A Proposal for Network Air Traffic Flow Management Incorporating Fairness and Airline Collaboration*

Dimitris Bertsimas† Shubham Gupta‡

April 25, 2011

Abstract

There has been significant research effort in the academic literature related to Air Traffic Flow Management (ATFM). Yet, the research has not been fully implemented in practice. In our opinion, the key reasons are - i) the existing network models approach the problem from the point of view of a central decision-maker without taking into account the airlines to a significant degree; and ii) the notions of fairness as introduced under the Collaborative Decision-Making (CDM) paradigm operate under a single-airport setting and do not address network effects (presence of multiple airports, sectors and various connectivity requirements). In this paper, we address these shortfalls by presenting a proposal which alleviates both these concerns. In stage I of our proposal, we present network models that incorporate a notion of fairness - controlling number of reversals and total amount of overtaking. In stage II, we allow for further airline collaboration by proposing a network model for slot reallocation. We provide empirical results of the proposed optimization models on national-scale, real world datasets spanning across six days that show interesting tradeoffs between fairness and efficiency. We report promising computational times of less than 30 minutes for up to 25 airports and provide theoretical evidence that illuminates the strength of our formulations.

1 Introduction

The sustained growth of the aviation industry has put a tremendous strain on the available resources of the air transporation system. This is evidenced by the steady increase in flight delays and severe congestion at airports. In 2008, approximately 22% of the flights in the United States were delayed by more than 15 minutes, while another 2% were cancelled (Bureau of Transportation Statistics [17]). Moreover, during the 12-month period ending in

*Research funded by the NSF Grant EFRI-0735905.

†Boeing Professor of Operations Research, Sloan School of Management, Co-director of the Operations Research Center, Massachusetts Institute of Technology, E40-147, Cambridge, MA 02139-4307, dbertsim@mit.edu

‡Operations Research Center, Massachusetts Institute of Technology, Cambridge, MA 02139-4307, shubhamg@mit.edu

September 2008, 138 million minutes of system delay led to an estimated $10 billion in costs for US airlines [1].

Air Traffic Flow Management (ATFM) refers to the set of strategic processes that try to reduce congestion costs and support the goal of safe, efficient and expeditious aircraft movement. ATFM procedures try to resolve local demand-capacity mismatches by adjusting the aggregate traffic flows to match scarce capacity resources. Ground Delay Programs (GDPs) are one of the most sophisticated ATFM initiatives currently in use that attempt to address airport arrival capacity reductions. Under this mechanism, delays are applied to flights at their origin airports that are bound for a common destination airport which is suffering from reduced capacity or excessive demand. The premise for this tool is that it is better to absorb delays for a flight while it is grounded at its origin airport rather than incurring air-borne delay near the affected destination airport which is both unsafe and more costly (in terms of fuel costs). Airspace-Flow Programs (AFPs) are another tool that were introduced recently in 2007 and are similar to GDPs in terms of operational details. FAA uses this tool to control arrival rate into a Flow Constrained Area (FCA), e.g., a weather affected segment of the airspace. Some of the other ATFM tools include assigning air-borne delays, dynamic re-routing and speed control. We briefly review the literature on the existing ATFM tools below.

Odoni [12] first conceptualized the problem of scheduling flights in real time in order to minimize congestion costs. Thereafter, several models have been proposed to handle different versions of the problem. The problem of assigning ground-delays in the context of a single-airport (*Single-Airport Ground-Holding Problem*) has been studied in Terrab and Odoni [15], Richetta and Odoni [13], [14]; and in the multiple airport setting (*Multi-Airport Ground-Holding Problem*) in Terrab and Paulose [16], Vranas et al. [20]. The problem of controlling release times and speed adjustments of aircraft while air-borne for a network of airports taking into account the capacitated airspace (*Air Traffic Flow Management Problem*) has been studied in Bertsimas and Stock [5], Helme [9], Lindsay et al. [11]. The problem with the added complication of dynamically re-routing aircrafts (*Air Traffic Flow Management Rerouting Problem*) was first studied by Bertsimas and Stock [6]. Recently, Bertsimas et al. [7] have presented a new mathematical model for the ATFM problem with dynamic re-routing which has superior computational performance. For a detailed survey of the various contributions and a taxonomy of all the problems, see Bertsimas and Odoni [4] and Hoffman et al. [10]. It is important to highlight that network formulations present significant challenges of computational tractability, and so far few network models have been able to incorporate equity considerations effectively. Despite the significant progress outlined here, this set of literature while addressing *network effects* (presence of multiple airports, sectors and various connectivity requirements), takes the point of view of a single decision-maker and does not address the issue of fairness among airlines.

We review next some other important concepts in the ATFM literature - namely, *Collaborative Decision-Making* (CDM) and *Ration-by-Schedule* (RBS). The decision-making responsibilities in ATFM initiatives are shared between a number of stakeholders (primarily, airlines and the FAA). This poses a major challenge as their actions are highly interdependent and demand real-time exchange of information between the FAA and the airlines. This realization of enhanced cooperation between the various stakeholders led to the adoption of Collaborative Decision-Making (CDM) philosophy (Ball et al. [2], Wambsganss [21]) by the FAA in the 1990s. Under CDM, all ATFM initiatives are conducted in a way that gives significant decision-making responsibilities to airspace users (see Hoffman et al. [10] for details on CDM). All recent efforts to improve ATFM have been guided by this philosophy. In the US, "Ration-by-Schedule" (RBS) is the fundamental principle for all the CDM initiatives. Under this paradigm - arrival slots at airports are assigned to flights in accordance with a first-scheduled, first-served (FSFS) priority discipline (see Ball et al. [2], Wambsganss [21] for details on rationing). In the case of GDP and AFP planning, all stakeholders have agreed that this principle is fair to all parties. This allocation process is followed by a *Compression algorithm*, which fills open slots created by flights that are canceled. The compression procedure gives airlines an incentive to report accurate flight information, by rewarding them for reporting cancellations. The combined process, RBS plus Compression (formally called RBS^{++}) is the policy currently in use for slot allocations. Despite the use of RBS in a GDP setting, there have been no network models that satisfy the RBS principle in a multi-airport setting. This is because, applying RBS to each of the airports individually might not lead to a schedule that preserves time, sector and flight connectivities. In addition, the imposition of a maximum permissible delay on each flight (as required by any tractable optimization model) would mean that a feasible solution under RBS might not even exist if the capacity reduction at some airports is significant. Hence, there is no straightforward extension of RBS from a single-airport setting to an airspace context. As part of the CDM philosophy, researchers have also explored dynamic interaction with airlines. Towards this aim, Vossen and Ball [19] [18] have studied opportunities for slot trading in a single-airport setting where the aim is to formalize an optimization problem for the FAA given the offers to trade from various airlines. In summary, the CDM concepts discussed here while addressing the issues of fairness among airlines in a single-airport setting, do not address network effects.

Our goal in this paper is to propose an optimization based approach that a) incorporates network effects and builds upon the ATFM literature; and b) takes into account fairness considerations among airlines by building upon the CDM philosophy. Specifically, our proposal consists of the following two stages:

Stage I - Network ATFM model incorporating fairness:
We generalize the classical ATFM models ([5]) to incorporate fairness considerations for airlines. The objective function used in the existing network ATFM models is to minimize the total delay costs across all flights, i.e., the focus is on overall system efficiency. A disadvantage of such an approach is that the solution to such models can have a large number of reversals, i.e., the resulting order of flight arrivals can be quite different as compared to the published flight schedules. Specifically, for two flights f and f' arriving at the same destination airport, a *reversal* occurs if f was scheduled to arrive before f' but f' arrives before f in the actual

schedule. Moreover, across these reversals, there might be different number of time-periods of overtaking. Concretely, the duration between the arrival of flights f' and f constitute the amount of *overtaking*. Consequently, the total overtaking across these reversals might be large. Because of this deviation from the original flight ordering, it becomes difficult to implement such a solution because of the coupling in the crew assignments and the use of hub and spoke networks. We propose discrete optimization models that add these fairness controls. The key output in this stage is the assignment of flights to different time periods.

Stage II - Slot reallocation through airline collaboration:
We generalize the notion of intra-airline exchange of arrival slots, a key component of the current CDM practice in a single-airport setting to network-wide slot reallocation among airlines. Specifically, we propose an optimization model which takes as input the assignment of flights to different time periods from Stage I and permits the airlines to trade these assigned slots across different airports, thereby, resulting in improved internal objective functions. The model proposed for Stage II of our proposal allows airlines to react to the schedule determined in Stage I by taking into account their flights in the entire network and making appropriate tradeoffs.

Contributions of this work

We feel our work makes the following contributions:

1. We present a proposal for the network ATFM problem which incorporates both fairness and airline collaboration while operating under a broad CDM paradigm. Specifically, we formulate discrete optimization models to impose fairness controls in Stage I and a model for slot reallocation in Stage II.

2. We provide empirical results of the proposed optimization models on national-scale, real world datasets spanning across six days. We report promising computational times of less than 30 minutes for up to 25 airports and provide theoretical evidence on the strength of our formulations.

Simultaneously, Barnhart et al. [3] develop an alternative way to address fairness in the context of ATFM. They develop a fairness metric that measures deviation from FSFS and propose a discrete optimization model that directly minimizes this metric. They further develop an exponential penalty approach, and report encouraging computational results using simulated regional and national scenarios. In contrast to this work, our paper differs in the following respects: a) we provide an exact method to model overtaking within a mathematical programming framework while [3] proposes an approximate method. In addition, we propose another, although related, metric of fairness - controlling the number of reversals; b) we explicitly model network effects that [3] does not; and c) our proposal consists of a slot reallocation phase (Stage II) which enables airlines to implicitly trade their internal objective functions to derive more utility. In summary, both papers contribute to the understanding of fairness in ATFM by approaching the problem from distinct perspectives.

Organization of the paper

Section 2 summarizes an adaptation of the Bertsimas Stock-Patterson model [5] in order to accommodate our proposal. Section 3 introduces models of fairness. Section 4 introduces our model of slot reallocation. Section 5 discusses how our proposal can be integrated with the current CDM practices. Section 6 reports computational results of the proposed optimization models on six days of national-scale, real world datasets. Section 7 summarizes our conclusions and the Appendix reports polyhedral analysis that illuminates the strength of our proposed formulations.

2 ATFM Problem: Notation, Bertsimas Stock-Patterson Model and Solutions

In this section, we reproduce the Bertsimas Stock-Patterson model [5] for the ATFM problem which provides the starting point for all the models presented herein. We use an extended version of the notation used in that paper in order to accommodate fairness and slot reallocation considerations. Finally, we illustrate difficulties relative to fairness considerations in the solutions obtained from this model.

The Decision Variables

The decision variables are:

$$w_{j,t}^{f} = \begin{cases} 1, & \text{if flight } f \text{ arrives at sector } j \text{ by time } t, \\ 0, & \text{otherwise.} \end{cases}$$

This definition of the decision variables, using "*by*" instead of "*at*", is critical to the understanding of the formulation. The variables are defined only for the set of sectors an aircraft may fly through on its route to the destination airports. In addition, variables are used for the departure and the arrival airports, in order to determine the optimal times for departure and for arrival. Since we do not consider flight cancellations, at least two variables can be fixed a priori for each flight: each aircraft has to take off by the end of a feasible time window and has to land, as well, within a feasible time window, which is determined by the time of departure.

