A Tractable Optimization Framework for Air Traffic Flow Management Addressing Fairness, Collaboration and Stochasticity

by

Shubham Gupta

B.Tech., Indian Institute of Technology, Kanpur (2007)
S.M., Massachusetts Institute of Technology (2010)

Submitted to the Sloan School of Management
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy in Operations Research

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 2012

Author .....
Sloan School of Management
May 18, 2012

Certified by .....
Dimitris J. Bertsimas
Boeing Professor of Operations Research
Thesis Supervisor

Accepted by .....
Patrick Jaillet
Dugald C. Jackson Professor of Electrical Engineering and Computer Science, Co-Director, Operations Research Center

A Tractable Optimization Framework for Air Traffic Flow Management Addressing Fairness, Collaboration and Stochasticity

by

Shubham Gupta

Submitted to the Sloan School of Management
on May 18, 2012, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy in Operations Research

Abstract

We propose a tractable optimization framework for network Air Traffic Flow Management (ATFM) with an eye towards the future. The thesis addresses two issues in ATFM research: a) fairness and collaboration amongst airlines; and b) uncertainty inherent in capacity forecasts. A unifying attraction of the overall dissertation is that the Collaborative Decision-Making (CDM) paradigm, which is the current philosophy governing the design of new ATFM initiatives, is treated as the starting point in the research agenda.

In the first part of the thesis, we develop an optimization framework to extend the CDM paradigm from a single-airport to a network setting by incorporating both fairness and airline collaboration. We introduce different notions of fairness emanating from a) First-Scheduled First-Served (FSFS) fairness; and b) Proportional fairness. We propose exact discrete optimization models to incorporate them. The first fairness paradigm which entails controlling number of reversals and total amount of overtaking is especially appealing in the ATFM context as it is a natural extension of Ration-By-Schedule (RBS). We allow for further airline collaboration by proposing discrete optimization models for slot reallocation. We provide empirical results of the proposed optimization models on national-scale, real world datasets that show interesting tradeoffs between fairness and efficiency. In particular, schedules close to the RBS policy (with single digit reversals) are possible for a less than 10% increase in delay costs. We utilize case studies to highlight the considerable improvements in the internal objective functions of the airlines as a result of slot exchanges. Finally, the proposed models are computationally tractable (running times of less than 30 minutes).

In the second part, we address the important issue of capacity uncertainty by presenting the first application of robust and adaptive optimization in the ATFM problem. We introduce a weather-front based approach to model the uncertainty inherent in airspace capacity estimates resulting from the impact of a small number of weather fronts. We prove the equivalence of the robust problem to a modified instance of the deterministic problem; solve the LP relaxation of the adaptive problem using

affine policies; and report extensive empirical results to study the inherent tradeoffs.

Thesis Supervisor: Dimitris J. Bertsimas
Title: Boeing Professor of Operations Research

Acknowledgments

Five years ago, I embarked upon a journey with loads of excitement and plenty of apprehension. By the grace of god, I have been able to successfully complete my endeavor of earning a PhD!

I met my advisor, Professor Dimitris Bertsimas during my first week at MIT and we have worked together ever since! I express my sincere gratitude for his support and wisdom throughout my doctoral studies. I especially thank him for keeping up with me during phases when I was undergoing "turbulence". His voracious appetite for challenging problems and energy for creative endeavors is simply amazing. This thesis is a testimony of my cordial and fruitful relationship with him over these multiple years!

I thank Prof. Amedeo Odoni who made me gel into the MIT culture early on and has been a wonderful mentor ever since. I thank the other two members of my PhD committee - Prof. Georgia Perakis with whom I've interacted often right from the first optimization class to the seminar coordination in the final year, and Prof. Hamsa Balakrishnan for useful suggestions on my research. It has been a privilege to learn and interact with the amazing faculty members at MIT - I'd like to especially acknowledge Professors John Tsitsiklis and Vivek Farias for some of the most enjoyable classes. I thank Guglielmo Lulli for an enjoyable and productive collaboration in my final year! I thank Juliane Dunkel for giving me an opportunity to be a TA for one of my favorite subjects.

I owe a lot to all my friends who shared the ebbs and flows of this journey with me. I express my sincere thanks to Shashi Mittal for the help and guidance (over many years now)! He was the reason why I applied to ORC and took the "PhD leap". A special thanks to Chaithanya Bandi for the uncountable squash sessions and dinners which kept me sane. Not withstanding the "Shubandi" debacle, there were plenty other memorable events. Our friendship over the last four years has been one of the non-academic highlights. I thank Mike Frankovich for the wonderful conversations over many "Indian buffets" and over multiple projects in office. I thank the fellow students of my incoming class: Allison, Claudio, Cong, Mallory, Matthew, Nikos, Phil, Shree, Ton and Wei for the warm camaraderie. Finally, I thank Dave Goldberg for the company during numerous all-nighters in the "Student Center" and Dan Iancu from whom I've learnt a lot.

On the family front, I would like to thank my parents for always providing me with the best of opportunities. They have never enforced their desires on what to do

in life and have fully stood by me when times were tough. I dedicate this thesis and all my successes to both of you. I thank my sister Swita and brother-in-law Prateek for their unconditional love and support. Finally, a special mention for the two jewels in my life: Shreya and Shaurya, who have given me plenty of joy!

To my parents

Ad majorem Dei gloriam

Contents

1 Introduction **19**

1.1 The Case for Optimization 20
1.2 Overview of Air Traffic Management 24
1.3 Collaborative Decision-Making 28
1.4 Taxonomy of Existing Models 32
1.5 Starting Point: Bertsimas Stock-Patterson Model 35
1.6 Contributions and Thesis Outline 38

2 Network Models that Incorporate Concepts of Fairness **43**

2.1 Introduction 43
2.2 Controlling the Total Amount of Overtaking 49
2.3 Controlling the Total Number of Reversals 54
2.4 Controlling the Difference between Per Flight Airline Delays 57
2.5 Controlling both Reversals and Difference in Per Flight Airline Delays 58
2.6 Prioritizing between Airport Reversals and Sector Reversals 61
2.7 Computational Results 64
2.8 Conclusions 76

3 Network Models that Incorporate Airline Collaboration **79**

3.1 Motivation 79
3.2 Input Data 83
3.3 Generalization of Vossen-Ball Single-Airport Model 87
3.4 A Model based on Monotone Variables 89
3.5 Computational Results 91
3.6 Integration and Comparison with Current CDM Practice 95
3.7 Conclusions 96

4 Addressing Capacity Uncertainty **97**
4.1 Motivation 97
4.2 Our Proposal 100
4.3 Model of Weather-front Induced Capacity Uncertainty 101
4.4 Characterization of Weather-front Induced Uncertainty Set 108
4.5 Solution Methodologies 115
4.6 Computational Results 118
4.7 Conclusions 125

5 Conclusion **127**
5.1 Thesis Summary 127
5.2 Directions for Future Research 128

A Fairness Models **131**
A.1 Strength of TFMP-Reversal 131

B Airline Collaboration **135**
B.1 Strength of Slot Reallocation Models 135
B.2 Integer Equivalence of Slot Trading Models 137

C Robust and Adaptive Optimization **141**
C.1 Optimization Under Discrete Uncertainty Sets 141
C.2 Power of Affine Policies in Multistage Adaptive Optimization 142
C.3 Number of Extreme Points 146

List of Figures

1-1 Airline profits during 2001-2010 (Data Source: IATA). 20
1-2 OPSNET yearly delays for 2001-2010. 21
1-3 OPSNET monthwise delays for 2010. 22
1-4 Delays attributable to weather during 2001-2010. 22
1-5 Components of ATM. 24
1-6 Objectives of ANSP and Airlines. 26
1-7 Snapshot of flights in the NAS (Source: [20]). 27
1-8 Optimization Tools available to FAA. 27
1-9 Schematic of a GDP under CDM. 31
1-10 An example demonstrating the current operational details of a GDP. 31
1-11 Three core topics of the thesis. 39

2-1 Mapping of our proposal with the current three-stage CDM practice. This chapter covers Stage I of the overall proposal whereas Chapter 3 covers Stage II. 45
2-2 Pictorial depiction of a reversal. 46
2-3 A reversible pair of flights $(f, f') \in \mathcal{R}^A$ $(\text{dest}_f = k)$. 47
2-4 Example demonstrating the utility of controlling reversals over the RBS solution. * denotes slots which remain unutilized despite being available. The capacity of Airport 1 gets reduced by two-thirds, whereas that of Airport 2 by one-third. 49
2-5 Convex hull of the integer points in Table 2.2 to model overtaking $(i = t - \underline{\mathrm{T}}_j^{f'})$. 53
2-6 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). 62

2-7 Illustration of a scenario where controlling sector reversals seems an overkill. There are multiple ATFM programs operating simultaneously. Specifically, there are GDPs operating at two nearby airports (BOS) and (LGA). Moreover, there is no ATFM program operating en-route. 63
2-8 Two scenarios to study the relevance of controlling sector reversals in addition to airport reversals. 63
2-9 Distribution of per flight airline delays from (TFMP) in units of 15 minutes. 66
2-10 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. 69
2-11 Distribution of per flight airline delays from (TFMP-Reversal) in units of 15 minutes. 69
2-12 Distribution of per flight airline delays from (TFMP-Dev) in units of 15 minutes. 71
2-13 (TFMP-Rev-Dev): Tradeoff between reversals and difference in airline delays. Note that the horizontal axis corresponds to $\sum_{w \in \mathcal{W}} |d_w - \gamma|$ (in units of 15 minutes). Specifically, the value 0.2 corresponds to 3 minutes. For each day, the five points correspond to (λ_a^r, λ_d) set to (0, 100), (10, 100), (100, 100), (100, 10) and (100, 0) respectively. 72
2-14 Distribution of per flight airline delays from (TFMP-Rev-Dev) in units of 15 minutes. 74
2-15 Impact of controlling sector reversals on i) airport reversals and total delay cost (Left); and ii) sector reversals (Right). 75
2-16 Impact of super-linear cost coefficients on i) reversals and total delay cost (Left); and ii) distribution of flight delays (Right). 76

3-1 Illustration of a Scenario Motivating Airline Collaboration. 81
3-2 Illustration of centralized and decentralized mechanisms for airline collaboration. 81
3-3 Motivation for slot trading in a single-airport setting. 83
3-4 Motivation for slot trading in a multi-airport (network) setting. 83
3-5 Illustration of the structure of an offer $(f_d, t_{d'}; f_u, t_{u'})$. $\text{dest}_{f_d} = A$, $\text{dest}_{f_u} = B$, $t_{d'} = D_{f_d} + 3$ and $t_{u'} = D_{f_u} - 4$. 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. 84

3-6 Example demonstrating the utility of our slot reallocation phase over the intra-airline substitution. 85
3-7 Example Illustration of AMAL offers. Using the terminology of set $\mathcal{O}$, the proposed trade offers are (A1, 3; A2, 4), (B1, 6; B2, 2) and (C1, 4; C2, 1). 85
3-8 Representation of trade offers on a directed graph. The set of nodes correspond to the six slots, whereas the arcs represent various possible flight plans. For example, arc (A1, 2) denotes the assignment of slot 2 to flight A1 (from the currently allotted slot 1). 86
3-9 One of the final assignments after the execution of a feasible combination of trades. 86
3-10 Illustration of network flow constraints in the single-airport model. (Taken from [36]) 87
3-11 Distribution of number of executed trades across airlines from TFMP-Trading-BG. Left: (*Objective 1*); Right: (*Objective 2*) 94

4-1 OPSNET monthwise delays for 2010. 98
4-2 Delays attributable to weather during 2001-2010. 98
4-3 Illustration of our overall proposal for solving multi-period ATFM problem addressing capacity uncertainty. 101
4-4 Depiction of the capacity profile under a *weather-front* based uncertainty set for a single affected airspace element. The plot is for a particular realization of the parameters (T_a, d, α). 103
4-5 Illustration of the applicability of step functions to model capacity profiles. **AC** denotes actual sector capacity and **SF** denotes the step function capacity. 103
4-6 Traversal of a *weather-front* across the NAS. It has three phases over the course of its existence. 104
4-7 Explanation of some of the notation ($\mathcal{T} = 5$). 105
4-8 Illustration of the characteristics of the deterministic, robust and adaptive routes of a flight. 108
4-9 Illustration of the geometry of the *randomized rounding* algorithm for proving the integrality of polyhedron $\mathcal{X}$. $\mathbf{y}^*, \mathbf{z}^*$ satisfy $y_t^* \leq y_{t+1}^*$, $z_t^* \leq z_{t+1}^*$ and $z_t^* \leq y_t^*$. 112
4-10 **Left:** A simplex uncertainty set in $\mathbb{R}^2$. **Right:** A polyhedral set in $\mathbb{R}^2$ broken into three simplices. 118

4-11 Flight traffic (departures and arrivals) as a) a function of time of day (Left); and b) a function of airport (Right). 119
4-12 *Characteristics of Robust Solutions.* **Left:** Price of robustness as a function of capacity reduction. **Right:** Schedule deviation as a function of capacity reduction. The red line corresponds to the best linear fit. 122
4-13 *Utility of Robust Solutions.* The plot depicts the relation between robust and new deterministic cost (for the protracted feasible schedule) for different scenarios of the uncertainty set. 123
4-14 *Running times for* i) deterministic and robust problems (**Left**); and ii) adaptive problem (**Right**). C denotes CPLEX solver time and T denotes Total time (including ROME parsing time). 125

C-1 Example uncertainty sets with and without $\mathbf{b}_{\min}$ (black filled circle denotes $\mathbf{b}_{\min}$). 147
C-2 **Left**: Upper bound on the number of extreme points in the weather-front based uncertainty set ($k = 1$). The bold numbers indicate the attractive cases from the point of view of obtaining a single affine policy; and **Right**: Plot of the upper bound on $\mathcal{E}$ ($k = 1, P = 2$). . . 149

List of Tables

1.1 Direct and indirect delay costs to economy. 21

1.2 Flight On-Time Statistics (Data Source: Bureau of Transportation Statistics [35]). 22

1.3 Taxonomy of ATFM Models. 32

2.1 Utility of our proposal of enforcing FSFS fairness. 50

2.2 Truth table for modeling the overtaking variables. 51

2.3 Upper bound on the size of the models. 60

2.4 Numerical Example: Upper bound on the size of the models. 60

2.5 Summary of the datasets. $\mathcal{CF}$ denotes the number of connecting flights. 64

2.6 Performance of TFMP. $\mathcal{RV}$ denotes the number of reversals and $\mathcal{OV}$ the amount of overtaking. 66

2.7 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. 68

2.8 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. 70

2.9 Computational Performance of (TFMP-Dev). 71

2.10 Computational performance of (TFMP-Rev-Dev) in units of 15 minutes. $\mathcal{RV}$ denotes the number of reversals. 72

2.11 Balancing Sector Reversals with Airport Reversals ($\mathcal{SR}$ denotes the number of sector reversals, $\mathcal{AR}$ denotes the number of airport reversals and $\mathcal{DC}$ denotes the delay cost). Fairness imposed in 5 sectors of the north-east region and 10 airports spatially close to these sectors. . . . 75

3.1 Relative benefits of a centralized and decentralized slot reallocation mechanism. 82
3.2 Left: Value of the slots to the two airlines; Right: Value proposition of various assignments. 82
3.3 Computational performance of the two trading models - Objective Function 1 (maximize the total number of executed trades). 92
3.4 Computational performance of the two trading models - Objective Function 2 (minimize the deviation in the number of trades executed for an airline from the mean). 92
3.5 Computational performance of TFMP-Trading-BG. 93
3.6 Comparison of TFMP-Trading-BG between single-airport and network-wide settings. SA denotes the results from SA-TRADING and NW denotes the results from TFMP-TRADING. 94

4.1 Computational Experience with Π_2. 121
4.2 Utility of Robust Solutions: Best fit lines. 123
4.3 Computational Experience with Π_6. 124

C.1 Number of extreme points for the weather-front based polytope. $\mathcal{AS}$ denotes airspace, $\mathcal{OP}$ denotes one-phase and $\mathcal{AP}$ denotes all-phases. $\bigoplus$ denotes polyhedron concatenation. 147

List of Abbreviations

AAR	Airport Acceptance Rate
AFP	Airspace Flow Program
AOC	Airline Operational Centers
ATC	Air Traffic Control
$ATFM$	Air Traffic Flow Management
BTS	Bureau of Transportation Statistics
CDM	Collaborative Decision-Making
CTD	Controlled Time of Departure
CTA	Controlled Time of Arrival
FAA	Federal Aviation Administration
FCA	Flow Controlled Area
FSM	Flight Schedule Monitor
GDP	Ground Delay Program
GHP	Ground-Holding Program
NAS	National Airspace System
$NASA$	National Aeronautics and Space Administration
$NEXTOR$	National Center of Excellence for Aviation Operations Research
$OPSNET$	Operations Network
RBS	Ration-By-Schedule

Chapter 1

Introduction

Each one of us has most likely experienced a canceled flight, missed a flight connection or arrived late in our destination. The agony caused due to missing an important personal function or not making a conference on time is inexplicable. Consequently, the smooth operation of the National Air Transportation System (NATS) is paramount and a critical component of societal welfare.

To put in perspective the uncertain operating conditions of airlines, consider the financial performance over the last 10 years. Figure 1-1 plots the airline profits for the last 10 years (period of 2001-2010). The oscillating pattern of losses and profits should be immediately apparent. The attacks of 9/11 meant that airlines suffered losses for five straight years after 2001. Subsequent resurgence in traffic coupled with stable fuel prices ensured profits during the next couple of years. But, the financial crisis of 2008 and very high fuel prices lead to one of the worst years with total losses exceeding $25 billion in 2008. Recently, airlines have posted a net profit of $10 billion in 2010 with a forecast of a positive year during 2011. Airlines, thus, typically operate in a very uncertain financial environment which is a consequence of factors such as fluctuating fuel prices, market competition, strongly stochastic airspace capacity (leading to delays) and a whole host of external events (like 9/11, financial crisis, etc.). The lack of sustained profits further exacerbates the quality of service provided by various airlines and leads to increased disruptions as airlines try to manage their operations in the most cost-effective manner.

Although, a lot of research effort has been undertaken in the last two decades, there is still a lack of a centralized optimization-based tractable framework which automatically reroutes planes (under reduced capacity), cancels flights (to maximize system efficiency) and dynamically adjusts schedules (in response to unexpected weather events). Our aspiration in this research effort is to propose models which bridge this

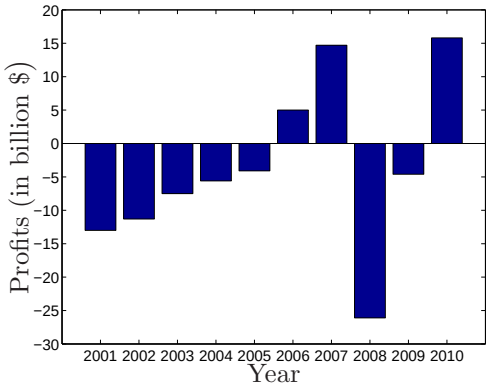


Figure 1-1: Airline profits during 2001-2010 (Data Source: IATA).

gap, at least from an Operations Research standpoint. The exact issues that this thesis addresses are elaborated later in Section 1.6. The implementation of these models still requires political will and consensus amongst various stakeholders to transition into practice. To drive home the point on the need for optimization models, we present concrete evidence on the cost of delays.

1.1 The Case for Optimization

The sustained growth of the aviation industry has put a tremendous strain on the available resources of the air transportation system. This is evidenced by the steady increase in flight delays and severe congestion at airports. In 2010, approximately 18% of the flights in the United States were delayed by more than 15 minutes, while another 2% were cancelled (Bureau of Transportation Statistics [35]). Moreover, during the 12-month period ending in September 2008, 138 million minutes of system delay led to an estimated $10 billion in costs for US airlines [28].

Figure 1-2 depicts the yearly trend on the commercial aviation delays over the last 10 years. There is a sharp dip in the delays in the aftermath of the terrorist attacks on Sept 11, 2001. But, after 2003, there has again been a steady increase in the delays until the economic recession in the year 2008 caused the traffic to plunge again. Nonetheless, the magnitude of delays in recent years (around 0.3 million minutes) highlights the widespread prevalence of delays.

To assess the impact of the cost of the delays to the economy (both direct and indirect), Table 1.1 reports the cost of delay numbers calculated by various recognized agencies. The reported numbers highlight the enormity of the economic impact of aviation delays which provide ample impetus for the development of optimization

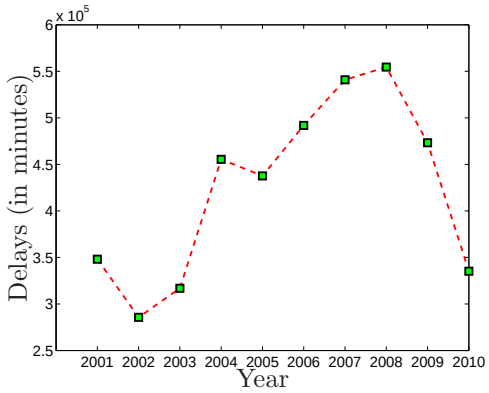


Figure 1-2: OPSNET yearly delays for 2001-2010.

tools.

Source	Year	Cost ($ billion)	
		Direct	Indirect
NEXTOR	2007	27	4
JEC	2007	31	10
IATA	2008	14	-

Table 1.1: Direct and indirect delay costs to economy.

Flight On-Time Statistics

To assess the performance of the current ATM system, Table 1.2 reports statistics on the percentage of flights that are delayed by more than 15 minutes and those that are cancelled altogether. As should be evident from these numbers, the percentage flights which don't arrive on-time consistently range between 18 and 25% which emphasizes that almost 1 in 4 flights don't arrive on time. Furthermore, typically 1-3% of all flights are cancelled altogether. This is significant as cancelled flights cause subsequent disruptions due to missed connections.

Impact of Weather

Weather accounts for the majority of the total air traffic delays caused due to terminal, en-route congestion and several other operational factors. To assess the impact

Year	Total Flight Operations	% Flights Delayed	% Cancelled
2006	2,685,218	21.66	1.50
2007	2,688,939	24.61	2.52
2008	2,644,820	24.11	2.26
2009	2,540,586	19.26	1.59
2010	2,517,616	18.45	2.08

Table 1.2: Flight On-Time Statistics (Data Source: Bureau of Transportation Statistics [35]).

of weather on the total aviation delays, we consider the OPSNET[1] delays data for the year 2010. As evidenced in the monthly delays plot in Figure 1-3, there is a significant spike in the delays for the summer months (May-July), when there is pronounced convective weather activity. Moreover, Figure 1-4 indicates that approximately 65-75% of total delays are attributable to weather in the last ten years. These two observations highlight the importance of addressing weather induced capacity uncertainty for mitigating aviation delays.

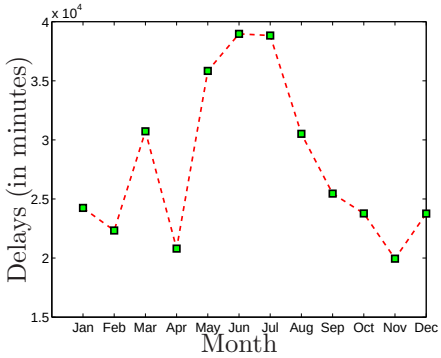


Figure 1-3: OPSNET monthwise delays for 2010.

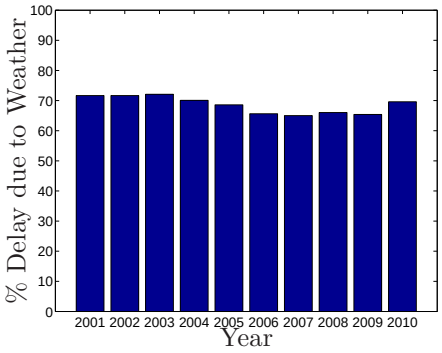


Figure 1-4: Delays attributable to weather during 2001-2010.

Challenges

The enormity of the aviation delays and the corresponding impact on the economy gives ample motivation for the development and deployment of optimization tools to mitigate them. In addition to the congestion problems being faced currently, there are

[1]The Operations Network (OPSNET) is one of the official sources of National Airspace (NAS) air traffic operations and delay data.

a number of other challenges which make the control of aviation delays an especially arduous task. We elaborate upon them in this section.

- **Current Airport Operations.** Many airports in the US are currently operating at or near capacity. There is typically a nonlinear relationship between delay and demand. This means that an incremental increase in demand causes a significant (viz., non-linear) increase in the delays when the operations are executed close to the system capacity. A consequence of this fact is that handling increasing demand needs more prudent use of the available resources in the NAS which manifests itself in the use of optimization tools.

- **Enhanced Airport Infrastructure.** A potential solution to mitigate the impacts of increased demand on airport congestion is to enhance the capacity by building new runways or increase the number of airports. Unfortunately, these tools suffer from various logistical complexities (of design and planning) and are further hindered by tedious legal approvals. Moreover, such projects typically take years (if not decades) to be implemented in totality, a timeframe not amenable for short-term congestion alleviation. Finally, such projects do not ease congestion in the en-route airspace. Therefore, it becomes paramount to seek alternative avenues for congestion management.

- **Future Demand Trends.** The FAA estimates that through 2025, demand will increase by more than 60% averaging 3.4% per year in the domestic carrier revenue passenger miles. This continued surge in demand seems daunting and provides added thrust to sophisticated ATFM initiatives of the form aspired in this thesis.

NextGen

Currently, in the United States, significant resources are being deployed in the development of a future Air Traffic Management (ATM) system called the Next Generation Air Transportation System (NGATS or NextGen) [29]. The expected key benefits of this new ATM system are improvements in various aspects of aviation encompassing ground and terminal operations, technological advancements in aircraft monitoring, efficient coordination of various ATM tools and improved tracking of actual flight paths. In particular, a major proposal of NGATS is improvements in 4DT (four-dimensional trajectory) uncertainty (4DT capability is defined as the ability to precisely fly an assigned 3D trajectory while meeting specified timing constraints on

arrival at waypoints [29]). This would lead to enhanced predictability and control over the trajectory of an aircraft with the resulting decrease in the delays in the system. These developments further motivate the design of optimization models as the NGATS technologies would give tighter control over the exact trajectories that aircraft fly and thereby lead to stricter adherence to optimized routes.

We now expand upon the tools currently in use by FAA for realizing the goal of safe and expeditious aircraft movements.

1.2 Overview of Air Traffic Management

A multitude of factors like inclement weather, operational outages and demand surges can collude together to cause severe disruptions to air traffic movements. This necessitates the use of tools that mitigate the potentially catastrophic impact on the system performance. *Air Traffic Management* (ATM) is a broad term used to refer to the composite of all such services. In terms of the exact objectives, ATM can be further classified into two components (as depicted in Figure 1-5):

- Air Traffic Control (ATC)
- Air Traffic Flow Management (ATFM)

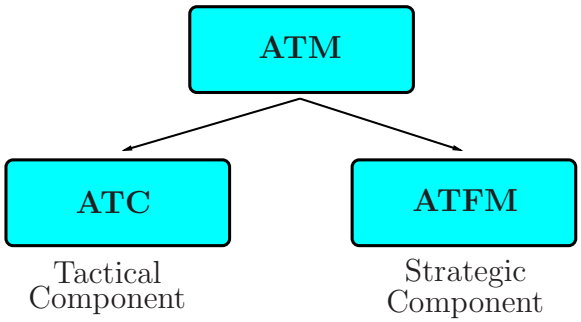


Figure 1-5: Components of ATM.

We discuss each of these in some detail now. *Air Traffic Control* (ATC) refers to the set of "tactical" processes that aim to ensure safe separation between individual flights. More precisely, these pertain to the separation services provided by human controllers looking after different sectors. Each of these controllers keeps track of the flights in their respective sectors and ensure that safe separation is maintained between all flights and that the traffic to the next sector is presented in an orderly manner. In contrast, *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 airborne delay near the affected destination airport which is both unsafe and more costly (in terms of fuel costs). Another recently introduced tool similar to a GDP is an Airspace-Flow Program (AFP) which is used to control arrival rate into a weather affected segment of the airspace, also known as a Flow Constrained Area (FCA). AFPs have been operational since 2007. Some of the other ATFM tools include assigning airborne delays, dynamic re-routing and speed control.

Stakeholders

There are essentially two sets of stakeholders in the air traffic domain:

1. Air Navigation Service Provider (ANSPs); and

2. Users (Airlines and general aviation)

The primary responsibility of ANSP is the smooth functioning of the ATM system. In the US, FAA is the primary ANSP, whereas EUROCONTROL assumes this responsibility for continental Europe. The users are comprised of the airlines, general aviation and military.

Figure 1-6 depicts the broad hierarchy of objectives that the ANSP and airlines try to achieve. *Safety* at every stage of operations forms the most critical objective and overrides every other goal. For the ANSPs, ensuring equitable workload amongst the available resources (like controllers) represents the next important objective. Finally, equitable distribution of resources amongst users is the least important aspect of the overall planning process. In contrast, for airlines, ensuring that irregular operations are managed appropriately and the total delays are minimized represent the next important goals.

