measure of equity for the single-resource TFM problem and calculate the inequity associated with flight exemptions. Chang et al. (2001) describe the collaborative decision making (CDM) approach with updated equity considerations that was incorporated into the FAA's GDP in the late 1990s. In a recent paper, Brennan (2007) describes how the CDM-enhanced GDP approach has been extended to the AFP.

On the multiresource side, Vranas, Bertsimas, and Odoni (1994) developed the first integer programming formulation for the multi-airport GDP. Bertsimas and Patterson (1998) extended this formulation to the full air traffic system using a novel variable definition. Subsequent research has focused primarily on computational efficiency and the incorporation of rerouting constraints (see Andreatta, Brunetta, and Guastalla 2000; Hoffman and Ball 2000; Bertsimas, Lulli, and Odoni 2011). The discussion of inequities inherent in a network formulation of the TFM problem by Lulli and Odoni (2007) provides a critical backdrop for our work. Bertsimas, Farias, and Trichakis (2011) show that under reasonable assumptions the theoretical price of fairness in TFM is bounded and typically quite low, which is consistent with our computational results. In recent and related work, Bertsimas and Gupta (2011) consider fairness and collaboration in the context of the nationwide TFM problem.

In this paper, we develop integer programming formulations for the multiresource TFM problem that incorporate fairness considerations. Unlike other network approaches, instead of including all airports and sectors, we restrict the problem to the coordination of multiple, conflicting TFM programs. We believe that by considering this restricted problem, our work

will help bridge the gap between the two divergent research paths described above.

1.4. Contributions

The contributions of this paper fall into four categories.

1. Demonstrating the potential inefficiencies of the TFM conflict resolutions approaches utilized in practice. The general TFM scheduling problem is NP-hard, so these inefficiencies are not surprising. Nonetheless, this result is not well understood. Most important, we demonstrate that these inefficiencies are not just theoretically plausible but are realized in historical scenarios.

2. Developing a fairness metric that extends to the multiresource setting, including analysis of the resulting fairness properties and relationship to current industry standards.

3. Developing two optimization approaches for coordinating TFM programs that balance the tradeoff between equity, as measured by category 2 above, and efficiency, as measured by aggregate system delay. The latter of these two models is computationally tractable for national-scale TFM problems.

4. Generating computational results and analysis using large-scale historical instances derived from 2007 data obtained from Flight Schedule Monitor, the tool used to manage these programs.

The structure of the paper follows these main points. In §2, we demonstrate the limitations of the current TFM conflict resolution approaches, discuss inherent fairness properties, and develop our fairness metric. In §3, we develop two integer programming formulations. Lastly, in §4, we report and discuss our computational results.

The starting point for our formulations is the model developed in Bertsimas and Patterson (1998) and the first-scheduled, first-served concept of fairness inherent in RBS, as described in §1.1. RBS has three salient features. First, it is algorithmically trivial to implement and has a linear running time with respect to the number of flight steps. Thus, the approach can be scaled to arbitrarily large problems. Second, for an isolated GDP or AFP, the RBS method always leads to a solution that minimizes system delay (Vossen and Ball 2005). Third, it maintains a first-scheduled, first-served invariant, which is the industry accepted notion of fairness endorsed by the primary stakeholders, that is, the FAA and the airlines. In particular, as should be apparent based on the example in §1.2, any multiresource extension of RBS will fail on a very important front: it will no longer provide delay-optimality guarantees as in the single-resource case. This is to be expected, because fairness may in general come at the expense of increased aggregate delays. The main modeling contribution of this paper

is precisely to address this deficiency. Specifically, we seek a formulation for fairness that has the following properties:

1. In the single-resource setting, it should reduce to (the accepted standard) RBS, which, as discussed above, is delay-optimal in this case;
2. Because there will typically be a tradeoff between aggregate system delay and any flight-based fairness criterion, the formulation should essentially consider a bicriterion approach that enables the study of the tradeoff curve between the two; and
3. The formulation should compare favorably to the approaches currently utilized in practice for the multiresource setting.

Using historical TFM scenarios, we demonstrate a computationally viable optimization formulation that satisfies all of these properties. We estimate that this model can reduce flight delays by 4% or more on some of the worst days, resulting in system-wide savings on the order of \$25 to \$50 million annually. The concepts and modeling approaches developed in our paper readily extend to the nationwide TFM problem, although we focus almost exclusively on the problems associated with coordinating GDPs and AFPs. We do so with the goal of having our work provide a bridge, both academically and practically, between current approaches and a long-term vision of nationwide TFM. That is, we hope that our work allows future TFM research, including our own, to build upon a foundation that has a high likelihood of being accepted in practice. Additionally, as the frequency and complexity of AFPs increases because of increasing en route congestion, we expect the inefficiencies we identify with current approaches to be exacerbated, providing further justification for an optimization-based approach.

2. Ration-by-Schedule and Fairness

As discussed in the introduction, understanding and incorporating industry-accepted views of fairness has been a significant roadblock to the implementation of optimization-based techniques for managing TFM programs. One of the more significant challenges is that the first-scheduled, first-served concept of fairness underlying RBS does not directly extend to the setting in which a single flight may interact with multiple TFM programs (e.g., a GDP plus one or more AFPs). With this in mind, we turn our attention to developing a measure of overall schedule fairness that (i) is consistent with first-scheduled, first-served in a single-resource environment and (ii) naturally extends to the setting in which there are interactions between TFM programs.

To provide additional context, we first illustrate problems with the multiresource RBS approaches utilized in practice to resolve conflicts between conflicting TFM programs (i.e., GDPs and AFPs). The main

advantage of these approaches is that they are simple extensions of RBS in the single-resource setting, and thus the resulting schedules are similar to the single-resource RBS schedules. Unfortunately, this simplicity can also lead to significant costs in terms of efficiency and therefore total delays. Next, we describe the properties we believe should underlie any measure of schedule fairness in a multi-resource setting, using simple examples to demonstrate the importance and significance of these properties. Lastly, we develop a robust measure of schedule fairness that incorporates the properties we outline. The purpose of this metric is to evaluate the relative fairness of competing scheduling approaches.

2.1. Problems with Multiresource Ration-by-Schedule

One downside of precedence RBS that is not a factor with the exemption alternative is that AFP capacities, specified in terms of controlled arrival rates, may be (and often are) violated. By examining the controlled schedule from the example in §1.2 (Table 5), we see that two flights (B and D) are scheduled to arrive at FCA1 simultaneously even though the controlled arrival rate was established at one flight every five minutes. It is difficult to measure the impact associated with this issue because, in practice, AFPs are constructed in a subjective fashion. That is, the parameters of each AFP, such as duration and arrival rate, are tweaked until the end result satisfies subjective criteria for safety. Additionally, with an AFP traffic flow is controlled through a line or region of air space that may be hundreds of miles long; thus, two flights that arrive at the same time may be very far apart geographically. Nonetheless, precedence RBS makes it difficult, if not impossible, to precisely control traffic flow through the air. Additionally, as air traffic congestion continues to increase, airspace controls are expected to become more common, exacerbating this problem.

A more significant issue with both multiresource RBS scheduling approaches is that either may lead to inefficient resource utilization. For instance, consider the planned flight schedule in Table 2 and the precedence RBS schedule in Table 5. In the precedence RBS schedule, we see that the 18:40 arrival slot for FCA1 is unused because flights C and D cannot depart earlier than planned. If we swap the order of flights A and B into LGA, flight B is then able to use the 18:40 slot, which allows flights C and D to depart on time (using the same capacity profile as the precedence RBS schedule). This sequence of exchanges reduces the total delay from 13 minutes (9 minutes for flight B and 4 minutes for flight D) to 10 minutes (all for flight A). Similarly, consider the exemption RBS schedule provided in Table 6. This

schedule results in 23 minutes of delay (9 minutes each for flights B and D, and 5 minutes for flight C). As with the precedence RBS example, the 18:40 slot into FCA1 is unused. If we swap the order of flights A and B into LGA, flight B is able to use the 18:40 slot, which allows flight C to depart on time and flight D to depart 4 minutes late, resulting in 14 minutes of delay (10 minutes for flight A and 4 minutes for flight B). **Note that the exemption RBS schedule results in more delay because, in contrast to the precedence RBS schedule, no AFP capacity violations are allowed.**

The last issue with the two approaches is that the expected RBS arrival order for certain resources may be violated based on the resolution of conflicting CTDs. In the precedence RBS example above, flight B was originally scheduled to arrive at FCA1 first but was instead scheduled second after resolution of the conflicting CTDs. Although the RBS order is violated in this case, it is likely not a fairness issue because LGA is a more congested resource along flight B's route. On the other hand, consider two flights, the first of which passes through a severely constrained FCA en route to a more mildly constrained arrival airport, the second just through the FCA. Assuming the GDP was implemented first, under either approach the GDP-based CTD will take priority for the first flight, allowing it to avoid the impact of the more severe AFP. The second flight will be affected only by the AFP and thus incur significantly greater, and therefore inequitable, delays. As should be apparent from this example, we can construct scenarios in which either multiresource RBS schedule is arbitrarily unfair.

2.2. Principles for Measuring Fairness

The challenge to incorporating fairness into the multiresource setting is that the link between original schedule order and delay optimality breaks down when some flights are affected by multiple TFM programs. Thus, in a multiresource setting we need to make a tradeoff between fairness relative to the original schedule order and efficiency in terms of total system delay. To find the appropriate tradeoff, we need a method of measuring the relative (un)fairness of competing schedules.