Notation

The model's formulation requires definition of the following notation:

$\mathcal{K}$: set of airports,

$\mathcal{F}$: set of flights,

$\mathcal{T}$: set of time periods,

$\mathcal{W}$: set of airlines,

$\mathcal{F}_w \subseteq \mathcal{F}$: set of flights belonging to airline w,

$\mathcal{S}$: set of sectors,

$\mathcal{S}^f \subseteq \mathcal{S}$: sequence of sectors flown by flight f,

$\mathcal{C}$: set of pairs of flights that are continued,

$\mathcal{R}^j$: set of pairs of flights that are reversible in resource j (definition below),

$\mathcal{R}^S$: set of pairs of flights that are reversible in sectors (definition below),

$\mathcal{R}^A$: set of pairs of flights that are reversible at airports (definition below),

$\mathcal{P}_i^f$: sector i's preceding sector in the path of flight f,

$\mathcal{L}_i^f$: sector i's subsequent sector in the path of flight f,

$D_k(t)$: departure capacity of airport k at time t,

$A_k(t)$: arrival capacity of airport k at time t,

$S_j(t)$: capacity of sector j at time t,

d_f : scheduled departure time of flight f,

a_f^j : scheduled arrival time of flight f in resource j,

s_f : turnaround time of an airplane after flight f,

orig_f : airport of departure of flight f,

dest_f : airport of arrival of flight f,

l_{fj} : minimum number of time units that flight f must spend in sector j,

M : maximum permissible delay for a flight,

T_j^f : set of feasible time periods for flight f to arrive in resource j,

$\underline{T}_j^f$: first time period in the set T_j^f,

$\overline{T}_j^f$: last time period in the set T_j^f,

$T_{f,f',j}^{\text{reversal}}$: set of time-periods common for flights f and f' where a reversal could occur in resource j (definition below),

$O_{f,f',j}^{\max}$: maximum amount of overtaking possible between flights f and f' in resource j (definition below),

$\mathcal{O}$: set of all possible trades (definition in Section 4),

$\mathcal{O}^f \subseteq \mathcal{O}$: set of offers containing flight f,

c_f : time period assigned to flight f from Stage I optimization.

The key additions relative to the notation used in [5] are $\mathcal{W}$, $\mathcal{F}_w$, $\mathcal{R}^j$, $\mathcal{R}^S$, $\mathcal{R}^A$, $T^{\text{reversal}}_{f,f',j}$, $O^{\max}_{f,f',j}$, $\mathcal{O}$, $\mathcal{O}^f \subseteq \mathcal{O}$ and c_f.

The sets $\mathcal{R}^j$, $\mathcal{R}^S$ and $\mathcal{R}^A$

We give next the definition of $\mathcal{R}^j$ (set of pairs of flights that are reversible in resource j). We make a distinction between the case when the resource j is a sector and when it is an airport.

Definition 2.1. For an airport $k \in \mathcal{K}$, a pair of flights (f, f') belongs to $\mathcal{R}^k$ if the following two conditions are satisfied:

1. $\text{dest}_f = \text{dest}_{f'} = k$, i.e., the destination airport of both flights f and f' is the same.
2. $a^k_f \leq a^k_{f'} \leq a^k_f + M$, i.e., the scheduled time of arrival of flight f' at the destination airport lies between the scheduled time of arrival of flight f and the last time period in the set of feasible time periods that the flight f can arrive at its destination airport.

Definition 2.2. For a sector $j \in \mathcal{S}$, a pair of flights (f, f') belongs to $\mathcal{R}^j$ if the following two conditions are satisfied:

1. $j \in \mathcal{S}^f$ and $j \in \mathcal{S}^{f'}$, i.e., sector j is common to the path of flights f and f'.
2. $a^j_f \leq a^j_{f'} \leq a^j_f + M$, i.e., the scheduled time of arrival of flight f' in sector j lies between the scheduled time of arrival of flight f and the last time period in the set of feasible time periods that the flight f can arrive in sector j.

For each pair of flights $(f, f') \in \mathcal{R}^j$, we count a *reversal*, if in the resulting solution, flight f' arrives before flight f in resource j (i.e., $\exists t$ such that $w^{f'}_{j,t} > w^f_{j,t}$). We call the reversals occurring in sectors as *sector reversals* and the reversals occurring at the airports as *airport reversals*. This clustering is motivated from fairness considerations in a network setting and is elaborated upon later in Section 3.

Definition 2.3. The set of pairs of flights that are reversible in sectors ($\mathcal{R}^S$) and at airports ($\mathcal{R}^A$) is defined as:

1. $\mathcal{R}^S = \bigcup_{j \in \mathcal{S}} \mathcal{R}^j$
2. $\mathcal{R}^A = \bigcup_{k \in \mathcal{K}} \mathcal{R}^k$

The set $T^{\text{reversal}}_{f,f',j}$ and parameter $O^{\max}_{f,f',j}$

Definition 2.4. The set $T^{\text{reversal}}_{f,f',j}$ (set of time-periods common for flights f and f' where a reversal could occur) is defined as $\{\underline{T}^{f'}_j, \ldots, \overline{T}^f_j - 1\}$.

To elaborate on Definition 2.4, $T^{\text{reversal}}_{f,f',j}$ is the set of time-periods t, such that it is possible to have the following assignment: $w^f_{j,t} = 0$ and $w^{f'}_{j,t} = 1$. This assignment would imply that a reversal occurs at time t.

Definition 2.5. The parameter $O^{\max}_{f,f',j}$ is defined as $|T^{\text{reversal}}_{f,f',j}|$ (cardinality of the set $T^{\text{reversal}}_{f,f',j}$), and hence is equal to $\overline{T}^f_j - \underline{T}^{f'}_j - 1$.

To elaborate on Definition 2.4, $O^{\max}_{f,f',j}$ is the maximum amount of overtaking possible between flights f and f' and would be attained when $w^f_{j,\overline{T}^f_j - 1} = 0$ and $w^{f'}_{j,\underline{T}^{f'}_j} = 1$.

Example 2.1. Figure 1 depicts a reversible pair of flights $(f, f') \in \mathcal{R}^A$. Let $\text{dest}_f = \text{dest}_{f'} = k$. In this example, the arrows correspond to the set of time-periods common for both flights. Moreover, the set of time-periods marked by these arrows (except for the last one) constitute $T^{\text{reversal}}_{f,f',k}$. This is because, the model enforces $w^f_{k,a^k_f+M} = 1$ at the outset and hence, it is not possible to have a reversal at $a^k_f + M$. Finally, $O^{\max}_{f,f',k} = |T^{\text{reversal}}_{f,f',k}| = 6$.

PSfrag replacements

a^k_f

$a^k_f + M$

$a^k_{f'}$

$a^k_{f'} + M$

Feasible times of arrival for flight f

Feasible times of arrival for flight f'

$T^{\text{reversal}}_{f,f',k}$

Figure 1: A reversible pair of flights $(f, f') \in \mathcal{R}^A$ ($\text{dest}_f = k$).

The Objective Function

We use an adapted expression for total delay (TD) cost for each flight introduced recently in Bertsimas et al. [7]. The total delay (TD) cost is a combination of the costs of air-borne delay (AD) and ground-holding delay (GD) ($TD = GD + \alpha \cdot AD$, where $\alpha > 1$ because air-borne delay is typically more costly than ground-holding delay). By substituting AD in terms of TD (i.e., $AD = TD - GD$), the objective can be rewritten as $\alpha \cdot TD - (\alpha - 1) \cdot GD$.

Consequently, the objective function is composed of two terms: a first term that takes into account the cost of the total delay assigned to a flight and a second term which accounts for the cost reduction obtained when a part of the total delay is taken as ground delay at the origin airport. The objective function cost coefficients are a super-linear function of the tardiness of a flight of the form $(t - a_f^k)^{1+\epsilon}$, with ϵ close to zero. Hence, for each flight f and for each time period t, we define the following two cost coefficients:

$$\begin{aligned} c_{\text{total}}^f(t) &= \alpha(t - a_f^k)^{1+\epsilon} : \text{total cost of delaying flight } f \text{ for } (t - a_f^k) \text{ units of time,} \\ c_g^f(t) &= (\alpha - 1)(t - d_f)^{1+\epsilon} : \text{cost reduction obtained by holding flight } f \text{ on the} \\ &\qquad \text{ground for } (t - d_f) \text{ units of time,} \end{aligned}$$

The motivation of using super-linear cost coefficients is that it will favor moderate assignment of total delays between two flights rather than assigning much larger delay to one as compared to the other flight. To elaborate, consider the following example:

Example 2.2. Suppose we wish to assign 2 units of delay to 2 flights. Then, an objective function with linear cost coefficients is equally likely to generate the following two assignments: i) 1 unit of delay to both flights and ii) 2 unit of delay to one flight and 0 to the other. In contrast, super-linear cost coefficients (with $\epsilon = 0.001$ for example) will assign 1 unit to both flights because $1^{1.001} + 1^{1.001} < 2^{1.001} + 0$.

In view of the above, the delay cost function for each flight f takes the following form:

$$\left(\sum_{t \in T_{\text{dest}_f}^f} c_{\text{total}}^f(t) \cdot (w_{\text{dest}_f,t}^f - w_{\text{dest}_f,t-1}^f) - \sum_{t \in T_{\text{orig}_f}^f} c_g^f(t) \cdot (w_{\text{orig}_f,t}^f - w_{\text{orig}_f,t-1}^f) \right)$$

The TFMP model

The complete description of the model, referred to as (TFMP), is as follows:

$$IZ_{\text{TFMP}} = \min \sum_{f \in \mathcal{F}} \left(\sum_{t \in T_{\text{dest}_f}^f} c_{\text{total}}^f(t) \cdot (w_{\text{dest}_f,t}^f - w_{\text{dest}_f,t-1}^f) - \sum_{t \in T_{\text{orig}_f}^f} c_g^f(t) \cdot (w_{\text{orig}_f,t}^f - w_{\text{orig}_f,t-1}^f) \right)$$

subject to:

$$\sum_{f\in\mathcal{F}:\text{orig}_f=k} (w^f_{k,t} - w^f_{k,t-1}) \leq D_k(t), \qquad \forall k \in \mathcal{K},\ \ t\in\mathcal{T}. \tag{1a}$$

$$\sum_{f\in\mathcal{F}:\text{dest}_f=k} (w^f_{k,t} - w^f_{k,t-1}) \leq A_k(t), \qquad \forall k \in \mathcal{K},\ \ t\in\mathcal{T}. \tag{1b}$$

$$\sum_{f\in\mathcal{F}:j\in\mathcal{S}_f, j'=\mathcal{L}^f_j} (w^f_{j,t} - w^f_{j',t}) \leq S_j(t), \qquad \forall j \in \mathcal{S},\ \ t\in\mathcal{T}. \tag{1c}$$

$$w^f_{j,t} - w^f_{j',t-l_{fj'}} \leq 0, \qquad \forall f\in\mathcal{F},\ \ t\in T^f_j,\ \ j\in\mathcal{S}^f\ \ : j\neq \text{orig}_f, j'=\mathcal{P}^f_j. \tag{1d}$$

$$w^f_{\text{orig}_f,t} - w^{f'}_{\text{dest}_{f'},t-s_f} \leq 0, \qquad \forall (f,f')\in\mathcal{C},\ \ \forall t\in T^f_k. \tag{1e}$$

$$w^f_{j,t-1} - w^f_{j,t} \leq 0, \qquad \forall f\in\mathcal{F},\ \ j\in\mathcal{S}^f,\ \ t\in T^f_j. \tag{1f}$$

$$w^f_{j,t} \in \{0,1\}, \qquad \forall f\in\mathcal{F},\ \ j\in\mathcal{S}^f,\ \ t\in T^f_j. \tag{1g}$$

The first three sets of constraints take into account the capacities of the various elements of the system. Constraints (1a) ensure that the number of flights which may take off from airport k at time t, will not exceed the departure capacity of airport k at time t. Likewise, Constraints (1b) ensure that the number of flights which may arrive at airport k at time t, will not exceed the arrival capacity of airport k at time t. Finally, Constraints (1c) ensure that the total number of flights which may feasibly be in Sector j at time t will not exceed the capacity of Sector j at time t.