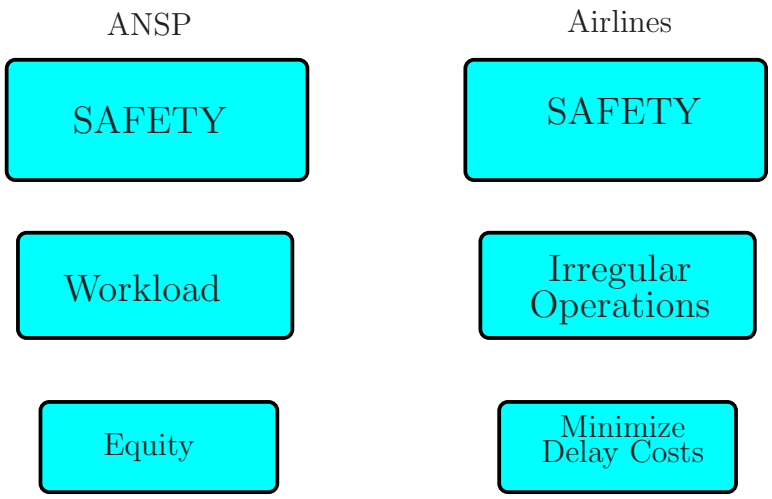


Figure 1-6: Objectives of ANSP and Airlines.

Organization of FAA's ATC Capabilities

Figure 1-7 shows an example snapshot of flights in the NAS at any point in time (there are around 5000 aircraft operating in the system during peak periods). The FAA utilizes a multi-layered hierarchy of entities to implement ATC services. At the apex of this pyramid is a state-of-the-art facility called Air Traffic Control System Command Center (ATCSCC) which is the centralized decision-making engine of the FAA. Operational since May 1994, this facility based near Washington DC houses latest technology and is the most sophisticated facility of its kind in the world. The primary responsibilities of ATCSCC include executing the strategic ATFM functions. This involves continuous monitoring of current and projected demand and simultaneously updating estimates of capacity limits (in the presence of bad weather and runway closures). As and when demand exceeds capacity, the ATCSCC decides on the strategies to be implemented to resolve congestion. At the second level of this hierarchy are entities that take a more localized view of traffic management. Thus, to have a more focused control on the traffic at a regional level, the FAA has 22 Air Route Traffic Control Centers (ARTCCs) encompassing the entire geographical landscape of continental US. Each of these entities is responsible for ensuring safe separation between aircraft that fall within its boundaries. The interaction amongst the various ARTCCs is coordinated by the ATCSCC. To have a more microscopic control, each ARTCC is further divided into a number of sectors which span a three-dimensional volume of the airspace. A small number of human air traffic controllers control the traffic within each of the sectors by communicating with the pilots. As an aircraft arrives close to the terminal area, the control is shifted from the human air traffic

controller of the sector to Terminal Radar Approach Control facilities (TRACONs). Finally, during the taxi in and taxi out within the airport runway area, the aircraft is controlled by airport towers. The third (and final) layer within this hierarchy is occupied by human controllers who ensure that the micro details of the overall schedule is implemented in their respective regions of control.

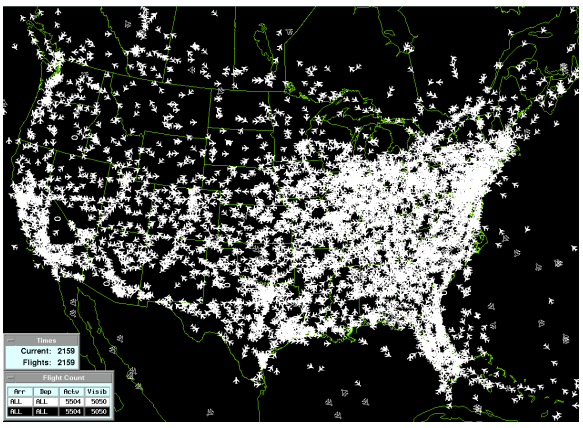


Figure 1-7: Snapshot of flights in the NAS (Source: [20]).

Optimization Tools

The FAA has a number of tools available at its disposal to achieve the goal of safe and expeditious aircraft movements. Some of the possible interventions are depicted in Figure 1-8.

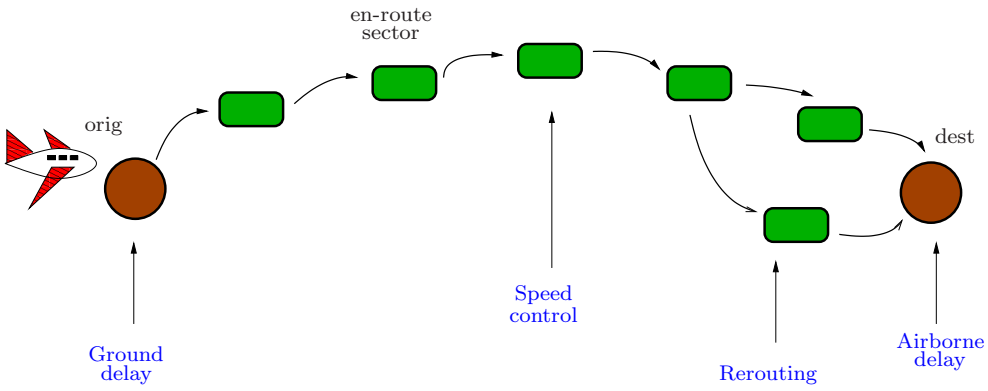


Figure 1-8: Optimization Tools available to FAA.

A brief description of each of these tools is as follows:

- **Ground Delay:** 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 airborne delay near the affected destination airport which is both unsafe and more costly (in terms of fuel costs).

- **Airborne Delay:** Delays are applied to flights near the affected destination airport by making them hover in the air because of landing capacity limitations at the destination airport. This option is less preferred to ground-delays because of the reasons outlined in the preceding bullet.

- **Speed Control:** Within a sector, aircraft speed can be controlled to meter its arrival into the subsequent sector to match the controller workload and capacity limitations in different sectors.

- **Rerouting:** FAA uses this tool to alter arrival rate into a Flow Constrained Area (FCA), e.g., a weather affected segment of the airspace by rerouting flights to different routes.

GDPs and AFPs are some of the most extensively used ATFM tools currently. As already mentioned, GDPs control arrival rate into a weather affected airport whereas AFPs control arrival rate into a weather affected segment of the airspace. Unfortunately, these tools don't take a network-wide view, thereby, not optimizing simultaneously disrupted operations at other affected airports due to delay propagation. In addition, rerouting is done based on the National Playback [4] in current practice (in the US). This database comprises of alternative routes that should be taken to avoid parts of the airspace routinely affected by severe weather. These are computed based on historical data. Although, it has been found that this playback-based rerouting process sometimes leads to local congestion in the regions through which traffic is rerouted. Thus, the current state of affairs beg for a more scientific mathematical based approach that identifies on a global basis promising reroutes (and other ATFM interventions), while remaining capacity-feasible throughout.

1.3 Collaborative Decision-Making

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. [6], Wambsganss [39]) 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. [23] for details on CDM). All recent efforts to improve ATFM have been guided by this philosophy. The overall objectives of CDM can be summarized as follows (reproduced verbatim as visualized by the designers, please refer to Ball et al. [6] for more details):

- generating better information, usually by merging flight data directly from the Airspace System with information generated by airspace users;
- creating common situational awareness by distributing the same information both to traffic managers and to airspace users; and
- creating tools and procedures that allow airspace users to respond directly to capacity/demand imbalances and to collaborate with traffic flow managers in the formulation of flow management actions.

In the US, "Ration-by-Schedule" (RBS) is the fundamental principle for GDPs and all the CDM initiatives. We now elaborate upon this important principle below:

Ration-by-Schedule

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. [6], Wambsganss [39] for details on rationing). In the case of GDP 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 during GDPs. The success of RBS in a single-airport GDP setting is a consequence of the following three salient features:

1. The practical implementation of this principle is trivial and has a linear running time with respect to the number of flight steps. Thus, the approach is attractive from a scalability standpoint.
2. For an isolated GDP or AFP, the RBS method always leads to a solution that minimizes the minutes of system delay [36].

3. This notion of fairness is endorsed by the primary stakeholders (i.e., the FAA and the airlines) and is the industry accepted paradigm.

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 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. Finally, Odoni and Lulli [26] discuss examples highlighting the intrinsic inequities in the network problem.

Operational Details of a GDP

Figure 1-9 depicts a schematic diagram of the various stages of executing a GDP under CDM practice. There are three key stages 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).

Figure 1-10 depicts an example illustrating the current operational details of a GDP. There are three airlines A, B and C operating seven flights between them. Airline A has 3 flights (A1, A2 and A3), Airline B has 2 flights (B1, B2) and Airline C has 2 flights (C1, C2) in the GDP. The original published ordering is shown in the leftmost table. The next table shows the output of applying the RBS principle to the original sequence. Assume that the capacity is reduced by half, thereby leading to an Airport

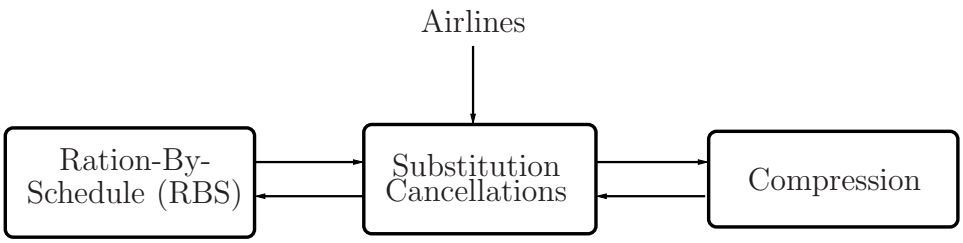


Figure 1-9: Schematic of a GDP under CDM.

Acceptance Rate (AAR) of one flight every two slots (compared to a flight every slot under nominal conditions). As a result, each flight is assigned a Controlled Time of Arrival (CTA) which is double the slot allotted in the original sequence. The RBS sequence is followed by substitution-cancellation phase wherein the airlines are given the flexibility of changing the order of the flights assigned to them as well as canceling flights as long as the resulting schedule remains capacity-feasible. In this example, Airline C substitutes flight C1 by C2 and Airline B cancels flight B1. The final step in this sequence is the application of the Compression procedure wherein flights are moved up to fill up the holes created by canceled flights (as long as the final allotment is compatible with the earliest arrival times of all flights).

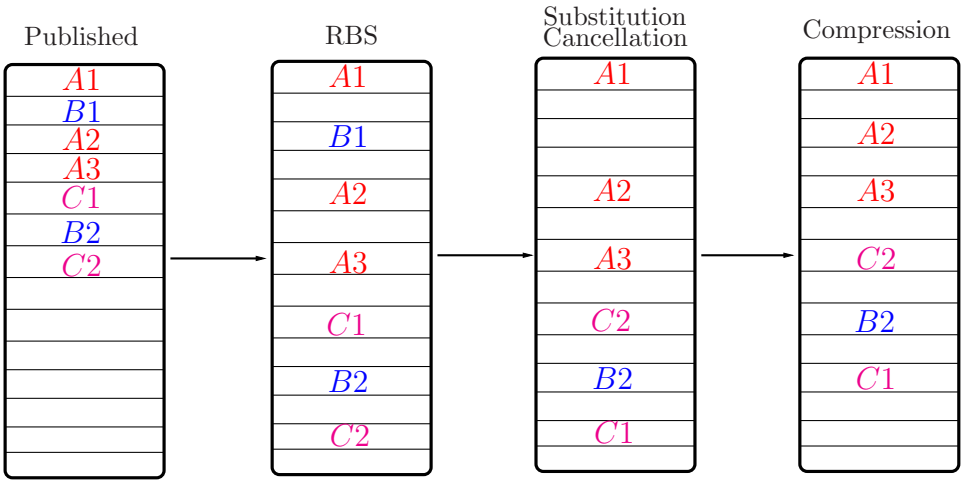


Figure 1-10: An example demonstrating the current operational details of a GDP.

1.4 Taxonomy of Existing Models

Table 1.3 summarizes a classification of the various ATFM models in terms of a broad set of characteristics that govern the exact setting of the problem being modeled. More precisely, factors like uncertainty, adaptability, connectivity, equity considerations and rerouting govern the clustering of various ATFM models. As should be evident from the table, the ATFM problem in full generality is as complex as an engineering problem can potentially get. The optimization paradigm needed is multi-stage dynamic and stochastic with the additional complexity of ensuring equity and operating within a collaborative setting. For a detailed survey of the various contributions and a taxonomy of all the problems, see Bertsimas and Odoni [12] and Hoffman et al. [23].

Characteristic	**Classification**	
Adaptability	Static	Dynamic
Connectivity	N/w of Capacitated Resources	Single Capacitated Resource
Control	Single Decision-Maker	Collaborative Decision-Making
Equity	Without Equity Considerations	With Equity Considerations
Rerouting	Without Rerouting	With Rerouting
Uncertainty	Deterministic	Stochastic

Table 1.3: Taxonomy of ATFM Models.

Odoni [27] 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. We classify the various models based on the characteristics listed in Table 1.3.

- **Connectivity:** Single Vs. Network.
  - **Single Capacitated Resource.** The problem of assigning ground-delays in the context of a single-airport (*Single-Airport Ground-Holding Problem*, SAGHP henceforth) has been studied in Terrab and Odoni [33], Richetta and Odoni [30], [31].
  - **Network of Capacitated Resources.** The problem of assigning ground-delays in the multiple airport setting (*Multi-Airport Ground-Holding Problem*) in Terrab and Paulose [34], Vranas et al. [38]. The problem of controlling release times and speed adjustments of aircraft while airborne for

a network of airports taking into account the capacitated airspace (*Air Traffic Flow Management Problem*) has been studied in Bertsimas and Stock-Patterson [13], Helme [22], Lindsay et al. [25].

- **Control:** Centralized Vs. Collaborative.
  - **Single Decision-Maker.** Most of the network models (e.g. Vranas et al. [38], Bertsimas and Stock-Patterson [13]) do not incorporate airlines into the decision-making.
  - **Collaborative Decision Making.** As part of the CDM philosophy, researchers have also explored dynamic interaction with airlines. Towards this aim, Vossen, Ball [36] [37] 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.
- **Equity.** Socially Optimum Vs. Equitable Distribution amongst Airlines.
  - **Without.** Most of the network models (e.g. Vranas et al. [38], Bertsimas and Stock-Patterson [13]) do not address fairness amongst airlines.
  - **With.** Barnhart et al. [8] develop a way to address fairness in the context of network 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. Please refer to the PhD thesis of Fearing [21] for extensive details.
- **Rerouting.** Fixed Trajectories Vs. Multiple Routes.
  - **Without.** Most of the network models (e.g. Vranas et al. [38], Bertsimas and Stock-Patterson [13]) do not incorporate rerouting.
  - **With.** The problem with the added complication of dynamically rerouting aircrafts (*Air Traffic Flow Management Rerouting Problem*) was first studied by Bertsimas and Stock-Patterson [14]. Recently, Bertsimas et al. [10] have presented a new mathematical model for the ATFM problem with dynamic re-routing which has superior computational performance.
- **Uncertainty.** Deterministic Vs. Stochastic.

– **Deterministic.** Most of the network models (e.g. Vranas et al. [38], Bertsimas and Stock-Patterson [13]) do not address uncertainty inherent in the capacity forecasts.

– **Stochastic.** One of the first attempts at dealing with Stochastic SAGHP was by Richetta and Odoni [30], [31]. Subsequently, Ball et al. [7] proposed another model for the same problem. Recently, Mukherjee and Hansen [3] study the SAGHP in a dynamic stochastic setting. Kotnyek and Richetta [5] present equitable models for the stochastic SAGHP and prove the equivalence of integrality and equity in the model presented in [31].

1.5 Starting Point: Bertsimas Stock-Patterson Model

In this section, we reproduce the Bertsimas Stock-Patterson model [13] for the ATFM problem which provides the starting point for all the models presented in this thesis.

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,

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

$\mathcal{S}^f \subseteq \mathcal{S}$: set of sectors that can be flown by flight f,

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

P_i^f : preceding sector of sector i in flight f's path,

L_i^f : subsequent sector of sector i in flight f's path,

$D_k(t)$: 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 : scheduled arrival time of flight f,

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,

T_j^f : set of feasible time periods for flight f to arrive in sector 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.

The Objective Function

The objective function minimizes the total delay costs (which is a combination of the costs of airborne delay (AH) and ground-holding delay (GH)). We use an adapted expression introduced recently in Bertsimas et al. [10].

The total delay (TD) cost is a combination of the costs of airborne delay (AD) and ground-holding delay (GD) ($TD = GD + \alpha \cdot AD$, where $\alpha > 1$ because airborne 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:

$$c^f_{\text{total}}(t) = \alpha(t - a_f^k)^{1+\epsilon} : \text{total cost of delaying flight } f \text{ for } (t - a_f^k) \text{ units of time,}$$

$$c^f_g(t) = (\alpha - 1)(t - d_f)^{1+\epsilon} : \text{cost reduction obtained by holding flight } f \text{ on the ground for } (t - d_f) \text{ units of time,}$$

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 1.5.1. 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 units 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 $1^{1.001} + 1^{1.001} < 2^{1.001} + 0$.

The TFMP model

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

$$IZ_{\text{TFMP}} = \min \sum_{f \in \mathcal{F}} \Big(\sum_{t \in T^f_{\text{dest}_f}} c^f_{\text{total}}(t) \cdot (w^f_{\text{dest}_f,t} - w^f_{\text{dest}_f,t-1}) - \sum_{t \in T^f_{\text{orig}_f}} c^f_g(t) \cdot (w^f_{\text{orig}_f,t} - w^f_{\text{orig}_f,t-1}) \Big)$$

subject to:

$$\sum_{f \in \mathcal{F}:\text{orig}_f = k} (w^f_{k,t} - w^f_{k,t-1}) \leq D_k(t), \qquad \forall k \in \mathcal{K},\ t \in \mathcal{T}. \tag{1.1a}$$

$$\sum_{f \in \mathcal{F}:\text{dest}_f = k} (w^f_{k,t} - w^f_{k,t-1}) \leq A_k(t), \qquad \forall k \in \mathcal{K},\ t \in \mathcal{T}. \tag{1.1b}$$

$$\sum_{f \in \mathcal{F}: j \in \mathcal{S}_f, j' = L^f_j} (w^f_{j,t} - w^f_{j',t}) \leq S_j(t), \qquad \forall j \in \mathcal{S},\ t \in \mathcal{T}. \tag{1.1c}$$

$$w^f_{j,t} - w^f_{j',t-l_{fj'}} \leq 0, \qquad \forall f \in \mathcal{F},\ t \in T^f_j,\ j \in \mathcal{S}^f\ : j \neq \text{orig}_f, j' = P^f_j. \tag{1.1d}$$

$$w^f_{\text{orig}_f,t} - w^{f'}_{\text{dest}_{f'},t-s_f} \leq 0, \qquad \forall (f,f') \in \mathcal{C},\ \forall t \in T^f_k. \tag{1.1e}$$

$$w^f_{j,t-1} - w^f_{j,t} \leq 0, \qquad \forall f \in \mathcal{F},\ j \in \mathcal{S}^f,\ t \in T^f_j. \tag{1.1f}$$

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

The first three sets of constraints take into account the capacities of the various elements of the system. Constraints (1.1a) 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 (1.1b) 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 (1.1c) 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 (1.1d) 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 (1.1e) 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 (1.1f) 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 ($t \geq \tilde{t}$).

In the remainder, we shall refer to the model just introduced as TFMP.

Remark 1. **Aircraft and Passenger 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 banks of flights arriving at a hub and then departing within a short duration of time. The presence of spare aircraft too inhibit the enforcement of strict aircraft connectivities. 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 thesis, 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. Finally, the model does not capture Minimum Connection Times (MCT) to permit pax transfers.

1.6 Contributions and Thesis Outline

Very broadly, the thesis comprises of three core topics as illustrated in Figure 1-11. The first topic of fairness (or equity) is of central importance if any optimization-based proposal has to be deployed online. This is a consequence of the competitive marketplace and mistrust between the airlines and the FAA. The second topic of airline collaboration is more relevant now than ever before because of the acceptance of CDM philosophy. Finally, addressing capacity uncertainty is the third important topic. As already emphasized that convective weather and thunderstorms account for majority (around 70%) of the total delays. Thus, optimization models that consider static deterministic capacity are of limited practical use.

Our overall aspiration therefore is to use these three pillars to propose a *tractable*

and *optimization-based* framework which will help FAA in identifying promising ATFM interventions *dynamically* while operating in a *stochastic* environment. These three topics are organized into two main parts of the thesis. A brief description of the two parts is as follows:

- **Part 1: Extending CDM to an airspace setting.** In the first part, we extend the Collaborative Decision-Making (CDM) paradigm from a single-airport setting to an airspace context. More precisely, we propose an optimization framework in a network setting which incorporates fairness and airline collaboration (covered in Chapters 2 and 3).
- **Part 2: Addressing capacity uncertainty.** In the second part, we address the important issue of capacity uncertainty lacking in deterministic models. Towards this goal, we present the first application of robust and adaptive optimization in the Air Traffic Flow Management (ATFM) problem (covered in Chapter 4).

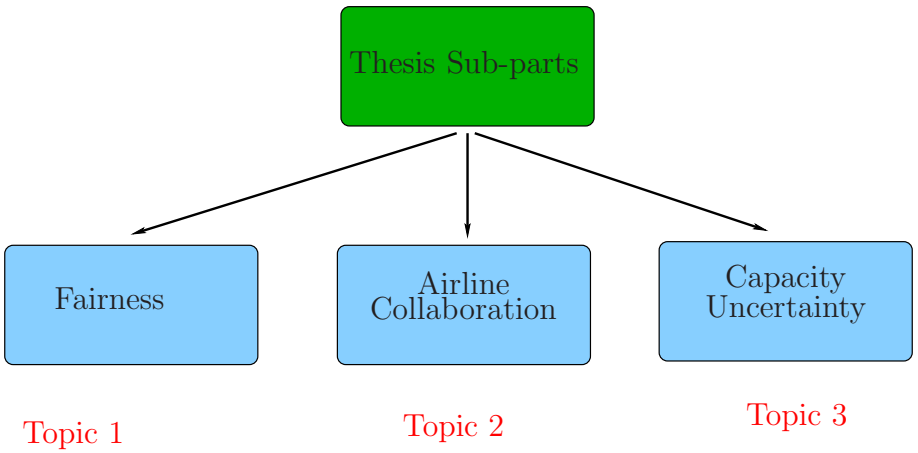


Figure 1-11: Three core topics of the thesis.

A brief summary of all the chapters is as follows:

- **Chapter 2. Fairness.** In this chapter, we present network models that incorporate different notions of fairness, namely i) FSFS fairness - controlling number of reversals and total amount of overtaking; ii) Proportional fairness - equalizing airline delays; and iii) a combination of the FSFS and Proportional fairness paradigms. 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. The important takeaways

are the possibility of generating schedules close to the RBS policy (at a less than 10% increase in delay costs) and the price of fairness for proportional fairness is negligible. We report promising computational times of less than 30 minutes on large-scale instances which are encouraging for real-time deployment.

- **Chapter 3. Airline Collaboration.** In this chapter, we allow for further airline collaboration by proposing network models for slot reallocation. This is a generalization of the intra-airline substitution phase of the current CDM practice to inter-airline reallocation across multiple airports. An attractive feature of this stage of our proposal is the airline input in the form of AMAL ("*at-most, at-least*") trade offers which enable multiple trade combinations to be possible without needing sophisticated data input. We develop two models: i) a generalization of the Vossen-Ball slot trading model to a network setting; and ii) a model based on the monotone variables used in the Bertsimas Stock-Patterson model. We study the polyhedral structure of the two models and design case studies to demonstrate the potential benefits to internal objective functions of airlines. Both optimization models solve to optimality in seconds.

- **Chapter 4. Capacity Uncertainty.** In this chapter, we address the issue of capacity uncertainty. We introduce a weather-front based approach to model the uncertainty inherent in airspace capacity estimates resulting from the impact of a small number of weather fronts moving across the National Airspace (NAS). The key advantage of our uncertainty set construction is its low-dimensionality (uncertainty in only two parameters govern the overall uncertainty set for each airspace element). We formulate the resulting ATFM problem under capacity uncertainty within the robust and adaptive optimization framework. We prove the equivalence of the robust problem to a modified instance of the deterministic problem and solve optimally the LP relaxation of the adaptive problem using affine policies. Finally, we report empirical results from the proposed models on real-world flight schedules augmented with simulated weather fronts that illuminate the merits of our proposal. The key takeaways are: a) the robust problem inherits all the attractive properties of the deterministic problem (e.g., strong integrality properties and fast computational times); b) the price of robustness is typically small; and c) adaptability leads to useful benefits.

Notation and Preliminaries

Throughout the thesis, we denote scalar quantities by lowercase, non-bold face symbols (e.g., $w \in \mathbb{R},\ k \in \mathbb{N}$), vector quantities by lowercase, boldface symbols (e.g., $\mathbf{w} \in \mathbb{R}^n,\ n > 1$), and matrices by uppercase, boldface symbols (e.g., $\mathbf{A} \in \mathbb{R}^{n \times n},\ n > 1$). We use prime ($'$) to denote transpose (e.g., $\mathbf{x}'\mathbf{y}$). We denote by $\mathbf{e}$ the unit vector $(1, \ldots, 1)$ comprising of all ones. The dimension is implicit in the context (e.g., $\mathbf{e}'\mathbf{w}$ where $\mathbf{w} \in \mathbb{R}^3 \implies \mathbf{e} \in \mathbb{R}^3$).

Chapter 2

Network Models that Incorporate Concepts of Fairness

The optimization model proposed by Bertsimas Stock-Patterson [13] (introduced in Chapter 1) took a viewpoint of a Centralized Decision-Maker (FAA) without addressing the preferences of individual airlines. Such a perspective does not take into account the possible disparity in the distribution of delays across airlines. This is an important consideration given the competitive marketplace and the mistrust between the airlines and FAA. Consequently, in this chapter, we augment the optimization model to address this critical issue of fairness among airlines. We present empirical and theoretical evidence that the models developed in this chapter provide high quality solutions on national-scale datasets in reasonable computational times.

2.1 Introduction

As mentioned in Chapter 1, the first goal in this thesis is to propose an optimization based approach that:

1. incorporates network effects and builds upon the ATFM literature; and
2. 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 ([13]) 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. Moreover, across these reversals, there might be different number of time-periods of overtaking. Hence, 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. Thus, this leads to the application of First-Scheduled, First-Served (FSFS) fairness paradigm which enables control on the number of reversals and amount of overtaking. Furthermore, the resulting allocations might have disparity in the distribution of delays across airlines. This motivates the application of Proportional fairness. We propose integer programming 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 Compression, 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 allow airlines to react to the schedule determined in Stage I by taking into account their flights in the entire network and making appropriate tradeoffs.

Given that each of these two stages are quite detailed, we devote a full chapter to each one of them. As a result, this chapter exclusively focuses on Stage I of our proposal, namely, models for the network ATFM fairness which incorporates notions of fairness. Chapter 3 delves into Stage II of our proposal, namely, models for slot reallocation, thereby facilitating airline collaboration.

Figure 2.1 depicts the mapping of our proposal to various stages of the current CDM practice (introduced in Chapter 1). Stage I outputs the next best alternative to RBS, i.e, a schedule that minimizes the number of reversals (and amount of overtaking). Stage II generalizes the Substitution+Compression phase of the current CDM practice by utilizing models for slot reallocation. We believe this proposal presents a natural framework to extend CDM from an airport to an airspace context.

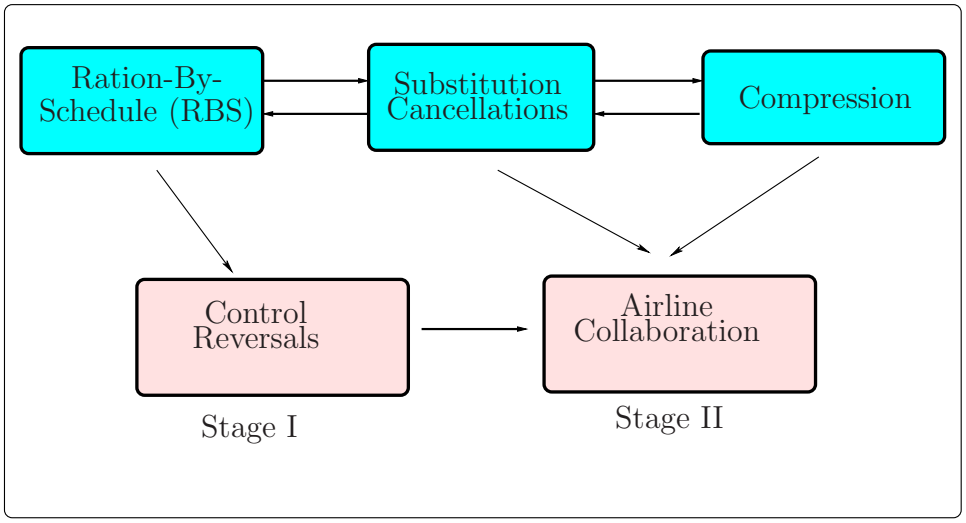


Figure 2-1: Mapping of our proposal with the current three-stage CDM practice. This chapter covers Stage I of the overall proposal whereas Chapter 3 covers Stage II.