The concept of fairness is by nature subjective and often domain specific. Even within air traffic, there are many plausible ways to measure schedule fairness, each leading to different results. For example, in a single-resource setting, one measure of fairness is the number of slots a flight deviates from its initial order position (e.g., if a flight scheduled to arrive fourth is instead allocated the twelfth arrival slot, we would say that flight's schedule was unfair by eight positions). Unfortunately, in the multiresource setting, using position-based metrics without considering delay can lead to imbalances in the fairness

penalty incurred between resources. Other proposals include measuring schedule fairness by comparing average or maximum flight delays between airlines. These types of measures ignore variation in congestion along flight routes and thus are also problematic in the multiresource setting. In this section, we describe properties that we believe are critical to measuring fairness in the multiresource setting. These properties are motivated primarily as extensions of the successful properties of RBS in the single-resource environment. In the following section, we use these properties to obtain a multiresource fairness deviation metric.

PROPERTY 1. The measure of schedule fairness should be determined with reference to the original schedule ordering. Because of the success of RBS in the single-resource setting, the concept of first-scheduled, first-served has come to be widely accepted by airlines and the FAA.

PROPERTY 2. The measure of schedule fairness should be applicable to a single flight as well as the overall schedule. That is, the measure should be able to determine the amount by which each flight's schedule varies from first-scheduled, first-served.

PROPERTY 3. The unit of fairness deviation and its relative magnitude should be consistent between resources. In a single-resource setting, position-based deviation is an accepted measure of fairness deviation. In the multiresource setting, this is confounded due to varying congestion levels between resources. An eight-position delay (going from fourth to twelfth) could mean 30 minutes of delay in a low-capacity airport, but only 10 minutes of delay in a higher capacity one.

PROPERTY 4. No flight should expect to receive less delay than what would be caused by the most congested resource along its route. That is, there should be a fairness penalty only if a flight incurs more delay than its original schedule order would indicate for each of the resources along its route.

PROPERTY 5. The measure of a flight's deviation from the original schedule should be calculated relative to the total delay assigned to the flight (ground delay plus air delay), not intermediate arrival times into controlled resources. This property is relevant if the scheduling approach allows both ground delays (by assigning CTDs) and en route delays (by mandating air speed reductions or arrival queuing) to be assigned. In practice, the schedule created by the FAA using RBS assumes that a flight will incur no delays en route and only assigns ground delays through CTDs. Airborne delays are subsequently managed by air traffic controllers en route or at the arrival airport. Network TFM models, such as the one described in Bertsimas and Patterson (1998), consider both of these problems simultaneously in order to improve efficiency and predictability.

2.3. Time-Order Deviation Metric

With these properties in mind, we now develop a measure for evaluating the unfairness of a controlled schedule. First, we use features of both the planned (predisruption) schedule and the controlled (postdisruption) schedule to determine a fair delay threshold for each flight, which we refer to as the *maximum expected delay*. Next, we calculate the *time-order deviation* for each flight as the amount by which the flight's delay in the controlled schedule exceeds this threshold.

We first determine the fair delay threshold for the case in which a flight, f , utilizes just one controlled resource. We let j be the flight's position in the planned arrival ordering for this resource. For example, if there is a GDP at LGA and flight f was planned to be the third arrival into LGA, j would equal three. We define flight f 's *expected delay* for the controlled resource as the difference between the j th controlled arrival time for the resource and the planned arrival time of the flight. Note that the order of flights could be swapped in the controlled schedule; thus, the j th controlled arrival for the resource might not be the same as the j th planned arrival. Continuing the example of a GDP at LGA, if flight f was planned to be the third flight to arrive into LGA at 19:15, and the third controlled arrival into LGA is scheduled for 19:20, we would say that flight f has a five-minute expected delay into LGA. As long as flight f incurs no more than five minutes of delay, we would consider the controlled schedule to be fair from flight f 's perspective.

The case in which a flight, f , utilizes multiple controlled resources is a bit more complicated. For each of the controlled resources flight f is scheduled to utilize, we could calculate the expected delay as in the example above. Each of these expected delay values would represent a fair delay threshold, assuming that flight f utilized no other controlled resources along its route. Thus, in the case in which a flight utilizes multiple controlled resources, we set the fair delay threshold to the maximum of the expected delay value across these controlled resources. As long as flight f incurs no more delay than this maximum expected delay, we consider the controlled schedule to be fair from flight f 's perspective. One could argue that the fair delay threshold should be higher than this value because the flight uses multiple controlled resources, but it clearly should not be any lower. Note that this general definition applies to flights that utilize just one controlled resource because for these cases the maximum expected delay equals the expected delay for the single resource.

Using these definitions, we define each flight's time-order deviation as the amount by which its total

delay in the controlled schedule exceeds the maximum expected delay along its route. In the event that the maximum expected delay exceeds the flight's total delay, we set the time-order deviation equal to zero. That is, a schedule is not fairer if a flight arrives earlier than expected, even though this might reduce the overall system delay. Time-order deviation can be considered a generalization of deviation from the ideal RBS allocation used to measure exemption bias in Vossen et al. (2003). The key differences are (i) time-order deviation is measured relative to the most congested resource along a flight's route and (ii) time-order deviation is measured relative to a feasible, controlled schedule instead of a potentially infeasible, idealized allocation. Next, we consider an example in which we calculate the time-order deviation for a flight that is affected by a GDP at LGA and an AFP at FCA1.

2.3.1. Time-Order Deviation Example. Consider a flight planned to depart from Boston Logan International Airport (BOS) at 18:00, arrive at FCA1 at 18:45, and arrive into LGA at 19:15. Prior to schedule disruption, the flight is planned to be the fourth arrival into FCA1 and the third arrival into LGA. Based on a GDP at LGA and an AFP at FCA1, the flight is subsequently given a CTD of 18:25 (corresponding to 25 minutes of ground delay). To calculate the time-order deviation for this flight, we need to additionally know the order of arrivals into FCA1 and LGA in the controlled schedule. Based on the controlled arrival orderings listed in Tables 7 and 8, we can calculate the time-order deviation as follows. First, we calculate the flight's expected delay into FCA1 as the time of the fourth controlled arrival into FCA1 (18:55) minus the flight's planned arrival time into FCA1 (18:45), which equals 10 minutes. Next, we calculate

Table 7 Controlled Flight Order for FCA1

Order	FCA1 Arrival
1	18:35
2	18:45
3	18:50
4	18:55
5	19:00
6*	19:05

Table 8 Controlled Flight Order for LGA

Order	LGA arrival
1	19:00
2	19:10
3	19:20
4	19:30
5*	19:40

the flight's expected delay into LGA as the time of the third controlled arrival into LGA (19:20) minus the flight's planned arrival time into LGA (19:15), which equals five minutes. The referenced arrival times in Tables 7 and 8 are highlighted in **bold**. The maximum expected delay for the flight is the 10 minutes from FCA1. In the controlled schedule, the total delay for the flight is 25 minutes (the difference between the 18:25 CTD and the 18:00 planned departure time). Thus, the time-order deviation for the flight is 15 minutes (25 minutes of total delay minus 10 minutes of maximum expected delay). Thus, we would say that 60% of the delay assigned to this flight is unfair as measured by time-order deviation. In Tables 7 and 8, the rows corresponding to the controlled arrival times for the referenced flight have been marked with an asterisk (*) although they are not used directly in the calculation of the flight's time-order deviation.

We define the time-order deviation for a controlled schedule as the sum of the time-order deviations for each flight represented in the schedule. If we divide the total time-order deviation by the total delay assigned, the result describes the average percentage of unfair delay assigned to each flight (relative to the individual fair delay thresholds). The time-order deviation of individual flights may vary significantly from the average, so in our results in §4.4 we also consider the distribution of flight delay. As expected, time-order deviation satisfies all of the principles laid out in the previous section. That is, (i) time-order deviation is calculated relative to the original schedule order, (ii) the measure can be applied to each flight in the controlled schedule, (iii) the unit of measure (i.e., time) is consistent between resources, (iv) the measure is calculated relative to the most restricted resource along each flight's route (i.e., relative to the maximum expected delay), and (v) the measure is based on the total delay and not intermediate arrival times. Note that for a single controlled resource, or for a set of independent controlled resources (such as multiple GDPs), the time-order deviation metric achieves zero if the controlled schedule matches the schedule that results from independent RBS allocations for each controlled resource. For this independent resource case, among the set of delay-minimal schedules and because of the uniqueness of the first-scheduled, first-served solution, the time-order deviation metric achieves zero if and only if the controlled schedule matches the RBS allocations.

3. Optimization Approaches

In this section, we describe two integer programming formulations, the solutions to which describe the ground holding that should be assigned to each flight. Each formulation allows for the flexible trade-off between a delay term and a fairness term in the

minimization objective. In the first model, the fairness term is a convex approximation of the fairness metric developed in the previous section. We call this the time-order deviation approximation (TODA) model. In the second model, we use an exponentially growing delay penalty to enforce fairness. This approach has considerable computational advantages yet sacrifices little in terms of fairness achieved according to time-order deviation. We refer to this model as the ration-by-schedule exponential penalty (RBS-EP) model.

In §3.1, we develop the common notation as well as define the input data used in both our formulations. In §3.2, we provide the portion of the optimization formulation that is common to both the TODA and RBS-EP models. Section 3.3 provides the formulation for the TODA model, and §3.4 the formulation for the RBS-EP model. We discuss some issues related to practical implementation in §3.5.