The next three sets of constraints capture the various connectivities - namely sector, flight and time connectivity. Constraints (1d) stipulate that a flight cannot arrive at Sector j by time t if it has not arrived at the preceding sector by time $t - l_{fj'}$. In other words, a flight cannot enter the next sector on its path until it has spent at least $l_{fj'}$ time units (the minimum possible) traveling through one of the preceding sectors on its current path. Constraints (1e) represent connectivity between flights. They handle the cases in which a flight is continued, i.e., the flight's aircraft is scheduled to perform a subsequent flight within some user-specified time interval. The first flight in such cases is denoted as f' and the subsequent flight as f, while s_f is the minimum amount of time needed to prepare flight f for departure, following the landing of flight f'. Constraints (1f) ensure connectivity in time. Thus, if a flight has arrived at element j by time $\tilde{t}$, then $w^f_{j,t}$ has to have a value of 1 for all later time periods ($t \geq \tilde{t}$).

Remark 1. **Including aircraft or flight connectivities**. In TFMP, flight connectivity constraints are included as hard constraints, i.e., they need to be satisfied a priori based on planned aircraft connections. In current practice, this is not exactly the approach taken due to the presence of hub and spoke networks which motivate the use of bank of flights arriving at a hub and then departing within a short duration of time. Nonetheless, we make a *modeling choice* to keep these constraints as they enable stronger polyhedral structure (thereby leading to shorter computational times) and are consistent with the original model proposed by Bertsimas and Stock-Patterson. Nonetheless, in all the models presented in this paper, these constraints can be removed and our proposal will still remain entirely consistent with its global objectives. Thus, this modeling is not a consequence of any other restrictions.

Solutions from (TFMP)

Here, we illustrate the difficulties relative to fairness considerations in the solutions obtained from the formulation (TFMP). We report a solution from (TFMP) for one of the six datasets on which we have performed our experiments in this paper. This dataset comprises of 5092 flights, 5 airlines, 55 airports and 100 sectors. The analysis is carried over 96 time-periods each of duration 15 minutes. First, the total number of reversals in the resulting solution is **915**. Moreover, there are **1492** units of overtaking across these reversals. To put the number of reversals in perspective, the maximum possible number of reversals is around 9500 (this is attained when maximum elements of $\mathcal{R}^A$ are reversed with the resulting schedule remaining feasible). This indicates that the sequence of flight arrivals in the solutions from (TFMP) differ significantly from the scheduled sequence of flight arrivals (note that if M was large enough, then the schedule with 0 reversals would be considered most fair). These two observations in the solutions obtained from the formulation (TFMP) is present across all the six datasets. In the next section, we introduce discrete optimization models that control the number of reversals and the amount of overtaking.

3 Network models that incorporate concepts of fairness

3.1 Minimizing the total amount of overtaking

A notion of fairness widely agreed upon by the airlines is to have a schedule that preserves the order of flight arrivals at an airport according to the published schedules. As mentioned in the Introduction, this is known as Ration-by-Schedule (RBS). But, given the capacity reductions at the airports, it might not always be possible to have a feasible solution under RBS in a network setting. A close equivalent to the RBS solution would be one which has a small amount of overtaking. Hence, in such a scenario, a plan which minimizes the total overtaking while keeping the total delay cost small might be a more desirable solution. As the first approach, we present a model which achieves this objective. It is appropriate to remark that this fairness criterion has considerable intuitive appeal as it naturally captures the priorities inherent in the original airport arrival schedules.

For every reversible pair of flights $(f, f') \in \mathcal{R}^j$, let $g_{f,f',j}$ denote the total amount of overtaking between flights f and f'. Then, we need to define the following set of variables to express $g_{f,f',j}$.

$$s^i_{f,f',j} = \begin{cases} 1, & \text{if flight } f' \text{ arrives but } f \text{ does not arrive by time } \underline{T}_j^{f'} + i \text{ in resource } j, \\ 0, & \text{otherwise.} \end{cases}$$

The definition above implies the following:

$$s^i_{f,f',j} = 1 \iff \left\{ w^f_{j,\underline{T}^{f'}_j+i} = 0,\ \ w^{f'}_{j,\underline{T}^{f'}_j+i} = 1 \right\}$$

Table 1 summarizes the various feasible combinations of these variables under the above definition. Thus, an alternative way to express $s^i_{f,f',j}$ is as follows:

$$s^i_{f,f',j} = \max\left\{ w^{f'}_{j,\underline{T}^{f'}_j+i} - w^f_{j,\underline{T}^{f'}_j+i}, 0 \right\}. \tag{2}$$

S.No.	$w^f_{j,\underline{T}^{f'}_j+i}$	$w^{f'}_{j,\underline{T}^{f'}_j+i}$	$s^i_{f,f',j}$
1	0	0	0
2	0	1	1
3	1	0	0
4	1	1	0

Table 1: Truth table for modeling the overtaking variables.

Equation (2) implies that the following constraints suffice to express $s^i_{f,f',j}$ in a mathematical programming framework if the objective is to minimize $s^i_{f,f',j}$:

$$s^i_{f,f',j} \geq w^{f'}_{j,\underline{T}^{f'}_j+i} - w^f_{j,\underline{T}^{f'}_j+i}, \tag{3a}$$

$$s^i_{f,f',j} \geq 0. \tag{3b}$$

The set of variables $s^i_{f,f',j}$ are defined for $i \in \{0, \ldots, O^{\max}_{f,f',j}\}$. Now, $g_{f,f',j} \in \{0, \ldots, O^{\max}_{f,f',j}\}$ can be defined as follows:

$$g_{f,f',j} = \sum_{i=0}^{O^{\max}_{f,f',j}} s^i_{f,f',j}. \tag{4}$$

We shall work with an alternative description of Equation (3a) to make the exposition on overtaking clearer. We substitute $i = t - \underline{T}^{f'}_j$ in Equation (3a) to rewrite it as follows:

$$w^{f'}_{j,t} \leq w^f_{j,t} + s^{t-\underline{T}^{f'}_j}_{f,f',j}. \tag{5}$$

We prove next that the following set of constraints are required to model overtaking between $(f, f') \in \mathcal{R}^j$ if we use an objective function to minimize $g_{f,f',j}$ in addition with $s^{t-\underline{T}^{f'}_j}_{f,f',j} \geq 0,\ \forall t \in T^{\text{reversal}}_{f,f',j}$:

$$w^{f'}_{j,t} \leq w^f_{j,t} + s^{t-\underline{T}^{f'}_j}_{f,f',j}, \qquad \forall t \in T^{\text{reversal}}_{f,f',j}. \tag{6}$$

Theorem 1. *If we use an objective function of minimizing* $g_{f,f',j}$ *(the total amount of overtaking for* $(f,f') \in \mathcal{R}^j$*) in addition with* $s_{f,f',j}^{t-\underline{T}_j^{f'}} \geq 0,\ \forall t \in T_{f,f',j}^{reversal}$*, then Constraint (6) correctly captures the semantics of overtaking.*

Proof. In case, there is no reversal, i.e.,

$$w_{j,t}^{f'} \leq w_{j,t}^{f},\ \ \forall t \in T_{f,f',j}^{\text{reversal}},$$

then Constraint (6) becomes reduntant. Since we minimize total amount of overtaking (and $s_{f,f',j}^{t-\underline{T}_j^{f'}} \geq 0,\ \ \forall t \in T_{f,f',j}^{\text{reversal}}$), it forces:

$$s_{f,f',j}^{t-\underline{T}_j^{f'}} = 0,\ \ \forall t \in T_{f,f',j}^{\text{reversal}},$$

leading to $g_{f,f',j} = 0$. On the contrary, if there are i units of overtaking, then $\exists t \in \{\underline{T}_j^{f'}, \ldots, \underline{T}_j^{f'} + O_{f,f',j}^{\max} - i\}$ such that:

$$w_{j,t-1}^{f'} = 0,\ \ w_{j,t}^{f'} = 1,$$

$$w_{j,t+i-1}^{f} = 0,\ \ w_{j,t+i}^{f} = 1.$$

Now, the time-connectivity constraints (Constraints (1f)) imply that:

$$w_{j,t+m}^{f'} = 1,\ \ w_{j,t+m}^{f} = 0,\ \forall\, 0 \leq m \leq i-1.$$

Constraint (6) then enforces

$$s_{f,f',j}^{k} = 1,\ \ \forall\, t \leq k \leq t+i-1.$$

Again, since we minimize total amount of overtaking (and $s_{f,f',j}^{t-\underline{T}_j^{f'}} \geq 0,\ \ \forall t \in T_{f,f',j}^{\text{reversal}}$), therefore,

$$s_{f,f',j}^{k} = 0,\ \ \forall\, 0 \leq k < t,\ \ t+i-1 < k \leq O_{f,f',j}^{\max},$$

leading to $g_{f,f',j} = i$. In summary, Constraint (6) (in addition with $s_{f,f',j}^{t-\underline{T}_j^{f'}} \geq 0$), correctly model overtaking if the objective function is to minimize $g_{f,f',j}$. □

The proof of Theorem 1 relied critically on the assumption that we use an objective function that minimizes $g_{f,f',j}$. Next, we propose a formulation to model overtaking which is independent of the objective function used. We propose a set of constraints that capture the convex hull of the four feasible integer points enumerated in Table 1, namely $(0,0,0), (0,1,1), (1,0,0)$ and $(1,1,0)$. Figure 2 depicts the convex hull of these four points.