Notation

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

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

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

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

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

$T^{\text{reversal}}_{f,f',j}$: set of time-periods common for flights f and f' where a reversal could occur in resource j,

$O^{\max}_{f,f',j}$: maximum amount of overtaking possible between flights f and f' in resource j,

M : maximum permissible delay for a flight.

The key additions relative to the notation used in [13] are $\mathcal{W}$, $\mathcal{F}_w$, $\mathcal{R}^j$, $\mathcal{R}^S$, $\mathcal{R}^A$, $T^{\text{reversal}}_{f,f',j}$, $O^{\max}_{f,f',j}$, M.

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 2.1.1. 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 2.1.2. 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 (elaborated upon later in the chapter). Figure 2-2 pictorially depicts a reversal. The shaded aircraft is scheduled to arrive before the non-shaded aircraft, but the opposite sequence is realized, thereby, leading to a reversal.

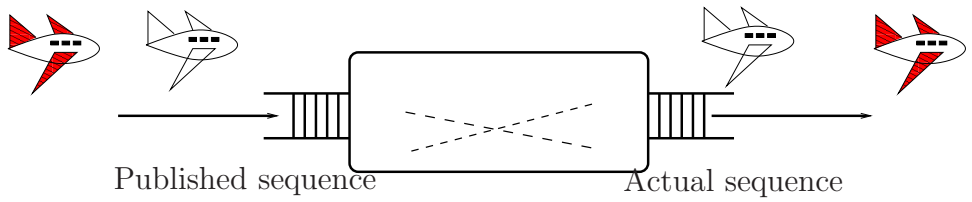


Figure 2-2: Pictorial depiction of a reversal.

Definition 2.1.3. 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 2.1.4. 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 2.1.4, $T^{\text{reversal}}_{f,f',j}$ 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 2.1.5. 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 2.1.4, $O^{\max}_{f,f',j}$ 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 2.1.1. Figure 2-3 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, $O^{\max}_{f,f',k} = |T^{\text{reversal}}_{f,f',k}| = 6$.

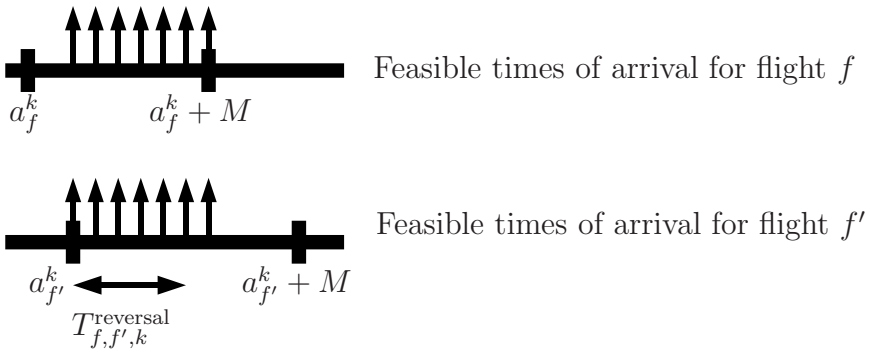


Figure 2-3: A reversible pair of flights $(f, f') \in \mathcal{R}^A$ $(\text{dest}_f = k)$.

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 chapter. 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). Moreover, the distribution of delays across airlines is non-uniform with one airline getting almost three times the delay of another airline. These observations in the solutions obtained from the formulation (TFMP) is present across all the six datasets.

Comparison with RBS

To demonstrate the utility of FSFS fairness, consider the scenario depicted in Figure 2-4. We have two airports and two airlines (A and B) operating a set of flights. Flight A1 (arriving at airport 1) is continued by flight A5 (arriving at airport 2). The turnaround time is 1 unit followed by 2 units of flight time for the subsequent flight. Similarly, flight B2 (arriving at airport 1) is continued by flight B5 (arriving at airport 2). The capacity at Airport 1 gets reduced by 66% meaning that we have 1 flight arriving every 3 slots as opposed to 1 flight every slot (under nominal conditions). In contrast, the capacity at Airport 2 gets reduced by 33% meaning that we have 2 flights arriving every 3 slots. Let us first consider the utility of a schedule which controls reversals compared to the RBS schedule. Because of the coupling between A5 and A1, flight A5 is unable to utilize the slot allotted to it under the RBS paradigm because it is not compatible with the earliest slot it can have given the slot assigned to A1. A similar situation holds for flight B5. Therefore, in the RBS schedule, there is available capacity which is not utilized. Even though RBS is delay-optimal in a single-airport setting (see [36]), this example illustrates that in a multi-airport setting, *the attractive property of delay optimality is not necessarily true.* By allowing reversals, these unused slots can be potentially filled, thereby, leading to a more efficient schedule. For instance, in this example, by reversing flights B1 and A1 at Airport 1, flight A5 can now utilize the slot assigned under the original RBS sequence. Using a similar argument, by reversing flights B2 and A2, flight B5 can

utilize the original RBS slot. Furthermore, by allowing reversals, the final schedule can overcome cancellations which might occur in the RBS schedule (for instance, flights B6 and A7 get cancelled in the RBS schedule).

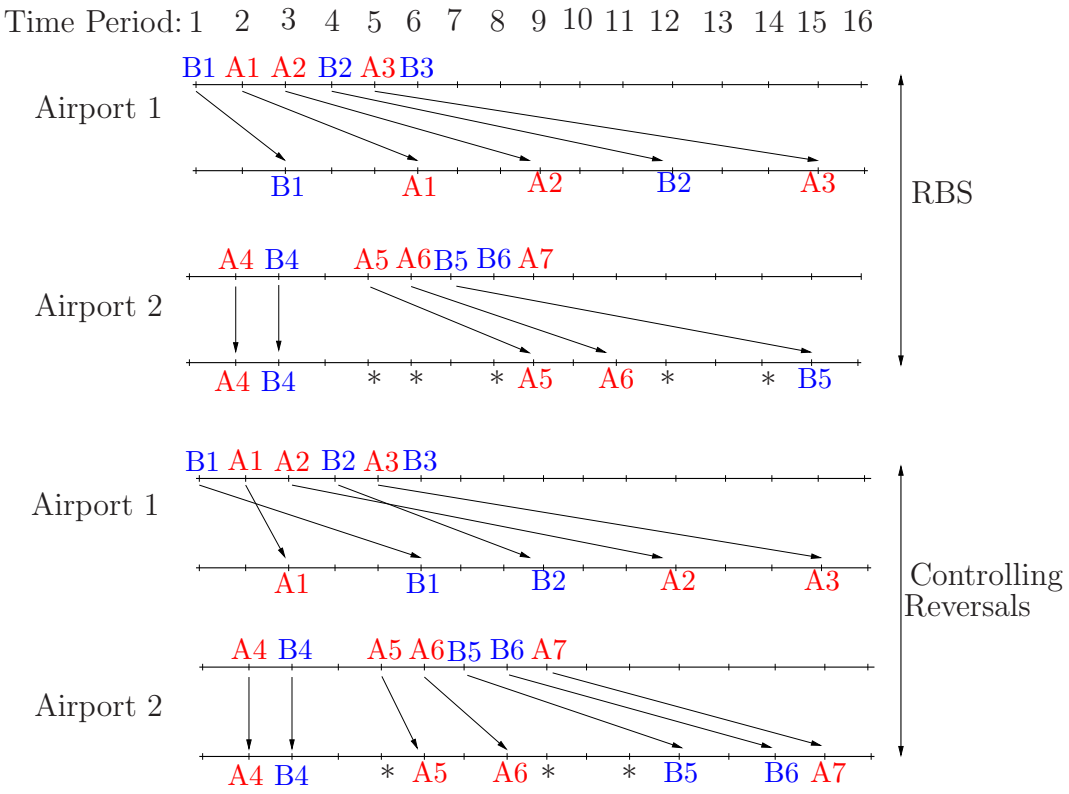


Figure 2-4: Example demonstrating the utility of controlling reversals over the RBS solution. * denotes slots which remain unutilized despite being available. The capacity of Airport 1 gets reduced by two-thirds, whereas that of Airport 2 by one-third.

Table 2.1 summarizes the advantages of our proposal relative to RBS^{++} (RBS + Compression) on the performance metrics of total delay, cancellations, capacity (under) utilization, reversals and on-time performance. As is evident, by allowing reversals, we get the dual benefits of reduced delays and better utilization of the available capacity relative to the RBS schedule. In addition, a by-product of reduced wastage of capacity is decrease in the number of cancellations.

2.2 Controlling 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 in the Online Airline Guide (OAG). As previously mentioned, this is known as Ration-

Metric	**Our Proposal**	**RBS^{++}**
Total Delay (no. of slots)	43	54
Cancellations	1	3
Slots Unused	3	5
Reversals	2	0

Table 2.1: Utility of our proposal of enforcing FSFS fairness.

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.

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 $g_{f,f',j}$.

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

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 2.2 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:

$$s^i_{f,f',j} = \max \left\{ w^{f'}_{j,\underline{T}^{f'}_j + i} - w^f_{j,\underline{T}^{f'}_j + i}, 0 \right\}. \tag{2.1}$$

Equation (2.1) implies that the following constraints suffice to express $s^i_{f,f',j}$ in a

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

Table 2.2: Truth table for modeling the overtaking variables.

mathematical programming framework if the objective is to minimize $s^{i}_{f,f',j}$:

$$s^{i}_{f,f',j} \geq w^{f'}_{j,\underline{T}_j^{f'}+i} - w^{f}_{j,\underline{T}_j^{f'}+i}, \tag{2.2a}$$

$$s^{i}_{f,f',j} \geq 0. \tag{2.2b}$$

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:

$$g_{f,f',j} = \sum_{i=0}^{O^{\max}_{f,f',j}} s^{i}_{f,f',j}. \tag{2.3}$$

We shall work with an alternative description of Equation (2.2a) to make the exposition on overtaking clearer. We substitute $i = t - \underline{T}_j^{f'}$ in Equation (2.2a) to rewrite it as follows:

$$w^{f'}_{j,t} \leq w^{f}_{j,t} + s^{t-\underline{T}_j^{f'}}_{f,f',j}. \tag{2.4}$$

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 $s^{t-\underline{T}_j^{f'}}_{f,f',j} \geq 0,\ \forall t \in T^{\text{reversal}}_{i,i',r}$:

$$w^{f'}_{j,t} \leq w^{f}_{j,t} + s^{t-\underline{T}_j^{f'}}_{f,f',j}, \qquad \forall t \in T^{\text{reversal}}_{i,i',r}. \tag{2.5}$$

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 $s^{t-\underline{T}_j^{f'}}_{f,f',j} \geq 0,\ \forall t \in T^{reversal}_{i,i',r}$, then Constraint (2.5) correctly captures the semantics of overtaking.*

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

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

then Constraint (2.5) becomes reduntant. Since we minimize total amount of over-

taking (and $s^{t-\underline{T}_j^{f'}}_{f,f',j} \geq 0, \ \forall t \in T^{\text{reversal}}_{i,i',r}$), it forces:

$$s^{t-\underline{T}_j^{f'}}_{f,f',j} = 0, \ \forall t \in T^{\text{reversal}}_{i,i',r},$$

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

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

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

Now, the time-connectivity constraints (Constraints (1.1f)) imply that:

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

Constraint (2.5) then enforces

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

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

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

leading to $g_{f,f',j} = i$. In summary, Constraint (2.5) (in addition with $s^{t-\underline{T}_j^{f'}}_{f,f',j} \geq 0$), 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 2.2, namely $(0,0,0), (0,1,1), (1,0,0)$ and $(1,1,0)$. Figure 2-5 depicts the convex hull

of these four points. We introduce the following set of constraints to model overtaking:

$$w^{f'}_{j,t} \leq w^{f}_{j,t} + s^{t-\underline{T}^{f'}_j}_{f,f',j}, \qquad \forall (f,f') \in \mathcal{R}^j,\ \ j \in \mathcal{S} \cup \mathcal{K},\ \ t \in T^{\text{reversal}}_{i,i',r}. \tag{2.6a}$$

$$w^{f}_{j,t} \leq w^{f'}_{j,t} + 1 - s^{t-\underline{T}^{f'}_j}_{f,f',j}, \qquad \forall (f,f') \in \mathcal{R}^j,\ \ j \in \mathcal{S} \cup \mathcal{K},\ \ t \in T^{\text{reversal}}_{i,i',r}. \tag{2.6b}$$

$$w^{f}_{j,t} + s^{t-\underline{T}^{f'}_j}_{f,f',j} \leq 1, \qquad \forall (f,f') \in \mathcal{R}^j,\ \ j \in \mathcal{S} \cup \mathcal{K},\ \ t \in T^{\text{reversal}}_{i,i',r}. \tag{2.6c}$$

$$-w^{f'}_{j,t} + s^{t-\underline{T}^{f'}_j}_{f,f',j} \leq 0, \qquad \forall (f,f') \in \mathcal{R}^j,\ \ j \in \mathcal{S} \cup \mathcal{K},\ \ t \in T^{\text{reversal}}_{i,i',r}. \tag{2.6d}$$

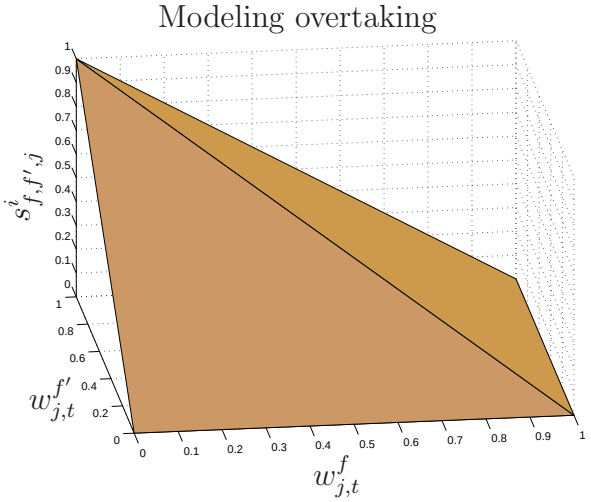


Figure 2-5: Convex hull of the integer points in Table 2.2 to model overtaking ($i = t - \underline{T}^{f'}_j$).

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

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

$$\min \sum_{f \in \mathcal{F}} \left(\sum_{t \in T^f_{\text{dest}_f}} c^f_{\text{total}}(t) \cdot (w^f_{\text{dest}_f,t} - w^f_{\text{dest}_f,t-1}) - \sum_{t \in T^f_{\text{orig}_f}} c^f_g(t) \cdot (w^f_{\text{orig}_f,t} - w^f_{\text{orig}_f,t-1}) \right) +$$

$$\lambda^o_s \cdot \left(\sum_{j \in \mathcal{S},\ (f,f') \in \mathcal{R}^j} \sum_{i=0}^{O^{\max}_{f,f',j}} s^i_{f,f',j} \right) + \lambda^o_a \cdot \left(\sum_{k \in \mathcal{K},\ (f,f') \in \mathcal{R}^k} \sum_{i=0}^{O^{\max}_{f,f',k}} s^i_{f,f',k} \right)$$

subject to:

$$(1.1a) - (1.1f).$$
$$(2.6a) - (2.6d).$$
$$s^i_{f,f',j} \in \{0,1\}, \qquad \forall (f,f') \in \mathcal{R}^j,\ \ j \in \mathcal{S} \cup \mathcal{K},\ \ i \in \{0,\ldots,O^{\max}_{f,f',j}\}.$$

2.3 Controlling the Total Number of Reversals

The model introduced in Section 2.2 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:

$$s_{f,f',j} = 1 \iff \Big\{\exists i \in \{0,\ldots,O^{\max}_{f,f',j}\},\ \ s^i_{f,f',j} = 1\Big\} \tag{2.7}$$

Building upon Equation (2.7), we have the following:

$$s_{f,f',j} = \max_{t \in T^{\text{reversal}}_{f,f',j}} \Big\{ s^{t-\underline{T}^{f'}_j}_{f,f',j} \Big\},$$
$$s_{f,f',j} = \max_{t \in T^{\text{reversal}}_{f,f',j}} \Big\{ \max\{w^{f'}_{j,t} - w^{f}_{j,t}, 0\} \Big\},$$

$$s_{f,f',j} = \max\Big\{ \max_{t \in T^{\text{reversal}}_{f,f',j}} \{w^{f'}_{j,t} - w^{f}_{j,t}\}, 0 \Big\}. \tag{2.8}$$

Equation (2.8) implies that the following constraints suffice to express $s_{f,f',j}$ in a mathematical programming framework if the objective is to minimize $s_{f,f',j}$:

$$s_{f,f',j} \geq w^{f'}_{j,t} - w^{f}_{j,t},\ \ \forall t \in T^{\text{reversal}}_{f,f',j}. \tag{2.9a}$$
$$s_{f,f',j} \geq 0. \tag{2.9b}$$

Equation (2.9a) can be rearranged as follows:

$$w_{j,t}^{f'} \leq w_{j,t}^{f} + s_{f,f',j}, \ \ \forall t \in T_{f,f',j}^{\text{reversal}}. \tag{2.10}$$

Theorem 2. *If we use an objective function of minimizing $s_{f,f',j}$, then Constraint (2.10) 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.,

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

then Constraint (2.10) 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 $\exists t \in T_{f,f',j}^{\text{reversal}}$ such that:

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

Constraint (2.10) then implies that $s_{f,f',j} \geq 1$. Again, minimizing $s_{f,f',j}$ makes $s_{f,f',j} = 1$ ensuring that Constraint (2.10) 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):

$$w_{j,t}^{f'} \leq w_{j,t}^{f} + s_{f,f',j}, \qquad \forall (f, f') \in \mathcal{R}^j, \ \ j \in \mathcal{S} \cup \mathcal{K}, \ \ t \in T_{f,f',j}^{\text{reversal}}. \tag{2.11a}$$

$$w_{j,t}^{f} \leq w_{j,t}^{f'} + 1 - s_{f,f',j}, \qquad \forall (f, f') \in \mathcal{R}^j, \ \ j \in \mathcal{S} \cup \mathcal{K}, \ \ t \in T_{f,f',j}^{\text{reversal}}. \tag{2.11b}$$

If there is a reversal between flights f and f' in resource j, i.e., $s_{f,f',j} = 1$, then Constraint (2.11a) becomes redundant and Constraint (2.11b) 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 (2.11b) becomes redundant and Constraint (2.11a) 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.

The TFMP model with the additional control on reversals is as follows:

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

$$\min \sum_{f \in \mathcal{F}} \left(\sum_{t \in T^f_{\text{dest}_f}} c^f_{\text{total}}(t) \cdot (w^f_{\text{dest}_f,t} - w^f_{\text{dest}_f,t-1}) - \sum_{t \in T^f_{\text{orig}_f}} c^f_g(t) \cdot (w^f_{\text{orig}_f,t} - w^f_{\text{orig}_f,t-1}) \right) +$$

$$\lambda^r_s \cdot \left(\sum_{j \in \mathcal{S},\ (f,f') \in \mathcal{R}^j} s_{f,f',j} \right) + \lambda^r_a \cdot \left(\sum_{k \in \mathcal{K},\ (f,f') \in \mathcal{R}^k} s_{f,f',k} \right)$$

subject to:

$$\begin{aligned} &(1.1a) - (1.1f). \\ &(2.11a) - (2.11b). \\ &s_{f,f',j} \in \{0,1\}, \qquad \forall (f,f') \in \mathcal{R}^j,\ \ j \in \mathcal{S} \cup \mathcal{K}. \end{aligned}$$

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 (2.11a) and (2.11b). We show in the Appendix that the polyhedron induced by Constraints (2.11a) and (2.11b) is the convex hull of solutions in $IP_{\text{Reversal}}(f, f', j)$.

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

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 (2.11a) and (2.11b) reduce to Constraint (2.12) which ensures that flight f' cannot arrive before flight f:

$$w^{f'}_{\text{dest}_{f'},t} \le w^f_{\text{dest}_f,t}, \qquad \forall (f,f') \in \mathcal{R}^A,\ \ t \in T^{\text{reversal}}_{f,f',\text{dest}_f}. \tag{2.12}$$

2.4 Controlling the Difference between Per Flight Airline Delays

The models presented so far did not include the airlines explicitly in the decision-making process, i.e., the carrier identity is not taken into account while assigning delays to flights, and hence, the distribution of delays among the airlines can still be non-uniform. In this section, we introduce a metric which enforces the per flight airline delay as close to each other as possible. The premise for such a metric is that the load (number of operating flights) that an airline imposes on a weather affected area should be taken into account while executing an ATFM program. To elaborate, suppose a weather front impacts 1000 flights of Airline A, but only 100 flights of Airline B. It would be unfair if both airlines get the same minutes of total delay as it would mean a ten-fold increase in the per flight delay of Airline A when compared to Airline B. Hence, this metric captures the general notion of fairness in resource allocation problems, namely, coming up with a solution which balances the net utility generated for every player. In this setup, the utility for each player (an airline) is taken to be the per flight airline delay on the set of impacted flights.

This metric, albeit, has some conceptual difficulties - consider an example where 100 flights of Airline A get affected but only 1 flight of Airline B gets affected, then requiring that both the airlines get the same per flight airline delay means that every flight of Airline A gets the same delay as that of Airline B, which, in turn implies an unnecessary increase in the delays for Airline A. Thus, in such a setting, this metric is unlikely to be useful. In summary, we feel the that this metric would be advantageous in a setting when an ATFM initiative impacts similar number of flights of the airlines involved.

In this chapter, we use this metric on national scale datasets to empirically quantify the price of fairness under this paradigm. This is because we do not have information on the exact weather conditions prevalent on the datasets we experiment upon, and hence, we are unable to filter in the set of flights that have been impacted due to capacity reductions at airports. Finally, we feel that such a metric also helps obtain computational insights on how our framework can accommodate alternative objective functions.

Let d_w denote the per flight delay for airline w and γ denote the mean of the per flight airline delay across all airlines.

$$d_w = \sum_{f \in \mathcal{F}_w} \Bigg(\sum_{t \in T^f_{\text{dest}_f}} c^f_{total}(t) \cdot (w^f_{\text{dest}_f,t} - w^f_{\text{dest}_f,t-1}) - \sum_{t \in T^f_{\text{orig}_f}} c^f_g(t) \cdot (w^f_{\text{orig}_f,t} - w^f_{\text{orig}_f,t-1}) \Bigg) \Bigg/ |\mathcal{F}_w|,$$

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

The TFMP model with the additional control on the per flight airline delays is as follows:

$$IZ_{\text{TFMP-Dev}} = \min \sum_{f \in \mathcal{F}} \Big(\sum_{t \in T^f_{\text{dest}_f}} c^f_{\text{total}}(t) \cdot (w^f_{\text{dest}_f,t} - w^f_{\text{dest}_f,t-1}) - \sum_{t \in T^f_{\text{orig}_f}} c^f_g(t) \cdot (w^f_{\text{orig}_f,t} - w^f_{\text{orig}_f,t-1}) \Big) + \lambda_d \cdot \Big(\sum_{w \in \mathcal{W}} |d_w - \gamma| \Big)$$

subject to: (1.1a)-(1.1f).

The absolute value terms $|d_w - \gamma|$ are linearized by replacing each such term by a new variable and enforcing that variable to be greater than both $d_w - \gamma$ and $\gamma - d_w$.

2.5 Controlling both Reversals and Difference in Per Flight Airline Delays

The final model we study is to control both the total number of reversals and the difference in per flight airline delays. Hence, the objective function in this case is a weighted sum of three components - total delay cost, total number of reversals and difference in per flight airline delays costs.

The TFMP model with control on both reversals and per flight airline delays is as follows:

$$IZ_{\text{TFMP-Rev-Dev}} =$$

$$\min \sum_{f \in \mathcal{F}} \left(\sum_{t \in T^f_{\text{dest}_f}} c^f_{\text{total}}(t) \cdot (w^f_{\text{dest}_f,t} - w^f_{\text{dest}_f,t-1}) - \sum_{t \in T^f_{\text{orig}_f}} c^f_g(t) \cdot (w^f_{\text{orig}_f,t} - w^f_{\text{orig}_f,t-1}) \right) +$$

$$\lambda^r_a \cdot \left(\sum_{k \in \mathcal{K},\ (f,f') \in \mathcal{R}^k} s_{f,f',k} \right) + \lambda_d \cdot \left(\sum_{w \in \mathcal{W}} |d_w - \gamma| \right)$$

subject to:

$$(1.1a) - (1.1f).$$
$$(2.11a) - (2.11b).$$
$$s_{f,f',k} \in \{0,1\} \qquad \forall (f,f') \in \mathcal{R}^k.$$

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 chapter 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 r_w denote the number of reversals per flight for airline w and γ denote the mean of the r_w's across all airlines.

$$r_w = \left(\sum_{f' \in \mathcal{F}_w} s_{f,f'} \right) \Big/ |\mathcal{F}_w|,$$
$$\gamma = \left(\sum_{w \in \mathcal{W}} r_w \right) \Big/ |\mathcal{W}|.$$

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

Size of the Formulations.
Denoting with

$$M = \max_{f \in \mathcal{F}, j \in P_f} |T_f^j|, \qquad N = \max_{f \in \mathcal{F}} |\mathcal{S}^f|,$$

the total number of decision variables and constraints for the various models can be bounded as listed in Table 2.3.

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 2.3: 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: $\mathcal{K} = 50$, $\mathcal{T} = 100$, $\mathcal{S} = 100$, $\mathcal{R} = 50000$, $\mathcal{F} = 10000$, $\mathcal{C} = 8000$, $M = 6$ and $N = 5$. For this example, the upper bound on the number of variables and constraints are listed in Table 2.4.

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 2.4: 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).

2.6 Prioritizing between Airport Reversals and Sector Reversals

Currently RBS is the principle used to maintain the sequencing of flights at capacitated elements of NAS (both for GDPs and AFPs) and indirectly allocate delays to airlines. By optimizing over reversals, we control the degree of disruptions to such sequences when capacity decreases. This allows us to decrease the overall delay (relative to RBS) at the expense of increasing reversals. That is, we allow more flexibility on the tradeoff between optimality (i.e., delays) and fairness (i.e., reversals). This tradeoff is controlled by the parameters λ_a^r and λ_s^r which control the airport and sector reversals respectively. In Section 2.7, we provide empirical evidence on the tradeoff between fairness and optimality. We now examine multiple contrasting scenarios to understand better the pros and cons of controlling sector reversals.

The operational details of AFPs (one of the recent ATFM programs) under the current CDM practice facilitate the control of sector reversals. 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 2.6.1. Consider the scenario depicted in Figure 2-6. 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.

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

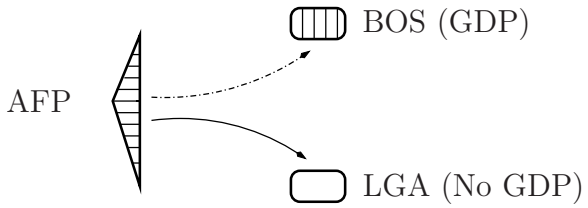


Figure 2-6: Illustration of a scenario where controlling sector reversals seems appropriate. There are multiple ATFM programs operating simultaneously. Specifically, an AFP is spatially followed by an airport with a GDP (BOS) and an airport with no GDP (LGA).

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 addition, it seems reasonable to expect that primary stakeholders (airlines and flying passengers) would be less concerned as to how the en-route resources are allocated if they are satisfied with the final arrival sequences at the airports. To elaborate, from a passengers' standpoint, the key objective is to reach the destination on-time (which is correlated with the concerned flight not overtaking other flights). The exact trajectory of the flight is of little importance as long as that objective is achieved. In a similar vein, an airline is primarily concerned with ensuring excellent on-time performance at destination airports. Therefore, in a setting where rerouting is not allowed, airlines would be less concerned with the way en-route resources are allocated (if FSFS guarantees are satisfied at the airports). Finally, this scheme of things would be attractive to the FAA as its key objective in designing various ATFM programs is to satisfy these primary stakeholders.

Example 2.6.2. Consider the scenario depicted in Figure 2-7. There are two airports spatially close where GDPs (BOS and LGA) are operational. There is no ATFM program operational in the en-route airspace. There are two streams of flights going through the airspace to one of these airports. In this scenario, we believe there is no compelling reason to control reversals in the en-route airspace. In fact, controlling sector reversals appears an overkill as it would increase the delay costs.