3.1. Data and Notation

We consider a set of discretized time intervals $\mathcal{T} = \{0, \dots, T-1\}$, where T represents the end of the day and each interval is defined to have equal duration, typically either five minutes or 15 minutes. We consider a set of controlled resources, \mathcal{R} , that will typically include arrival airports (for GDPs) and FCAs (for AFPs). System resources that are not capacity controlled provide no binding constraints on the system and are excluded from \mathcal{R} . For each resource, $r \in \mathcal{R}$, and each time interval, $t \in \mathcal{T}$, we specify a capacity of b_{rt} , which can be either the allowable arrival rate or the maximum number of flights allowed to occupy the resource during the interval. For GDPs and AFPs, resource capacities are specified in terms of the allowable arrival rate.

Additionally, we consider a set of flight legs, \mathcal{F} . For each flight leg, $f \in \mathcal{F}$, we define the controlled flight plan to be the sequence of controlled resources scheduled to be utilized over the course of the flight. For instance, consider the flight from Portland International Airport (PWM) to New York John F. Kennedy Airport (JFK) depicted in Figure 2, with TFM programs in place at FCA1 and JFK. For this flight, the controlled flight plan would be a sequence containing FCA1 followed by the arrival resource for JFK. Notationally, we let $|f|$ represent the number of steps

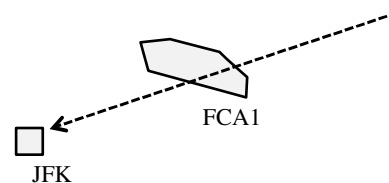


Figure 2 PWM to JFK Flight Path Intersecting Two Controlled Resources, FCA1 and JFK

Table 9 Data Values for PWM to JFK Controlled Flight Plan Based on a 07:00 Scheduled Departure

Scheduled time	i	r	α	δ
07:35	1	FCA1	31	2
08:15	2	JFK	39	1

in the controlled flight plan for flight f and use the shorthand $\mathcal{J}(f)$ to represent the set of step indices $\{1, \dots, |f|\}$. For each step in the controlled flight plan, in addition to the resource, r , we must specify the earliest start time, α , and the processing time, δ . That is, $\alpha \in \mathcal{T}$ represents the first time interval at which the step can be scheduled and $\delta \in \mathbb{N}^+$ the number of time intervals the step needs to be processed (i.e., landing time at an arrival airport or dwell time in an occupancy-controlled FCA). Notationally, we let $r(f, i)$, $\alpha(f, i)$, and $\delta(f, i)$ refer to the appropriate values for step i of the flight plan for flight f . In our formulation, $\alpha(f, i+1) - \alpha(f, i)$ represents the minimum number of time intervals between the starts of steps i and $i+1$. Thus, we require $\alpha(f, i) + \delta(f, i)$ to be less than or equal, rather than strictly equal, to $\alpha(f, i+1)$. For example, if the resources for two sequential steps are not geographically adjacent, $\alpha(f, i+1) - \alpha(f, i) - \delta(f, i)$ would represent the travel time between the boundaries of the two resources. In Table 9, we provide sample values for these fields based on the example described above (see Figure 2), with five minute time intervals starting at 05:00. In our example, the referenced flight is scheduled to occupy FCA1 for 10 minutes en route to JFK.

For each resource, r , we assume there is a preferred ordering of tasks (i.e., flight steps) corresponding to the original schedule. That is, for resource r we would prefer to start the task indexed by j before the task indexed by $j+1$, where each task corresponds to a flight step, (f, i) . Using this notation, we let $j(f, i)$ represent the task index of flight step (f, i) for the corresponding resource, $r(f, i)$. Additionally, we let (r, j) represent the time interval task j would be assigned based on performing single-resource RBS for r .

Summarizing the above, we have the following model inputs:

- \mathcal{T} = set of discrete time intervals;
- \mathcal{R} = set of capacity-controlled resources;
- b_{rt} = capacity of resource r over time interval t ;
- \mathcal{F} = set of flights;
- $|f|$ = number of steps in controlled flight plan for flight f ;
- $\mathcal{J}(f)$ = set of step indices in controlled flight plan for flight f ;
- $r(f, i)$ = resource required by flight step i for flight f ;
- $\alpha(f, i)$ = earliest start time for flight step i for flight f ;

$\delta(f, i)$ = processing time of flight step i for flight f ;

$J(r)$ = number of tasks (i.e., flight steps) assigned to resource r ;

$\mathcal{J}(r)$ = set of task indices $\{1, \dots, J(r) - 1\}$;

$j(f, i)$ = the task index of flight step i for flight f ; and

$\text{RBS}(r, j)$ = RBS start interval for task j of resource r .

3.2. Model Foundation

In this section, we describe the components of the deterministic, multiresource TFM formulation that provide the foundation for the two models we develop. This formulation is derived from the Bertsimas and Patterson (1998) nationwide TFM model.

3.2.1. Decision Variables. For both formulations, we use the following variable definitions:

$$y_{fit} = \begin{cases} 1, & \text{if flight plan step } i \text{ for flight } f \text{ has started} \\ & \text{by time } t \text{ and} \\ 0 & \text{otherwise.} \end{cases}$$

3.2.2. Constraints. We first ensure that the sequence $[y_{fi0}, \dots, y_{fi(T-1)}]$, which we refer to as $[y_{fi}]$, is monotonically increasing:

$$y_{fit} \leq y_{fi(t+1)} \quad \forall f \in \mathcal{F}, \forall i \in \mathcal{J}(f), \forall t \in \{0, \dots, T-2\}. \quad (1)$$

Next, we guarantee that each flight step is scheduled and that no flight step is scheduled before its minimum start time:

$$y_{fi(T-1)} = 1 \quad \forall f \in \mathcal{F}, \forall i \in \mathcal{J}(f). \quad (2)$$

$$y_{fi(\alpha(f, i)-1)} = 0 \quad \forall f \in \mathcal{F}, \forall i \in \mathcal{J}(f)$$

$$\text{s.t. } \alpha(f, i) > 0. \quad (3)$$

We also enforce the appropriate order between flight steps in a controlled flight plan as follows:

$$y_{f(i+1)t} = y_{fi(t-\alpha(f, i+1)+\alpha(f, i))} \quad \forall f \in \mathcal{F}, \forall i \in \mathcal{J}(f) \setminus \{|f|\}. \quad (4)$$

We require strict equality in (4) to disallow allocation of airborne delays. To allow the allocation of airborne delays, we could replace this with an inequality (i.e., $y_{f(i+1)t} \leq \dots$). The last set of constraints is to ensure that resource capacities are not violated:

$$\sum_{\{(f, i): r(f, i)=r\}} (y_{fit} - y_{fi(t-\delta(f, i))}) \leq b_{rt} \quad \forall r \in \mathcal{R}, \forall t \in \mathcal{T}. \quad (5)$$

Note that $y_{fit} - y_{fi(t-\delta(f, i))}$ represents whether flight f is performing flight plan step i at time t .

3.2.3. Objective Function. The delay term in the objective function of each formulation represents the aggregate costs associated with flight delay, which we model as follows. First, we note that the start time of flight plan step i for flight f , $s(f, i)$, can be written as

$$s(f, i) = T - \sum_{t=0}^{T-1} y_{fit}.$$

The total delay for flight f , $d(f)$, is equivalent to the delay accumulated through the last step in the flight plan, $|f|$, which can be written as

$$d(f) = s(f, |f|) - \alpha(f, |f|).$$

In the base formulation, the objective is to minimize total delay:

$$\min \sum_{f \in \mathcal{F}} d(f). \quad (6)$$

Constraints (4) ensure that the total delay for flight f is equivalent to the delay assigned before the first step in the controlled flight plan, which allows us to allocate all of the flight delay as ground holding.

3.3. Time-Order Deviation Approximation (TODA) Model

Using the notation described in the previous section, we first provide the mathematical definition of time-order deviation. Letting $\hat{s}(r, j)$ represent the start time for the j th flight step to utilize resource r in the controlled schedule, we have

(Maximum Expected Delay)

$$\text{MED}(f) \triangleq \max_{i \in \mathcal{J}(f)} \{\hat{s}(r(f, i), j(f, i)) - \alpha(f, i)\}; \text{ and } \quad (7)$$

(Time-Order Deviation)

$$\text{TOD}(f) \triangleq (d(f) - \text{MED}(f))^+. \quad (8)$$

In Equation (8), $d(f)$ represents the total delay assigned to flight f , as in the objective function for the base formulation (6). There are two challenges to calculating time-order deviation within a mathematical programming model. The first is that to calculate expected delay, we need the sorted list of scheduled start times for each resource. That is, in addition to maintaining a view of the schedule from each flight's perspective, we also need to maintain a view of the schedule from each resource's perspective. We address this by creating schedule variables that maintain a fixed relative order for each resource and are bound to the original flight-centric schedule variables. The second challenge is that time-order deviation is a nonconvex function because of the inner maximum from Equation (7). Thus, time-order deviation cannot be represented directly in a

linear minimization objective. Instead, we approximate time-order deviation by replacing the maximum over all flight steps in Equation (7) with an average over the flight steps that we estimate a priori will lead to the most delay. We do so by computing which steps, i , would be assigned the most delay based on independent RBS allocations performed for each resource. This gives us an estimate of congestion from capacity-demand imbalances, although it ignores delays introduced due to interactions between resources.