We introduce the following set of constraints to model overtaking:

$$w_{j,t}^{f'} \leq w_{j,t}^{f} + s_{f,f',j}^{t-\underline{T}_j^{f'}}, \qquad \forall (f,f') \in \mathcal{R}^j, \;\; j \in \mathcal{S} \cup \mathcal{K}, \;\; t \in T_{f,f',j}^{\text{reversal}}. \tag{7a}$$

$$w_{j,t}^{f} \leq w_{j,t}^{f'} + 1 - s_{f,f',j}^{t-\underline{T}_j^{f'}}, \qquad \forall (f,f') \in \mathcal{R}^j, \;\; j \in \mathcal{S} \cup \mathcal{K}, \;\; t \in T_{f,f',j}^{\text{reversal}}. \tag{7b}$$

$$w_{j,t}^{f} + s_{f,f',j}^{t-\underline{T}_j^{f'}} \leq 1, \qquad \forall (f,f') \in \mathcal{R}^j, \;\; j \in \mathcal{S} \cup \mathcal{K}, \;\; t \in T_{f,f',j}^{\text{reversal}}. \tag{7c}$$

$$-w_{j,t}^{f'} + s_{f,f',j}^{t-\underline{T}_j^{f'}} \leq 0, \qquad \forall (f,f') \in \mathcal{R}^j, \;\; j \in \mathcal{S} \cup \mathcal{K}, \;\; t \in T_{f,f',j}^{\text{reversal}}. \tag{7d}$$

PSfrag replacements

$w_{j,t}^{f}$

$w_{j,t}^{f'}$

$s_{f,f',j}^{i}$

Modeling overtaking

Figure 2: Convex hull of the integer points in Table 1 to model overtaking ($i = t - \underline{T}_j^{f'}$).

The TFMP model with the additional control on total amount of overtaking (referred to as TFMP-Overtake henceforth) is as follows:

$IZ_{\text{TFMP-Overtake}} =$

$$\min \sum_{f \in \mathcal{F}} \left(\sum_{t \in T_{\text{dest}_f}^{f}} c_{\text{total}}^{f}(t) \cdot (w_{\text{dest}_f,t}^{f} - w_{\text{dest}_f,t-1}^{f}) - \sum_{t \in T_{\text{orig}_f}^{f}} c_g^{f}(t) \cdot (w_{\text{orig}_f,t}^{f} - w_{\text{orig}_f,t-1}^{f}) \right) +$$

$$\lambda_s^o \cdot \left(\sum_{j \in \mathcal{S},\ (f,f') \in \mathcal{R}^j} \sum_{i=0}^{O_{f,f',j}^{\max}} s_{f,f',j}^{i} \right) + \lambda_a^o \cdot \left(\sum_{k \in \mathcal{K},\ (f,f') \in \mathcal{R}^k} \sum_{i=0}^{O_{f,f',k}^{\max}} s_{f,f',k}^{i} \right)$$

subject to:

$$(1a) - (1g).$$

$$(7a) - (7d).$$

$$s_{f,f',j}^{i} \in \{0,1\}, \qquad \forall (f,f') \in \mathcal{R}^j, \;\; j \in \mathcal{S} \cup \mathcal{K}, \;\; i \in \{0, \ldots, O_{f,f',j}^{\max}\}.$$

3.2 Minimizing the total number of reversals

The model introduced in Section 3.1 took into account the magnitude of overtaking within each reversal. In this section, we introduce a model which controls the total number of reversals.

For each element $(f, f') \in \mathcal{R}^j$, we introduce the following new variable:

$$s_{f,f',j} = \begin{cases} 1, & \text{if there is a reversal,} \\ 0, & \text{otherwise.} \end{cases}$$

Next, we relate the variables $s_{f,f',j}$ (used to model a reversal) to the variables $s^i_{f,f',j}$ (used to model overtaking). It is evident that a reversal occurs if and only if there is at least one time-period of overtaking. Mathematically, it translates to the following:

$$s_{f,f',j} = 1 \iff \left\{ \exists i \in \{0, \ldots, O^{\max}_{f,f',j}\},\ \ s^i_{f,f',j} = 1 \right\} \tag{8}$$

Building upon Equation (8), we have the following:

$$s_{f,f',j} = \max_{t \in T^{\text{reversal}}_{f,f',j}} \left\{ s^{t - T^{f'}_j}_{f,f',j} \right\},$$

$$s_{f,f',j} = \max_{t \in T^{\text{reversal}}_{f,f',j}} \left\{ \max\{w^{f'}_{j,t} - w^{f}_{j,t}, 0\} \right\},$$

$$s_{f,f',j} = \max \left\{ \max_{t \in T^{\text{reversal}}_{f,f',j}} \{w^{f'}_{j,t} - w^{f}_{j,t}\}, 0 \right\}. \tag{9}$$

Equation (9) implies that the following constraints suffice to express $s_{f,f',j}$ in a mathematical programming framework if the objective is to minimize $s_{f,f',j}$:

$$s_{f,f',j} \geq w^{f'}_{j,t} - w^{f}_{j,t}, \ \ \forall t \in T^{\text{reversal}}_{f,f',j}. \tag{10a}$$

$$s_{f,f',j} \geq 0. \tag{10b}$$

Equation (10a) can be rearranged as follows:

$$w^{f'}_{j,t} \leq w^{f}_{j,t} + s_{f,f',j}, \ \ \forall t \in T^{\text{reversal}}_{f,f',j}. \tag{11}$$

Theorem 2. *If we use an objective function of minimizing $s_{f,f',j}$, then Constraint (11) in addition with $s_{f,f',j} \geq 0$ correctly captures the semantics of modeling a reversal.*

Proof. In case, there is no reversal, i.e.,

$$w^{f'}_{j,t} \leq w^{f}_{j,t}, \ \ \forall t \in T^{\text{reversal}}_{f,f',j},$$

then Constraint (11) becomes redundant. Since we minimize $s_{f,f',j}$ (and $s_{f,f',j} \geq 0$), it forces $s_{f,f',j} = 0$. On the contrary, if there is a reversal, then $\exists t \in T^{\text{reversal}}_{f,f',j}$ such that:

$$w^{f'}_{j,t} = 1, \ \ w^{f}_{j,t} = 0.$$

Constraint (11) then implies that $s_{f,f',j} \geq 1$. Again, minimizing $s_{f,f',j}$ makes $s_{f,f',j} = 1$ ensuring that Constraint (11) indeed models a reversal correctly. □

The proof of Theorem 2 relied critically on the assumption that we use an objective function that minimizes $s_{f,f',j}$. Here, we present a formulation that models a reversal correctly independently of the objective function used. For each element $(f,f') \in \mathcal{R}^j$, we introduce the following constraints to (TFMP):

$$w_{j,t}^{f'} \le w_{j,t}^{f} + s_{f,f',j}, \qquad \forall (f,f') \in \mathcal{R}^j,\ \ j \in \mathcal{S} \cup \mathcal{K},\ \ t \in T_{f,f',j}^{\text{reversal}}. \tag{12a}$$

$$w_{j,t}^{f} \le w_{j,t}^{f'} + 1 - s_{f,f',j}, \qquad \forall (f,f') \in \mathcal{R}^j,\ \ j \in \mathcal{S} \cup \mathcal{K},\ \ t \in T_{f,f',j}^{\text{reversal}}. \tag{12b}$$

If there is a reversal between flights f and f' in resource j, i.e., $s_{f,f',j} = 1$, then Constraint (12a) becomes redundant and Constraint (12b) stipulates that if flight f has arrived by time t, then flight f' has to arrive by that time, hence ensuring that flight f cannot arrive before flight f'. Similarly, if there is no reversal, i.e., $s_{f,f',j} = 0$, then Constraint (12b) becomes redundant and Constraint (12a) stipulates that if flight f' has arrived by time t, then flight f has to arrive by that time, hence ensuring that flight f' cannot arrive before flight f. Thus, we are able to model a reversal with the addition of only one variable ($s_{f,f',j}$).

Given this additional set of constraints, the model then minimizes a weighted combination of total delay costs and total number of reversals. The parameters λ_s^r and λ_a^r are chosen appropriately to control the degree of fairness in sector reversals and airport reversals respectively.

(TFMP) extended with reversals: (TFMP-Reversal).

The TFMP model with the additional control on reversals is as follows:
$IZ_{\text{TFMP-Reversal}} =$

$$\min \sum_{f\in\mathcal{F}} \left(\sum_{t\in T_{\text{dest}_f}^f} c_{\text{total}}^f(t) \cdot (w_{\text{dest}_f,t}^f - w_{\text{dest}_f,t-1}^f) - \sum_{t\in T_{\text{orig}_f}^f} c_g^f(t) \cdot (w_{\text{orig}_f,t}^f - w_{\text{orig}_f,t-1}^f) \right) +$$

$$\lambda_s^r \cdot \left(\sum_{j\in\mathcal{S},\ (f,f')\in\mathcal{R}^j} s_{f,f',j} \right) + \lambda_a^r \cdot \left(\sum_{k\in\mathcal{K},\ (f,f')\in\mathcal{R}^k} s_{f,f',k} \right)$$

subject to:

$$\begin{aligned} &(1a) - (1g). \\ &(12a) - (12b). \\ &s_{f,f',j} \in \{0,1\}, \qquad \forall (f,f') \in \mathcal{R}^j,\ \ j \in \mathcal{S} \cup \mathcal{K}. \end{aligned}$$

For each element $(f,f') \in \mathcal{R}^j$, let $IP_{\text{Reversal}}(f,f',j)$ denote the set of all feasible binary vectors satisfying Constraints (12a) and (12b). We show in the Appendix that the polyhedron induced by Constraints (12a) and (12b) is the convex hull of solutions in $IP_{\text{Reversal}}(f, f', j)$.

$$\begin{aligned} IP_{\text{Reversal}}(f, f', j) = \{ & w^f_{j,t} \in \{0,1\}, \;\; s_{f,f',j} \in \{0,1\} | \\ & w^{f'}_{j,t} - w^f_{j,t} - s_{f,f',j} \leq 0, \qquad t \in T^{\text{reversal}}_{f,f',j}, \\ & w^f_{j,t} - w^{f'}_{j,t} + s_{f,f',j} \leq 1, \qquad t \in T^{\text{reversal}}_{f,f',j}. \} \end{aligned}$$

Remark 2. **RBS Policy - a special case of (TFMP-Reversal).** When there is sufficient capacity at all airports, such that a feasible solution under RBS exists (i.e., there are no reversals), this model is capable of generating that solution (using a sufficiently high λ^r_a) while minimizing the total delay costs. *Hence, a solution under RBS policy is a special case of our model.* Since, a solution under RBS preserves the order of flight arrivals, therefore, for every pair of flights $(f, f') \in \mathcal{R}^A$, the variable $s_{f,f',\text{dest}_f} = 0$, and Constraints (12a) and (12b) reduce to Constraint (13) which ensures that flight f' cannot arrive before flight f:

$$w^{f'}_{\text{dest}_{f'},t} \leq w^f_{\text{dest}_f,t}, \qquad \forall (f, f') \in \mathcal{R}^A, \;\; t \in T^{\text{reversal}}_{f,f',\text{dest}_f}. \tag{13}$$

3.3 Prioritizing between airport reversals and sector reversals

In Section 3.2, we introduced a discrete optimization model with the capability of controlling sector reversals in addition to airport reversals. In this section, we discuss the relative merits of controlling sector reversals. In Section 6.6, we report empirical evidence to validate our conclusions.