To analyze the relevance of controlling sector reversals in scenarios where competing schedules have the same number of airport reversals, we study the following example:

Example 2.6.3. Consider the two settings as depicted in Figure 2-8. In Scenario 1, there are two separate sequence of flights traveling between two distinct origin-destination pairs. Specifically, there is one sequence traveling from airport A to

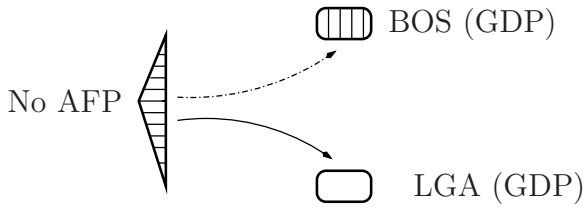


Figure 2-7: Illustration of a scenario where controlling sector reversals seems an overkill. There are multiple ATFM programs operating simultaneously. Specifically, there are GDPs operating at two nearby airports (BOS) and (LGA). Moreover, there is no ATFM program operating en-route.

airport $A^{'}$ through the congested sector S_1. There is a separate sequence of flights from airport B to airport $B^{'}$ through the same sector S_1. Consider two competing schedules as follows: i) all flights from A to $A^{'}$ go before flights from B to $B^{'}$; and ii) all flights from B to $B^{'}$ go before flights from A to $A^{'}$. It is evident that both these schedules have 0 airport reversals, but it is difficult to conclude which one is more equitable. Thus, there is some truth to the premise of controlling sector reversals in sector S_1. In Scenario 2, there are two streams of flights traveling from the same origin airport C to the destination airport $C^{'}$ but through different sectors S_2 and S_3 respectively. In this setting, it is difficult to comprehend the rationale behind controlling sector reversals. The reason being that this is similar to a single-airport GDP scenario (at airport $C^{'}$), and currently there are no equity considerations on the en-route resources when executing a GDP. Moreover, controlling reversals in sectors S_2 or S_3 will strictly lead to increase in total delay which seems unnecessary.

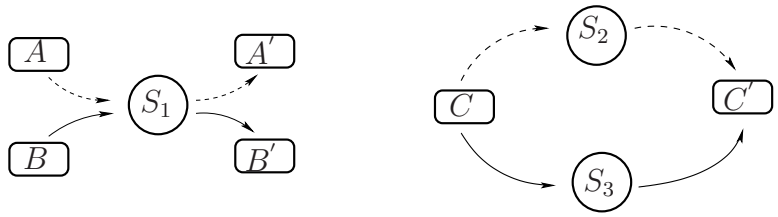


Figure 2-8: Two scenarios to study the relevance of controlling sector reversals in addition to airport reversals.

In summary, the discussion in this section leads to the following conclusions:

- Controlling airport reversals (where GDPs are implemented) should be the first order objective.
- Simultaneously, sector reversals should only be controlled in the en-route airspace

where AFPs (or other ATFM programs) are operational. We propose that in the model TFMP-Reversal, the tradeoff parameter λ_s^r used to control sector reversals be set to a relatively lower value as compared to the parameter λ_a^r used to control airport reversals.

2.7 Computational Results

In this section, we report computational results from the optimization models introduced in the previous section 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.

Statistics of the Datasets

Table 2.5 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 distribution of delays across airlines.

Day	**No. of Flights**	$\mathcal{CF}$	**Total Delay** (units of 15 min.)	**Reversals**	**Overtaking**
1	5092	2691	4438	9944	30652
2	5844	3298	4926	7756	19003
3	5780	3310	3079	4797	12231
4	4590	2301	3907	6488	18825
5	5128	2728	3326	6399	19998
6	4781	2504	3101	4351	12954

Table 2.5: Summary of the datasets. $\mathcal{CF}$ denotes the number of connecting flights.

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. Each of the

55 major airports of the US is 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 the 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.

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 capacities at all the airports by noting the actual times of departure and arrival of the flights and we use these values as capacity inputs to run the optimization models. 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 airport capacities on that day has to be higher than the actual number of operations. Finally, we set values for the nominal sector capacities that lead to no delays when these airport capacities are used.

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 through $(a_f + M)$. The size of all the optimization models and hence, the computational times, are sensitive to the value of M. 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.

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 chapter 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).

Performance of TFMP

Table 2.6 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**	**Delay** (15 min.)		**Actual**		**TFMP**	
	Flights	**Actual**	**TFMP**	$\mathcal{RV}$	$\mathcal{OV}$	$\mathcal{RV}$	$\mathcal{OV}$
1	5092	4438	3385	9944	30652	915	1492
2	5844	4926	3492	7756	19003	924	1426
3	5780	3079	2242	4797	12231	753	1191
4	4590	3907	3053	6488	18825	769	1235
5	5128	3326	2648	6399	19998	801	1291
6	4781	3101	2542	4351	12954	526	822

Table 2.6: Performance of TFMP. $\mathcal{RV}$ denotes the number of reversals and $\mathcal{OV}$ the amount of overtaking.

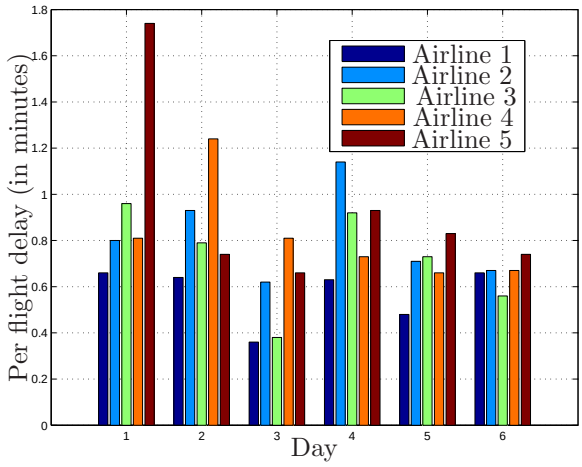


Figure 2-9: Distribution of per flight airline delays from (TFMP) in units of 15 minutes.

Performance of TFMP-Overtake

Table 2.7 reports the computational performance of (TFMP-Overtake) model on the six datasets. These results pertain to the parameter λ_o^a 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.

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 paramater

Figure 2-10 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 λ_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.

The effect of the number of airports

Table 2.8 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. The capacity input used for the results in Table 2.8 is 20% higher than the exact number of aircraft operations that happened on the day under consideration. 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

Day (# of Flights)	$\mathcal{AF}$	**Solution Time (sec.)**	$\mathcal{RV}$	$\mathcal{OV}$	**Delay Cost (15 min.)**	**% Increase in Delay Cost over (TFMP)**
	0	261	915	1492	3525	
1	5	186	2	4	3690	4.68
(5092)	15	727	13	27	4243	20.36
	25	3073	26	39	4662	32.25
	30	3600	39	65	4815	36.59
	0	108	924	1426	3604	
2	5	206	1	2	3802	5.49
(5844)	15	596	6	9	4029	11.79
	25	806	9	12	4080	13.20
	30	3530	16	18	4510	25.13
	0	311	753	1191	2313	
3	5	170	3	6	2401	3.80
(5780)	15	397	8	18	2584	11.71
	25	295	11	22	2651	14.61
	30	3394	17	28	3096	33.85
	0	51	769	1235	3173	
4	5	150	1	1	3628	14.33
(4590)	15	746	5	6	4201	32.39
	25	691	13	18	4452	40.30
	30	3600	29	33	4743	49.47
	0	178	801	1291	2744	
5	5	116	0	0	2871	4.62
(5128)	15	492	10	18	3319	20.95
	25	1983	17	26	3505	27.73
	30	3600	25	36	3804	38.62
	0	49	526	822	2637	
6	5	143	5	7	2826	7.16
(4781)	15	378	9	15	3070	16.42
	25	479	15	22	3145	19.26
	30	1305	28	49	3383	28.28

Table 2.7: 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.

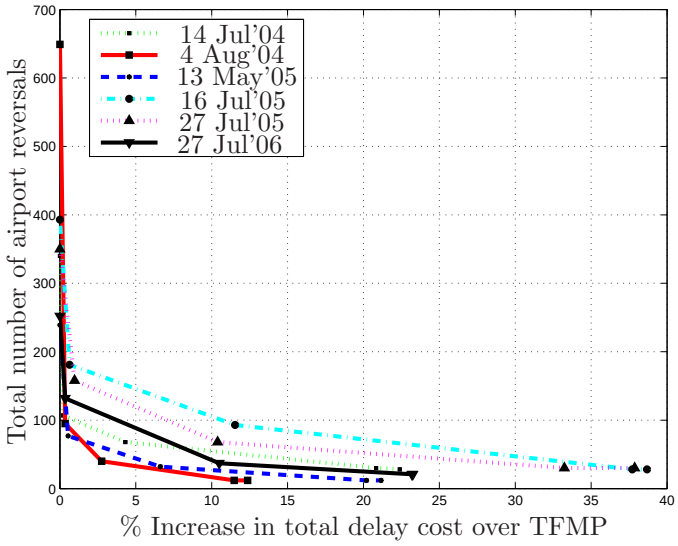


Figure 2-10: Effect of the tradeoff parameter λ_a^r. The five points for each day correspond to the result from (TFMP-Reversal) with $\lambda_a^r = 0$, 1, 10, 100 and 1000.

for up to 25 airports. As expected, the total amount of reversals reported in Table 2.8 is always less than the corresponding number in Table 2.7, whereas the opposite is true for the amount of overtaking.

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.

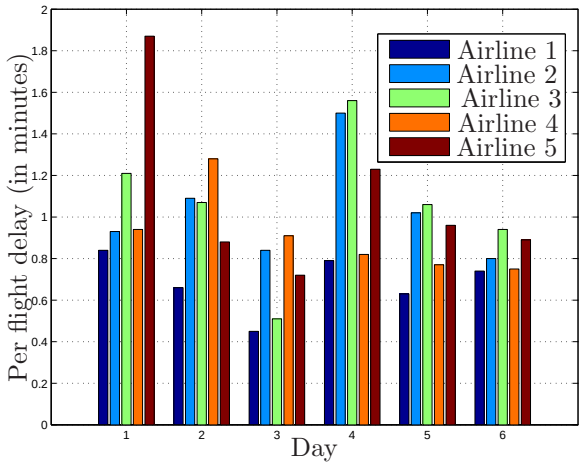


Figure 2-11: Distribution of per flight airline delays from (TFMP-Reversal) in units of 15 minutes.

Day (# of Flights)	$\mathcal{AF}$	**Solution Time (sec.)**	$\mathcal{RV}$	$\mathcal{OV}$	**Delay Cost (15 min.)**	**% Increase in Delay Cost over (TFMP)**
	0	261	915	1492	3525	
1	5	228	2	4	3691	4.70
(5092)	15	899	13	30	4246	20.45
	25	3600	25	53	4402	24.87
	30	3600	37	84	4586	30.09
	0	108	924	1426	3604	
2	5	213	1	2	3785	5.02
(5844)	15	837	6	9	4028	11.76
	25	1024	9	12	4077	13.12
	30	3600	16	30	4403	22.16
	0	311	753	1191	2313	
3	5	219	3	6	2406	4.02
(5780)	15	282	8	19	2583	11.67
	25	384	11	24	2648	14.48
	30	3600	15	35	3048	31.77
	0	51	769	1235	3173	
4	5	165	1	1	3627	14.30
(4590)	15	1367	5	10	4138	30.41
	25	2292	12	22	4468	40.81
	30	3600	29	55	4558	43.64
	0	178	801	1291	2744	
5	5	225	0	0	2866	4.44
(5128)	15	654	10	24	3319	20.95
	25	1556	16	35	3504	27.69
	30	3600	25	38	3925	43.03
	0	49	526	822	2637	
6	5	213	5	7	2836	7.54
(4781)	15	751	9	16	3059	16.00
	25	654	15	25	3125	18.50
	30	3600	28	51	3387	28.44

Table 2.8: 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.

Performance of TFMP-Dev

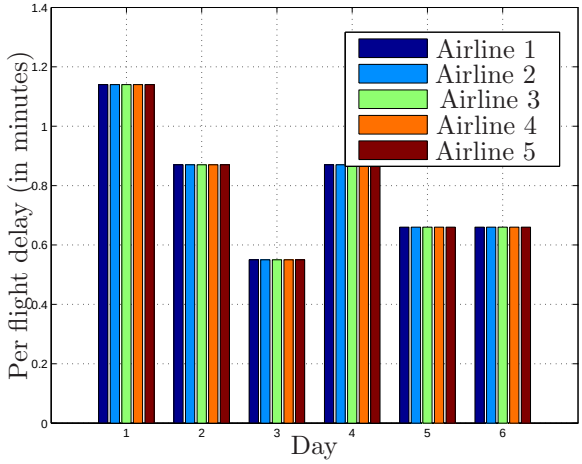


Figure 2-12: Distribution of per flight airline delays from (TFMP-Dev) in units of 15 minutes.

Day	**No. of Flights**	**Total Delay Cost** (units of 15 min.)	**Sol. Time** (in sec.)	**No. of Reversals**	**% Increase over** (TFMP)
1	5092	4835	1459	1317	6.35
2	5844	5086	1862	1256	0.24
3	5780	3184	1484	1100	1.30
4	4590	4024	1050	1061	0.54
5	5128	3422	1079	882	0.41
6	4781	3170	1010	509	0.00

Table 2.9: Computational Performance of (TFMP-Dev).

Table 2.9 reports computational performance of the (TFMP-Dev) model on the six datasets. These results pertain to the parameter λ_d set to 100. The average increase in total delay costs across the different days over the (TFMP) solution is 1.46% which suggests that this fairness criterion is satisfied at a small cost. Since the increase in total delay costs over the (TFMP) solution is small, we do not explore the degradation in total delay costs as a function of λ_d. It is evident from Figure 2-12 that the distribution of per flight delays across all airlines is nearly the same. But, the number of reversals with this model is large in all cases (1021 on average). (TFMP-Dev) takes an average time of 1324 seconds (less than 25 minutes) to come up with an optimal solution.

Day	**No. of Flights**	**Total Delay Cost** (15 min.)	**Sol. Time** (in sec.)	$\mathcal{RV}$	**% Increase over TFMP**	**Gap** %
1	5092	5366	7200	39	18.03	13.56
2	5844	5786	3298	13	14.07	2.02
3	5780	3845	7200	12	20.99	2.65
4	4590	5505	7200	34	37.55	8.04
5	5128	4504	5737	32	24.28	0.57
6	4781	3833	7200	26	20.91	13.17

Table 2.10: Computational performance of (TFMP-Rev-Dev) in units of 15 minutes. $\mathcal{RV}$ denotes the number of reversals.

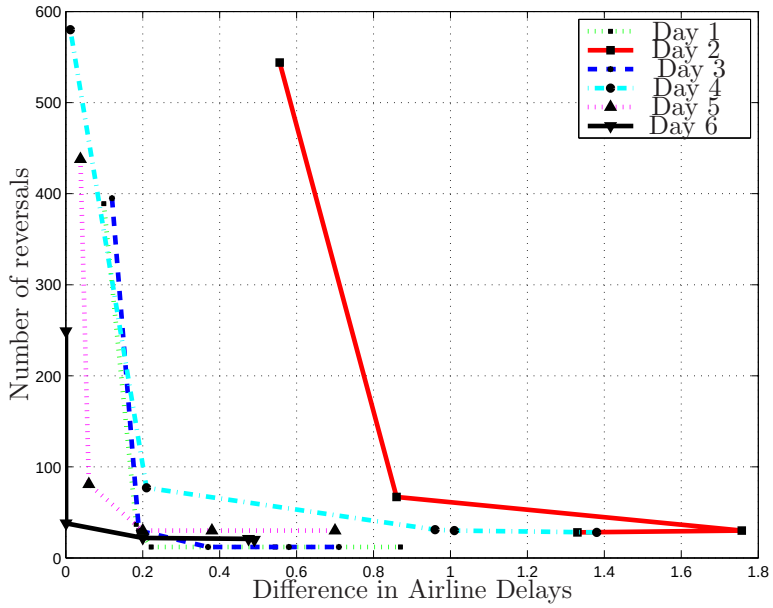


Figure 2-13: (TFMP-Rev-Dev): Tradeoff between reversals and difference in airline delays. Note that the horizontal axis corresponds to $\sum_{w \in \mathcal{W}} |d_w - \gamma|$ (in units of 15 minutes). Specifically, the value 0.2 corresponds to 3 minutes. For each day, the five points correspond to (λ_a^r, λ_d) set to (0, 100), (10, 100), (100, 100), (100, 10) and (100, 0) respectively.

Performance of TFMP-Rev-Dev

The preceding discussions suggest that (TFMP-Reversal) can give a solution with a small number of reversals but the per flight airline delays are, in general, quite different, whereas (TFMP-Dev) model can give a solution with nearly the same per flight airline delays but, a large number of reversals. This gives motivation for studying the model (TFMP-Rev-Dev) so as to satisfy both the fairness properties. Table 2.10 reports the results from the model (TFMP-Rev-Dev) with the parameters $\lambda_a^r = 100$ and $\lambda_d = 100$. The model is able to achieve a small number of reversals (between 10 and 40). Moreover, the distribution of delays across airlines (shown in Figure 2-14) is better than (TFMP) and (TFMP-Reversal). Thus, the model is able to satisfy both the objectives well. The computational times from this model are less attractive than the previous models (in four cases, the model does not reach provable optimality even after 7200 seconds).

Effect of the tradeoff parameters

Figure 2-13 shows the tradeoff between reversals and difference in per flight airline delays as a function of λ_a^r and λ_d. The five points on the graph for each day correspond to the results from (TFMP-Rev-Dev) for the parameters (λ_a^r, λ_d) set to (0, 100), (10, 100), (100, 100), (100, 10) and (100, 0). The tradeoff frontier is sharp. It falls off quickly and thereafter become constant. There is a narrow band where both objectives (number of reversals and difference in per flight airline delays) take a small value. The average increase in total delay costs over the (TFMP) solution for the parameters (λ_a^r, λ_d) set to (10, 100) is 8.34% and for (100, 10), it is 24.38%. Since the solution corresponding to the weight parameters (10, 100) has a small number of reversals (less than 100) and small difference in airline delays (less than 3 minutes), it suggests that we can obtain solutions satisfying both the fairness criteria for less than 10% increase in delay costs.

Controlling Sector Reversals and Balancing with Airport Reversals

We believe that controlling airport reversals (and overtaking) should be the primary objective, and controlling sector reversals the secondary goal. In this section, we study the interaction of the two objectives with the aim to quantify the price of controlling

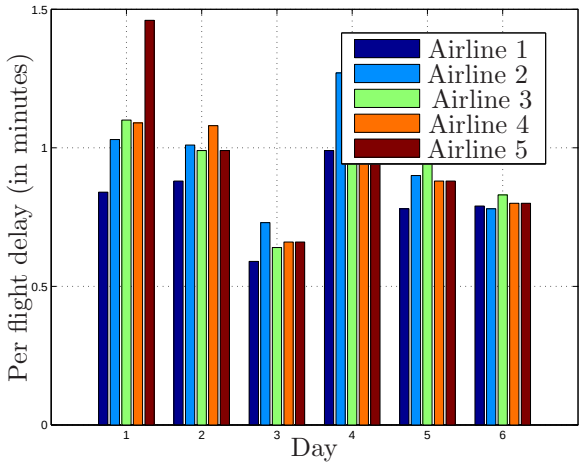


Figure 2-14: Distribution of per flight airline delays from (TFMP-Rev-Dev) in units of 15 minutes.

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 2.11 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 2-15 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 2-15 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)).

Day	$\lambda_r^s = 0$, $\lambda_r^a = 100$			$\lambda_r^s = 1$, $\lambda_r^a = 100$			$\lambda_r^s = 10$, $\lambda_r^a = 100$		
	$\mathcal{SR}$	$\mathcal{AR}$	$\mathcal{DC}$	$\mathcal{SR}$	$\mathcal{AR}$	$\mathcal{DC}$	$\mathcal{SR}$	$\mathcal{AR}$	$\mathcal{DC}$
1	3174	8	3780	253	11	3865	127	12	4261
2	3455	6	3860	169	8	3944	59	8	4235
3	1289	9	2486	58	9	2493	14	10	2631
4	3242	5	3663	326	6	3874	136	4	4292
5	1646	0	2897	268	0	2958	75	1	3456
6	1615	10	2832	144	10	2907	46	9	3235

Table 2.11: Balancing Sector Reversals with Airport Reversals ($\mathcal{SR}$ denotes the number of sector reversals, $\mathcal{AR}$ denotes the number of airport reversals and $\mathcal{DC}$ denotes the delay cost). Fairness imposed in 5 sectors of the north-east region and 10 airports spatially close to these sectors.

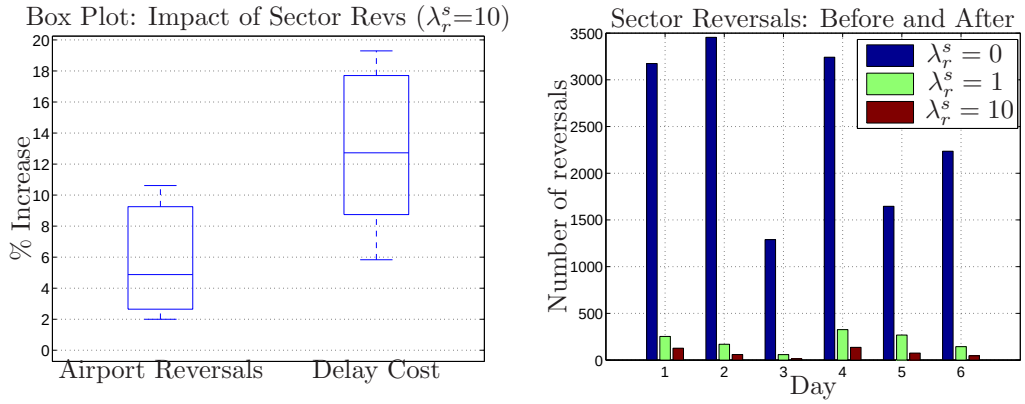


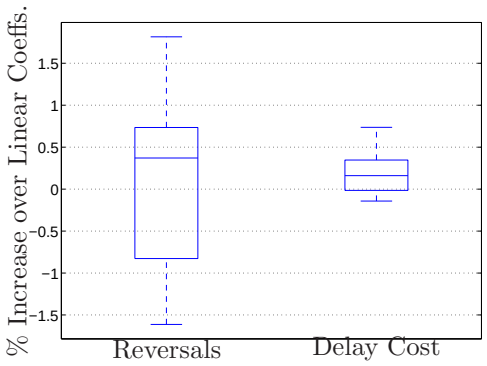
Figure 2-15: Impact of controlling sector reversals on i) airport reversals and total delay cost (Left); and ii) sector reversals (Right).

Interaction with Super-linear Cost Coefficients

We used super-linear cost coefficients in the overall objective function as additional means to impose equity as it 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.

The left plot in Figure 2-16 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 2-16 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.



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 2-16: Impact of super-linear cost coefficients on i) reversals and total delay cost (Left); and ii) distribution of flight delays (Right).

2.8 Conclusions

In this chapter, we propose integer programming models that add fairness controls to the network ATFM problem. The issue of fairness is critical, as is evident from the fact that no optimization model (in a network setting) has been deployed online inspite of the significant number of models proposed in the research literature. We propose different metrics of fairness such as i) FSFS fairness (control of reversals and overtaking); ii) Proportional fairness (equalizing airline delays); and iii) a combination of both. The numerical results we report indicate that 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). We obtain solutions that are able to control the total reversals and overtaking up to less than 100 (from the models TFMP-Reversal and TFMP-Overtake respectively) with a less than 10% increase in delay costs. In contrast, the price of fairness for equalizing delays is relatively small, 1.46% on average. 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.

Chapter 3

Network Models that Incorporate Airline Collaboration

In this chapter, we present Stage II of our proposal to extend CDM from an airport to an airspace setting. We generalize the intra-airline flight substitution and compression phase in current CDM practice to inter-airline slot exchanges across airports. We introduce two models to achieve this functionality. In the first model, we extend the single-airport model (a network flow model with side constraints) proposed by Vossen and Ball [36] to a network setting. In the second model, we propose a new formulation that introduces only one additional variable per offer above the TFMP model. We let airlines submit offers to trade across various airports. Finally, we report extensive empirical results to highlight the utility of the proposed models.

3.1 Motivation

The motivation for encouraging slot reallocation amongst airlines follows from the observation that each airline typically has very different internal objective functions. These are opaque to other airlines and to the FAA. In fact, significant uncertainty (at every step of the planning process) and schedule disruptions imply that airline objectives change dramatically from one period to another. Thus, it is reasonable to assume that the exact motives of individual airlines should not be expected to be known to anyone but the airline itself. To elaborate upon this discussion, consider the setting depicted in Figure 3-1. Three airlines A, B, C have contrasting objective functions that they are trying to internally optimize. Airline A is concerned with

maximizing the on-time performance of its flights. In contrast, Airline B faces strong coupling between the incoming and outgoing flights at the hub airport. Consequently, its main concern is to ensure that flights bound to its hub airport have minimum delays. Yet another Airline C aspires to minimize the number of flight cancellations. Starting from an initial schedule, it is possible that exchange of slots amongst these three airlines leads to a schedule wherein each of the three airline derives more utility with respect to its objective functions. Therefore, a market mechanism wherein airlines propose options for flight reallocations to the FAA would be attractive because the airlines are not revealing their exact motives and the FAA can ensure a solution with greater global welfare without making any airline unhappier from the status quo. There are two key ingredients in the practical implementation of such a mechanism from an operational standpoint:

1. **Input from Airlines.** The first pertains to the exact structure of the input needed from the airlines that describes their substitution preferences. A critical requirement for this is that it should be simple enough for the Airline Operational Centers (AOCs) to be able to decide it in real-time. Furthermore, simplicity in input data would enable the FAA to execute the matching problem between different airlines in a more expeditious manner.

2. **Mediation from FAA.** Given airline preferences, the FAA then needs to mediate and solve the optimization problem resulting from matching the offers. The critical requirement in this phase is that the method should be efficient to be run in near real-time. This way a new schedule can be announced to the airlines so that multiple iterations can be performed successively.

Possible Mechanisms

A myriad of different schemes could be implemented to achieve the desired functionality of slot trading (depicted in Figure 3-2). As a decentralized mechanism, one could envision airlines buying and selling slots amongst themselves (once they have an initial allocation from RBS). As an alternative, there could be a market mechanism to achieve the same functionality. In contrast, a centralized mechanism could be designed wherein airlines propose offers which are appealing to them to exchange slots. The FAA would then act as the centralized clearing agency tasked with evaluating and selecting possible trades. The pros-and-cons of the two approaches are relatively straightforward to visualize. The decentralized mechanism suffers from

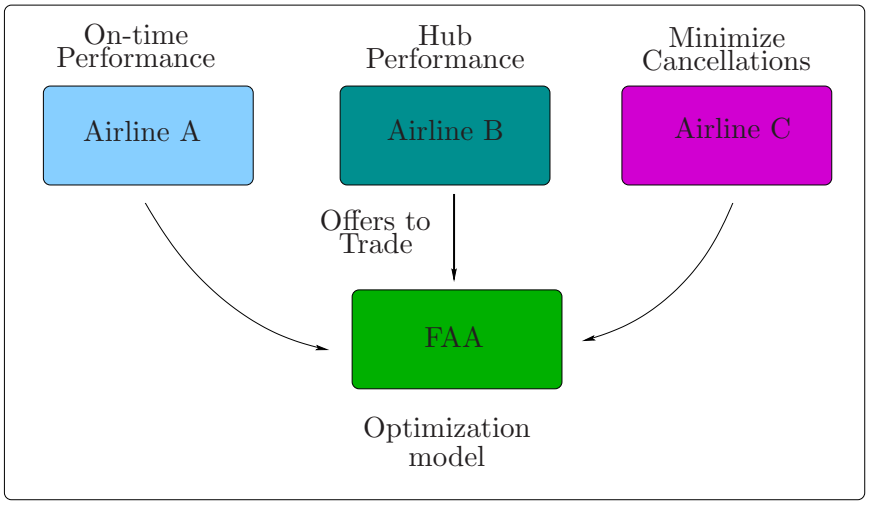


Figure 3-1: Illustration of a Scenario Motivating Airline Collaboration.

various pitfalls. For starters, given the significant uncertainty and dynamic environment, exchanges carried amongst airlines could interfere with other ATFM initiatives, thereby, resulting in potential operational infeasibility. Furthermore, there has to be a notion of equity built-in into the new schedule that is generated as a result of slot exchanges. In fact, a decentralized mechanism could facilitate monopolistic markets. Whereas, the centralized approach is attractive from two counts: i) the proposals from the airlines can be quite complex which allows for multiple possible exchanges; and ii) since the FAA can now solve a formal mediation problem, the critical aspect of ensuring equity can be explicitly incorporated in determining the overall set of executed trades. Given the competitive marketplace, the latter advantage is especially appealing as the fundamental tenet of CDM paradigm is ensuring equity in the allocation of resources at every step of the planning process.