3.3.1. Model Adjustments. We first define the ordered auxiliary variables described above:

$$u_{rjt} = \begin{cases} 1, & \text{if } j \text{ tasks for resource } r \text{ have been} \\ & \text{scheduled to start by time } t \text{ and} \\ 0 & \text{otherwise.} \end{cases}$$

Based on this definition, $(u_{rjt} - u_{rj(t-1)})$ will indicate when task j of resource r starts in the optimized schedule. Note that this might or might not be the same as the start time of the task originally scheduled to occupy position j .

Next, we add the following constraints to the model to ensure that the variables maintain the definition above:

$$u_{rjt} \leq u_{rj(t+1)} \quad \forall r \in \mathcal{R}, \forall j \in \mathcal{J}(r), \quad \forall t \in \{0, \dots, T-2\}, \text{ and } \quad (9)$$

$$u_{rj(T-1)} = 1 \quad \forall r \in \mathcal{R}, \forall j \in \{1, \dots, J(r)\}. \quad (10)$$

Constraints (9) and (10) ensure that the sequence of ordered auxiliary variables $[u_{rj}]$ maintains the same monotonically increasing form as the sequence of flight step variables $[y_{fi}]$. We also need to ensure that the appropriate order for the auxiliary variables is maintained; that is, task $(j+1)$ cannot start before task j :

$$u_{r(j+1)t} \leq u_{rjt} \quad \forall r \in \mathcal{R}, \forall j \in \mathcal{J}(r), \forall t \in \mathcal{T}. \quad (11)$$

The last, and most important, set of constraints ensures that, by each interval, the number of scheduled flights according to the ordered auxiliary variables and the flight step variables coincide:

$$\sum_{j=1}^{J(r)} u_{rjt} = \sum_{\{(f, i): r(f, i)=r\}} y_{fit} \quad \forall r \in \mathcal{R}, \forall t \in \mathcal{T}. \quad (12)$$

That is, constraints (12) ensure that when a flight step is scheduled within an interval, one of the sequences of ordered auxiliary variables must flip from zero to one in that same interval.

With these definitions in mind, we can calculate the expected delay (ED) for flight step (f, i) and resource $r(f, i)$, denoted $\text{ED}(f, i)$:

$$\text{ED}(f, i) = \sum_{\alpha(f, i)}^{T-1} (1 - u_{r(f, i)j(f, i)t}). \quad (13)$$

The right-hand side measures the number of intervals from the earliest start time for flight step (f, i) until the $j(f, i)$ th task starts for resource $r(f, i)$.

As discussed in the introduction to this section, we estimate which resources for each flight f will maximize expected delay by computing which steps i would be assigned the most delay according to independent RBS allocations. For flight f , we denote the maximum RBS delay as $d_{\text{MAX}}^{\text{RBS}}(f)$ and the set of steps achieving the maximum RBS delay as $\mathcal{J}_{\text{MAX}}(f)$:

$$\begin{aligned} d^{\text{RBS}}(f, i) &= \text{RBS}(r(f, i), j(f, i)) - \alpha(f, i); \\ d_{\text{MAX}}^{\text{RBS}}(f) &= \max_{i \in \mathcal{J}(f)} \{d^{\text{RBS}}(f, i)\}; \quad \text{and} \end{aligned} \quad (14)$$

$$\mathcal{J}_{\text{MAX}}(f) = \{i \in \mathcal{J}(f) : d^{\text{RBS}}(f, i) = d_{\text{MAX}}^{\text{RBS}}(f)\}.$$

We now have the tools needed to describe the fairness term we add to (6) to calculate the approximate time-order deviation in our TODA objective function:

$$\min \sum_{f \in \mathcal{F}} d(f) + \psi \left(d(f) - \sum_{i \in \mathcal{J}_{\text{MAX}}(f)} \frac{\text{ED}(f, i)}{|\mathcal{J}_{\text{MAX}}(f)|} \right)^+. \quad (15)$$

Within the sum, the second term represents the approximate time-order deviation scaled by a factor of $\psi > 0$, which controls the tradeoff between system delay and approximate time-order deviation. Within the approximate time-order deviation term, the inner sum calculates the average expected delay across the flight steps that achieved maximum RBS delay according to independent RBS allocations. When the set $\mathcal{J}_{\text{MAX}}(f)$ equals the steps that achieve the maximum expected delay in the optimized schedule, our approximate time-order deviation will equal the true time-order deviation as described in §2.3. Otherwise, the average expected delay across $\mathcal{J}_{\text{MAX}}(f)$ will be less than the maximum expected delay, and the approximate time-order deviation will be strictly larger than the true time-order deviation. The $(\cdot)^+$ ensures that we add the approximate time-order deviation to our objective only if the total delay exceeds this average expected delay.

3.4. Ration-by-Schedule Exponential Penalty (RBS-EP) Model

Unlike the TODA model, the RBS-EP model requires the introduction of no new variables or constraints to the foundational model described in §3.2. The only change required is modifying the functional form of the objective function. The intuition behind the RBS-EP model has two parts. The first is that no flight should expect to incur less delay than its worst-case RBS delay, $d_{\text{MAX}}^{\text{RBS}}(f)$ as defined in Equation (14) in the previous section. But because of interactions between resources, it is unlikely that each flight will be able to achieve this exactly. So to provide flexibility, we penalize each interval of delay beyond $d_{\text{MAX}}^{\text{RBS}}(f)$ by an exponentially increasing amount.

3.4.1. Model Adjustments. One of the nice properties of discrete scheduling models is that we can associate different objective coefficients with each possible start time for a task. To achieve an exponentially increasing penalty, we need only to determine the appropriate coefficients for each flight and potential start interval. Thus, we let c_{ft} be the coefficient associated with the last step of flight f starting at time t :

$$\begin{aligned} c_{ft} &= \min \left\{ t - \alpha(f, |f|), d_{\text{MAX}}^{\text{RBS}}(f) \right\} \\ &\quad + \sum_{\epsilon=1}^{t - \alpha(f, |f|) - d_{\text{MAX}}^{\text{RBS}}(f)} \lambda^{\epsilon}. \end{aligned} \quad (16)$$

Based on the definition above, we have $c_{ft} - c_{f(t-1)} = \lambda^{(t - \alpha(f, |f|) - d_{\text{MAX}}^{\text{RBS}}(f))}$ assuming $t > \alpha(f, |f|) + d_{\text{MAX}}^{\text{RBS}}(f)$. Thus, assuming $\lambda > 1$, the incremental cost of each additional interval of delay beyond $d_{\text{MAX}}^{\text{RBS}}(f)$ increases exponentially. A sample plot of this cost function is represented in Figure 3 for $\lambda = 2$, $d_{\text{MAX}}^{\text{RBS}}(f) = 4$, and $\alpha(f, |f|) = 0$.

With the cost coefficients c_{ft} defined as above, the objective function for the RBS-EP model is

$$\min \sum_{f \in \mathcal{F}} \left[\sum_{t=\alpha(f, |f|)}^{T-1} c_{ft} (y_{f|f|t} - y_{f|f|(t-1)}) \right]. \quad (17)$$

The difference $(y_{f|f|t} - y_{f|f|(t-1)})$ equals 1 if and only if the last step for flight f begins at time t , thus applying a penalty of c_{ft} as desired. In the exponential penalty model, the base of the exponent, $\lambda > 1.0$, used in defining c_{ft} , implicitly controls the tradeoff between aggregate system delay and fairness.

3.5. Integration Issues

As noted, these computational models build on Bertsimas and Patterson (1998) and Andreatta, Brunetta, and Guastalla (2000). Beyond the fairness considerations, however, there are two key differences in the approach we outline. First, in each of the referenced models, flights can be assigned air delay en route. Because of the limited number of AFPs implemented in practice, and the deterministic nature of

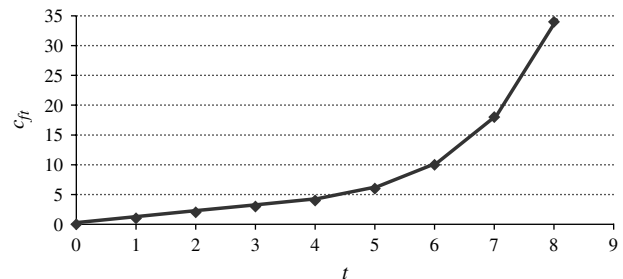


Figure 3 Plot of Delay Cost Function for $\lambda = 2$, $d_{\text{MAX}}^{\text{RBS}}(f) = 4$, and $\alpha(f, |f|) = 0$

our formulations, allowing air delay provides little value for the historical scenarios we consider in our results. Thus, to simplify exposition of the model as well as maintain consistency with current practice, we consider only ground delay. Additionally, in the referenced models, planned aircraft connections between flight legs are maintained in the controlled schedule (i.e., planned aircraft connections are represented as constraints in the formulation). Our models do not include connectivity constraints between flight legs. Note that both of our models could include these constraints and remain entirely consistent; thus, it is an explicit *modeling choice* to omit them. We have made the decision to exclude constraints of this type because, again, this change leads to an approach that is consistent with current practice. Most important, our models can utilize the same inputs as existing TFM programs, allowing for direct comparison as well as easier integration. Additionally, because of each airline's ability to swap aircrafts and cancel flights, it is unclear whether *strict* connectivity constraints are in the best interest of the airlines. Including these constraints increases the amount of delay assigned, under the assumption that airlines have less flexibility to respond than they do in practice. On the other hand, excluding these constraints leaves the full burden of resolving infeasibilities to the airlines. We believe that understanding the impact of aircraft connectivity in this context is an important open research question.