To start with, it seems reasonable to expect that primary stakeholders (airlines and flying passengers) won't be bothered as to how the en-route resources are allocated if they are satisfied with the final outcome of the arrival sequences at the airports. In fact, this scheme of things would be acceptable to the FAA as its key objective in designing various ATFM programs is to satisfy these primary stakeholders. In addition to this observation, we feel there are two key arguments in favor of controlling sector reversals:

1. **AFPs.** The first pertains to the operational details of AFPs (one of the recent ATFM programs) under the current CDM practice. The recently introduced AFPs operate much like GDPs, i.e., the arrival slots in the affected airspace are allotted using the RBS principle. Therefore, in a scenario where multiple AFPs and GDPs operate simultaneously, a natural extension of fairness is controlling reversals in the en-route airspace affected by an AFP. To elaborate, we give the following example:

   Example 3.1. Consider the scenario depicted in Figure 3. There is an AFP operational in the en-route airspace followed by a GDP at BOS (Boston). Simultaneously, there is another airport LGA (New York LaGuardia) nearby with no GDP. There are two streams of flights going through the AFP to one of these airports. In a scenario where sector reversals in the AFP are not controlled, a schedule with no airport reversals might be such that all flights going to BOS are allowed to go first before any other flight to

LGA through the AFP. Such a plan might not be acceptable to the stakeholders of LGA because all flights destined to LGA are assigned large delays even though they are part of an ATFM program (the AFP in this case). This might get further exacerbated in a scenario where LGA (the non-GDP airport) is a hub airport for a particular airline, in which case this airline is clearly not treated equitably.

2. **Stochastic capacities.** The second argument is a result of a stochastic setting, wherein the future realized capacity can be quite different compared to the deterministic capacity inputs used in the optimization model. In this scenario, it is possible that controlling sector reversals might lead to a schedule which is able to mitigate the impacts caused due to stochastic nature of future capacities.

PSfrag replacements

BOS (GDP)

LGA (No GDP)

AFP

Figure 3: Illustration of a scenario where controlling sector reversals seems appropriate. There are multiple ATFM programs operating simultaneously. Specifically, an AFP is spatially followed by an airport with a GDP (BOS) and an airport with no GDP (LGA).

On the negative side, we believe that imposing additional constraints of controlling sector reversals will lead to two key impacts: i) increase in system delays over a solution which only controls airport reversals; and ii) potential change in the number of airport reversals (due to downstream effects). Thus, in this balancing act, these two consequences have to be carefully mitigated to ensure neither one gets exacerbated.

In summary, the discussion in this section culminates with the conclusion that *controlling sector reversals should be a secondary goal.* Consequently, we propose that in the model TFMP-Reversal, the tradeoff parameter λ_s^r used to control sector reversals be set to a significantly lower value as compared to the parameter λ_a^r used to control airport reversals.

Extensions. We now elaborate on how the proposed models (TFMP-Reversal and TFMP-Overtake) can be extended to accommodate alternative objective functions.

- *Incorporating alternative objective functions:*
  Although the models presented in this paper minimize the number of reversals and amount of overtaking, it is possible to extend them to accommodate alternative objective functions. For instance, suppose we want to equalize the resulting reversals and overtaking among airlines taking into account the number of flights they operate. This can be achieved as follows:
  Let d_w denote the number of reversals per flight for airline w and γ denote the mean

of the d_w's across all airlines.

$$d_w = \left(\sum_{f' \in \mathcal{F}_w} s_{f,f'} \right) \Big/ |\mathcal{F}_w|,$$

$$\gamma = \left(\sum_{w \in \mathcal{W}} d_w \right) \Big/ |\mathcal{W}|.$$

Then, we add $\left| d_w - \gamma \right|$ term to the objective function of minimizing the total delay cost with an appropriate tradeoff parameter.

Size of the Formulations. Denoting with

$$N = \max_{f \in \mathcal{F}} |\mathcal{S}^f|,$$

the total number of decision variables and constraints for the various models can be bounded as listed in Table 2.

Model	No. of Decision Variables	No. of Constraints
TFMP	$\lvert\mathcal{F}\rvert MN$	$2\lvert\mathcal{K}\rvert\lvert\mathcal{T}\rvert + \lvert\mathcal{S}\rvert\lvert\mathcal{T}\rvert + 2\lvert\mathcal{F}\rvert MN + 2\lvert\mathcal{F}\rvert N + M\lvert\mathcal{C}\rvert$
TFMP-Reversal	$\lvert\mathcal{F}\rvert MN + \lvert\mathcal{R}^A\rvert$	$2\lvert\mathcal{K}\rvert\lvert\mathcal{T}\rvert + \lvert\mathcal{S}\rvert\lvert\mathcal{T}\rvert + 2\lvert\mathcal{F}\rvert MN + 2\lvert\mathcal{F}\rvert M + M\lvert\mathcal{C}\rvert + 2\lvert\mathcal{R}^A\rvert M$
TFMP-Overtake	$\lvert\mathcal{F}\rvert MN + M\lvert\mathcal{R}^A\rvert$	$2\lvert\mathcal{K}\rvert\lvert\mathcal{T}\rvert + \lvert\mathcal{S}\rvert\lvert\mathcal{T}\rvert + 2\lvert\mathcal{F}\rvert MN + 2\lvert\mathcal{F}\rvert N + M\lvert\mathcal{C}\rvert + \lvert\mathcal{R}^A\rvert M$

Table 2: Upper bound on the size of the models.

In order to get a feeling of the size of the formulations, let us consider an example that adequately represents the U.S. network.

Example 3.2. Let $|\mathcal{K}| = 50$, $|\mathcal{T}| = 100$, $|\mathcal{S}| = 100$, $|\mathcal{R}^A| = 50000$, $|\mathcal{F}| = 10000$, $|\mathcal{C}| = 8000$, $M = 6$ and $N = 5$. The duration of each period is 15 minutes implying a planning horizon of almost a day. For this example, the upper bound on the number of variables and constraints are listed in Table 3.

Model	No. of Decision Variables	No. of Constraints
TFMP	300,000	780,000
TFMP-Reversal	350,000	1,380,000
TFMP-Overtake	600,000	1,080,000

Table 3: Numerical Example: Upper bound on the size of the models.

Since we introduce only one class of variables $s_{f,f',j}$ for all elements $(f, f') \in \mathcal{R}^j$, the number of variables in the model (TFMP-Reversal) are comparable to the original model (TFMP).

4 Network model for slot reallocation that incorporates airline collaboration

In this section, we present a model (called TFMP-Trading) for slot reallocation in a network setting that introduces only one additional variable per offer above the TFMP model. We let airlines submit offers to trade slots assigned to its flights across different airports. The executed set of trades should ensure that the resulting schedule is still feasible taking into account all kinds of network connectivities and airspace capacities.

The set $\mathcal{O}$ (Set of Airline Offers)

We give next the definition of $\mathcal{O}$ (set of airline offers). We use a structure proposed by Vossen and Ball [19] that allows the airlines to submit so-called "at-least, at-most" offers. Airlines submit offers of the following kind: $(f_d, t_{d'}; f_u, t_{u'})$ which means that the airline is willing to move flight f_d to a later time-period, but no later than $t_{d'}$; in return for moving flight f_u to an earlier time-period, but no later than $t_{u'}$. The destination airports of flights f_d and f_u are allowed to be distinct. Figure 4 gives an example to illustrate the semantics of the offer structure. The set $\mathcal{O}$ contains all such four-tuples $(f_d, t_{d'}; f_u, t_{u'})$ submitted by the airlines after a schedule is generated from Stage I of our proposal. Note that c_{f_d} and c_{f_u} denote the slots allotted to the two flights from Stage I, and hence, for such an offer to be useful, we must have $c_{f_d} < t_{d'}$ and $c_{f_u} > t_{u'}$. Finally, $\mathcal{O}^f \subseteq \mathcal{O}$ defines the set of offers containing flight f.

PSfrag replacements

c_{f_d}

$t_{d'}$

$t_{u'}$

c_{f_u}

Airport A

Airport B

Figure 4: Illustration of the structure of an offer $(f_d, t_{d'}; f_u, t_{u'})$. $\text{dest}_{f_d} = A$, $\text{dest}_{f_u} = B$, $t_{d'} = c_{f_d} + 3$ and $t_{u'} = c_{f_u} - 4$. The offer states that the airline is willing to delay flight f_d by at-most 3 slots if in return flight f_u is moved earlier by at-least 4 slots.

4.1 (TFMP-Trading): A model for slot reallocation in a network setting

Here, we introduce our model of slot reallocation in a network setting. The model only introduces one additional variable per offer in addition to $w^f_{j,t}$, which are the variables used in the TFMP model of Bertsimas Stock-Patterson [5].

The Decision Variables

- $o_{dd'uu'} \in \{0,1\} = 1$ if offer $(f_d, t_{d'}; f_u, t_{u'})$ is executed.

Constraints

$$
\begin{aligned}
& (1a) - (1g). \\
o_{dd'uu'} &\leq w^{f_d}_{\text{dest}_{f_d}, t_{d'}}, && \forall (f_d, t_{d'}; f_u, t_{u'}) \in \mathcal{O}. && \text{(14a)} \\
o_{dd'uu'} &\leq w^{f_u}_{\text{dest}_{f_u}, t_{u'}}, && \forall (f_d, t_{d'}; f_u, t_{u'}) \in \mathcal{O}. && \text{(14b)} \\
\sum_{j \in \mathcal{O}^f} o_j &\leq 1, && \forall f \in \mathcal{F}. && \text{(14c)} \\
w^f_{\text{dest}_f, c_f} - w^f_{\text{dest}_f, c_f - 1} &\geq 1 - (\sum_{j \in \mathcal{O}^f} o_j), && \forall f \in \mathcal{F}. && \text{(14d)} \\
w^f_{j,t} &\in \{0,1\}, && \forall f \in \mathcal{F}, \;\; j \in \mathcal{S}^f, \;\; t \in T^f_j. && \text{(14e)} \\
o_{dd'uu'} &\in \{0,1\}, && \forall (f_d, t_{d'}; f_u, t_{u'}) \in \mathcal{O}. && \text{(14f)}
\end{aligned}
$$

Constraints (14a) and (14b) enforce that when an offer $o_{dd'uu'}$ is executed (i.e., $o_{dd'uu'} = 1$), then $w^{f_d}_{\text{dest}_{f_d}, t_{d'}} = 1$ and $w^{f_u}_{\text{dest}_{f_u}, t_{u'}} = 1$, i.e., flights f_d and f_u cannot arrive after the respective time-periods in the offer, namely, t'_d and t'_u. This ensures that the semantics of the structure of an offer are satisfied. Constraint (14c) enforces that for each flight, at most one offer can get executed. Moreover, constraint (14d) stipulates that if no offer for a flight f is executed (i.e., $o_j = 0, \; \forall j \in \mathcal{O}^f$), then the flight will arrive at the time-period allotted from Stage I (c_f).

Objective Function

In the model presented above, we have not explicitly stated the objective function that should be used. It is evident that fairness in the number of executed offers across airlines would again be relevant in this stage of our proposal.

Let n_w denote the number of trades executed corresponding to airline w, and let γ denote the mean of the trades executed across all airlines.

$$
\begin{aligned}
n_w &= \sum_{f \in \mathcal{F}_w, j \in \mathcal{O}^f} o_j, \\
\gamma &= \Big(\sum_{w \in \mathcal{W}} n_w \Big) \Big/ |\mathcal{W}|.
\end{aligned}
$$

In the next section, we report computational results based on the following two objective functions:

- *Objective 1*: maximize the total number of trades (max $\sum_{(f_d,t_{d'};f_u,t_{u'})\in\mathcal{O}} o_{dd'uu'}$).
- *Objective 2*: minimize the difference in the number of trades executed for each airline from the mean (min $\sum_{w\in\mathcal{W}} |n_w - \gamma|$).