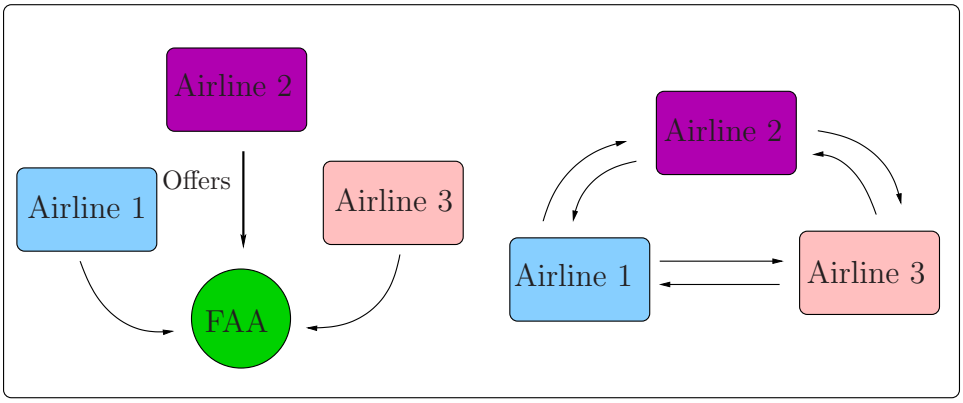


Figure 3-2: Illustration of centralized and decentralized mechanisms for airline collaboration.

	Centralized	Decentralized
Ensuring Equity	Possible	Impossible
Network Implementability	Possible	Impossible
Running Times	Less Quick	Very Quick

Table 3.1: Relative benefits of a centralized and decentralized slot reallocation mechanism.

Example Illustration

To concretely motivate the utility of slot trading, consider the following examples. Figure 3-3 depicts a scenario motivating reallocation in a single-airport setting. There are five slots that need to be assigned between two Airlines A and B at this airport. The slots marked with red are the ones currently allotted to Airline A, whereas the ones marked with green are currently allotted to Airline B. The priority (based on delay cost levels) associated with the current assignment of flights to slots might be such that the total value to both airlines might be greater as a result of a 2-for-2 trade. Suppose the value of the four slots s_1, s_2, s_3 and s_4 to both airlines A and B is as depicted in Table 3.2. Then, in case, a 2-for-2 trade occurs (i.e., A trades its current assignment of slots s_1 and s_2 for Airline B slots s_3 and s_4), the net gain to A is: 2500 - 100 + 300 - 2000 = \$700 and to B is: 500 - 600 + 2500 - 2000 = \$400. Thus, even though both airlines place positive values on all of the slots, both airlines benefit from the trade.

	Value for Airline	
Slot	**A**	**B**
s_1	\$2000	\$500
s_2	\$100	\$2500
s_3	\$2500	\$600
s_4	\$300	\$2000

	Airline	
Total Value	**A**	**B**
Before trade	\$2100	\$2600
After trade	\$2800	\$3000
Gain	\$700	\$400

Table 3.2: Left: Value of the slots to the two airlines; Right: Value proposition of various assignments.

Figure 3-4 depicts a scenario motivating slot reallocation in a network setting. The initial slot assignment is similar to the one shown in Figure 3-3 except for the fact that the two favorable slots (for each airline) are at different airports. The only added complication in this case is that if a trade happens, then the resulting

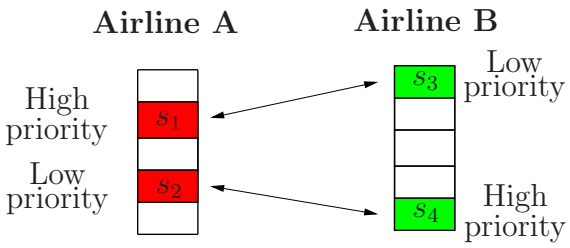


Figure 3-3: Motivation for slot trading in a single-airport setting.

assignment should also be feasible (in the network setting), i.e., it should satisfy the various connectivity constraints.

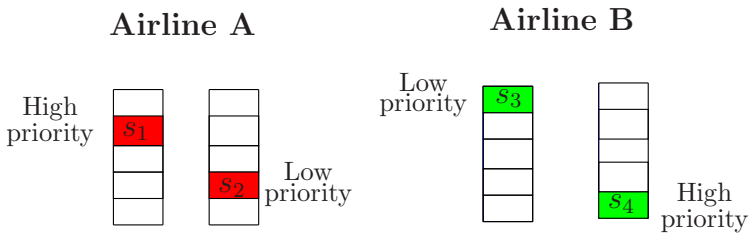


Figure 3-4: Motivation for slot trading in a multi-airport (network) setting.

3.2 Input Data

In this section, we elaborate upon the data required from airlines. First, we introduce some notation:

$\mathcal{O}$: set of all possible airline offers (definition below),
$\mathcal{O}^f \subseteq \mathcal{O}$: set of offers containing flight f,
D_f : time period assigned to flight f from Stage I optimization.

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, Ball [36] that allows the airlines to submit so-called AMAL, "*at-most, at-least*" 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. 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 D_{f_d} and D_{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 $D_{f_d} < t_{d'}$ and $D_{f_u} > t_{u'}$. Please see Figure 3-5 for an example.

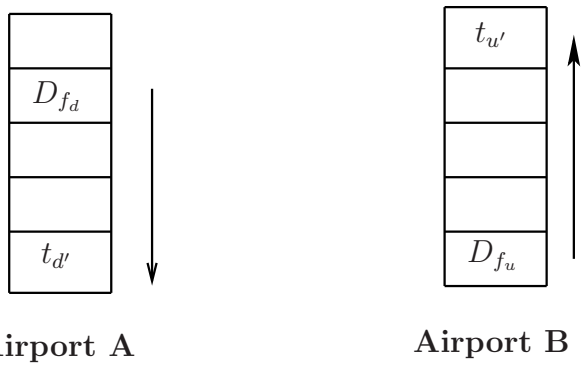


Figure 3-5: Illustration of the structure of an offer $(f_d, t_{d'}; f_u, t_{u'})$. $\text{dest}_{f_d} = A$, $\text{dest}_{f_u} = B$, $t_{d'} = D_{f_d} + 3$ and $t_{u'} = D_{f_u} - 4$. The offer states that the airline is willing to delay flight f_d by at-most 3 slots if in return flight f_u is moved earlier by at-least 4 slots.

Utility of Slot Reallocation

Given the allocation from the fairness stage, we now contrast the utility of the inter-airline slot reallocation phase of our proposal compared to the intra-airline substitution phase in current CDM practice. We build upon the scenario depicted in Figure 2-4 (in Chapter 2). We use the RBS solution and the solution controlling reversals as the starting point for Stage II. Figure 3-6 depicts the resulting schedule. In the example we consider, each airline is concerned with maximizing the number of flights arriving on-time (defined as arrival within one time-period) at its hub airport. Airport 1 is the hub of A and Airport 2 is the hub of B. Then, if Airline A submits an offer (A6, 6; A3, 10) and Airline B submits another offer (B1, 9; B6, 6), then a trade gets executed. The key observation is that by allowing inter-airline substitution (with airline offers across multiple airports), we are able to have 2 more flights (B6, A3) arrive on-time. It is not difficult to envision potential benefits that might ensue because of trades across multiple airlines. In contrast, with the intra-airline substitution phase, it is not possible to have flights A3 and B6 arrive on-time.

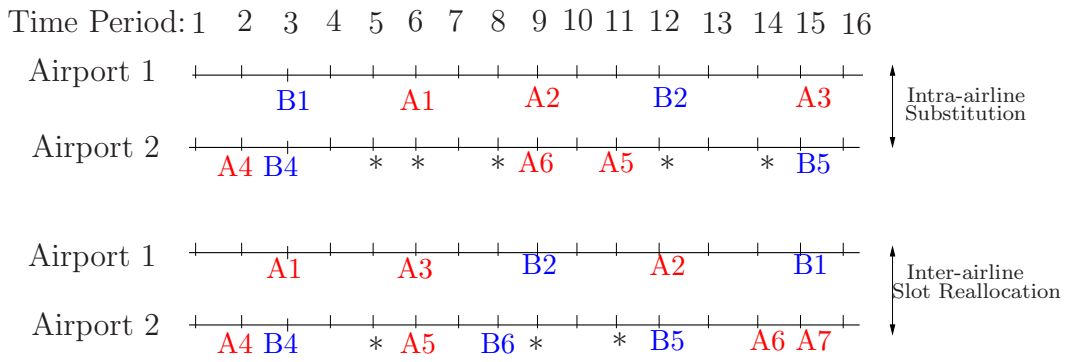


Figure 3-6: Example demonstrating the utility of our slot reallocation phase over the intra-airline substitution.

A Graph Theoretic Interpretation of Trade Offers

Based on the structure of the AMAL airline offers, the resulting set of trade offers has an insightful interpretation using the toolset of graph theory. This interpretation was first noted in Sherali et al. [32]. This facilitates modeling the slot exchange functionality in terms of network flow constraints in the graph. More precisely, consider a graph where the set of nodes are the slots and the arcs represent flights being assigned a new slot (pointed to by the arc head) from the currently allotted one (given by tail of the arc). Thus, the input airline offers can be viewed on a directed network (see Figure 3-8 for example). The set of acceptable trades correspond to directed cycles in this graph.

As a concrete example, Figure 3-7 depicts a setting in an airspace region involving three airlines A, B and C. There are six slots that need to be assigned between six flights (two belonging to each of the three airlines). Most importantly it shows examples of AMAL trade offers proposed by the three airlines. For instance, Airline A inputs the offer (A1, 3; A2, 4) meaning that it is willing to delay flight A1 to at-most slot 3, if in return, flight A1 is moved to at-least slot 4.

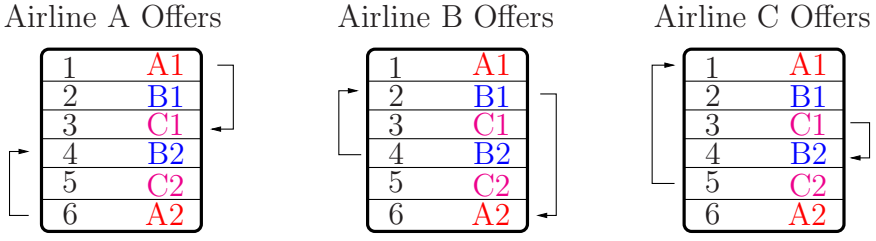


Figure 3-7: Example Illustration of AMAL offers. Using the terminology of set $\mathcal{O}$, the proposed trade offers are (A1, 3; A2, 4), (B1, 6; B2, 2) and (C1, 4; C2, 1).

Figure 3-8 shows the directed graph network representation of the example de-

scribed by Figure 3-7. Any feasible set of trades corresponds to directed cycles in the network. Although, the opposite is not true, because every cycle does not necessarily conform to the trade restrictions (e.g., the cycle {(2, 4, 2)} does not correspond to feasible trades). Some examples of cycles that correspond to sets of possible trades resulting in overall feasibility in this graph are $\{(2,5,2),(3,4,3)\}$ and $\{(1,2,6,5,1)\}$.

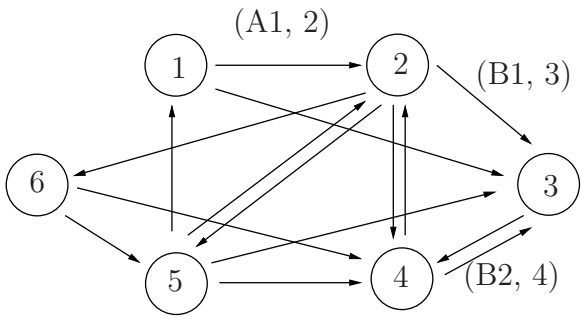


Figure 3-8: Representation of trade offers on a directed graph. The set of nodes correspond to the six slots, whereas the arcs represent various possible flight plans. For example, arc (A1, 2) denotes the assignment of slot 2 to flight A1 (from the currently allotted slot 1).

Starting from the initial RBS allottment shown on the left in Figure 3-9, the right side shows the final arrival sequence resulting from the execution of a feasible set of trades. The new sequence corresponds to the cycle {(1, 2, 6, 5, 1)} in the directed graph representation.

Initial Assignment

1	A1
2	B1
3	C1
4	B2
5	C2
6	A2

After Trades

1	C2
2	A1
3	B2
4	C1
5	A2
6	B1

Figure 3-9: One of the final assignments after the execution of a feasible combination of trades.

There are two insightful observations that emanate from the preceding example. First, the benefits of inter-airline exchanges is immediately apparent. This is because in the current permissible CDM procedures, only intra-airline swapping of flights is allowed. But, it is possible that Airline A is not willing to swap flights A1 and A2, but still wants to (somehow) decrease the delay of flight A2. This scenario is at least made probable with the inter-airline exchanges. Second, the final sequence of flight arrivals is not a simple consequence of a two-for-two trade between two airlines. In fact, it

is a more complex transaction involving all three airlines interacting simultaneously amongst themselves. To elaborate, Airline A's slot (held initially by A1) is taken by Airline C (flight C2); Airline B's slot (held initially by B1) is taken by Airline A (flight A1); and finally, Airline C's slot (held initially by C1) is taken by Airline B (flight B2). In this sense, the presence of multiple airlines and AMAL trade offers facilitate more synergistic trade executions.

3.3 Generalization of Vossen-Ball Single-Airport Model

We now present the first model for slot reallocation in a network setting by extending the single-airport model of Vossen and Ball [36]. For each flight ($f_i \in \mathcal{F}$), the offers to trade determine a sequence of classes, which can be represented by the indices of the corresponding time-periods. Specifically, let k_i be the number of unique classes corresponding to all offers for flight f_i. Then, the offers specify a sequence $i_1, ..., i_{k_i}$ of classes, where $t_{i_k} < t_{i_{k'}}$, if $k < k'$ (please see Figure 3-10 for an example illustration). Moreover, within these k_i classes, there exists a class which corresponds to the originally allotted time-slot, so that if no trade is executed, then the flight will be assigned to this time-slot (we will call the execution of such an offer as a default offer). Let D_T contain the classes that correspond to downward moves for each flight and U_T contain the classes that correspond to upward moves for each flight. Finally, for each flight $f_i \in \mathcal{F}$, let p_i denote the index of the class corresponding to the default offer.

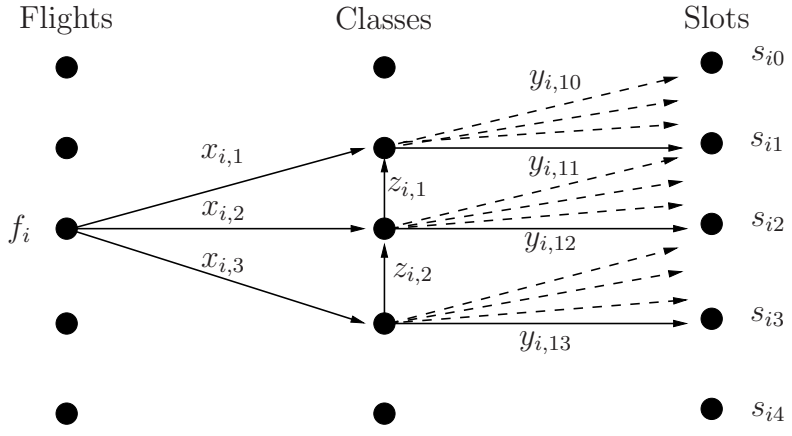


Figure 3-10: Illustration of network flow constraints in the single-airport model. (Taken from [36])

Let us give a numerical example to concretize the above notions of a class and

default offer. Suppose, flight $f_8 \in \mathcal{F}$ occurs in the following set of offers: $(f_8, 6; f_1, 3), (f_8, 7; f_2, 5), (f_8, 9; f_3, 1), (f_4, 10; f_8, 4), (f_5, 12; f_8, 4)$ and $(f_6, 22; f_8, 2)$. Moreover, suppose flight f_8 was assigned time-period 5 from Stage I, then, $k_8 = 6$ and the unique sequence of classes for flight f_8 would be 2, 4, 5, 6, 7, 9. In the above example, the elements $(f_8, 6)$, $(f_8, 7)$ and $(f_8, 9)$ belong to the set D_T, while the elements $(f_8, 4)$ and $(f_8, 2)$ belong to the set U_T. Moreover, $p_8 = 3$.

The Decision Variables

- $x_{ik} \in \{0, 1\} = 1$ if flight f_i is assigned to class i_k.

- $y_{ij} \in \{0, 1\} = 1$ if flight f_i is assigned to time-period t_j.

- $z_{ik} \in \{0, 1\} = 1$ if flight f_i has been assigned to a class with index lower than k in the sequence of classes for f_i but receives at least slot t_{i_k}.

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

Note that the variables y_{ij} are naturally captured in terms of the fundamental variables $w^f_{j,t}$.

- $y_{ij} = w^{f_i}_{\text{dest}_{f_i}, t_j} - w^{f_i}_{\text{dest}_{f_i}, t_j - 1}$.

Constraints

$$(1.1a) - (1.1f).$$

$$\sum_{k=1}^{k_i} x_{ik} = 1 \qquad \forall f_i \in \mathcal{F}. \tag{3.1a}$$

$$x_{i1} + z_{i1} = w^{f_i}_{\text{dest}_{f_i},t_{i_1}} - w^{f_i}_{\text{dest}_{f_i},t_{i_0-1}} \qquad \forall f_i \in \mathcal{F}. \tag{3.1b}$$

$$x_{ik} + z_{ik} - z_{i(k-1)} = w^{f_i}_{\text{dest}_{f_i},t_{i_k}} - w^{f_i}_{\text{dest}_{f_i},t_{i_{k-1}}} \qquad \forall f_i \in \mathcal{F},\ 1 < k \leq k_i. \tag{3.1c}$$

$$x_{dd'} = \sum_{(f_d,t_{d'},f_u,t_{u'})\in\mathcal{O}} o_{dd'uu'} \qquad \forall (f_d, t_{d'}) \in D_T. \tag{3.1d}$$

$$x_{uu'} = \sum_{(f_d,t_{d'},f_u,t_{u'})\in\mathcal{O}} o_{dd'uu'} \qquad \forall (f_u, t_{u'}) \in U_T. \tag{3.1e}$$

$$w^{f_i}_{\text{dest}_{f_i},D_{f_i}} - w^{f_i}_{\text{dest}_{f_i},D_{f_i}-1} \geq x_{ip_i} \qquad \forall f_i \in \mathcal{F}. \tag{3.1f}$$

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

$$o_{dd'uu'} \in \{0,1\} \qquad \forall (f_d, t_{d'}; f_u, t_{u'}) \in \mathcal{O}.$$

$$x_{ik} \in \{0,1\} \qquad \forall f_i \in \mathcal{F},\ 1 \leq k \leq k_i.$$

$$z_{ik} \in \{0,1\} \qquad \forall f_i \in \mathcal{F},\ 1 \leq k \leq k_i.$$

Constraint (3.1a) represents the assignment of flights to classes, i.e., it enforces that every flight will be assigned to exactly one class. Constraints (3.1b) and (3.1c) represents the subsequent assignment of classes to time-slots, i.e., once a flight is assigned to a class, this set of constraints assigns the flight to one of the time-slots feasible for this class. In addition to these network flow constraints, the side constraints (3.1d) and (3.1e) ensure that the resulting trades only include offers proposed by the airlines, i.e., if a flight is assigned to a class, it should correspond to the execution of one of the offers corresponding to this class. Constraint ensures that if the default offer gets executed, then, the flight is assigned the slot in the initial assignment. The above set of constraints define the set of feasible trades.

3.4 A Model based on Monotone Variables

The second model we propose introduces only one additional variable per offer above the variables used in the TFMP model of Bertsimas-Stock [13], namely, $w^f_{j,t}$.

The Decision Variables

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

Constraints

$$(1.1a) - (1.1f).$$

$$o_{dd'uu'} \leq w^{f_d}_{\text{dest}_{f_d},t_{d'}} \qquad \forall (f_d, t_{d'}; f_u, t_{u'}) \in \mathcal{O}. \tag{3.2a}$$

$$o_{dd'uu'} \leq w^{f_u}_{\text{dest}_{f_u},t_{u'}} \qquad \forall (f_d, t_{d'}; f_u, t_{u'}) \in \mathcal{O}. \tag{3.2b}$$

$$\sum_{j \in \mathcal{O}^f} o_j \leq 1 \qquad \forall f \in \mathcal{F}. \tag{3.2c}$$

$$w^f_{\text{dest}_f,D_f} - w^f_{\text{dest}_f,D_f-1} \geq 1 - (\sum_{j \in \mathcal{O}^f} o_j) \qquad \forall f \in \mathcal{F}. \tag{3.2d}$$

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

$$o_{dd'uu'} \in \{0,1\} \qquad \forall (f_d, t_{d'}; f_u, t_{u'}) \in \mathcal{O}.$$

Constraints (3.2a) and (3.2b) 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 (3.2c) enforces that for each flight, at most one offer can get executed. Moreover, constraint (3.2d) 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 (D_f).

Objective Function

In the models presented, 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.

$$n_w = \sum_{f \in \mathcal{F}_w, j \in \mathcal{O}^f} o_j,$$

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

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 ($\min \sum_{w \in \mathcal{W}} |n_w - \gamma|$).

3.5 Computational Results

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). To elaborate, 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, flight f_2 will arrive on-time (given the definition of on-time performance).

Table 3.5 reports the results from the two network slot reallocation models when the objective function used is to maximize the total number of executed trades, while Table 3.4 reports the results when the objective function used is to minimize the difference in the number of trades executed across different airlines. We show in Appendix B that the set of feasible binary integer vectors of the two formulations is the same, and hence, as expected, both models give the same optimal solution in all cases. The computational times of the model TFMP-Trading-VB is consistently

smaller than the model TFMP-Trading-BG, although the solution times for both the models are less than a minute.

		TFMP-Trading-VB		**TFMP-Trading-BG**	
Day	**No. of Flights**	$\mathcal{OE}$	**Time** (in sec.)	$\mathcal{OE}$	**Time** (in sec.)
1	5092	283	13	283	22
2	5844	256	14	256	26
3	5780	150	5	150	15
4	4590	278	11	278	19
5	5128	194	9	194	20
6	4781	153	6	153	13

Table 3.3: Computational performance of the two trading models - Objective Function 1 (maximize the total number of executed trades).

		TFMP-Trading-VB		**TFMP-Trading-BG**	
Day	**No. of Flights**	$\mathcal{OE}$	**Time** (in sec.)	$\mathcal{OE}$	**Time** (in sec.)
1	5092	190	20	190	29
2	5844	180	20	180	31
3	5780	140	8	140	19
4	4590	215	19	215	27
5	5128	140	12	140	22
6	4781	90	9	90	16

Table 3.4: Computational performance of the two trading models - Objective Function 2 (minimize the deviation in the number of trades executed for an airline from the mean).

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

In this section, we contrast the performance of TFMP-Trading-BG 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. A local offer is one where both the flights involved have the same destination airport. On the contrary, a network offer contains flights whose destination airports are distinct. Hence, the single-airport model will not have any executed trades that correspond to the network offers. Table 3.6 reports the results from the two versions under the two objectives described before. The number reported under TFMP-Trading-BG is the number of offers executed from the network model proposed in this chapter, whereas, SA-Trading reports the results obtained from TFMP-Trading-BG 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-BG emphasizing that network connectivities are indeed relevant, whereas, for higher fraction of network offers, TFMP-Trading-BG performs better reinforcing the utility of network offers.

		TFMP-Trading-BG			
		Objective 1		Objective 2	
Day	**No. of Flights**	**Offers Executed**	**Sol. Time** (in sec.)	**Offers Executed**	**Sol. Time** (in sec.)
14 Jul'04	5092	283	13	190	29
4 Aug'04	5844	256	14	180	31
13 May'05	5780	150	5	140	19
16 Jul'05	4590	278	11	215	27
27 Jul'05	5128	194	9	140	22
27 Jul'06	4781	153	6	90	16

Table 3.5: Computational performance of TFMP-Trading-BG.

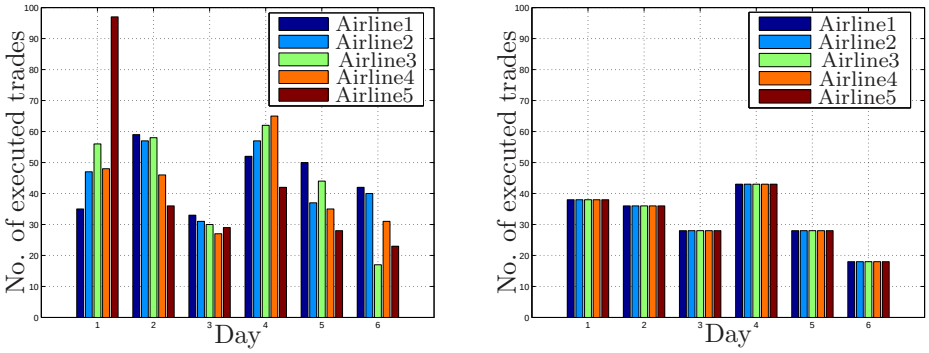


Figure 3-11: Distribution of number of executed trades across airlines from TFMP-Trading-BG. Left: (*Objective 1*); Right: (*Objective 2*)

		No. of Offers Executed			
% Local	**% Network**	Objective 1		Objective 2	
Offers	**Offers**	**SA**	**NW**	**SA**	**NW**
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 3.6: Comparison of TFMP-Trading-BG between single-airport and network-wide settings. SA denotes the results from SA-TRADING and NW denotes the results from TFMP-TRADING.

3.6 Integration and Comparison with Current CDM Practice

Given that our aspiration in the first part of the thesis (extending CDM from an airport to an airspace setting) 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.

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.

3.7 Conclusions

In this chapter, we propose models for slot reallocation in a network setting. The key advantages of the models were the simplicity of the data input requirement from the airlines and the fast running times which are attractive for potential deployment in real-time settings. In the numerical results, 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 (the objective function in our case study).

[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

Chapter 4

Addressing Capacity Uncertainty

As alluded to in Chapter 1, weather accounts for the majority of the total air traffic delays caused due to terminal, en-route congestion and several other operational factors. In this chapter, we study the first application of robust and adaptive optimization in the Air Traffic Flow Management (ATFM) problem to address capacity uncertainty. Finally, we report extensive empirical results from the proposed models on real-world flight schedules augmented with simulated weather fronts that illuminate the merits of our proposal.

4.1 Motivation

To assess the impact of weather on the total aviation delays, we consider the OPSNET[1] delays data for the year 2010. As evidenced in the monthly delays plot in Figure 4-1, there is a significant spike in the delays for the summer months (May-July), when there is pronounced convective weather activity. Moreover, Figure 4-2 indicates that approximately 65-75% of total delays is attributable to weather in the last ten years. These two observations highlight the importance of addressing weather induced capacity uncertainty for mitigating aviation delays.

During GDP planning in current practice, AARs are determined based on forecasted capacity estimates. But, these forecasts are rarely achieved in practice as it is difficult to predict exact airport operating conditions several hours in advance. An overly optimistic estimate would result in air-borne delays for flights close on arrival which is much more undesirable than ground-holds at the departing airport.

[1]The Operations Network (OPSNET) is one of the official sources of National Airspace (NAS) air traffic operations and delay data.

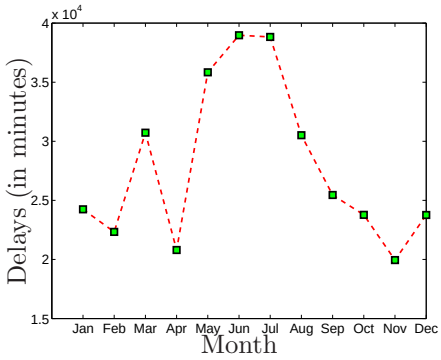


Figure 4-1: OPSNET monthwise delays for 2010.

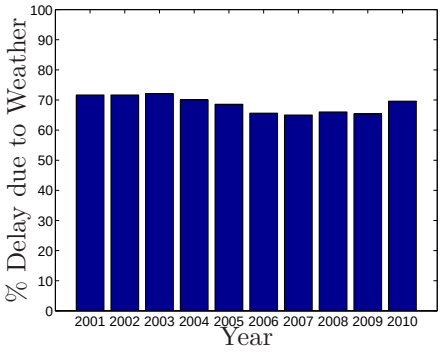


Figure 4-2: Delays attributable to weather during 2001-2010.

In contrast, a pessimistic estimate results in avoidable ground delays for some flights and the subsequent propagated delays to connecting flights. To partly mitigate the impact of this uncertainty, FAA exempts long-distance flights from the set of flights considered in the GDP. This is done by defining a geographical region around the destination airport (where the GDP is being implemented) and limiting the scope of this initiative to flights within this area. This modification still suffers from the pitfall that in case of low realized capacity, the flights within the locally defined area will get much larger delays as priority is given to the long-haul flights to use the limited slots. Thus, it should be evident that a optimization-based dynamic approach might overcome these challenges for the GDP setting.

Even though convective weather en-route and reduced visibility in terminal areas cause significant disruption, there is a lack of network-wide ATFM optimization tools that address the stochastic nature of the airspace capacity. In fact, the existing academic models for network ATFM assume a deterministic estimate on the available capacities at airports and sectors (airspace elements henceforth). Not accounting for capacity uncertainty may lead to suboptimal and possibly infeasible solutions. This state of affairs invites a new mathematical approach that incorporates the uncertainty inherent in the estimates of the airspace resources to come up with a robust schedule.