Another integration consideration is how these approaches fit into the three-stage collaborative decision making (CDM) framework described in §1.1. The key feature to note is that the output of our models can be easily translated into a slot assignment for the corresponding programs. In this sense, we maintain the same output format as that of existing approaches (corresponding to the first stage in the CDM process). To determine a single program to manage each flight (for stages two and three of the CDM process), we could simply choose the program that would be assigned under the current approaches. Alternatively, we propose that each flight be assigned to the program that maximizes the expected delay in the resulting schedule, as defined in Equation (7) in §3.3. Using this approach, each flight would be controlled by the program corresponding to the resource that creates the most congestion along its route.

A level of complexity not considered in our work is the dynamic nature of TFM programs. Over the course of the day, TFM programs are created and modified as more precise weather forecasts are revealed. These dynamics create challenges for any deterministic allocation approach (including RBS) because they introduce potential inequities and inefficiencies into the system. Effectively addressing these

challenges remains an interesting, important, and, to our knowledge, open research question.

4. Computational Results

Here we provide computational experiments to demonstrate the practical value of the RBS-EP model. We highlight three key results from our historical scenarios. The first is that under a conservative comparison between RBS-EP and current practice, the RBS-EP model improves efficiency, as measured by total delays, while maintaining equivalent levels of equity. The second is that the RBS-EP model closely tracks the tighter TODA approximation of the efficient frontier between aggregate delay and fairness, calculated according to our time-order deviation metric. Lastly, the RBS-EP model is computationally efficient, allowing for the solution of even complex, national-scale problems within reasonable computing times.

4.1. An Apples-to-Apples Comparison

One challenge in comparing our optimization-based approaches to current approaches is that precedence RBS allows the specified capacities to be violated (as discussed in §2.1). An optimization-based approach, on the other hand, ensures that all resource capacity constraints are strictly satisfied. Thus, if the same capacities are utilized as inputs to both procedures, precedence RBS would typically perform better because of its ability to arbitrarily exceed FCA capacity constraints (and the inability of our optimization-based approaches to do so). In most cases, exemption RBS does not exhibit this same characteristic because exempted flights reduce the effective capacity of future programs.

To level the playing field when comparing to precedence RBS, we first perform a precedence RBS allocation. Next, we compress the precedence RBS schedule, attempting to fill any gaps created in the resource schedules. For resources and time intervals for which the resulting allocation exceeds the specified capacity, we update the corresponding capacity as an input into each optimization-based approach. By updating the capacity, we ensure that our optimization-based approaches do not exceed the initial capacity *any more than* the precedence RBS schedule. For instance, based on the four flight example in §2.1 and five minute discretization intervals, we would increase the capacity of FCA1 to two flights for the five-minute interval from 18:45 through 18:49, keeping the capacity at one for all other intervals. Although this leads to a fairer comparison between the two approaches, the playing field is still tilted toward precedence RBS. Because of the inherent limitations of precedence RBS, we can only perform comparisons for the capacity allocations that directly correspond to a precedence RBS schedule. Fortunately, as demonstrated in §2.1,

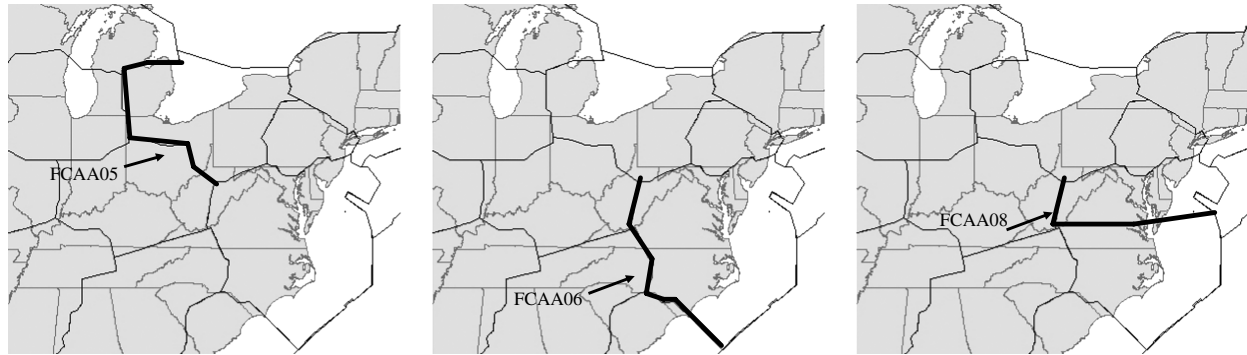


Figure 4 Depiction of Airspace Boundaries for FCAA05, FCAA06, and FCAA08 (Federal Aviation Administration 2006)

this still leaves some inefficiencies that optimization-based approaches are capable of exploiting.

4.2. Construction of Historical Scenarios

To construct each of our scenarios, we start with flight schedule data that correspond to a single day of relatively clear weather operations (April 23, 2007). This schedule includes estimated entry and exit times for each sector along each flight's route. The schedule data were obtained from Flight Schedule Monitor, the TFM decision support tool developed for the FAA by Metron Aviation (Metron Aviation 2009a). For purposes of our experiments, we treat this schedule as representing the Official Airline Guide (i.e., the planned airline flight schedules). Thus, the defining characteristics of each scenario are the set of controlled resources and corresponding capacities.

To construct the capacity reduction scenarios, we use historical TFM program data, also obtained from Metron Aviation's Flight Schedule Monitor. These data include reporting times, effective times and durations, and TFM program capacities for each 15-minute interval. From these data, we choose 10 representative days on which both GDPs and AFPs were implemented. For each of these days, we create two scenarios, one to reproduce historical behavior and a second to analyze the hypothetical impact of further reductions in FCA capacities. To create the historical scenarios, we reduce all hourly arrival capacities by 7.5% relative to the historical data. This reduction compensates for our clear weather day, April 23, 2007, having fewer flight operations than days during the summer (when all of the capacity reduction scenarios occurred). We calculated this reduction level by comparing flight frequencies in the Airline On-Time Performance Database, which includes flights for the 20 largest domestic airlines. For the hypothetical scenarios, we reduce airport capacities by 7.5% and FCA capacities by 25%, both relative to the historical data. We utilize these hypothetical scenarios to understand how efficiency improvements might change as en route congestion increases and AFPs are used more

heavily going forward. Each of the approaches utilized in practice is sensitive to the order of program implementation; thus, we use the historical reporting times to determine this order, reproducing historical behavior as accurately as possible.

The historical AFPs we utilize affect traffic heading into the Northeast corridor through one or more of the boundary-based flow constrained areas: FCAA05, FCAA06, and FCAA08. Figure 4 depicts each of these boundaries. Because our schedule data include only sector entry and exit times, in our scenarios we replace FCAA08 with FCAA06. Both FCAA06 and FCAA08 are used to address weather in the Ohio Valley region or in the Washington D.C. airspace (Federal Aviation Administration 2006). In Table 10, we report scenario details for the 10 days of capacity reductions we overlay on the clear weather flight schedule. The Conflicts column reports the percentage of flights affected by more than one TFM program. By construction, each of these values is the same for the historical and hypothetical scenarios described above. For each of the scheduling approaches we test, including the two multiresource RBS approaches, we discretize time into five-minute intervals.

Table 10 Details on Controlled Resources for Historical and Hypothetical Scenarios

Date	Flights	Airports	FCAs	Conflicts (%)
5/2/2007	1,858	(2) LGA SFO	A06	4.3
5/9/2007	1,572	(2) IAD JFK	A05	8.7
6/19/2007	5,191	(8) ATL DCA EWR IAD JFK LGA SFO PHL	A05 A06	15.6
6/27/2007	4,682	(4) CYYZ JFK LGA ORD MDW	A05 A06	7.5
6/28/2007	3,583	(5) EWR IAD JFK LGA SFO	A05 A06	15.7
7/5/2007	2,585	(2) CYYZ EWR	A05 A06	5.1
7/16/2007	1,858	(2) LGA SFO	A06	4.3
7/18/2007	3,705	(5) EWR JFK LGA PHL SFO	A05 A06	12.5
7/27/2007	3,944	(4) EWR LGA ORD SFO	A05 A06	6.6
9/27/2007	3,953	(6) ATL CYYZ EWR JFK LGA PHL	A05	5.7

4.3. The Tradeoff Between Efficiency and Fairness

In this section, we demonstrate the tradeoff between efficiency, as measured by aggregate delay, and fairness, as measured by the time-order deviation of the resulting schedule for each of the 20 scenarios described in the previous section (10 historical and 10 hypothetical). We create tradeoff curves by adjusting ψ and λ , the parameters that control the relative tradeoff for each of the optimization-based approaches developed in §3. First, we compare the TODA model to the RBS-EP model for the less complex scenarios. Because the TODA model is not computationally tractable for the more complex scenarios, we evaluate only the RBS-EP model for these problems.

We employ two computational heuristics to solve the TODA and RBS-EP optimization problems. First, because we discretize time into five minute intervals, it is computationally intractable to allow each flight to be delayed indefinitely. Doing so would result in well over a million binary decision variables for some of our instances. Instead, we restrict the amount of allowable delay on a flight-by-flight basis. Specifically, we allow each flight f to be assigned up to $d_{\text{MAX}}^{\text{RBS}}(f)$ plus an additional 15 or 30 minutes of delay, where $d_{\text{MAX}}^{\text{RBS}}(f)$ is the maximum independent RBS delay allocation as defined in Equation (14) in §2.3. Although there is suboptimality associated with this approach, the resulting schedule is, by construction, quite fair because it is close to the accepted RBS allocation. Second, we use a greedy integer rounding heuristic to convert each solution of the linear relaxation into a feasible flight schedule. We do so by greedily scheduling flights in order, based on the relaxed start time to the first step in each flight plan, $s(f, 1)$. This heuristic ensures that after solving the root node relaxation

during branch-and-bound search, we always have a good feasible solution. This is critically important in the TFM setting, in which we must be able to guarantee a solution in a relatively short amount of time (preferably one minute or less).