5 Integration and comparison with current CDM practice

Given that our aspiration in this research effort is to bridge the gap between theory and practice, we now elaborate on how our proposal can be integrated within the broader CDM paradigm currently in use. Our goal in this section is to demonstrate the practical viability of our procedural framework and verify the compatibility of the input and output data requirements of our models with the available operational technology.

We start by revisiting the CDM paradigm alluded to in the Introduction and provide a detailed description of the phases involved in the coordination of various ATFM initiatives (like GDPs and AFPs). There are three key phases involved in the decision-making process:

1. **RBS for each ATFM program.** FAA invokes the RBS policy to allocate arrival slots to the airlines for each ATFM program based on the original schedule ordering.
2. **Airline response to schedule disruption.** Based on the slots allotted, an airline is allowed to make changes to the schedule by canceling flights and swapping the slots of two or more of its own flights if they are compatible with the scheduled departure times.
3. **Final coordination by the FAA.** FAA accepts the relevant changes proposed by the airlines to come up with a overall feasible schedule. This is further complemented by Compression (wherein the FAA attempts to fill in any holes created by cancellations to further optimize the final schedule).

We now demonstrate how our proposal can be integrated within the three-stage CDM framework described above. To start with, we propose to filter in the flights affected by all ATFM programs that the FAA intends to use. The trajectories of all other flights are deterministically fixed and the capacity corresponding to them is removed from the capacity inputs to our optimization models.

- **Stage 1: Control reversals/overtaking.** This stage of our proposal interfaces with phase 1 of the CDM framework. Rather than applying RBS to each ATFM program, we control the reversals and overtaking in the resulting flight sequences by using TFMP-Reversal and TFMP-Overtake. The input requirements for our models, namely, the set of feasible times that a flight can be in a sector and the capacity inputs are readily available from Flight Schedule Monitor (FSM)[1]. Furthermore, the output of our models can be easily converted to a slot assignment for each flight (by following the scheduled order of the slots allotted for flights during each time-period) and thus is compatible with the current operational practice.

- **Stage 2: Airline collaboration.** This stage of our proposal interfaces with the last two phases of the CDM framework. The input required from the airlines on the offers to trade are readily available as the airlines know the slots allotted to them from Stage 1. Moreover, airlines also propose flights cancellations (emanating from operational infeasibility) after considering their slot assignments. The final exercise of Compression goes through as is done presently to fill in gaps.

These two stages can be repeated (if necessary) to enable program revisions. In summary, we believe that our proposal fits well within the CDM framework used currently and the data input and output requirements are compatible with operational feasibility.

6 Computational Results

In this section, we report computational results from the optimization models introduced in Section 3 and Section 4 on national-scale, real world datasets spanning across six days. The dataset for each day encompasses 55 major airports of the US and covers operations of the top five airlines. Each dataset contains data on the actual flight arrival and departure times for that particular day which lets us compute the actual delays.

6.1 Statistics of the Datasets

Table 4 summarizes the statistics of six days of flights data. These correspond to the operations at the 55 major airports of the US. We filter in the flights corresponding to the operations of the top 5 airlines (measured by the number of flight operations) - Southwest (SWA), American (AAL), Delta (DAL), United (UAL) and Northwest (NWA) to enable us to better analyze the results.

6.2 Experimental Setup

In our experimental setup, the airspace is subdivided into sectors of equal dimensions (10 by 10) that form a grid, thereby, having a total of 100 sectors. The 55 major airports of the

[1]FSM is an important software technology currently used by FAA that provides users with up-to-date, real-time information on the future flight behavior and provides common system-wide situational awareness for all stakeholders

Day	No. of Flights	No. of Connecting Flights	Total Delay (units of 15 min.)
14th Jul'04	5092	2691	4438
4th Aug'04	5844	3298	4926
13th May'05	5780	3310	3079
16th Jul'05	4590	2301	3907
27th Jul'05	5128	2728	3326
27th Jul'06	4781	2504	3101

Table 4: Summary of the datasets.

US are then mapped to one of these 100 sectors based on its geographical coordinates. For each flight, we fix its flight trajectory (i.e., the sequence of sectors in its path) based on a randomized shortest path from the origin to the destination airport. Using the information on its flight time, we compute the minimum amount of time that each flight has to spend in a sector. This value is then used to calculate the set of feasible times that a flight can be in a sector. By tracking the tail number of an aircraft, we form the set of connecting flights. We believe this setup represents a reasonably approximate testbed of the actual NAS operations to study the implications of incorporating various fairness criteria and the resulting impacts on efficiency.

For a sample day, we know the scheduled departure and arrival times of a flight as well as what actually happened on that day. We use this to compute its ground and air-hold delays. Further, we compute the base capacities at all the resources (airports and sectors) by noting the actual times of departure and arrival of the flights. It is important to note that this capacity corresponds to the exact number of flight departures and arrivals that happened on that particular day and hence, is the most conservative estimates of the capacity. The available capacities on that day has to be higher than the actual number of operations. Although, data difficulties make it hard to infer the exact capacities available for the concerned day. Nonetheless, we use educated augmented values of base capacity as capacity inputs to run the optimization models. Specifically, we increase the capacity by a certain percentage over the number of aircraft operations that occurred for each capacitated resource (we use a slack of 20%). This ensures that the weather impact across the airspace is captured.

A critical parameter of the optimization models is the maximum permissible delay for a flight (M). This value is used to define the set of feasible times that a flight can be in a particular sector. For example, the set of feasible times that a flight f can arrive at its destination airport is given by all values in a_f^k through $(a_f^k + M)$, where $\text{dest}_f = k$. The size of all the optimization models and hence, the computational times, are directly proportional to the value of M (e.g., for $M \geq 10$, the number of variables and constraints quickly scale up to the order of millions). The use of this parameter is not motivated from equity considerations and has no consequences on the fairness properties of the resulting schedules. We use a value of $M = 6$, which corresponds to 6 time periods (each of length 15 minutes), hence permitting a maximum delay of 90 minutes.

Remark 3. While a small value of M will facilitate the computational efficiency of the optimization models, the framework is generic enough to allow for a larger value of M for a select set of flights for which higher delays are expected (or desired).

To compute optimal solutions, we use the CPLEX-MIP solver 11.0, implemented using AMPL as a modeling language on a laptop with 2 GB RAM and Linux Ubuntu OS. The instances reported in this paper have a typical size of the order of 300,000 variables (this increases significantly for TFMP-Overtake) and 800,000 constraints (this increases significantly for TFMP-Reversal).

6.3 Performance of (TFMP)

Table 5 reports solutions from the (TFMP) model for the case when the capacity used is 20% over and above the actual number of aircraft operations. There is an average reduction of 23% in the total absolute delays. This illustrates the benefits that could be achieved by using a centralized optimization-based approach. Furthermore, the number of reversals consistently range between 500 and 1000 and amount of overtaking range between 800 and 1500 across all days. This confirms that, although, there can be significant benefits in the total delay costs by using the model (TFMP), the number of reversals and overtaking might be high.

Day	No. of Flights	Actual Delay (units of 15 min.)	TFMP Delay (units of 15 min.)	Number of Reversals	Amount of Overtaking
14 Jul'04	5092	4438	3385	915	1492
4 Aug'04	5844	4926	3492	924	1426
13 May'05	5780	3079	2242	753	1191
16 Jul'05	4590	3907	3053	769	1235
27 Jul'05	5128	3326	2648	801	1291
27 Jul'06	4781	3101	2542	526	822

Table 5: Performance of (TFMP).

6.4 Performance of (TFMP-Overtake)

Table 6 reports the computational performance of (TFMP-Overtake) model on the six datasets. These results pertain to the parameter λ_a^o set to 100. The number reported under 'Total overtaking' takes into account the relative magnitudes of overtaking within each reversal, i.e., the number of time periods by which a flight overtakes its preceding flight when a reversal occurs. The degradation in total delay costs from (TFMP-Overtake) model over the (TFMP) solution range between 13% and 41% for fairness at 25 airports, the average being 24.5%. The model on average takes less than 30 minutes to converge to optimality for up to 25 airports.

Day (# of Flights)	No. of Airports with Fairness	Solution Time (sec.)	No. of Reversals	Total Overtaking	Total Delay Cost (units of 15 min.)	% Increase in Delay Cost over (TFMP)
	0	261	915	1492	3525	
14 Jul	5	186	2	4	3690	4.68
2004	15	727	13	27	4243	20.36
(5092)	25	3073	26	39	4662	32.25
	30	3600	39	65	4815	36.59
	0	108	924	1426	3604	
4 Aug	5	206	1	2	3802	5.49
2004	15	596	6	9	4029	11.79
(5844)	25	806	9	12	4080	13.20
	30	3530	16	18	4510	25.13
	0	311	753	1191	2313	
13 May	5	170	3	6	2401	3.80
2005	15	397	8	18	2584	11.71
(5780)	25	295	11	22	2651	14.61
	30	3394	17	28	3096	33.85
	0	51	769	1235	3173	
16 Jul	5	150	1	1	3628	14.33
2005	15	746	5	6	4201	32.39
(4590)	25	691	13	18	4452	40.30
	30	3600	29	33	4743	49.47
	0	178	801	1291	2744	
27 Jul	5	116	0	0	2871	4.62
2005	15	492	10	18	3319	20.95
(5128)	25	1983	17	26	3505	27.73
	30	3600	25	36	3804	38.62
	0	49	526	822	2637	
27 Jul	5	143	5	7	2826	7.16
2006	15	378	9	15	3070	16.42
(4781)	25	479	15	22	3145	19.26
	30	1305	28	49	3383	28.28

Table 6: Computational performance of (TFMP-Overtake). Note that the row with k airports corresponds to imposing fairness at k airports and no fairness at the remaining $|\mathcal{K}|-k$ airports. In particular, $k = 0$ corresponds to the (TFMP) solution.

6.5 Performance of (TFMP-Reversal)

The (TFMP-Reversal) model minimizes a weighted combination of total delay costs and total number of reversals where λ_a^r is the weight parameter. We study the tradeoff inherent in these conflicting objectives in two ways - a) as a function of the tradeoff parameter λ_a^r and b) as a function of the number of airports where this fairness criterion is imposed.

The effect of the tradeoff parameter

Figure 5 plots the tradeoff in the number of reversals with the total delay cost as a function of λ_a^r for fairness based on controlling total reversals imposed at 25 airports. The five points on the plot for each day correspond to the result from (TFMP-Reversal) with $\lambda_a^r = 0$, 1, 10, 100 and 1000. Initially, there is a significant reduction in the number of reversals at the cost of a small increase in total delay cost, but the subsequent benefits in the number of reversals come at a high cost. For all days, the model is able to achieve less than 100 reversals for a degradation of at most 10% in the total delay cost. To achieve reversals between 10 and 30, the degradation in total delay costs range between 10% and 40% across all days.

PSfrag replacements

ncrease in total delay cost over TFMP

Total number of airport reversals

n airport reversals and total delay cost

14 Jul'04

4 Aug'04

13 May'05

16 Jul'05

27 Jul'05

27 Jul'06

Figure 5: Effect of the tradeoff parameter λ_a^r. The five points for each day correspond to the result from (TFMP-Reversal) with $\lambda_a^r = 0$, 1, 10, 100 and 1000.