Background

There are primarily two approaches in the literature to address decision-making under uncertainty, namely, i) *Stochastic Programming*; and ii) *Robust Optimization*. Dantzig [19] proposed the approach of stochastic programming which entails generating scenarios for uncertain data with appropriate probabilities. Unfortunately, this

approach suffers significantly from the practical difficulty of not knowing the exact distribution of the data to generate relevant scenarios. Furthermore, it generally becomes intractable quickly as the number of scenarios increases, thereby posing substantial computational challenges. In the last two decades, an alternative approach by the name of robust optimization has been studied to overcome these challenges (see Bertsimas et al. [16] and the recent book by Ben Tal et al. [2] for extensive literature review and references). The key idea of robust optimization is to construct appropriate uncertainty sets for the uncertain parameters that captures the probabilistic properties of the problem. The goal subsequently is to construct a solution that is feasible for all outcomes of the uncertainty set and which optimizes the worst-case objective. The key advantage of robust optimization is that it presents a tractable framework to model optimization problems under uncertainty. Specifically, the robust counterpart of a linear optimization problem (LOP) is still a LOP [16].

Although, robust optimization has been successful in tractably solving many classes of optimization problems with uncertainty, it may suffer from a pitfall of the possibility of conservative solutions. This is a consequence of the optimization over the worst-case realization of the uncertain parameters. This drawback is further aggravated in multi-stage problems as robust optimization produces a single (static) solution. Consequently, there is an alternative paradigm for multi-period decision-making called *adaptive optimization* wherein decisions are adapted to capture the progressive information revealed over time. There are two classes of models within adaptive optimization: i) *policy-based full adaptability*; and ii) *finite adaptability*. The former class (also referred to as adjustable robust policies) concerns itself with solutions which depend on realizations of the uncertainty in past stages and optimizes the worst-case objective (see Ben Tal et al. [2] for extensive details). In contrast, finite adaptability (introduced in Bertsimas and Caramanis [9]) takes the middle ground between fully-adaptable and the robust (single static) solution and uses a finite number of possible outcomes for each stage a priori such that at least one outcome is feasible for each stage. Within the fully-adaptable framework, the most extensively studied class of policies is *affine*[2] or *linear decision rules* (LDRs). The success of affine policies is due to its computational tractability and strong empirical evidence reported in a variety of application settings. This is further complemented by recent theoretical results, for instance, Bertsimas et al. [17] show that affine policies are optimal for two-stage adaptive optimization problem for simplex uncertainty sets.

[2]Affine policies have been studied extensively in control theory (please refer to survey by Bemporad et al. [1] for details).

Assuming we have a methodology to deal with the two issues (corresponding modules) outlined above, we now describe our proposal for solving the multi-period uncertain ATFM problem. Figure 4-3 plots a block diagram of our macro proposal to incorporate the two modules into actual operations. Each time an ATFM initiative is sought, the multitude of FAA weather technologies are called upon to provide information to a central coordinator to model the capacity uncertainties under Module I. Subsequently, this object is passed on to Module II where in conjunction with the current flight information and trajectories (as extracted from Flight Schedule Monitor[3]), an optimized schedule is realized. This interplay between the two modules continues every time a new loop of optimization is sought to exploit the progressive information revealed over time.

4.2 Our Proposal

Our overall strategy to address capacity uncertainty in the network-wide ATFM problem under the robust and adaptive paradigm consists of two inter-related issues:

- **Issue I**: *Model of weather-induced uncertain capacity.* We propose a low-dimensional uncertainty model (viz., which can be constructed by few underlying parameters) that intuitively captures the dynamics of moving weather fronts. Normally, the most severe disruptions occur due to the presence of an ongoing weather front which traverses through some part of the NAS. We believe that the uncertainty governing the movement of this weather front is dictated by the uncertainty in the time of arrival, the duration of impact and the reduction in capacity commensurate with the intensity of the front. The overall uncertainty model is the result of the impact of a small number of weather fronts moving across different parts of the NAS. We refer to this model as a *weather-front based approach*.

- **Issue II**: *Tractable solution methodologies for the robust and adaptive ATFM problem.* Given the model for capacity uncertainty, we invoke the recent advances in the theory of robust and adaptive optimization as surveyed earlier to solve the uncertain ATFM problem. Specifically, we prove the equivalence of the robust problem to a modified instance of the deterministic problem and solve the linear relaxation of the adaptive problem using affine policies.

[3]Flight Schedule Monitor (FSM) is an important FAA tool that creates a common situational awareness among all users and service providers in the National Airspace System.

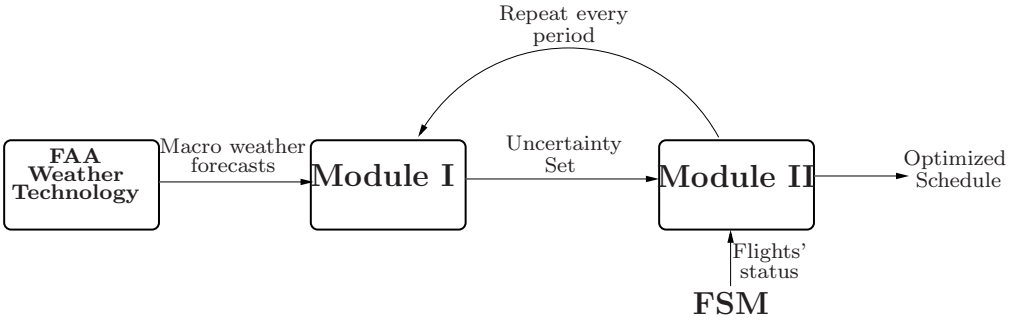


Figure 4-3: Illustration of our overall proposal for solving multi-period ATFM problem addressing capacity uncertainty.

Preliminaries. In this chapter, *vertical vector concatenation* is denoted by the comma (,) operator, e.g., $\mathbf{u} = (u_1, \ldots, u_n) \in \mathbb{R}^n$ and $\mathbf{v} = (v_1, \ldots, v_n) \in \mathbb{R}^n$, then $(\mathbf{u}, \mathbf{v}) \triangleq (u_1, \ldots, u_n, v_1, \ldots, v_n)$. Vectors indexed by a particular time-period are quite pervasive. Therefore, we refer to vectors specific to time-period t in two ways: i) including the index in parenthesis, e.g., $\mathbf{w}(t)$, or ii) as a subscript, e.g., $\mathbf{w}_t$. To refer to the i-th component of a vector at time t, we always use the parenthesis notation for time, e.g., $\mathbf{w}_i(t)$. Moreover, we also work extensively with quantities which depend on the entire history of available information at a given time t. Therefore, we define, for any time-varying vector quantity $\{\mathbf{b}_t \in \mathbb{R}^n\}_{t=1,\ldots,T}$, the following *stacked vector* $\mathbf{b}_{[t]} \triangleq (\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_{t-1}) \in \mathbb{R}^{n \times (t-1)}$, which represents measurements available at the beginning of period t. We use $\bigoplus$ to denote *polyhedron concatenation*. More precisely, for two polyhedron P_1 and P_2, $P_1 \bigoplus P_2 = \{(x, y) |\, x \in P_1, y \in P_2\}$.

4.3 Model of Weather-front Induced Capacity Uncertainty

In this section, we model capacity uncertainty by considering the impact of a small number of weather fronts moving across the airspace. An important advantage of this approach is that a few parameters can be used to capture the dynamics of a weather-front; thereby, leading to a low-dimensional description of the uncertainty set.

To motivate our approach, consider a day in which it is expected with high conviction that a storm would start during a certain time-period of the day (say, 4 to 6 pm). But, within this two hour interval, it is difficult to predict the exact time of arrival of storm, and thus it seems prudent to consider uncertainty in the time

of arrival. Furthermore, the resulting drop in capacity over the planning period is not known exactly. An overly pessimistic estimate might result in loss of capacity, whereas, an optimistic estimate might impact the subsequent connections if the actual reduction is worse. This suggests that there is uncertainty associated with the intensity of weather disruption. Finally, the last time of the impact of the storm (or alternatively, the duration of the storm) is not known precisely and should be captured in the uncertainty set. We believe, that a reasonable set of parameters which correctly model the dynamics of a weather front are the time of arrival, the duration and the capacity reduction.

Exact description for single airspace element. Following this discussion, we next describe the uncertainty set for a single airspace element. The key parameters of the weather-front affecting an airspace element are:

1. Time of arrival: $T_a \in \{\underline{T}_a, \ldots, \overline{T}_a\}$.
2. Duration: $d \in \{\underline{d}, \ldots, \overline{d}\}$.
3. Reduction in capacity: $\alpha \in \{\underline{\alpha}, \ldots, \overline{\alpha}\}$.

Subsequently, we use T_b to denote $T_a + d$ (time of revival of the weather-front), which implies that given the bounds on T_a and d, we have $T_b \in \{\underline{T}_a + \underline{d}, \ldots, \overline{T}_a + \overline{d}\}$.
For a particular realization of the three parameters, T_a, d and α, the capacity vector $\mathbf{b}$ can be written as follows:

$$\mathbf{b} = (C, \ldots, C, \alpha C, \ldots, \alpha C, C, \ldots, C),$$

that is (see also Figure 4-4),

$$b_k = \begin{cases} \alpha C, & k \in \{T_a, \ldots, T_b\}, \\ C, & k \in \mathcal{T} \setminus \{T_a, \ldots, T_b\}. \end{cases}$$

Assumption 1. We assume that the value of α is such that αC is an integer. The reason for this assumption is that the integrality proofs presented in Section 4.4 simplify as a result of this assumption.

We now give a concrete example of the uncertainty set we are trying to model.

Example 4.3.1. Suppose, we have a time-horizon of 5 periods, i.e, $\mathcal{T} = 5$. Moreover, let $T_a \in \{3, 4\}$, $d \in \{1, 2\}$, $\alpha \in \{0.6, 0.8\}$ and $C = 30$. Then, $\mathcal{U} = \big\{(30, 30, 18, 30, 30),$

Figure 4-4: Depiction of the capacity profile under a *weather-front* based uncertainty set for a single affected airspace element. The plot is for a particular realization of the parameters (T_a, d, α).

$(30, 30, 18, 18, 30), (30, 30, 30, 18, 30), (30, 30, 30, 18, 18), (30, 30, 24, 30, 30),$
$(30, 30, 24, 24, 30), (30, 30, 30, 24, 30), (30, 30, 30, 24, 24)\}$.

To concretely motivate the appropriateness of our proposal for modeling the capacity uncertainty, we present practical evidence of the applicability of fitting step functions to actual capacity profiles. Figure 4-5 plots the capacity profiles for two sectors from data obtained from Lincoln Labs. It should be evident from the plot that fitting these profiles by step functions is an appropriate approximation. A similar trend is true across many other sectors.

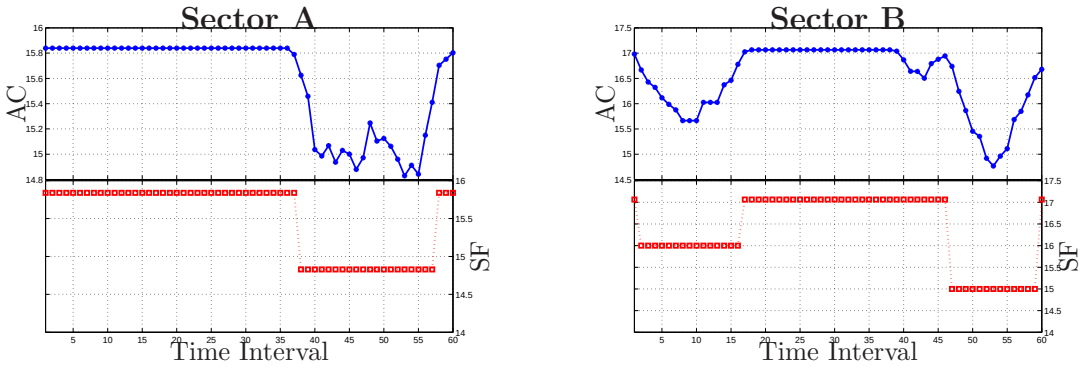


Figure 4-5: Illustration of the applicability of step functions to model capacity profiles. **AC** denotes actual sector capacity and **SF** denotes the step function capacity.

Remark 3. Although, we use the parameter set (T_a, d, α) to construct the uncertainty set, in some regions of the airspace, the parameter T_a may not be uncertain. For instance, in the SFO (San Fransisco) bay area, it is known with reasonable certainty, that weather disruptions due to fog will start daily from 5 am. Therefore, in this case, the uncertainty is only governed by the duration d and capacity reduction α.

Extending to an airspace setting. We now extend the uncertainty set proposed for a single airspace element to multiple airspace elements. We define a weather front as

an entity which is independent of all weather disturbances from other fronts. Suppose we have k weather fronts (denoted by $\mathcal{W}_1, \mathcal{W}_2, \ldots, \mathcal{W}_k$) over the course of a day. On a typical day, k would take a small value (say $k = 3, 4$). We model the traversal of a weather front through the day by dividing it into a number of *phases*, where a phase transition materializes when the spatial or temporal composition of the front changes substantially. This approach also leverages air controllers' human expertize and gives significant flexibility to the modeler in capturing the subtle details of the entire front profile. Each weather front $\mathcal{W}_i$ is further decomposed into phases: $(\mathcal{W}_i^1, \ldots, \mathcal{W}_i^{p_i})$, where p_i is the number of phases of the i^{th} weather front. $\mathcal{W}_i^j$ captures the snapshot of the weather front during phase j. In particular, the following details are expected to be known during each phase:

- $\mathcal{W}_i^j(\mathcal{S})$: set of sectors impacted by the i^{th} weather front during phase j;
- $\mathcal{W}_i^j(\mathcal{K})$: set of airports impacted by the i^{th} weather front during phase j.

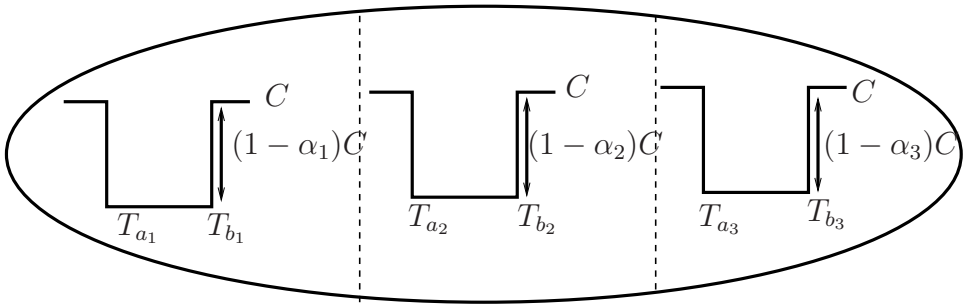


Figure 4-6: Traversal of a *weather-front* across the NAS. It has three phases over the course of its existence.

Model of weather front propagation. A key distinction between the modeling approach for a single airspace element and multiple airspace elements is the augmented set of parameters governing the uncertainty set for the latter case. Specifically, we introduce another parameter L which captures the lag between the front's time of arrival across consecutive phases. During phase I, the parameters (T_{a_1}, d_1, α_1) govern the capacity realizations for $\mathcal{W}_i^1$. Subsequently, the lag (denoted by L_1) dictates the set of parameters (T_{a_2}, d_2, α_2) of the weather front in phase II, and so on so forth. This way we model the strong correlation of T_{a_2} with T_{a_1}. More generally,

$$T_{a_k} = T_{a_{k-1}} + L_{k-1} = T_{a_1} + \sum_{j=1}^{k-1} L_j$$

We model the uncertainty inherent in L_j in an analogous way as T_a, i.e., $L_j \in \{\underline{L}_j, \ldots, \overline{L}_j\}$. Subsequently, the range of possible values for T_{a_k} is given by $\{\underline{T}_{a_{k-1}} + \underline{L}_{k-1}, \ldots, \overline{T}_{a_{k-1}} + \overline{L}_{k-1}\}$. Note that, even though we use the additional parameter L_j for front propagation, the eventual description of the uncertainty set is described completely by (T_a, d, α). But, we emphasize that modeling the propagation with our proposed way (i.e., using L_j) is quite intuitive and also facilitates developing the uncertainty intervals in which the parameters lie.

Models

We now formalize the setting in which we intend to solve the robust and adaptive ATFM problem. Using the discretization $\mathcal{T} = \{1, \ldots, T\}$ introduced in the formulation of TFMP (the deterministic ATFM problem), the right-hand side capacity vector $\mathbf{b}$ can be decomposed into time indexed sub-vectors, i.e., it can be written as:

$$\mathbf{b} = \big(\mathbf{b}_1, \ldots, \mathbf{b}_T\big),$$

where $\mathbf{b}_t$ has the following description:

$$\mathbf{b}_t = \big(\mathbf{b}_t^{\mathcal{A}}, \mathbf{b}_t^{\mathcal{D}}, \mathbf{b}_t^{\mathcal{S}}\big).$$

$\mathbf{b}_t^{\mathcal{A}}, \mathbf{b}_t^{\mathcal{D}}, \mathbf{b}_t^{\mathcal{S}}$ correspond to the components of $\mathbf{b}_t$ pertaining to the arrival capacity at airports, departure capacity at airports and sector capacities, respectively.

- $\mathbf{b}_t^{\mathcal{A}} = \Big(A_1(t), \ldots, A_{|\mathcal{K}|}(t)\Big)$; $\mathbf{b}_t^{\mathcal{D}} = \Big(D_1(t), \ldots, D_{|\mathcal{K}|}(t)\Big)$;
- $\mathbf{b}_t^{\mathcal{S}} = \Big(S_1(t), \ldots, S_{|\mathcal{S}|}(t)\Big)$;

To make the notation clearer, Figure 4-7 depicts the stacked capacity vector ($\mathbf{b}_{[t]}$) and the recourse variables ($w(\mathbf{b})$):

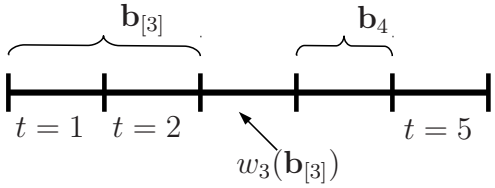


Figure 4-7: Explanation of some of the notation ($\mathcal{T} = 5$).

To make the dependence of time on $\mathbf{b}$ explicit, we define the following sub-vector of $\mathbf{b}^t$ indexed by time t:

$$\mathbf{b}^t = \big(A_1(t), \ldots, A_K(t)\big) \cup \big\{D_k(t) \mid k \in \mathcal{K}\big\} \cup \big\{S_j(t) \mid j \in \mathcal{S}\big\}$$

Before formulating the adaptive problem, we introduce the notation specific to recourse variables. We use $w_i(\mathbf{b}_{[i]})$ to denote the decision variables during period i which take a functional form of the capacity $\mathbf{b}_{[i]}$ (i.e., capacities during time-periods $1, \ldots, i-1$).

Multi-stage models. We first present the multi-stage adaptive and robust optimization models (over T periods):

$$\begin{aligned} IZ_{\text{Adapt}} = \min_{\mathbf{w}_1, \mathbf{w}_i(\mathbf{b}_{[i]})} \; & \Bigg[\mathbf{c}_1'\mathbf{w}_1 + \max_{\mathbf{b}\in\mathcal{U}} \; \Big[\mathbf{c}_2'\mathbf{w}_2(\mathbf{b}_{[2]}) + \cdots + \\ & \qquad \max_{\mathbf{b}\in\mathcal{U}} \; \big[\mathbf{c}_{T-1}'\mathbf{w}_{T-1}(\mathbf{b}_{[T-1]}) + \max_{\mathbf{b}\in\mathcal{U}} \; \mathbf{c}_T'\mathbf{w}_T(\mathbf{b}_{[T]})\big]\Big]\Bigg] \\ & \text{s.t. } \mathbf{A}_1\mathbf{w}_1 + \sum_{i=2}^{T} \mathbf{A}_i\mathbf{w}_i(\mathbf{b}_{[i]}) \leq \mathbf{b}, \;\; \forall \mathbf{b} \in \mathcal{U}, \qquad (\Pi^T_{\text{Adapt}}) \\ & \mathbf{w}_i(\mathbf{b}_{[i]}) \in \{0,1\}^{n_i}. \end{aligned}$$

In the problem Π^T_{Adapt}, the uncertainty only affects the right-hand side capacity vector $\mathbf{b}$. The constraint matrices $\mathbf{A}_i$ and the cost vectors $\mathbf{c}_i$ are assumed certain. The only restriction on the uncertainty set $\mathcal{U}$ is that $\mathcal{U} \subseteq \mathbb{R}^n_+$ or $\mathcal{U} \subseteq \mathbb{Z}^n_+$ (this corresponds to the requirement of non-negative, possibly integer capacities). The decisions $\mathbf{w}_i$ in every time-period i depend on the realizations of the uncertain parameters during the earlier periods (viz. $\mathbf{b}_{[i]}$). This requirement is known as *non-anticipativity*, and is a manifestation of the fact that while making decisions during a particular stage, the uncertainty realizations of later periods is not known and therefore, cannot be used. Finally, n_i is the number of decision variables in period i. The equivalent robust

version of the multi-stage adaptive problem is as follows:

$$
\begin{aligned}
IZ_{\text{Rob}} = \min_{\mathbf{w}} \ & \sum_{i=1}^{T} \mathbf{c}_i \mathbf{w}_i \\
\text{s.t.} \ & \sum_{i=1}^{T} \mathbf{A}_i \mathbf{w}_i \leq \mathbf{b}, \ \ \forall \mathbf{b} \in \mathcal{U}, \qquad (\Pi_{\text{Rob}}) \\
& \mathbf{w}_i \in \{0,1\}^{n_i}.
\end{aligned}
$$

Let Z_{Adapt} and Z_{Rob} denote the objective functions for the linear relaxation of Π^T_{Adapt} and Π_{Rob} respectively. Note that any feasible solution of Π_{Rob} is also feasible for Π^T_{Adapt}. Hence, $IZ_{\text{Adapt}} \leq IZ_{\text{Rob}}$ and this re-emphasizes the benefit of adaptability for mitigating the possibility of highly conservative robust solutions.

Remark 4. Although, we include the requirement of *non-anticipativity* in the most general form in the formulations above (i.e., the decisions $\mathbf{w}_i$ in every time-period i depend on the entire history $\mathbf{b}_{[i]}$), it is important to remark that in our application setting, this is a rather strong requirement. More precisely, since weather propagation is typically a Markovian phenomenon, it suffices to have $\mathbf{w}_i$ depend only on $\mathbf{b}_{i-1}$ (viz., the capacity realizations in the immediately previous time-period) as opposed to the entire history $\mathbf{b}_{[i]}$. This has an added advantage of easing the required computational effort as the optimal recourse policy requires the knowledge of fewer coefficients.

Two-stage models. We now give a special case of the multi-stage model with $T = 2$. We divide the variables $\mathbf{w}$ into two parts $\mathbf{w} := \left(\begin{smallmatrix}\mathbf{u}\\ \mathbf{v}\end{smallmatrix}\right)$.

$$
\begin{aligned}
IZ_{\text{Adapt}} = \min_{\mathbf{u}, \mathbf{v}(\mathbf{b})} \ & \mathbf{c}'\mathbf{u} + \max_{\mathbf{b} \in \mathcal{U}} \mathbf{d}'\mathbf{v}(\mathbf{b}) \\
\text{s.t.} \ & \mathbf{A}\mathbf{u} + \mathbf{B}\mathbf{v}(\mathbf{b}) \leq \mathbf{b}, \ \ \forall \mathbf{b} \in \mathcal{U}, \qquad (\Pi^2_{\text{Adapt}}) \\
& \mathbf{u} \in \{0,1\}^{n_1}, \mathbf{v}(\mathbf{b}) \in \{0,1\}^{n-n_1}.
\end{aligned}
$$

Note that the problem Π^2_{Adapt} is also referred to as an adjustable robust problem in the literature (see Ben Tal et al. [2]). n_1 corresponds to the number of variables in

$\mathbf{u}$. The equivalent robust version of the two-stage adaptive problem is as follows:

$$\begin{aligned} IZ_{\text{Rob}} = \min_{\mathbf{u},\mathbf{v}} \ & \mathbf{c}'\mathbf{u} + \mathbf{d}'\mathbf{v} \\ \text{s.t. } & \mathbf{A}\mathbf{u} + \mathbf{B}\mathbf{v} \leq \mathbf{b}, \ \ \forall \mathbf{b} \in \mathcal{U}, \\ & \mathbf{u} \in \{0,1\}^{n_1}, \mathbf{v} \in \{0,1\}^{n-n_1}. \end{aligned} \tag{Π_{Rob}}$$

Before proceeding further, it would be beneficial to present a qualitative picture of the relative merits of the robust and adaptive solutions. Figure 4-8 depicts a scenario comprising of a flight traversing through a reduced capacity segment of the airspace. Assume that the deterministic capacity estimate is non-zero but the set of possible capacity scenarios include the zero capacity realization. As a result, the deterministic route of the flight goes directly through this affected air-segment. In contrast, the robust route corresponds to ignoring the segment altogether as the worst-case scenario corresponds to 0 capacity and the robust route needs to be feasible (for all scenarios). These two extreme possibilities are bridged by the adaptive solution which consists of multiple paths depending on the exact realization of the uncertain capacity (in fact, the path might traverse through the segment or miss it depending on the actual capacity reduction).

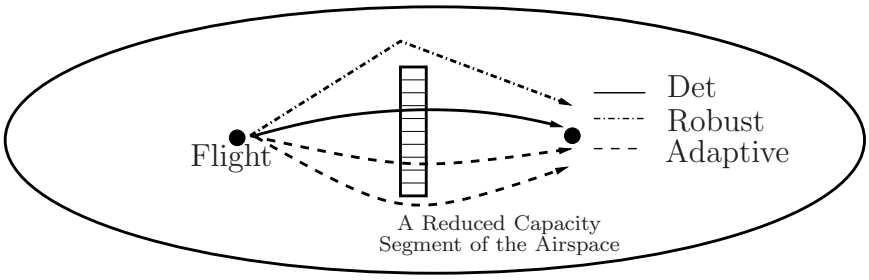


Figure 4-8: Illustration of the characteristics of the deterministic, robust and adaptive routes of a flight.

4.4 Characterization of Weather-front Induced Uncertainty Set

We now proceed towards incorporating the discrete uncertainty set introduced in Section 4.3 within a mathematical programming framework to solve the robust and adaptive ATFM problems.

Transitioning from $\mathcal{U}$ to $\text{conv}(\mathcal{U})$

Consider the following general optimization problem (OptDU denotes the optimization problem under discrete uncertainty set $\mathcal{U}_0$):

$$\begin{aligned} Z_{\text{OptDU}} = \min_{\mathbf{x}} \quad & \mathbf{c}'\mathbf{x} \\ \text{s.t.} \quad & \mathbf{A}\mathbf{x} \leq \mathbf{b}, \ \ \forall \mathbf{b} \in \mathcal{U}_\mathbf{0}. \end{aligned} \tag{4.1}$$

Although, the robust equivalent of OptDU is still an LP (with the right-hand side capacity $\mathbf{b}$ replaced by the component-wise minimum vector), the solution methodology for the adaptive problem is not straightforward. We need an alternative tractable line of attack for the case of discrete uncertainty sets. A standard approach to overcome this hurdle is to work with its convex hull to enable a polyhedral description. We formalize below the equivalence of solving over a discrete set and over the convex hull of this set (please refer to Appendix C.1 for the proof).

Proposition 1. *Let $\mathcal{U}_0$ denote a discrete set. Then, Z_{OptDU} can be calculated as:*

$$\begin{aligned} Z_{\text{OptDU}} = \min_{\mathbf{x}} \quad & \mathbf{c}'\mathbf{x} \\ \text{s.t.} \quad & \mathbf{A}\mathbf{x} \leq \mathbf{b}, \ \ \forall \mathbf{b} \in \text{conv}(\mathcal{U}_\mathbf{0}). \end{aligned} \tag{4.2}$$

where $\text{conv}(\mathcal{U}_0)$ denotes the convex hull of $\mathcal{U}_0$.

We next characterize $\text{conv}(\mathcal{U})$ for the discrete uncertainty set introduced in Section 4.3.

Polyhedral description of $\text{conv}(\mathcal{U})$

We use the following auxiliary variables:

$$y_t = \begin{cases} 1, & \text{if capacity drops } \textit{by} \text{ time } t, \\ 0, & \text{otherwise.} \end{cases}$$

$$z_t = \begin{cases} 1, & \text{if capacity revives } \textit{by} \text{ time } t, \\ 0, & \text{otherwise.} \end{cases}$$

The variables y_t and z_t are defined for $t \in \{\underline{T}_a, \ldots, \overline{T}_b\}$. But, given that the capacity drop cannot occur after $\overline{T}_a$ and capacity revival cannot occur before $\underline{T}_b$, we can set a few variables as parameters at the outset:

- $y_t = 1,\ t \in \{\overline{T}_a, \ldots, \overline{T}_b\}$;
- $z_t = 0,\ t \in \{\underline{T}_a, \ldots, \underline{T}_b - 1\};\ \ z_t = 1,\ t \in \{\overline{T}_b\}$

The definition of the auxiliary variables y_t and z_t as "*by*" rather than "*at*" enables a natural polyhedral description of the capacity profile.