In Figures 5 and 6, we compare the tradeoff curves generated by the TODA model to the tradeoff curves generated by the RBS-EP model. In these plots, each point represents a schedule generated using a specific value of ψ or λ , plotting average flight delay in the corresponding schedule against the percentage of unfair delay as measured by time-order deviation. For each model, we allow flights to be assigned up to $d_{\text{MAX}}^{\text{RBS}}(f)$ plus 30 minutes of delay. In general, we find that the RBS-EP model closely tracks the approximate efficient frontier between fairness and delay as estimated by the TODA model. The RBS-EP model does not, however, allow us to fully explore the lower end of this curve. For more complex scenarios, determining a baseline according to the TODA model is not computationally tractable, although we have verified this general relationship on smaller constructed scenarios outside of the ones shown here. When there is more significant network-based congestion, there is typically a small gap between these two curves. Even though the RBS-EP model does not directly minimize time-order deviation or its approximation, there is a fairly consistent trend between an increasing λ , the base of the exponential penalty, and a decreasing time-order deviation of the resulting schedule. That is, by simply adjusting the functional form of the delay penalty, we have created a model that closely tracks the more complex time-order deviation metric.

One thing to note in these charts is that a relatively large benefit in terms of time-order deviation

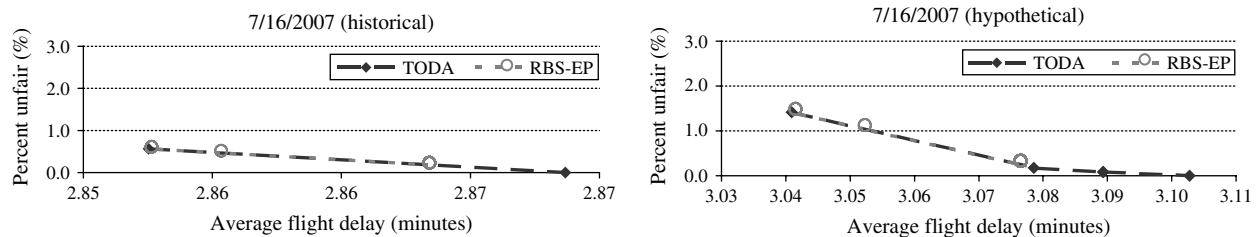


Figure 5 Historical and Hypothetical Precedence RBS Scenarios for 7/16/2007

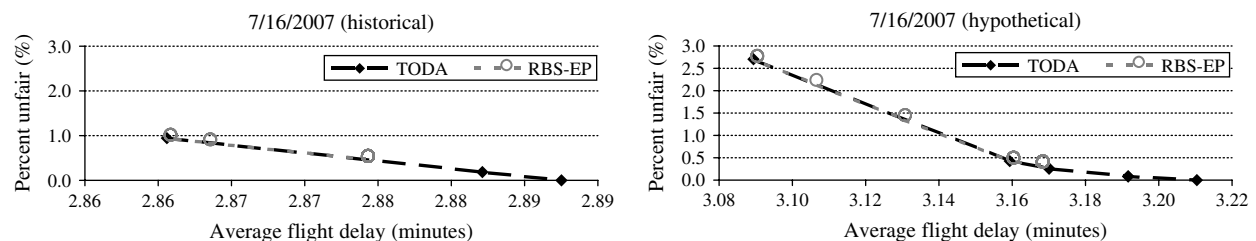


Figure 6 Historical and Hypothetical Exemption RBS Scenarios for 7/16/2007

Table 11 Comparison of RBS-EP Model to Precedence RBS for Historical Scenarios

Scenario	Precedence RBS		λ	RBS-EP (Fair, 30)		RBS-EP (1.001, 15)	
	Average flight delay	Unfair delay (%)		Delay reduction (%)	Unfair delay (%)	Delay reduction (%)	Unfair delay (%)
5/2/2007	6.8	0.0	N/A	—	—	3.3	1.9
5/9/2007	3.9	3.8	1.001	0.9	1.2	0.8	1.1
6/19/2007	19.7	0.5	1.501	0.5	0.3	0.8	0.6
6/27/2007	18.5	0.0	N/A	—	—	3.4	1.4
6/28/2007	24.5	0.0	N/A	—	—	2.2	1.3
7/5/2007	3.2	5.4	1.001	12.2	1.0	12.2	1.0
7/16/2007	2.9	0.0	N/A	—	—	0.6	0.6
7/18/2007	17.5	0.1	N/A	—	—	2.6	1.3
7/27/2007	4.5	0.5	N/A	—	—	7.7	3.0
9/27/2007	7.9	0.1	N/A	—	—	1.5	1.1
Summary	12.8	0.3		0.4	0.15	2.5	1.2

is gained by sacrificing a relatively small amount in terms of total or average flight delay. This is consistent with both the computational results that follow and the theoretical results developed in Bertsimas, Farias, and Trichakis (2011). In general, a substantial improvement in equity can be gained with little sacrifice in terms of efficiency.

In Tables 11, 12, 13, and 14 we summarize the results of our 40 test instances (10 days \times 2 scenarios \times 2 multiresource RBS approaches). For each instance, we compare the average flight delay and percentage of unfair delay (as measured by time-order deviation) of the multiresource RBS schedules to schedules generated using two different approaches based on the RBS-EP model. In each table, the summary row provides the results averaged across all affected flights over the 10 days. In the first approach, we allow each flight to be delayed up to $d_{\text{MAX}}^{\text{RBS}}(f)$ plus 30 minutes. We then choose and report the smallest parameter value for the exponential penalty base, λ , that leads to a solution at least as fair as the corresponding multiresource RBS solution, as measured by time-order deviation. For some instances we are unable to

find a schedule as fair using the RBS-EP model, typically because the multiresource RBS solution is almost perfectly fair. In these cases, we report N/A for the RBS-EP parameter value and list “—” for the percentage of delay reduction and percentage of unfair delay. When summarizing the results, we use the multiresource RBS solution in place of these values. We refer to this approach as RBS-EP (Fair, 30). In the second approach, we allow each flight to be delayed only 15 minutes beyond $d_{\text{MAX}}^{\text{RBS}}(f)$. For this approach, we report the average flight delay and percentage of unfair delay for the RBS-EP solution using $\lambda = 1.001$. We refer to this approach as RBS-EP (1.001, 15). Of the feasible schedules that minimize total delay, this approach selects the one that is most fair according to the exponential penalty. Although this schedule may not be as fair as the multiresource RBS schedule, no flight is likely to incur more than 15 minutes of unfair delay by construction. Across the 20 precedence scenarios, this second approach allocates 1.8% of unfair delay, on average, as compared to 1.5% for the precedence RBS schedules. For the 20 exemption scenarios, the second approach allocates 2.2% of

Table 12 Comparison of RBS-EP Model to Exemption RBS for Historical Scenarios

Scenario	Exemption RBS		λ	RBS-EP (Fair, 30)		RBS-EP (1.001, 15)	
	Average flight delay	Unfair delay (%)		Delay reduction (%)	Unfair delay (%)	Delay reduction (%)	Unfair delay (%)
5/2/2007	7.2	13.0	1.001	7.3	3.2	6.8	2.9
5/9/2007	4.5	0.0	N/A	—	—	5.7	1.1
6/19/2007	20.9	25.4	1.001	4.8	1.4	4.4	1.0
6/27/2007	19.0	5.6	1.001	5.5	2.5	4.5	1.4
6/28/2007	24.6	0.0	N/A	—	—	2.6	1.6
7/5/2007	4.3	0.0	N/A	—	—	14.5	0.9
7/16/2007	2.9	4.3	1.001	0.5	0.9	0.5	0.9
7/18/2007	18.1	5.9	1.001	5.4	2.4	5.2	2.7
7/27/2007	4.7	7.7	1.001	7.7	5.0	5.7	2.8
9/27/2007	8.2	3.5	1.001	4.5	1.6	4.1	1.2
Summary	13.4	9.3		4.0	1.6	4.5	1.6

Table 13 Comparison of RBS-EP Model to Precedence RBS for Hypothetical Scenarios

Scenario	Precedence RBS		λ	RBS-EP (Fair, 30)		RBS-EP (1.001, 15)	
	Average flight delay	Unfair delay (%)		Delay reduction (%)	Unfair delay (%)	Delay reduction (%)	Unfair delay (%)
5/2/2007	13.3	0.4	2.001	1.4	0.0	3.0	1.7
5/9/2007	5.5	6.0	1.001	0.7	1.2	0.7	1.2
6/19/2007	22.0	2.7	1.251	4.0	2.3	4.0	2.4
6/27/2007	22.6	1.0	1.751	5.7	1.0	6.9	2.1
6/28/2007	27.7	0.0	N/A	—	—	6.8	3.0
7/5/2007	14.7	7.6	1.001	5.0	0.4	5.0	0.4
7/16/2007	3.1	0.1	N/A	—	—	2.1	1.4
7/18/2007	24.6	4.4	1.001	13.4	4.8	11.1	3.1
7/27/2007	7.1	4.6	1.001	10.6	4.6	8.9	2.8
9/27/2007	9.0	2.0	1.001	3.5	1.8	2.8	1.1
Summary	16.7	2.5		5.3	1.8	6.4	2.3

unfair delay on average, compared to 6.5% for the exemption RBS schedules. Thus, we believe that this approach represents a reasonable and fair alternative, especially in an aggregate sense.