The effect of the number of airports

Table 7 reports the computational performance of the (TFMP-Reversal) model on the six datasets as a function of the number of airports where this fairness criterion is imposed. These results pertain to the tradeoff parameter λ_a^r set to 100. As is evident from the results reported across all days, the number of reversals can be controlled up to 10-30. The degradation in total delay costs from (TFMP-Reversal) model over the (TFMP) solution range between

13% and 40% for fairness at 25 airports, the average being 24.5%. The model on average takes less than 30 minutes to converge to optimality for up to 25 airports. As expected, the total amount of reversals reported in Table 7 is always less than the corresponding number in Table 6, whereas the opposite is true for the amount of overtaking.

Day (# of Flights)	No. of Airports with Fairness	Solution Time (sec.)	No. of Reversals	Total Overtaking	Total Delay Cost (units of 15 min.)	% Increase in Delay Cost over (TFMP)
	0	261	915	1492	3525	
14 Jul	5	228	2	4	3691	4.70
2004	15	899	13	30	4246	20.45
(5092)	25	3600	25	53	4402	24.87
	30	3600	37	84	4586	30.09
	0	108	924	1426	3604	
4 Aug	5	213	1	2	3785	5.02
2004	15	837	6	9	4028	11.76
(5844)	25	1024	9	12	4077	13.12
	30	3600	16	30	4403	22.16
	0	311	753	1191	2313	
13 May	5	219	3	6	2406	4.02
2005	15	282	8	19	2583	11.67
(5780)	25	384	11	24	2648	14.48
	30	3600	15	35	3048	31.77
	0	51	769	1235	3173	
16 Jul	5	165	1	1	3627	14.30
2005	15	1367	5	10	4138	30.41
(4590)	25	2292	12	22	4468	40.81
	30	3600	29	55	4558	43.64
	0	178	801	1291	2744	
27 Jul	5	225	0	0	2866	4.44
2005	15	654	10	24	3319	20.95
(5128)	25	1556	16	35	3504	27.69
	30	3600	25	38	3925	43.03
	0	49	526	822	2637	
27 Jul	5	213	5	7	2836	7.54
2006	15	751	9	16	3059	16.00
(4781)	25	654	15	25	3125	18.50
	30	3600	28	51	3387	28.44

Table 7: Computational performance of (TFMP-Reversal). Note that the row with k airports corresponds to imposing fairness at k airports and no fairness at the remaining $|\mathcal{K}|-k$ airports. In particular, $k = 0$ corresponds to the (TFMP) solution.

The computational times of both (TFMP-Reversal) and (TFMP-Overtake) are consistently less than 30 minutes for up to 25 airports, but they break down when we impose fairness at 30 airports and above. We believe that this is due to memory limitations. Given the fact that for some instances, the computational times are still prohibitively large, we study the times needed to achieve a solution that is within 20% of the optimal. In all cases, we could achieve feasible solutions satisfying this optimality criterion in less than 15 minutes.

6.6 Controlling sector reversals and balancing with airport reversals

As concluded in Section 3, we believe that controlling airport reversals (and overtaking) is our primary objective, and controlling sector reversals represents a secondary goal. In this section, we study the interaction of the two objectives with the aim to quantify the price of controlling sector reversals (on airport reversals and total delay cost). Our setup for this exercise comprises of controlling reversals in 5 sectors of the north-east region of the US and a set of 10 airports that lie spatially close to these sectors (we believe this would be the typical setting with multiple AFPs and GDPs operating simultaneously). Table 8 reports the sector reversals, airport reversals and total delay cost for different combinations of the tradeoff parameters λ_r^s and λ_r^a.

The left plot in Figure 6 is a box plot quantifying the percentage increase in the number of airport reversals and the delay cost by enforcing the additional control on sector reversals for the tradeoff parameter $\lambda_r^s = 10$. The percentage increase in delay cost lies between 5% and 20% with a mean of around 13%, whereas the percentage increase in airport reversals is around 5%. Moreover, the sector reversals can be reduced from the order of 1000s to 10s. In contrast, for $\lambda_r^s = 1$, the average increase in the total delay cost is 3% on average but the sector reversals are still in the 100s. The right plot in Figure 6 depicts the reduction in number of sector reversals possible by this explicit control (potential reduction from four digit reversals to two digit reversals).

Consequently, our overall conclusion regarding control of sector reversals is as follows:

1. We have developed a model capable of controlling sector reversals in conjunction with airport reversals.

2. We believe controlling sector reversals is a secondary goal after achieving the primary objective of controlling airport reversals. This is validated by our computational experiments of the form presented here (the impact on total delay cost is substantial (13% on average) for large λ_r^s (=10) and is relatively small (3% on average) for small λ_r^s (=1)).

6.7 Interaction with super-linear cost coefficients

We used super-linear cost coefficients in the overall objective function as additional means to impose equity. As explained earlier in Section 2, this eliminates flights with extreme delays. Since, our primary fairness proposal is controlling reversals and overtaking, we study the interaction of super-linear cost coefficients with this fairness criteria.

Day	$\lambda_r^s = 0,\ \lambda_r^a = 100$			$\lambda_r^s = 1,\ \lambda_r^a = 100$			$\lambda_r^s = 10,\ \lambda_r^a = 100$		
	Sector Revs	Airport Revs	Delay Cost	Sector Revs	Airport Revs	Delay Cost	Sector Revs	Airport Revs	Delay Cost
14 Jul'04	3174	8	3780	253	11	3865	127	12	4261
4 Aug'04	3455	6	3860	169	8	3944	59	8	4235
13 May'05	1289	9	2486	58	9	2493	14	10	2631
16 Jul'05	3242	5	3663	326	6	3874	136	4	4292
27 Jul'05	1646	0	2897	268	0	2958	75	1	3456
27 Jul'06	1615	10	2832	144	10	2907	46	9	3235

Table 8: Balancing Sector Reversals with Airport Reversals (Revs stand for Reversals). Fairness imposed in 5 sectors of the north-east region and 10 airports spatially close to these sectors.

PSfrag replacements

% Increase

Sector Revs (λ_r^s=10)

Airport Reversals

Delay Cost

Reversals

Delay Cost

PSfrag replacements

Day

Number of reversals

Sector Reversals: Before and After

$\lambda_r^s = 0$

$\lambda_r^s = 1$

$\lambda_r^s = 10$

Figure 6: Impact of controlling sector reversals on i) airport reversals and total delay cost (Left); and ii) sector reversals (Right).

The left plot in Figure 7 is a box plot quantifying the percentage increase in the number of airport reversals and the total delay by using super-linear coefficients over the solution obtained using linear coefficients (i.e., $\epsilon = 0$). The percentage increase in reversals consistently lies between -2% and 2% with a mean of around 0.3%. Furthermore, the total delays never differ by more than 1% with a mean difference of 0.1%. This highlights that using super-linear coefficients causes insignificant changes to the fairness and delay characteristics of the resulting schedules. The right side of Figure 7 is a table depicting the distribution of delays across flights. As is evident, the use of super-linear coefficients induces a more moderate assignment of delays in contrast to linear coefficients which lead to more flights with either 0 (minimum) or 6 (maximum) units of delay.

In summary, the use of super-linear cost coefficients achieves its objective of reducing flights with extreme delays while not causing any material changes to the number of reversals and total delay cost.

PSfrag replacements

e over Linear Coeffs.

f Super-linear coeffs.

Reversals

Delay Cost

Revesals

Delay Cost

Delay (units of 15 min)	Linear Coeffs	Super-linear Coeffs
0	4288	4280
1	398	402
2,3,4	340	346
5	51	52
6	15	12

Figure 7: Impact of super-linear cost coefficients on i) reversals and total delay cost (Left); and ii) distribution of flight delays (Right).

6.8 Performance of slot reallocation model (TFMP-Trading)

Given the assignment of flights to various time-periods from Stage I, we generate offers to trade for each airline which maximizes its on-time performance. In other words, each airline tries to maximize the number of flights with delay less than one time-unit (15 minutes). We give an example to elaborate on this.

Example 6.1. Suppose two flights f_1 and f_2 (belonging to the same airline) have been assigned two time-units of delay each from Stage I optimization. Moreover, let t_1 and t_2 be the time-periods assigned to the two flights respectively. Then, the owner airline generates an offer to trade which says that it is willing to delay flight f_1 further by three time-units if in return flight f_2 can arrive within one time-unit of delay, i.e., it generates the offer $(f_1, t_1 + 3; f_2, t_2 - 1)$. Thus, in case, this trade is executed, then flight f_2 will arrive on-time, thereby improving the internal objective function of the airline (which is to maximize the on-time performance).

Table 9 reports the results from our network slot reallocation model (TFMP-Trading). As is evident, the computational times of the model TFMP-Trading is consistently less than a minute which is encouraging for practical deployment. Figure 8 plots the distribution of the number of executed trades across airlines from TFMP-Trading for the two objectives. Not surprisingly, when fairness is not imposed (*Objective 1*), the distribution is quite skewed (e.g., for Day 1, Airline 5 gets almost three times more trades executed than Airline 1). In contrast, when fairness is explicitly incorporated (*Objective 2*), all airlines get exactly the same number of trades executed.

Comparison of TFMP-Trading between single-airport and network-wide settings

In this section, we contrast the performance of TFMP-Trading between single-airport and network-wide settings. It is evident that in the network version, there is a tradeoff between the flexibility of trading slots at different airports versus the added constraint of satisfying all network connectivities. In contrast, in a single-airport setting, there is the advantage of not having to satisfy the network connectivity requirements at the expense of losing on trades across different airports. So, to compare the two settings, we divide the set of generated offers between local and network offers as defined below.

Definition 6.1. An offer $(f_d, t_{d'}; f_u, t_{u'}) \in \mathcal{O}$ is

- a **local offer** if both the flights involved have the same destination airport, i.e., $\text{dest}_{f_d} = \text{dest}_{f_u}$.
- a **network offer** if the destination airports of the two flights are distinct, i.e., $\text{dest}_{f_d} \neq \text{dest}_{f_u}$.

Hence, the single-airport model will not have any executed trades that correspond to the network offers. Our overall goal is to demonstrate the utility of TFMP-Trading in executing trades across offers containing distinct airports. Table 10 reports the results from the two versions under the two objectives described in Section 4.1. The number reported under TFMP-Trading is the number of offers executed from the network model proposed in this paper, whereas, SA-Trading (SA stands for single-airport) reports the results obtained from TFMP-Trading after removing all the network satisfiability constraints and only taking into account the local offers. The numbers reported highlight the tradeoffs inherent in ignoring the network effects vis-a-vis trading slots at different airports. As the percentage of local offers increases, SA-Trading outperforms TFMP-Trading emphasizing that network connectivities are indeed relevant, whereas, for higher fraction of network offers, TFMP-Trading performs better reinforcing the utility of network offers. Finally, the number of trades executed with Objective 2 (when fairness is included) are less than 1 Objective which is along expected lines and demonstrate the flexibility of choosing alternative objective functions.

Day	**Objective 1**			**Objective 2**		
	Objective Cost	Offers Executed	Sol. Time (in sec.)	Objective Cost	Offers Executed	Sol. Time (in sec.)
14 Jul'04	283	283	13	0	190	29
4 Aug'04	256	256	14	0	180	31
13 May'05	150	150	5	0	140	19
16 Jul'05	278	278	11	0	215	27
27 Jul'05	194	194	9	0	140	22
27 Jul'06	153	153	6	0	90	16

Table 9: Computational performance of TFMP-Trading.