We give a mathematical description of the uncertainty set for the capacity profile as depicted in Figure 4-4. We start by providing a description for a particular realization of α:

$$\begin{aligned}\mathcal{U}_\alpha = \big\{\mathbf{b} \in \mathbb{Z}_+^m |\ & b_t = C(1-y_t) + \alpha C y_t + (1-\alpha)Cz_t,\ \forall t \in \{\underline{T}_a, \ldots, \overline{T}_b\};\\ & b_t = C,\ \forall t \in \mathcal{T} \setminus \{\underline{T}_a, \ldots, \overline{T}_b\};\\ & y_t \le y_{t+1};\ z_t \le z_{t+1};\ z_t \le y_t;\\ & y_{\overline{T}_a} = 1;\ z_{\underline{T}_b - 1} = 0;\ z_{\overline{T}_b} = 1;\ y_t, z_t \in \{0,1\}\big\}\end{aligned}$$

The three components $C(1-y_t), \alpha C y_t, (1-\alpha)Cz_t$ comprising b_t have the following semantics:

$$b_t = \underbrace{\underbrace{C(1-y_t)}_{\text{before drop}} + \overbrace{\alpha C y_t}^{\text{before revival}}}_{\text{after drop}} \overbrace{\quad\quad +(1-\alpha)Cz_t}^{\text{after revival}}$$

Let $|\mathcal{U}_\alpha| = K$, where $K = (\overline{T}_a - \underline{T}_a + 1) \times (\overline{d} - \underline{d} + 1)$. The K elements of $\mathcal{U}_\alpha$ are indexed by (i,j) where $i \in \{\underline{T}_a, \ldots, \overline{T}_a\},\ j \in \{i + \underline{d}, \ldots, i + \overline{d}\}$. Formally, $\mathbf{b}^{i,j}$ is defined as follows:

$$b_k^{i,j} = \begin{cases} \alpha C, & k \in \{i, \ldots, j\}, \\ C, & k \in \mathcal{T} \setminus \{i, \ldots, j\}. \end{cases} \tag{4.3}$$

A polyhedral description of conv($\mathcal{U}_\alpha$). Our goal now is to come up with a polyhedral description of conv($\mathcal{U}_\alpha$). We claim that the following polyhedron is precisely

$\text{conv}(\mathcal{U}_\alpha)$:

$$\begin{aligned}\mathcal{P}_\alpha = \big\{\mathbf{b} \in \mathbb{R}^m_+ |\; & b_t = C(1-y_t) + \alpha C y_t + (1-\alpha) C z_t,\ \forall t \in \{\underline{T}_a, \ldots, \overline{T}_b\}; \\ & b_t = C,\ \forall t \in \mathcal{T} \setminus \{\underline{T}_a, \ldots, \overline{T}_b\}; \\ & y_t \le y_{t+1};\ z_t \le z_{t+1};\ z_t \le y_t; \\ & y_{\overline{T}_a} = 1;\ z_{\underline{T}_b - 1} = 0;\ z_{\overline{T}_b} = 1;\ 0 \le y_t, z_t \le 1\big\}\end{aligned}$$

Remark 5. In the description of $\mathcal{U}_\alpha$ and $\mathcal{P}_\alpha$, it is useful to note that if $\overline{T}_a < \underline{T}_b$, then, the constraints $z_t \le y_t$ are redundant and hence, can be removed from the polyhedral description.

Theorem 3. *Suppose Assumption 1 holds. Then, $\mathcal{P}_\alpha$ is integral and is exactly the convex hull of $\mathcal{U}_\alpha$, i.e.,*

$$\mathcal{P}_\alpha = \text{conv}(\mathcal{U}_\alpha) \tag{4.4}$$

Proof. We use a technique to prove integrality of a polyhedron based on randomization known as *Randomized Rounding* (please refer to Chapter 3 of Bertsimas and Weismantel [15] for details). Let $\mathbf{y}^*, \mathbf{z}^*$ be an optimal solution of the following problem:

$$\begin{aligned} Z_{\text{LO}} = \min_{\mathbf{y},\mathbf{z}} \quad & \mathbf{c}'\mathbf{y} + \mathbf{d}'\mathbf{z} \\ \text{s.t.} \quad & (\mathbf{y}, \mathbf{z}) \in \mathcal{X}. \end{aligned} \tag{4.5}$$

$$\begin{aligned}\text{where } \mathcal{X} = \big\{\mathbf{y}, \mathbf{z} |\; & y_t \le y_{t+1},\ \ z_t \le z_{t+1},\ \ \forall t \in \{\underline{T}_a, \ldots, \overline{T}_b - 1\}; \\ & z_t \le y_t;\ \ 0 \le y_t, z_t \le 1,\ \ \forall t \in \{\underline{T}_a, \ldots, \overline{T}_b\}\big\}.\end{aligned}$$

Let Z_{IP} denote the optimal value of the optimization problem above with additional integrality constraints on y_t and z_t (i.e., $y_t, z_t \in \{0, 1\}$). From $\mathbf{y}^*, \mathbf{z}^*$, we create a new random integer solution $(\mathbf{y, z})$, that is feasible in $\mathcal{X}$. The randomization we use is as follows: sort the values $\mathbf{y}^*, \mathbf{z}^*$ from smallest to largest in the interval $[0, 1]$. Subsequently, generate a random variable U distributed uniformly in $[0, 1]$. The rounding

is then done as follows:

$$y_t = \begin{cases} 1, & y_t^* \geq U, \\ 0, & y_t^* < U. \end{cases} \qquad z_t = \begin{cases} 1, & z_t^* \geq U, \\ 0, & z_t^* < U. \end{cases}$$

Note that the solution produced is clearly feasible because $y_t \leq y_{t+1}$ and $z_t \leq z_{t+1}$ is trivially satisfied. Moreover, $z_t \leq y_t$ because $z_t^* \leq y_t^*$ (the optimal solution is feasible), and after rounding, z_t can never become 1 unless y_t becomes 1 (please see Figure 4-9 for an easy visualization).

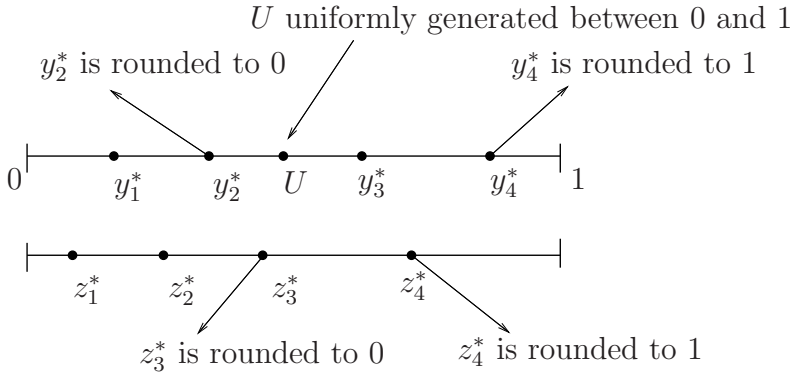


Figure 4-9: Illustration of the geometry of the *randomized rounding* algorithm for proving the integrality of polyhedron $\mathcal{X}$. $\mathbf{y}^*, \mathbf{z}^*$ satisfy $y_t^* \leq y_{t+1}^*$, $z_t^* \leq z_{t+1}^*$ and $z_t^* \leq y_t^*$.

Let Z_{H} be the value of the solution produced. The expected value of the solution is:

$$\begin{aligned} \mathbb{E}[Z_{\mathrm{H}}] &= \sum_{t \in \{\underline{T}_a, \ldots, \overline{T}_b\}} c_t \mathbb{P}(y_t = 1) + d_t \mathbb{P}(z_t = 1) \\ &= \sum_{t \in \{\underline{T}_a, \ldots, \overline{T}_b\}} c_t \mathbb{P}(y_t^* \leq U) + d_t \mathbb{P}(z_t^* \leq U) \\ &= \sum_{t \in \{\underline{T}_a, \ldots, \overline{T}_b\}} c_t y_t^* + d_t z_t^* \\ &= Z_{\mathrm{LO}}, \end{aligned}$$

and thus $Z_{\mathrm{IP}} = Z_{\mathrm{LO}}$. Since, $\mathbf{c}$ and $\mathbf{d}$ are arbitrary, the polyhedron $\mathcal{X}$ is integral. This implies that $\mathcal{P}_\alpha$ is integral. $\square$

The overall uncertainty set for the capacity profile as depicted in Figure 4-4 is as follows:

$$\mathcal{U} = \left(\bigcup_{\alpha \in \{\underline{\alpha}, \ldots, \overline{\alpha}\}} \mathcal{U}_\alpha \right)$$

A polyhedral description of conv($\mathcal{U}$). We show that $\mathcal{P}_{\underline{\alpha}}$ is precisely conv($\mathcal{U}$). For this, we need the following intermediate result:

Theorem 4.

$$\text{conv}(\mathcal{U}) = \text{conv}(\mathcal{U}_{\underline{\alpha}}) \tag{4.6}$$

Proof. Note the following relation,

$$\bigcup_{\alpha \in \{\underline{\alpha}, \ldots, \overline{\alpha}\}} \text{conv}(\mathcal{U}_\alpha) \subseteq \text{conv}\Big(\bigcup_{\alpha \in \{\underline{\alpha}, \ldots, \overline{\alpha}\}} \mathcal{U}_\alpha \Big) = \text{conv}(\mathcal{U})$$

In particular, for $\alpha = \underline{\alpha}$,

$$\text{conv}(\mathcal{U}_{\underline{\alpha}}) \subseteq \text{conv}(\mathcal{U}) \tag{4.7}$$

Next, we show that $\mathcal{U} \subseteq \text{conv}(\mathcal{U}_{\underline{\alpha}})$ by using the equivalence of conv($\mathcal{U}_{\underline{\alpha}}$) and $\mathcal{P}_{\underline{\alpha}}$ from Theorem 3. For a particular value of α, let $\mathbf{b}^{i,j} \in \mathcal{U}_\alpha$. Consider the following[4]:

$$y_k = \begin{cases} 0, & \text{if } k \in \big\{\underline{T}_a, \ldots, \underline{T}_a + i - 1\big\}, \\ \frac{1-\alpha}{1-\underline{\alpha}}, & \text{if } k \in \big\{\underline{T}_a + i, \ldots, \min(\overline{T}_a - 1,\ \underline{T}_a + j - 1)\big\}, \\ 1, & \text{otherwise.} \end{cases}$$

[4]The reason for the complicated indexing is that $y_{\overline{T}_a}$ is set to 1 in the polyhedral description $\mathcal{P}_{\underline{\alpha}}$. Therefore, the definitions of y_k and z_k need to account for whether the capacity drop occurs before or after $\overline{T}_a$. Hence, the need for $\min(\overline{T}_a - 1,\ \underline{T}_a + j - 1)$.

$$z_k = \begin{cases} 0, & \text{if } k \in \big\{\underline{T}_a, \ldots, \min(\overline{T}_a - 1,\, \underline{T}_a + j - 1)\big\}, \\ \frac{\alpha - \underline{\alpha}}{1 - \underline{\alpha}}, & \text{if } k \in \big\{\min(\overline{T}_a,\, \underline{T}_a + j), \ldots, \underline{T}_a + j - 1\big\}, \\ 1, & \text{otherwise.} \end{cases}$$

Using these y_k and z_k, we expand b_k using the definition of an element in $\mathcal{P}_{\underline{\alpha}}$ (note that the definitions of y_k and z_k imply $y_k \leq y_{k+1}$, $z_k \leq z_{k+1}$ and $z_k \leq y_k$):

$$b_k = C(1 - y_k) + \underline{\alpha} C y_k + (1 - \underline{\alpha}) C z_k$$

for all possible combinations of y_k and z_k over the various indices as shown below:

$$b_k = \begin{cases} C + 0 + 0 = C, & \text{if } k \in \big\{\underline{T}_a, \ldots, \underline{T}_a + i - 1\big\}, \\ \big(\frac{\alpha - \underline{\alpha}}{1 - \underline{\alpha}}\big) C + \underline{\alpha}\big(\frac{1 - \alpha}{1 - \underline{\alpha}}\big) C + 0 = \alpha C, & \text{if } k \in \big\{\underline{T}_a + i, \ldots, \\ & \quad \min(\overline{T}_a - 1,\, \underline{T}_a + j - 1)\big\}, \\ 0 + \underline{\alpha} C + (1 - \underline{\alpha})\big(\frac{\alpha - \underline{\alpha}}{1 - \underline{\alpha}}\big) C = \alpha C, & \text{if } k \in \big\{\min(\overline{T}_a,\, \underline{T}_a + j), \ldots, \\ & \quad \underline{T}_a + j - 1\big\}, \\ 0 + \underline{\alpha} C + (1 - \underline{\alpha}) C = C, & \text{if } k \in \big\{\underline{T}_a + j, \ldots, \overline{T}_b\big\}. \end{cases} \tag{4.8}$$

But, (4.8) has exactly the form of (4.3), the element $\mathbf{b}^{i,j} \in \mathcal{U}_\alpha$. Hence, $\mathbf{b}^{i,j} \in \mathcal{P}_{\underline{\alpha}}$. This implies $\mathcal{U}_\alpha \subseteq \mathcal{P}_{\underline{\alpha}}$. Since α is arbitrary, therefore,

$$\mathcal{U} \subseteq \underbrace{\mathcal{P}_{\underline{\alpha}} = \text{conv}(\mathcal{U}_{\underline{\alpha}})}_{\text{Using Theorem 3}}$$

Thus, $\text{conv}(\mathcal{U}_{\underline{\alpha}})$ is a convex set (a polyhedron). This implies that all convex combinations of the set $\mathcal{U}$ belongs to $\text{conv}(\mathcal{U}_{\underline{\alpha}})$. Hence,

$$\text{conv}(\mathcal{U}) \subseteq \text{conv}(\mathcal{U}_{\underline{\alpha}}) \tag{4.9}$$

The proposition subsequently follows from (4.7) and (4.9). □

We now have the equipment to state the main theorem of this section:

Theorem 5.

$$\mathcal{P}_{\underline{\alpha}} = \text{conv}(\mathcal{U}) \tag{4.10}$$

Proof. The proof follows trivially from Theorem 3 and Theorem 4. □

Note that only $\underline{\alpha}$ is needed to characterize conv($\mathcal{U}$). This implies that only two parameters T_a and d govern the uncertainty set for each airspace element. Let $\mathcal{P}^e_{\underline{\alpha}_{ij}}$ denote the uncertain realizations for airspace element e (where airspace element e belongs to the set of airspace elements affected by weather front i during phase j). The overall uncertainty set spanning all airspace elements is given by (please refer to Preliminaries in the Introduction for the meaning of the notation $\bigoplus$):

$$\mathcal{P}^{\text{overall}} = \bigoplus_{\substack{i=1,\ldots,k, j=1,\ldots,p_i, \\ e \in \mathcal{W}_i^j(\mathcal{S}) \cup \mathcal{W}_i^j(\mathcal{K})}} \mathcal{P}^e_{\underline{\alpha}_{ij}}$$

Note that the total number of uncertain parameters in the weather-front based modeling approach is $2 * k * P$, where k is the number of weather fronts and P is the maximum number of phases across all weather fronts.

4.5 Solution Methodologies

In this section, we propose solution approaches for the robust and adaptive versions of the TFMP problem under the uncertainty sets constructed in Section 4.4.

Robust TFMP

We explicitly characterize an equivalent form of the robust TFMP problem as a specific instance of the deterministic TFMP problem for an arbitrary uncertainty set. First, we introduce a definition on the component-wise minimum capacity vector relative to an uncertainty set to facilitate our proofs.

Definition 4.5.1. $\mathbf{b}_{\min} = (\underline{b}_1, \underline{b}_2, \ldots, \underline{b}_m)$, where $\underline{b}_i = \min\{b_i | \, \mathbf{b} \in \mathcal{U}\}$.

Theorem 6. *For an arbitrary uncertainty set $\mathcal{U}$ for the right hand side capacity, the robust problem is equivalent to solving the following modified* TFMP *instance:*

$$\begin{aligned} \min_{\mathbf{w}} \quad & \mathbf{c}'\mathbf{w} \\ \text{s.t.} \quad & \mathbf{A}\mathbf{w} \leq \mathbf{b}_{\min}, \\ & \mathbf{w} \in \{0,1\}^n. \end{aligned} \tag{4.11}$$

Proof. We show the equivalence of the set of feasible solutions of the robust model (denoted by $\mathcal{W}$) with the instance of TFMP where the right-hand side is $\mathbf{b}_{\min}$ (denoted by $\overline{\mathcal{W}}$). Specifically, $\mathcal{W} = \left\{\mathbf{w} \in \{0,1\}^n | \ \mathbf{A}\mathbf{w} \leq \mathbf{b},\ \mathbf{b} \in \mathcal{U}\right\}$ and $\overline{\mathcal{W}} = \left\{\mathbf{w} \in \{0,1\}^n | \ \mathbf{A}\mathbf{w} \leq \mathbf{b}_{\min}\right\}$.

Suppose $\mathbf{w} \in \mathcal{W}$, then for the i^{th} constraint, we have:

$$\begin{aligned} & \mathbf{a}_i'\mathbf{w} \leq b_i,\ \forall \mathbf{b} \in \mathcal{U}, \\ \Rightarrow\ & \mathbf{a}_i'\mathbf{w} \leq \min_{\mathbf{b}\in\mathcal{U}} b_i = \underline{b}_i. \end{aligned} \tag{4.12}$$

Since i is arbitrary, (4.12) further implies the following:

$$\begin{aligned} & \mathbf{a}_i'\mathbf{w} \leq \min_{\mathbf{b}\in\mathcal{U}} b_i,\ i = 1, \ldots, m, \\ \Rightarrow & \mathbf{A}\mathbf{w} \leq \mathbf{b}_{\min}, \\ \Rightarrow & \mathbf{w} \in \overline{\mathcal{W}}. \end{aligned} \tag{4.13}$$

Suppose $\mathbf{w} \in \overline{\mathcal{W}}$, then by definition of $\mathbf{b}_{\min}$:

$$\begin{aligned} & \mathbf{b}_{\min} \leq \mathbf{b},\ \forall \mathbf{b} \in \mathcal{U}, \\ \Rightarrow & \mathbf{A}\mathbf{w} \leq \mathbf{b},\ \forall \mathbf{b} \in \mathcal{U}, \\ \Rightarrow & \mathbf{w} \in \mathcal{W}. \end{aligned} \tag{4.14}$$

From (4.13) and (4.14), we have $\mathcal{W} = \overline{\mathcal{W}}$. The proposition subsequently follows, as the set of feasible solutions of the two problems and the objective function being minimized is exactly the same. □

Since TFMP is efficiently solvable in practice and has strong integrality properties, Theorem 6 suggests that the robust problem is no harder to solve than TFMP. Consequently, it is expected to be computationally efficient and have majority of the solutions from the linear relaxation integral. We report extensive computational results in Section 4.6 which will validate this remark.

Adaptive TFMP

In this section, we utilize affine policies to solve the linear relaxation of the adaptive problem. The adaptive problem with affine recourse enforced a priori becomes as follows:

$$
\begin{aligned}
IZ_{\text{Adapt}} = \min_{\mathbf{w}_1, \mathbf{P}_i, \mathbf{q}_i} \Bigg[& \mathbf{c}_1' \mathbf{w}_1 + \max_{\mathbf{b} \in \mathcal{U}} \Big[\mathbf{c}_2'(\mathbf{P}_2 \mathbf{b} + \mathbf{q}_2) + \cdots + \\
& \max_{\mathbf{b} \in \mathcal{U}} \Big[\mathbf{c}_{T-1}'(\mathbf{P}_{T-1} \mathbf{b} + \mathbf{q}_{T-1}) + \max_{\mathbf{b} \in \mathcal{U}} \mathbf{c}_T'(\mathbf{P}_T \mathbf{b} + \mathbf{q}_T) \Big] \Big] \Bigg] \\
& \text{s.t. } \mathbf{A}_1 \mathbf{w}_1 + \sum_{i=2}^{T} \mathbf{A}_i(\mathbf{P}_i \mathbf{b} + \mathbf{q}_i) \leq \mathbf{b}, \ \ \forall \mathbf{b} \in \mathcal{U}, \qquad (\Pi^T_{\text{Adapt}}) \\
& 0 \leq \mathbf{P}_i \mathbf{b} + \mathbf{q}_i \leq 1 \\
& \mathbf{P}_i \in \mathbb{R}^n_+, \mathbf{q}_i \in \mathbb{R}^{n_i}_+.
\end{aligned}
$$

The above formulation can be converted to a single linear program (the exact description for a two-stage adaptive problem is given in Appendix C.2). Although, we enforce affine recourse in the adaptive problem, it is not always optimal. As a result, we now explicitly characterize the cases when an affine policy is optimal. The key property which governs this is the number of extreme points in the uncertainty set (relative of m).

Characterizing the Optimality of Affine Policies

Bertsimas and Goyal [18] study the two-stage adaptive optimization problem Π^2_{Adapt} and prove the optimality of affine policy for a simplex[5] uncertainty set. Since, the TFMP problem is inherently a multi-period problem, it is appropriate that we extend

[5]A simplex is a set which is generated by $m+1$ extreme points, where m is the number of constraints.

the result for multiple periods as done below (please refer to Appendix C.2 for the proof):

Theorem 7. *Consider the problem* $\Pi^T_{\text{Adapt}}(\mathcal{U})$ *such that* $\mathcal{U}$ *is a simplex. Then, there is an optimal multi-stage solution* $\hat{\mathbf{w}}_i(\mathbf{b})$ *such that* $\hat{\mathbf{w}}_i(\mathbf{b})$ *are affine functions of* $\mathbf{b}$*, i.e., for all* $\mathbf{b} \in \mathcal{U}$*,*

$$\hat{\mathbf{w}}_i(\mathbf{b}) = \mathbf{P}_i\mathbf{b} + \mathbf{q}_i, \tag{4.15}$$

where $\mathbf{P}_i \in \mathbb{R}^{n_i \times m}$*,* $\mathbf{q}_i \in \mathbb{R}^{n_i}$*.*

Figure 4-10 shows two examples of uncertainty sets (in $\mathbb{R}^2$) and classifies cases where an affine policy is optimal. On the left is a simplex uncertainty set for which an affine policy is optimal. In contrast, the right side shows a a polyhedron with more than 3 extreme points for which a piece-wise affine policy (with three pieces) is optimal.

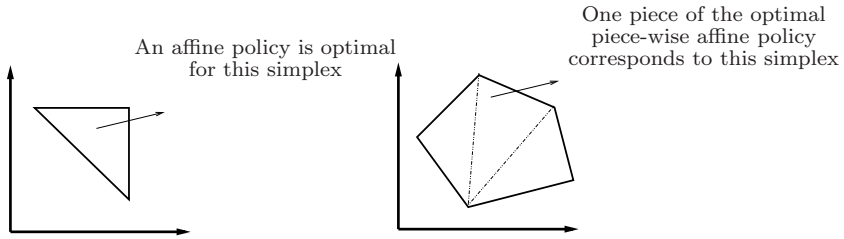


Figure 4-10: **Left:** A simplex uncertainty set in $\mathbb{R}^2$. **Right:** A polyhedral set in $\mathbb{R}^2$ broken into three simplices.

Remark 6. Although, the number of extreme points in the uncertainty set governs the cases when an affine recourse is optimal, they do not have any impact on the computational performance of both the robust and adaptive models. This is because the robust model is simply solving another deterministic instance, whereas the adaptive problem becomes an equivalent LP whose size is independent of the number of extreme points.

4.6 Computational Results

In this section, we report numerical results from the solution approaches introduced in Section 4.5.

Data and Setup

We utilize a national-scale database of real flight delays data over the past few years (also used for the fairness and collaboration chapters) to present proof-of-concept of the usefulness of the optimization methodologies proposed in this chapter.

The basic setup of the airspace is exactly the same as used in the fairness chapter. In addition, the entire airspace is subsequently divided into three regions, namely, north-east, south-west and central. Each of these regions is then subjected to simulated weather-fronts. The time of arrival, duration and capacity reduction of each weather-front is generated randomly from appropriate intervals. This enables the construction of uncertainty sets with varying magnitude of extreme points to study the implications on the complexity of solving the robust and adaptive problems and the corresponding price. Finally, the capacity inputs used for all the instances are at the "*infeasibility border*", i.e., values which when perturbed slightly on the conservative side lead to infeasibility of the overall problem.

Figure 4-11 plots the flight traffic (departures and arrivals). The left plot shows the aggregate demand as a function of time of day whereas the right plot is as a function of airport of operation. As expected, the distribution across airports is quite non-uniform as there are considerably more operations at hub airports.

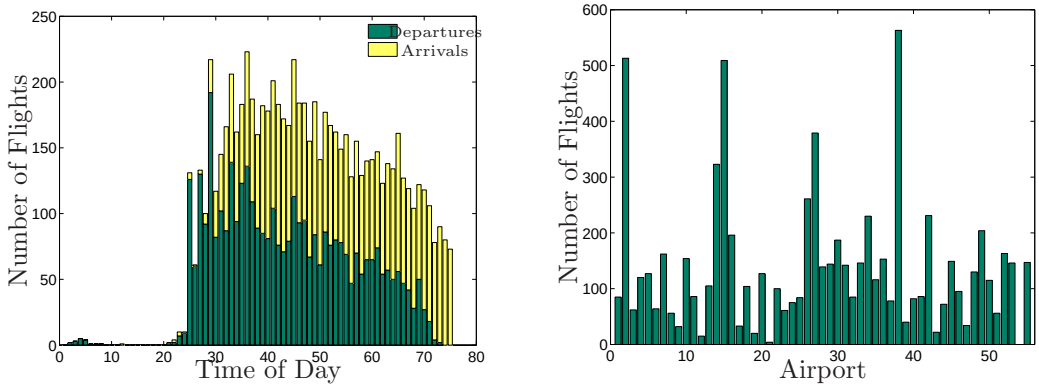


Figure 4-11: Flight traffic (departures and arrivals) as a) a function of time of day (Left); and b) a function of airport (Right).

To solve the robust and adaptive models (which use various uncertainty sets), we use a new tool for robust optimization - *Robust Optimization Made Easy* (ROME) [24]. We use the same tool to compute optimal solutions for the deterministic problem. ROME is a MATLAB-based algebraic toolbox which solves robust optimization problems in its full generality and provides support for adaptive routines by implementing affine policies.

Before undertaking an extensive set of computational experiments, we now put forth the following questions we wish to answer:

1. Are the running times obtained suitable for practical deployment?
2. How favorable are the integrality properties of the robust and adaptive solutions?
3. What is the price for incorporating robustness and adaptability?
4. How different are the robust and adaptive schedules compared to the deterministic counterpart (in terms of the resulting flight sequences)?

Results and Discussion

Table 4.1 reports the performance of the robust model. The running times reported in the table correspond to both the solver time and the ROME input parsing time (which accounts for majority of the total time). The integrality properties of the robust equivalent closely mimics the deterministic version which is known to have strong integrality properties (as evidenced by the fact that in all cases the % non-integral solutions is less than 1%). Furthermore, the running times for the robust model compares favorably with the deterministic counterpart.

Characteristics of Robust Schedules. To study the nature of robust schedules, we define two quantities which qualitatively govern the factors impacting the price of robustness and the deviation in the robust schedule (relative to the deterministic counterpart). The first is the percentage capacity reduction in the capacity vector used to solve the robust problem (viz. $\mathbf{b}_{\min}$) relative to the deterministic capacity estimate $\mathbf{b}_{\text{det}}$ (mathematically, this is $100 * \frac{\mathbf{e}'(\mathbf{b}_{\text{det}} - \mathbf{b}_{\min})}{\mathbf{e}'\mathbf{b}_{\text{det}}}$ and will be denoted by C_{Red} henceforth). This quantity is expected to govern the price of robustness in that higher the reduction, higher is the objective cost for the robust problem ($Z_{\text{Rob}} \rightarrow \infty$ as $\mathbf{b}_{\min} \rightarrow 0$). The second quantity (somewhat correlated with the first one) is the difference in the deterministic and robust flight schedules. It is defined as $\mathbf{e}'(\mathbf{a}^{\text{Det}} - \mathbf{a}^{\text{Rob}})$ where the vector $\mathbf{a}$ consists of the time of arrival of all flights at its destination airport (Det corresponds to deterministic and Rob to the robust problem). Figure 4-12 plots the price of robustness ($\mathcal{POR}$) and schedule deviation ($\mathcal{SD}$) as a function of C_{Red}. The key insight from the plot is that when $\mathbf{b}_{\min}$ is "close" to the deterministic capacity, both metrics $\mathcal{POR}$ and $\mathcal{SD}$ are small. Furthermore, the price of robustness and schedule deviation follow an approximately linear relationship with

	Deterministic		**Robust**		
Region (# of Flights)	Cost	Time (sec.)	Cost	Time (sec.)	% Nonint
	208	378	229	407	0
North-East	227	515	230	643	0
(500-1000)	145	636	161.5	638	0.15
	281	466	281	533	0
	935	4839	971.5	4721	0
	650	1996	677	1984	0
Central	193	103	193	95	0
(500-1000)	647	1770	675	1779	0
	935	4685	1013	4607	0
	240	401	248	402	0.10
	642	1937	647	1977	0
South-West	137	578	144	566	0.10
(500-1000)	208	358	208	387	0
	51	105	52	103	0
	935	4540	1044	4713	0

Table 4.1: Computational Experience with Π_2.