For the historical scenarios in Table 12, the days during which exemption RBS is the most unfair are 6/19/2007, 6/27/2007, 7/18/2007, and 7/27/2007. On each of these days, AFPs are reported earlier in the day than GDPs that represent possible sources of conflict. When the reverse ordering occurs, on days such as 6/28/2007 and 7/5/2007, the resulting exemption RBS schedule is extremely fair. Because AFPs affect a large geographic region, the relative capacity reductions are typically mild when compared to those resulting from GDPs. Thus, flights that are affected first by AFPs and exempted from subsequent GDPs are able to skirt the largest source of congestion along their routes. This, in turn, pushes further delays down to flights affected only by GDPs, creating large time-order deviations. From a fairness perspective, this demonstrates the sensitivity of exemption RBS to the ordering of program implementation. In Table 11, we see that for the historical scenarios precedence RBS

remains quite fair even for the days mentioned above. This is not particularly surprising because precedence RBS mitigates potential fairness issues by creating additional FCA capacity as needed. The costs of this additional capacity are likely realized downstream in terms of interventions en route, which makes the costs difficult to evaluate. Thus, in terms of calculating the cost benefits in \$4.5, we compare our RBS-EP approaches to exemption RBS.

Under our hypothetical scenarios, in which we further reduce the capacities of each AFP, we find that, in general, significantly greater delays are introduced using both the precedence and exemption RBS scheduling approaches. Although true in general, this statement does not hold for all scenarios. For some scenarios we are able to achieve a greater percentage of delay reductions in the less constrained historical scenarios. This demonstrates that the opportunity for delay reduction is not strictly increasing as capacities decrease, even though this appears to be true in aggregate. In these scenarios, precedence RBS remains more fair and allocates fewer delays than does exemption RBS. Nonetheless, even when using precedence

Table 14 Comparison of RBS-EP Model to Exemption RBS for Hypothetical Scenarios

Scenario	Exemption RBS		λ	RBS-EP (Fair, 30)		RBS-EP (1.001, 15)	
	Average flight delay	unfair delay (%)		Delay reduction (%)	Unfair delay (%)	Delay reduction (%)	Unfair delay (%)
5/2/2007	13.8	3.4	1.001	6.2	3.3	4.7	1.7
5/9/2007	6.7	0.0	N/A	—	—	5.3	1.3
6/19/2007	29.4	11.0	1.001	19.0	5.4	17.1	3.2
6/27/2007	24.0	3.0	1.251	6.8	2.3	6.8	2.3
6/28/2007	28.3	0.0	N/A	—	—	8.1	4.1
7/5/2007	16.9	0.0	N/A	—	—	4.3	0.4
7/16/2007	3.2	3.6	1.001	3.0	2.7	2.5	2.2
7/18/2007	28.5	4.7	1.001	11.4	3.9	9.3	2.6
7/27/2007	7.7	1.9	1.501	4.9	1.2	6.4	2.9
9/27/2007	9.8	2.7	1.001	7.6	1.8	6.9	1.0
Summary	19.0	4.4		8.8	2.6	9.6	2.6

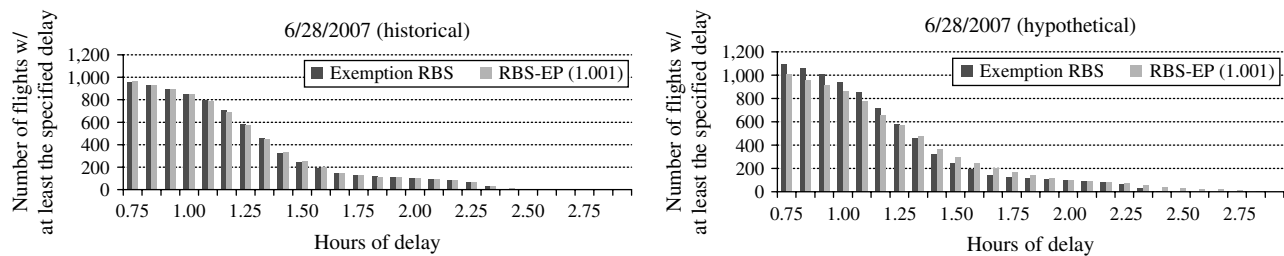


Figure 7 Allocated Flight Delay Distributions with RBS-EP λ Value of 1.001

RBS as a baseline, we are still able to realize delay reductions ranging from 5.3% for schedules that are just as fair, to 6.4% if we relax this restriction. These results suggest the importance of implementing an optimization-based approach to TFM program coordination before the prevalence and complexity of AFP utilization increases.

4.4. Flight Delay Distribution

In addition to the summary statistics listed in the tables in the previous section, it is important to consider the distribution of delays for affected flights. Airlines typically build slack into their flight schedules to preserve connections between aircraft, crews, and passengers. Delay that is less than the planned slack can be absorbed without schedule modifications. Delay that exceeds the planned slack often requires costly recovery operations. Thus, we need to ensure that our approach does not lead to a heavy tail of flight delays (i.e., a larger number of flights incurring a large amount of delay).

Consider the flight delay distributions charted in Figure 7 for the historical and hypothetical 6/28/2007 exemption RBS scenarios. These figures plot the number of flights that incur at least the specified number of hours of delay (starting at 45 minutes), comparing the exemption RBS schedule to the RBS-EP schedule allowing $d_{MAX}^{RBS}(f)$ plus 30 minutes of delay with $\lambda = 1.001$. Although the distributions are similar, the RBS-EP schedules have a longer tail, with many flights incurring at least 2.5 hours of delay, more than the maximum delay assigned in the exemption RBS schedule. Based on the discussion in the preceding paragraph, this is likely a significant issue.

Fortunately, the RBS-EP model provides an obvious mechanism for resolving these issues. By increasing the value of λ , the base of the exponential penalty, it puts additional pressure on the tail of the flight delay distribution. For example, consider the updated charts in Figure 8, for which we utilize λ values of 2.001 for each of the RBS-EP solutions, as compared to 1.001 in the previous charts. By increasing the value of λ , we have increased the total delay assigned from 1,429 hours to 1,443 hours in the historical scenario and from 1,527 hours to 1,581 hours in the hypothetical scenario. In so doing, we have managed to shrink the tails of the delay distribution, with the resulting schedules still more efficient than the exemption RBS schedules. This tradeoff between aggregate delay and the distribution of delay is another important consideration in choosing an appropriate value of λ in practice.

4.5. The Value of Efficiency

In Table 12, we see that exemption RBS allocates a total of 7,339 hours of delay across the 10 scenarios, whereas the RBS-EP (Fair, 30) approach allocates 7,046 hours of delay and the RBS-EP (1.001, 15) approach 7,008 hours of delay. Thus, we estimate that the RBS-EP model would lead to an overall delay reduction of 4.0% to 4.5% across days with conflicting TFM programs. Next, we would like to determine how much cost reduction, for airlines, passengers, and related industries, can be attributed to these delay reductions.

As mentioned in the introduction, the U.S. Congress Joint Economic Committee estimates that arrival delays cost the U.S. economy \$25.7 billion in 2007 (Joint Economic Committee 2008). The Bureau of Transportation Statistics estimates that in 2007, 37.7%

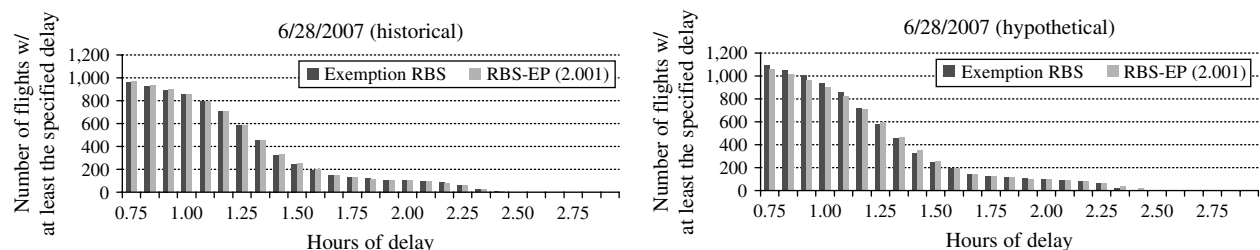


Figure 8 Updated Flight Delay Distributions Based on RBS-EP λ Value of 2.001

of flight delays resulted from the previous flight arriving late; thus, we estimate that the remaining 62.3% of flight delays are due to direct impacts (Research and Innovative Technology Administration 2009). These direct impacts, to which we attribute the full delay costs, led to 73.5 million minutes of delay, of which we estimate that 21.3 million minutes were due to ground holding programs using the Airline On-time Performance Database (U.S. Bureau of Transportation Statistics 2007). This represents approximately 30% of the direct impact delay; thus, we attribute 30% of the total delay costs—\$3.7 billion in increased airline operating costs, \$2.2 billion in passenger time lost, and \$1.8 billion in costs for related industries—to TFM programs. Of the delay assigned through these initiatives, approximately 13% was on days during which domestic GDPs and AFPs were both implemented. We consider these days our baseline for improvement because the multiresource RBS schedules are delay optimal when there are no conflicts between TFM programs. A 1% delay reduction on these days would save airlines \$4.8 million, passengers \$2.9 million, and related industries \$2.4 million annually, a total of just over \$10 million annually. Combining this with the above, we estimate that implementation of the RBS-EP model for coordinating TFM programs would lead to annual cost savings on the order of \$25 to \$50 million.