PSfrag replacements

Day

o. of executed trades

Airline 1

Airline 2

Airline 3

Airline 4

Airline 5

PSfrag replacements

Day

No. of executed trades

Airline 1

Airline 2

Airline 3

Airline 4

Airline 5

Figure 8: Distribution of number of executed trades across airlines from TFMP-Trading. Left: (*Objective 1*); Right: (*Objective 2*)

% Local Offers	% Network Offers	No. of Offers Executed (Obj. 1)		No. of Offers Executed (Obj. 2)	
		SA-Trading	**TFMP-Trading**	**SA-Trading**	**TFMP-Trading**
0	100	0	281	0	190
25	75	152	255	130	90
50	50	221	241	175	125
75	25	269	238	195	120
100	0	308	258	240	175

Table 10: Comparison of TFMP-Trading between single-airport and network-wide settings.

6.9 Summary of computations

To summarize, TFMP (the optimization model without fairness) is able to reduce the total delays by 23% on average when we take reasonable estimates on the available capacity (increasing it by 20% from the actual number of flight operations). In Stage I of our proposal, we obtain solutions that are able to control the total reversals and overtaking at airports up to less than 100 (from the models TFMP-Reversal and TFMP-Overtake respectively). In addition, we report promising computational times of less than 30 minutes for up to 25 airports from both models which make them well suited for online use. Furthermore, we also demonstrate the ability to control sector reversals and conclude that this should be a secondary objective because of the prohibitively large increase in delay cost (13% on average in the setting analyzed in this paper). In Stage II of our proposal, there are 220 trades executed on average when the objective function used is to maximize the number of trades, and 160 when we impose fairness. This reinforces the utility of the slot reallocation phase of our proposal as the airlines are able to increase the number of flights arriving on-time.

7 Conclusions

In this paper, we present a proposal that extends the CDM framework from an airport setting to an airspace context while explicitly incorporating fairness and airline collaboration. Specifically, in Stage I of our proposal, we introduce models for the ATFM problem that incorporate notions of fairness. Given a schedule from this stage, in Stage II, we propose an optimization model for slot reallocation to mimic the current substitution-cancellation process in a GDP setting. Further, we report empirical results of the proposed models on national-scale, real world datasets. We feel the key advantages of our proposal are high-quality of solutions, consideration of network effects and promising computational times which ensure practical tractability in real-time settings.

Acknowledgement

This research has been funded by the NSF Grant EFRI-0735905. We thank Bill Moser and Mark Weber of Lincoln Labs for providing us with the data.

Appendix

Strength of TFMP-Reversal:
Let us denote the polyhedron induced by the the additional set of constraints to model a reversal for each element $(f, f') \in \mathcal{R}^j$ as $P_{\text{Reversal}}(f, f', j)$.

Proposition 1. *The polyhedron $P_{\text{Reversal}}(f, f', j)$ is integral.*

Proof. $P_{\text{Reversal}}(f, f', j)$ can be written as follows:

$$
\begin{aligned}
P_{\text{Reversal}}(f, f', j) = \Big\{ x = (w^f_{j,t},\ s_{f,f',j}) |\ & 0 \le w^f_{j,t} \le 1,\ \ 0 \le s_{f,f',j} \le 1, \\
& w^{f'}_{j,t} - w^f_{j,t} - s_{f,f',j} \le 0, \qquad t \in T^{\text{reversal}}_{f,f',j}, \\
& w^f_{j,t} - w^{f'}_{j,t} + s_{f,f',j} \le 1, \qquad t \in T^{\text{reversal}}_{f,f',j}. \Big\}
\end{aligned}
$$

We make use of the following two facts from discrete optimization [8]:

Fact 1. Let $\mathbf{A}$ be an integral matrix. $\mathbf{A}$ is totally unimodular if and only if $\{\mathbf{x} |\ \mathbf{a} \le \mathbf{A}\mathbf{x} \le \mathbf{b}, \mathbf{l} \le \mathbf{x} \le \mathbf{u}\}$ is integral, for all integral vectors $\mathbf{a}$, $\mathbf{b}$, $\mathbf{l}$, $\mathbf{u}$.

Fact 2. A matrix $\mathbf{A}$ is totally unimodular if and only if each collection Q of rows of $\mathbf{A}$ can be partitioned into two parts so that the sum of the rows in one part minus the sum of the rows in the other part is a vector with entries only 0, +1 and -1.

Consider the following polyhedron P and let $\mathbf{A}$ be the matrix such that $P = \{\mathbf{x} |\ \mathbf{A}\mathbf{x} \le \mathbf{b}\}$:

$$
\begin{aligned}
P = \Big\{ & x = (w^f_{j,t},\ s_{f,f',j}) | \\
& w^{f'}_{j,t} - w^f_{j,t} - s_{f,f',j} \le 0, \qquad t \in T^{\text{reversal}}_{f,f',j}, \\
& w^f_{j,t} - w^{f'}_{j,t} + s_{f,f',j} \le 1, \qquad t \in T^{\text{reversal}}_{f,f',j}. \Big\}
\end{aligned}
$$

$$\mathbf{A} = \begin{bmatrix} 1 & -1 & 0 & 0 & 0 & \cdots & \cdots & 0 & -1 \\ -1 & 1 & 0 & 0 & 0 & \cdots & \cdots & 0 & 1 \\ 0 & 0 & 1 & -1 & 0 & \cdots & \cdots & 0 & -1 \\ 0 & 0 & -1 & 1 & 0 & \cdots & \cdots & 0 & 1 \\ \vdots & \vdots & \ddots & \ddots & \ddots & \ddots & \vdots & \vdots & -1 \\ \vdots & \vdots & \ddots & \ddots & \ddots & \ddots & \vdots & \vdots & 1 \\ 0 & 0 & \cdots & \cdots & 0 & 0 & 1 & -1 & -1 \\ 0 & 0 & \cdots & \cdots & 0 & 0 & -1 & 1 & 1 \end{bmatrix} \tag{15}$$

The matrix $\mathbf{A}$ has a special structure. If we remove the last column, the remaining matrix is a network matrix.

Let $B_1, B_2, \ldots, B_n$ be consecutive blocks of two rows each, i.e., block B_k contains the rows $2k-1$ and $2k$. For any collection Q of rows of the matrix $\mathbf{A}$, we show how to partition it into two parts J_1 and J_2 so that the sum of the rows in J_1 minus sum of the rows in J_2 is a vector with entries 0, +1 and -1 only. Suppose Q contains both the rows of some block B_i, then, put both these rows in J_1. The remaining rows (say m) in Q then come from different blocks, call them $R_{j_1}, R_{j_2}, \ldots, R_{j_m}$. These m rows are partitioned as follows:

Let Q_+ be the subset of these m rows where the last element is +1 and Q_- be those rows where the last element is -1. Then, put $\lceil \frac{|Q_+|}{2} \rceil$ rows of Q_+ in J_1 and the remaining rows in J_2. Similarly, put $\lceil \frac{|Q_-|}{2} \rceil$ rows of Q_- in J_1 and the remaining rows in J_2. Since the sum of two rows in the same block is all zeroes, therefore, all such blocks in J_1 do not affect the sum of all the rows in J_1. Let T denote the vector resulting from the sum of the rows in J_1 minus the sum of the rows in J_2. All the elements except the last one in T is exactly 0, +1 or -1 because of the structure of the matrix A. The contribution of the rows from Q_+ to the last element of T is either 0 or +1. Similarly, the contribution of the rows from Q_- to the last element of T is either -1 or 0. This implies that the last element of T which is the sum of these two contributions can either be +1, -1 or 0.

This shows that the matrix A is totally unimodular. Using Fact 1, we conclude that $P_{\text{Reversal}}(f, f', j)$ is integral. □

References

[1] Air Transport Association of America Inc., 2008. *http://www.airlines.org/economics/cost+of+delays*

[2] Ball M., Chen C-Y., Hoffman R., and Vossen T. Colloborative decision making in air traffic management: Current and future research directions. In Proceedings of ATM'99, Workshop on Advanced Technologies and Their Impact on Air Traffic Management in the 21st Century. 1999.

[3] Barnhart C., Bertsimas D., Caramanis C., Fearing D. (2009). Equitable and Efficient Coordination of Traffic Flow Management Programs. Submitted to Transportation Science.

[4] Bertsimas D., Odoni A.R. (1997) A critical survey of optimization models for tactical and strategic aspects of air traffic flow management. Technical report, NASA.

[5] Bertsimas D., Stock S. (1998) The Air Traffic Management Problem with Enroute Capacities. Operations Research **46**, 406-422.

[6] Bertsimas D., Stock S. (2000). The Traffic Flow Management Rerouting Problem in Air Traffic Control: A Dynamic Network Flow Approach, Transportation Science **34**, 239-255.

[7] Bertsimas D., Lulli G., Odoni A.R. (2008). The Air Traffic Flow Management Problem: An Integer Optimization Approach. IPCO 2008.

[8] Bertsimas D., Weismantel R. (2005). Optimization Over Integers. Dynamic Ideas.

[9] Helme M. (1992). Reducing air traffic delay in a space-time network, IEEE International Conference on Systems, Man and Cybernetics **1**, 236-242.

[10] Hoffman R., Mukherjee A. and Vossen T. (2007). Air Traffic Flow Management. Working Paper.

[11] Lindsay K., Boyd E., Burlingame R. (1993). Traffic flow management modeling with the time assignment model. Air Traffic Control Quarterly **1**, (3), 255-276.

[12] Odoni A.R. (1987). The Flow Management Problem in Air Traffic Control, in *Flow Control of Congested Networks*, A.R. Odoni, L. Bianco and G. Szego eds., 269-288, Springer-Verlag, Berlin.

[13] Richetta O. and Odoni A.R. (1993). Solving Optimally the Static Ground-Holding Policy Problem in Air Traffic Control. Transportation Science **27**, 228-238.

[14] Richetta O. and Odoni A.R. (1994). Dynamic Solution to the Ground-Holding Policy Problem in Air Traffic Control. Trans. Res. **28a**, 167-185.

[15] Terrab M. and Odoni A.R. (1993). Strategic Flow Control on an Air Traffic Network. Operations Research **41**, 138-152.

[16] Terrab M. and Paulose S. (1993). Dynamic Strategic and Tactical Air Traffic Flow Control. RPI Technical Report.

[17] U.S. Dept. of Transportation, Bureau of Transportation Statistics (2008). Airline On-Time Statistics. *http://www.transtats.bts.gov/HomeDrillChart.asp*

[18] Vossen T., Ball M. Slot Trading Opportunities in Collaborative Ground Delay Programs. Transportation Science. 2006.

[19] Vossen T., Ball M. Optimization and Mediated Bartering Models for Ground Delay Programs. Naval Research Logistics. 2006.

[20] Vranas P.B., Bertsimas D., Odoni A.R. (1994). The Multi-Airport Ground Holding Problem in Air Traffic Control. Operations Research, **42**, 249–261.

[21] Wambsganss M. Collaborative decision making through dynamic information transfer. Air Traffic Control Quarterly, **4**:107-123, 1996.