It is worth noting that the attribution approach utilized likely underestimates the value in at least two ways. First, by focusing our analysis on direct impact delays we are assuming propagated delay costs are allocated proportionally among different root causes. This likely underestimates the costs associated with TFM programs because these programs typically result in larger magnitudes of delays that are more likely to exceed schedule slack, leading to delay propagation. Second, the Joint Economic Committee estimates passenger delays by multiplying the number of passengers by the corresponding flight delays. This approach does not include the impact of missed connections or flight cancellations, both of which are prevalent during TFM initiatives.

4.6. Computational Performance

To be implemented in practice, an optimization-based approach to coordinating TFM programs must be extremely fast, preferably returning a good solution within a minute or less to support the subsequent CDM procedures. Fortunately, this is not a concern for either of our RBS-EP approaches, which use less than 10 seconds of CPLEX solver time per instance. As a reference, in Table 15 we list the CPLEX solver times for the hypothetical exemption RBS scenarios. We allow either 15 or 30 minutes of delay beyond $d_{\text{MAX}}^{\text{RBS}}(f)$ and compare the solution times for $\lambda =$

Table 15 CPLEX Times for Hypothetical Exemption RBS Scenarios Using RBS-EP Model

Date	CPLEX solver times (sec.)			
	RBS-EP (15)		RBS-EP (30)	
	$\lambda = 1.001$	$\lambda = 2.001$	$\lambda = 1.001$	$\lambda = 2.001$
5/2/2007	0.16	0.15	0.47	0.46
5/9/2007	0.12	0.12	0.45	0.44
6/19/2007	8.10	8.05	9.03	6.59
6/27/2007	2.15	3.37	3.70	7.69
6/28/2007	4.19	4.54	6.58	8.57
7/5/2007	0.39	0.39	1.32	1.32
7/16/2007	0.24	0.24	0.67	0.66
7/18/2007	1.34	3.54	3.68	7.68
7/27/2007	0.65	0.63	1.68	1.62
9/27/2007	0.51	0.51	2.09	2.06
Total	17.83	21.53	29.65	37.08

1.001 and $\lambda = 2.001$. Allowing a smaller amount of delay in the model reduces the number of decision variables, leading to a roughly 40% improvement in CPLEX times, on average. The performance measurements utilize the greedy integer rounding heuristic described in §4.3, with a CPLEX relative optimality gap of 0.01%. The computational tests are performed on a PC with dual Xeon 3220 Quad-Core processors and 16 Gigabytes of RAM, running Ubuntu v8.04 and CPLEX v11.2 through the Java interface.

5. Conclusion

In this research, we develop an optimization-based formulation that could be readily incorporated in practice by the FAA. Specifically, based on principles that have made RBS successful, we have developed a time-order deviation metric for schedule fairness that extends to the multiresource setting. This metric allows us to evaluate optimization-based scheduling approaches relative to each other, but more importantly, it allows us to compare these approaches to the approaches currently utilized in practice. Using this metric, we have demonstrated that in each of the approaches used in practice to resolve conflicts between TFM programs, an implicit tradeoff is made between equity and efficiency. We have also demonstrated that our two formulations, the TODA and RBS-EP models, can improve efficiency while maintaining a consistent level of fairness. Lastly, we have shown that the RBS-EP model is computationally tractable in practice, even for complex, national-scale problems.

Introducing optimization into the FAA's practices has been a significant challenge, as should be apparent from the literature review in §1.3. The RBS-EP model addresses many of these challenges and should thus provide a strong foundation for future research. Our goal is to have the RBS-EP model represent the

first step in an ongoing sequence of practical enhancements to the FAA's TFM procedures.

Acknowledgments

The authors gratefully acknowledge Professor Amedeo Odoni for being a supportive and enthusiastic advisor. His suggestions and guidance have heavily influenced this work. Additionally, the authors thank Robert Hoffman, Bert Hackney, Jason Burke, and Stephen Augustine of Metron Aviation for their time and expertise. Lastly, they thank the Associate Editor and three anonymous referees whose insightful comments have helped shape the paper. This work has been supported by NASA Next generation (NEXTGEN) and NSF's Emerging Frontiers in Research and Innovation (EFRI) office.

References

- Air Transportation Association. 2008. 2008 annual report. Retrieved April 2009, <http://www.airlines.org/Documents/economicreports/2008AnnualReport.pdf>.
- Andreatta, G., L. Brunetta, G. Guastalla. 2000. From ground holding to free flight: An exact approach. *Transportation Sci.* **34**(4) 394–401.
- Ball, M. O., C. Barnhart, G. Nemhauser, A. Odoni. 2007. Air transportation: Irregular operations and control. C. Barnhart, G. Laporte, eds. *Handbooks in operations Research and Management Science*, vol. 14. Elsevier B. V., Amsterdam, 23–28.
- Ball, M. O., R. Hoffman, A. Odoni, R. Rifkin. 2003. A stochastic integer program with dual network structure and its application to the ground-holding problem. *Oper. Res.* **51**(1) 167–171.
- Bertsimas, D., S. Gupta. 2011. A proposal for network air traffic flow management incorporating fairness and airline collaboration. Working paper.
- Bertsimas, D., S. S. Patterson. 1998. The air traffic flow management problem with enroute capacities. *Oper. Res.* **46**(3) 406–422.
- Bertsimas, D., V. Farias, N. Trichakis. 2011. The price of fairness. *Oper. Res.* **59**(1) 17–31.
- Bertsimas, D., G. Lulli, A. Odoni. 2011. An integer optimization approach to large-scale air traffic flow management. *Oper. Res.* **59**(1) 211–227.
- Brennan, M. 2007. Airspace flow programs—A fast path to deployment. *J. Air Traffic Control* **49**(1) 51–55.
- Chang, K., K. Howard, R. Oiesen, L. Shishler, M. Tanino, M. C. Wambsganns. 2001. Enhancements to the FAA ground-delay program under collaborative decision making. *Transportation Sci.* **34**(1) 57–76.
- de Neufville, R., A. Odoni. 2003. Airport Systems: Planning, Design, and Management. *Air Traffic Management*, Chap. 13. McGraw-Hill Professional, 525–545.
- Federal Aviation Administration. 2005. Operational concept for airspace flow programs (AFP). Retrieved January 2010, http://cdm.fly.faa.gov/Workgroups/flow_eval/OperationalConceptforAirspaceFlowPrograms.doc.
- Federal Aviation Administration. 2006. Graphical depiction of AFP. Retrieved January 2010. http://www.fly.faa.gov/What_s_New/AFP.ppt.
- Hoffman, R., M. O. Ball. 2000. A comparison of formulations for the single-airport ground-holding problem with banking constraints. *Oper. Res.* **48**(4) 578–590.
- Joint Economic Committee. 2008. Your flight has been delayed again: Flight delays cost passengers, airlines, and the U.S. economy billions. Retrieved May 2009, http://jec.senate.gov/public/?a=Files.Serve&File_id=47e8d8a7-661d-4e6b-ae72-0f1831dd1207.
- Kotnyek, B., O. Richetta. 2006. Equitable models for the stochastic ground-holding problem under collaborative decision-making. *Transportation Sci.* **40**(2) 133–146.
- Lulli, G., A. Odoni. 2007. The European air traffic flow management problem. *Transportation Sci.* **41**(4) 431–443.
- Metron Aviation. 2009a. Flight Schedule Monitor. Retrieved May 2009, <http://www.metronaviation.com/fsm.php>.
- Metron Aviation. 2009b. FSM version 8.80 new features. Retrieved January 2010, http://cdm.fly.faa.gov/products/FSM_Docs/FSM_8.80/FSM_8.80_Training_Guide_v1.3.pdf.
- Odoni, A. 1987. The flow management problem in air traffic control. A. Odoni, G. Szego, eds. *Flow Control of Congested Networks*. Springer Verlag, Berlin, 269–288.
- Research and Innovative Technology Administration. 2009. Understanding the reporting causes of flight delays and cancellations. Retrieved April 2009 from the Bureau of Transportation Statistics, <http://www.bts.gov/help/aviation/html/understanding.html>.
- Richetta, O., A. Odoni. 1993. Solving optimally the static ground-holding policy problem in air traffic control. *Transportation Sci.* **27**(3) 228–238.
- Richetta, O., A. Odoni. 1994. Dynamic solution to the ground-holding policy problem in air traffic control. *Transportation Res. A* **28**(3) 167–185.
- Terrab, M., A. Odoni. 1993. Strategic flow management for air traffic control. *Oper. Res.* **41**(1) 138–152.
- U.S. Bureau of Transportation Statistics. 2007. Airline on-time performance database. Retrieved April 2009, <http://www.transtats.bts.gov/>.
- U.S. Department of Transportation. 2004. Airport capacity benchmark report 2004. Joint work with the Federal Aviation Administration, the MITRE Corporation, and the Center for Advanced Aviation Systems.
- Vossen, T., M. O. Ball. 2005. Optimization and mediated bartering models for ground delay programs. *Naval Res. Logist.* **53**(1) 75–90.
- Vossen, T., M. O. Ball, R. Hoffman, M. C. Wambsganns. 2003. A general approach to equity in traffic flow management and its application to mitigating exemption bias in ground delay programs. *Air Traffic Control Quart.* **11**(4) 277–292.
- Vranas, P. B., D. Bertsimas, A. Odoni. 1994. The multi-airport ground-holding problem in air traffic control. *Oper. Res.* **42**(2) 249–261.