C_{Red} (the "*goodness of fit*" measure R^2 is greater than 0.8 for all plots which implies statistical significance). The results validate our intuitive expectation.

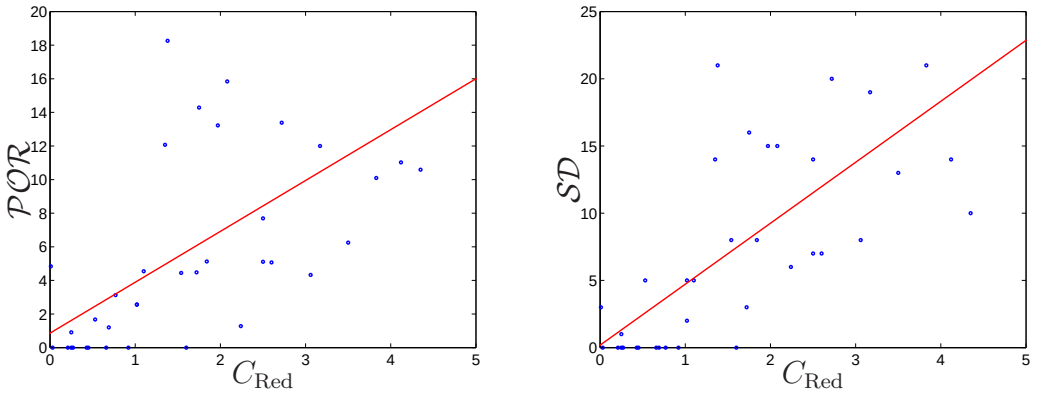


Figure 4-12: *Characteristics of Robust Solutions.* **Left:** Price of robustness as a function of capacity reduction. **Right:** Schedule deviation as a function of capacity reduction. The red line corresponds to the best linear fit.

While the preceding discussion sheds light on the relation between the deterministic and robust schedules, the price of robustness depicted in Figure 4-12 does not paint the complete picture. The reason being that the deterministic solution would most likely not even be feasible for a particular scenario of the uncertainty set, and so $\mathcal{POR}$ is not really the price the decision-maker has to pay for achieving robustness. Therefore, to better assess the utility of robust schedules, we propose the following thought experiment: we fix a particular scenario from the uncertainty set and use this as the actual capacity profile that materialized (we use the scenarios corresponding to $\alpha = \underline{\alpha}, \frac{\underline{\alpha}+\overline{\alpha}}{2}$ and $\overline{\alpha}$). Subsequently, we construct a feasible schedule (starting from the optimal deterministic solution) by protracting flight arrivals as and when capacity becomes available. The delays so obtained for the new deterministic schedule enable a more apt comparison with the robust schedule (which is guaranteed to be feasible for this capacity realization). Figure 4-13 plots the results of this exercise. The important takeaways are: i) robust schedules come at a small cost (as evidenced by the closeness of the red line to the 45 degree blue line); and ii) the cost of the protracted deterministic solution (corresponding to a scenario from the uncertainty set) is higher (sometimes substantially) than the robust solution in most cases, thereby, emphasizing the utility of deterministic schedules.

Characteristics of Adaptive Schedules. Table 4.3 reports solutions for the adaptive problem across a random set of instances. The worst-case adaptive cost (min-max objective) is the same as the robust cost in all cases. But, to demonstrate the utility of

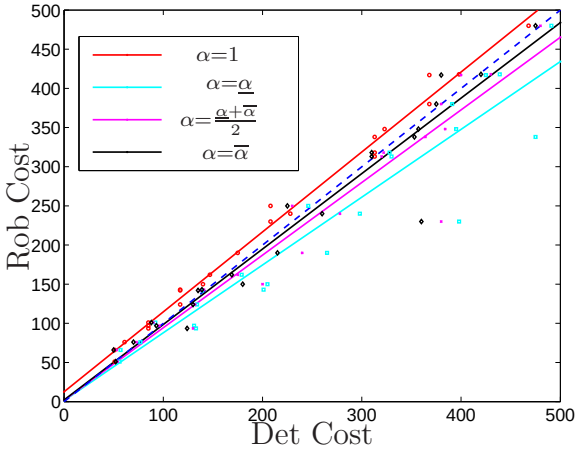


Figure 4-13: *Utility of Robust Solutions.* The plot depicts the relation between robust and new deterministic cost (for the protracted feasible schedule) for different scenarios of the uncertainty set.

Capacity Scenario	**Best Linear Fit**	$\mathbf{R^2}$
$\alpha = \underline{\alpha}$	y = 0.80x + 11.61	0.92
$\alpha = \frac{\underline{\alpha}+\overline{\alpha}}{2}$	y = 0.89x + 8.02	0.94
$\alpha = \overline{\alpha}$	y = 0.93x + 6.89	0.95

Table 4.2: Utility of Robust Solutions: Best fit lines.

adaptive solutions, we implement the adaptive solution on a rolling horizon basis. We implement the first-stage decisions of the optimal adaptive solution. Subsequently, we fix a scenario from the uncertainty set (corresponding to $\alpha = \frac{\underline{\alpha}+\overline{\alpha}}{2}$) and use this as the actual capacity profile that materialized. Finally, we re-optimize the second-stage decisions using this capacity scenario as deterministic input and refer to the aggregate cost as "Adjusted". The first observation pertaining to the adaptive schedules is the viability of generating integral first-stage variables (while enforcing affine dependence of the recourse variables). Some of other remarks are as follows:

- The results for adaptive problem indicate that in terms of running times, adaptive is quite expensive compared to the robust model (this is expected as the transformed LOP for the adaptive problem is of a much larger size).

- In all cases, the (worst-case) adaptive cost is exactly the same as the robust cost. But, the ("adjusted") cost obtained by re-optimizing over the intermediate realization of the uncertainty set leads to improvements in the cost.

	Robust		**Adaptive**		
Region	Cost	Time	Cost		Time
(# of Flights)		(sec.)	Worst-case	Adjusted	(sec.)
	92	84	92	84	265
North-East	119	73	119	110	491
(300-500)	54	55	54	48	369
	112	127	112	109	326
	87	43	87	83	813
Central	125	34	125	120	287
(300-500)	86	98	86	80	257
	35	63	35	33	382
	75	86	75	72	250
South-West	42	87	42	35	431
(300-500)	123	133	123	120	519
	99	106	99	95	294

Table 4.3: Computational Experience with Π_6.

Finally, Figure 4-14 plots the box plots of the running times for the robust and adaptive problems.

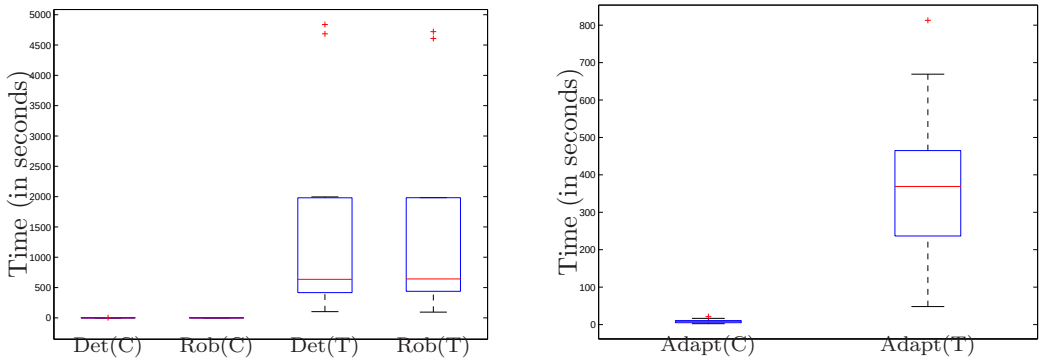


Figure 4-14: *Running times for* i) deterministic and robust problems (**Left**); and ii) adaptive problem (**Right**). C denotes CPLEX solver time and T denotes Total time (including ROME parsing time).

4.7 Conclusions

This chapter presents a framework for solving network air traffic flow management problems in a stochastic setting under the robust and adaptive optimization paradigm. To address uncertainty, we introduce a new weather-front based approach to construct the uncertainty set of possible capacity realizations. The key benefits of this approach are the low-dimensionality of the resulting discrete sets (as uncertainty in only two parameters govern the uncertainty set for each airspace element) and the intuitive appeal in the modeling approach. We propose a polyhedral description of the convex hull of the discrete uncertainty set to formulate the stochastic ATFM problem meaningfully. Subsequently, we formulate robust and adaptive optimization equivalent of the deterministic problem and propose tractable solution methodologies. We prove the equivalence of the robust problem to a new instance of the deterministic problem. For the adaptive counterpart, we utilize affine policies to solve the linear relaxation. Extensive empirical results highlight the utility of robust and adaptive solutions.

Chapter 5

Conclusion

In this chapter, we summarize the overall contributions of the thesis and point out directions for future research.

5.1 Thesis Summary

Our aspiration in this dissertation was to propose a tractable optimization framework which can provide decision engines to control air traffic while addressing the practical issues of equity, collaboration amongst airlines and capacity uncertainty. In doing so, we tried to adhere to the CDM philosophy (the governing set of principles for all ATM initiatives), which is to involve airlines at every step of the planning process. More precisely, we attacked three critical issues in ATFM research, namely:

1. **Fairness.** We introduced different notions of fairness, namely i) FSFS fairness - controlling number of reversals and total amount of overtaking; ii) Proportional fairness - equalizing airline delays; and iii) a combination of the FSFS and Proportional fairness paradigms in the ATFM problem. The FSFS fairness paradigm, in particular, is intuitively appealing as it represents the next best alternative to RBS (the currently agreed upon and implemented principle). We reported extensive empirical results of the proposed optimization models on national-scale, real world datasets that showed interesting tradeoffs between fairness and efficiency. The important takeaways were the possibility of generating schedules with single-digit reversals and overtaking (close to RBS) for less than 10% increase in the delay costs. Finally, computational times were less than 30 minutes which are encouraging for real-time deployment.

2. **Airline Collaboration.** Subsequently, we allowed for further collaboration amongst airlines by proposing network models for slot reallocation. This was a generalization of the intra-airline substitution phase of the current CDM practice to inter-airline reallocation across multiple airports. A structure of airline offers called AMAL (at-most, at-least) offers were used as data input. The attractive features of this input was the simplicity of offer structure which still made possible multiple combinations of feasible trades. Based on this data, two models were developed for slot reallocation flexibility. Finally, we reported on case studies which highlighted the considerable improvements in the internal objective functions of the airlines as a result of this flexibility. Both models were fast (solving to optimality in seconds), thereby being suited for real-time deployment.

3. **Capacity Uncertainty.** We addressed the issue of capacity uncertainty by studying the first application of the robust and adaptive optimization paradigm to the deterministic ATFM problem. We developed a weather-front based approach to model the uncertainty inherent in airspace capacity estimates resulting from the impact of a small number of weather fronts moving across the National Airspace (NAS). The key advantage of our uncertainty set construction was its low-dimensionality (uncertainty in only two parameters govern the overall uncertainty set for each airspace element). We proved that the resulting robust ATFM problem is just another deterministic instance. The adaptive problem, in contrast, was solved using affine policies. We reported extensive empirical results from the proposed models on real-world flight schedules augmented with simulated weather fronts that demonstrated interesting tradeoffs between deterministic, robust and adaptive solutions.

5.2 Directions for Future Research

We conclude by providing some directions for future research which builds upon the research done in this thesis.

1. **ATFM Problems.** First and foremost, within the ATFM context, there still remain important challenges and opportunities from a research standpoint.

    - **Robust ATFM Problem under Capacity Uncertainty.** In this thesis, we showed that the robust ATFM problem corresponds to solving a new

instance of the deterministic problem with the component-wise minimum capacity vector. A consequence of this result is that the robust problem may suffer from the possibility of over-conservative schedules. Therefore, to mitigate this, we need robust models that incorporate budgets of uncertainty on the simultaneous realizations of worst-case capacity for different weather-fronts. This problem potentially has ramifications for a generic tractable robust model under right hand side uncertainty.

- **Rerouting under Stochastic Capacity.** Although, we design a new framework for the ATFM problem under capacity uncertainty by invoking the robust and adaptive paradigm, the added complication of rerouting deserves more focused attention. In particular, there is a significant tractability challenge in addressing rerouting in the presence of capacity uncertainty. Therefore, a model which outputs different rerouting plans for various capacity realizations would be immensely appealing from a controller standpoint.
- **Market Mechanisms facilitating Airline Collaboration.** Even though this thesis proposes models for slot reallocation by allowing airlines to propose offers to trade slots at different airports, there still remains opportunities for more complex models which mimic trading just as in a financial marketplace. Thus, there is the opportunity for a more generalized model which involves side-payments for giving up slots voluntarily and allows for more sophisticated airline input.

2. **Robust and Adaptive Optimization.** Robust and Adaptive Optimization has been successful in presenting an alternative tractable paradigm for decision-making under uncertainty. There has been prolific research over the past decade on the theoretical foundations as well as in its application to a diverse set of problems. Nonetheless, there still remain challenges which deserve focused research attention.

   - **Robust Optimization under Discrete Uncertainty Sets.** The ATFM problem under capacity uncertainty has to account for the discreteness of the possible capacity estimates. This extends itself to the opportunity for developing a generic tractable robust model under discrete uncertainty sets.
   - **Complexity of Recourse Policies for Adaptive Optimization.** It would be appropriate to characterize the complexity of solving the prob-

lems resulting from using a particular class of recourse (for example, affine, polynomial, etc.). This would classify certain problems as being intrinsically intractable compared to those for which polynomial time algorithms can be designed.

Appendix A

Fairness Models

A.1 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 2. *The polyhedron $P_{\text{Reversal}}(f, f', j)$ is integral.*

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

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

We make use of the following two facts from discrete optimization [15]:

Fact 1. *Let* $\mathbf{A}$ *be an integral matrix.* $\mathbf{A}$ *is totally unimodular if and only if* $\{\mathbf{x} |\ \mathbf{a} \leq \mathbf{A}\mathbf{x} \leq \mathbf{b}, \mathbf{l} \leq \mathbf{x} \leq \mathbf{u}\}$ *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 $P =$

$\{\mathbf{x}|\ \mathbf{A}\mathbf{x} \leq \mathbf{b}\}$:

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

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

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

Let $\mathbf{B}_1, \mathbf{B}_2, \ldots, \mathbf{B}_n$ be consecutive blocks of two rows of $\mathbf{A}$ each, i.e., block $\mathbf{B}_k$ contains the rows $2k-1$ and $2k$. For any collection $\mathbf{Q}$ of rows of the matrix $\mathbf{A}$, we show how to partition it into two parts $\mathbf{J}_1$ and $\mathbf{J}_2$ so that the sum of the rows in $\mathbf{J}_1$ minus sum of the rows in $\mathbf{J}_2$ is a vector with entries 0, +1 and -1 only. Suppose $\mathbf{Q}$ contains both the rows of some block $\mathbf{B}_i$, then, put both these rows in $\mathbf{J}_1$. The remaining rows (say m) in $\mathbf{Q}$ then come from different blocks, call them $\mathbf{R}_{j_1}, \mathbf{R}_{j_2}, \ldots, \mathbf{R}_{j_m}$. These m rows are partitioned as follows:

Let $\mathbf{Q}_+$ be the subset of these m rows where the last element is +1 and $\mathbf{Q}_-$ be those rows where the last element is -1. Then, put $\lceil \frac{|\mathbf{Q}_+|}{2} \rceil$ rows of $\mathbf{Q}_+$ in $\mathbf{J}_1$ and the remaining rows in $\mathbf{J}_2$. Similarly, put $\lceil \frac{|\mathbf{Q}_-|}{2} \rceil$ rows of $\mathbf{Q}_-$ in $\mathbf{J}_1$ and the remaining rows in $\mathbf{J}_2$. Since the sum of two rows in the same block is all zeroes, all such blocks in $\mathbf{J}_1$ do not affect the sum of all the rows in $\mathbf{J}_1$. Let $\tilde{\mathbf{J}}$ denote the vector resulting from the sum of the rows in $\mathbf{J}_1$ minus the sum of the rows in $\mathbf{J}_2$. All the elements except

the last one in $\tilde{\mathbf{J}}$ is exactly 0, +1 or -1 because of the structure of the matrix $\mathbf{A}$. The contribution of the rows from $\mathbf{Q}_+$ to the last element of $\tilde{\mathbf{J}}$ is either 0 or +1. Similarly, the contribution of the rows from $\mathbf{Q}_-$ to the last element of $\tilde{\mathbf{J}}$ is either -1 or 0. This implies that the last element of $\tilde{\mathbf{J}}$ which is the sum of these two contributions can only be $+1$, -1 or 0.

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

Appendix B

Airline Collaboration

B.1 Strength of Slot Reallocation Models

We analyze the polyhedral structure of the two network formulations, namely, TFMP-Trading-VB and TFMP-Trading-BG. Let IP_{VB} and IP_{BG} denote the set of all feasible binary vectors in the two formulations respectively. Moreover, let P_{VB} and P_{BG} denote the polyhedra induced by the LP relaxations of the two models respectively. Let us denote the decision space of the two models by $\mathbf{u}$ and $\mathbf{v}$ respectively, $\mathbf{u} = (w^f_{j,t},\ o_{dd'uu'},\ x_{ik},\ z_{ik})$ and $\mathbf{v} = (w^f_{j,t},\ o_{dd'uu'})$. Note that $\mathbf{v} \subseteq \mathbf{u}$. Let $\mathbf{A}$ and $\mathbf{D}$ be the constraint matrices of the two formulations respectively and let $\mathbf{b}$ and $\mathbf{d}$ be the right-hand vectors.

The polyhedra P_{VB} and P_{BG} can be written as follows:

$$P_{VB} = \{\mathbf{u} = (w^f_{j,t},\ o_{dd'uu'},\ x_{ik},\ \ z_{ik})|\ 0 \leq w^f_{j,t} \leq 1,\ \ 0 \leq o_{dd'uu'} \leq 1,\\ 0 \leq x_{ik} \leq 1,\ \ 0 \leq z_{ik} \leq 1,\ \mathbf{Au} \leq \mathbf{b}\}$$

$$P_{BG} = \{\mathbf{v} = (w^f_{j,t},\ o_{dd'uu'})|\ 0 \leq w^f_{j,t} \leq 1,\ 0 \leq o_{dd'uu'} \leq 1, \mathbf{Dv} \leq \mathbf{d}\}$$

The set of all feasible binary vectors in the two polyhedra (denoted by IP_{VB} and IP_{BG}) can be written as follows:

$$IP_{VB} = \{\mathbf{u} \in \{0,1\}^n|\ \mathbf{Au} \leq \mathbf{b}\}$$

$$IP_{BG} = \{\mathbf{v} \in \{0,1\}^n|\ \mathbf{Dv} \leq \mathbf{d}\}$$

Let $\pi_v(P_{VB})$ be the projection of the polyhedron P_{VB} onto the v space. Moreover, let $\pi_v(IP_{VB})$ be the projection of the set IP_{VB} onto the v space. We need to answer the following questions:

1. Are the polyhedra P_{VB} and P_{BG} integral?
2. Is $\pi_v(P_{VB}) = P_{BG}$?
3. Is $\pi_v(IP_{VB}) = IP_{BG}$?
4. Are the two formulations equivalent?

Since the capacity and connectivity constraints are common to both formulations, they don't matter in the analysis.

Proposition 3. *The polyhedra P_{VB} and P_{BG} are not integral.*

Proof. We provide the following counter-example: there are two flights departing from a single-airport. The data of the problem is: $\mathcal{F} = \{f_1, f_2\}$, $\mathcal{K} = \{a\}$, $\mathcal{T} = \{1, 2, 3\}$, $\mathcal{O} = \{o_1\}$, $\mathcal{C} = \emptyset$. The capacities are: $D(1) = D(2) = D(3) = 1$. The slot assignment from Stage I is $D_{f_1} = 1$, $D_{f_2} = 3$ and the submitted offer to trade is $o_1 = \{f_1, 2;\ f_2, 2\}$.

Non-integrality of P_{VB}:

The decision variables are: $x = (w^1_{a,1},\ w^1_{a,2},\ w^1_{a,3},\ w^2_{a,1},\ w^2_{a,2},\ w^2_{a,3},\ x_{11},$

$x_{12}\, x_{21},\ x_{22},\ z_{11},\ z_{12},\ z_{21},\ z_{22},\ o_1)$. Letting $\mathbf{b}_{\text{ineq}} = (1, 1, 1, 0, 0, 0, 0, 0, 0)$ and $\mathbf{b}_{eq} = (1, 1, 0, 0, 0, 0, 0, 0)$, the feasible space can be written as $\mathbf{A}_{ineq}x \leq \mathbf{b}_{ineq}$, $\mathbf{A}_{eq}x = \mathbf{b}_{eq}$. We omit the description of $\mathbf{A}_{ineq}$ and $\mathbf{A}_{eq}$ for brevity. The objective function min: $w^1_{a,1} + w^1_{a,2} + w^1_{a,3} + w^2_{a,1} + w^2_{a,2} + w^2_{a,3} + 2o_1$ gives an optimal solution of $w^1_{a,1} = 0.4179$, $w^1_{a,2} = 1$, $w^1_{a,3} = 1$, $w^2_{a,1} = 0$, $w^2_{a,2} = 0.4179$, $w^2_{a,3} = 1$, $x_{11} = 0.4179$, $x_{12} = 0.5821$, $x_{21} = 0.4179$, $x_{22} = 0.5821$, $z_{11} = 0$, $z_{12} = 0$, $z_{21} = 0$, $z_{22} = 0$, $o_1 = 0.5821$, which shows that the polyhedron P_{VB} is not integral.

Non-integrality of P_{BG}:

The decision variables are: $x = (w^1_{a,1},\ w^1_{a,2},\ w^1_{a,3},\ w^2_{a,1},\ w^2_{a,2},\ w^2_{a,3},\ o_1)$. The resulting formulation (TFMP-Trading-BG) is:

$w^1_{a,1} + w^2_{a,1} \leq 1;\ w^1_{a,2} - w^1_{a,1} + w^2_{a,2} - w^2_{a,1} \leq 1;\ w^1_{a,3} - w^1_{a,2} + w^2_{a,3} - w^2_{a,2} \leq 1;\ w^1_{a,1} - w^1_{a,2} \leq$

$0;\ w^1_{a,2} - w^1_{a,3} \leq 0;\ w^2_{a,1} - w^2_{a,2} \leq 0;\ w^2_{a,2} - w^2_{a,3} \leq 0;\ o_1 - w^1_{a,2} \leq 0;\ o_1 - w^2_{a,3} \leq 0;\ -o_1 - w^1_{a,1} \leq -1;\ -o_1 - w^2_{a,3} + w^2_{a,2} \leq -1.$

Letting $b = (1, 1, 1, 0, 0, 0, 0, 0, 0, -1, -1)$, and

$$\mathbf{A} = \begin{bmatrix} 1 & 0 & 0 & 1 & 0 & 0 & 0 \\ -1 & 1 & 0 & -1 & 1 & 0 & 0 \\ 0 & -1 & 1 & 0 & -1 & 1 & 0 \\ 1 & -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & -1 & 0 \\ 0 & -1 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & -1 & 0 & 1 \\ -1 & 0 & 0 & 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 & 1 & -1 & -1 \end{bmatrix}$$

the feasible space can be written as $\mathbf{Ax} \leq \mathbf{b}$.

The objective function min: $w^1_{a,1} + w^1_{a,2} + w^1_{a,3} + w^2_{a,1} + w^2_{a,2} + w^2_{a,3} + o_1$ gives an optimal solution of $w^1_{a,1} = 1/2,\ w^1_{a,2} = 1/2,\ w^1_{a,3} = 1/2,\ w^2_{a,1} = 0,\ w^2_{a,2} = 1/2,\ w^2_{a,3} = 1,\ o_1 = 1/2$ which shows that the polyhedron P_{BG} is not integral. □

B.2 Integer Equivalence of Slot Trading Models

Proposition 4. $\pi_v(IP_{VB}) = IP_{BG}$.

Proof. **Claim 1.** $IP_{BG} \subseteq \pi_v(IP_{VB})$. Let $\tilde{v} \in IP_{BG}$, we will construct a vector $\tilde{u} \in IP_{VB}$ such that $\pi_v(\tilde{u}) = \tilde{v}$.

For each flight $f_i \in \mathcal{F}$, there are exactly two possibilities: either i) no trade will get executed, or ii) exactly one trade will be executed. Let $\tilde{u}_i \subseteq \tilde{u}$ and $\tilde{v}_i \subseteq \tilde{v}$ denote the subsets of the respective decision spaces that contain variables corresponding to flight f_i. Under both scenarios, we construct a vector $\tilde{u}_i \in IP_{VB}$ such that $\pi_v(\tilde{u}_i) = \tilde{v}_i$.

In case, no trade is executed (i.e., $\tilde{o}_j = 0, \forall j \in \mathcal{O}^{f_i}$), then constraint (3.2d) implies that $\tilde{w}^{f_i}_{\text{dest}_{f_i}, D_{f_i}} = 1$ and $\tilde{w}^{f_i}_{\text{dest}_{f_i}, D_{f_i}-1} = 0$. Constraints (3.1d) and (3.1e) further imply

that $\tilde{x}_{ik} = 0, \forall k \neq p_i$. This makes $\tilde{x}_{ip_i} = 1$ because of constraint (3.1a). From constraint (3.3), this implies $\tilde{w}^{f_i}_{\text{dest}_{f_i}, D_{f_i}} = 1$ and $\tilde{w}^{f_i}_{\text{dest}_{f_i}, D_{f_i}-1} = 0$. Putting $\tilde{z}_{ik} = 0, \forall k$ makes constraints (3.1b) and (3.1c) feasible and hence, the constructed vector $\tilde{u}_i$ is feasible for TFMP-Trading-VB. In case, exactly one trade is executed (i.e., $\exists s \in \mathcal{O}^{f_i}$ such that $\tilde{o}_s = 1$ and $\tilde{o}_r = 0, \forall r \neq s$), then constraints (3.1d) and (3.1e) imply that $\exists l \neq p_i$ such that $\tilde{x}_{il} = 1$. This makes $\tilde{x}_{ip_i} = 0$ (from constraint (3.1a)). Again, putting $\tilde{z}_{ik} = 0, \forall k$ makes constraints (3.1b) and (3.1c) feasible and hence, the constructed vector $\tilde{u}_i$ is feasible for TFMP-Trading-VB.

Claim 2. $\pi_v(IP_{VB}) \subseteq IP_{BG}$.

Now, let $\tilde{u} = (\tilde{w}^f_{j,t},\ \tilde{o}_{dd'uu'},\ \tilde{x}_{ik},\ \tilde{z}_{ik}) \in IP_{VB}$. We will show that $\tilde{v} = \pi_v(\tilde{u}) \in IP_{BG}$. Suppose for a flight f_i, $\tilde{x}_{ip_i} = 1$. Then, constraint (3.3) implies $\tilde{w}^{f_i}_{\text{dest}_{f_i}, D_{f_i}} = 1$ and $\tilde{w}^{f_i}_{\text{dest}_{f_i}, D_{f_i}-1} = 0$ and constraints (3.1d) and (3.1e) imply that $\tilde{o}_j = 0, \forall j \in \mathcal{O}^{f_i}$ (because $\tilde{x}_{ik} = 0, \forall k \neq p_i$ from constraint (3.1a)). This satisfies all constraints of TFMP-Trading-BG. On the contrary, if $\exists k \neq p_i$ such that $\tilde{x}_{ik} = 1$, then from constraints (3.1d) and (3.1e), there exists an offer containing flight f_i that will be executed. This satisfies constraints (3.2c) and (3.2d), and constraints (3.2a) and (3.2b) are also satisfied because of the structure of the network flow constraints.

From the claims above, it follows that $\pi_v(IP_{VB}) = IP_{BG}$. □

Since the set of feasible binary vectors of the two formulations is the same, this implies that the optimal objective function value (as long as it is a function of the variables $o_{dd'uu'}$) will be the same. This fact corroborates the exact same number of trades executed from the two models as reported in the computational results section.

Proposition 5. $\pi_v(P_{VB}) \neq P_{BG}$.

Proof. **Claim 1.** $P_{BG} \not\subseteq \pi_v(P_{VB})$. Let $\tilde{v} = (\tilde{w}^f_{j,t},\ \tilde{o}_{dd'uu'}) \in P_{BG}$. Suppose there are two offers that contain flight f_i (let us denote the offers by j_1 and j_2 and assume the time-periods corresponding to the two classes is the same, say i_1). Moreover, let $\tilde{o}_{j_1} = 0.3$ and $\tilde{o}_{j_2} = 0.7$ so that $\sum_{j \in \mathcal{O}^{f_i}} \tilde{o}_j = 1$. This satisfies constraints (3.2c) and (3.2d). Further, constraints (3.2a) and (3.2b) ensure that $\tilde{w}^{f_i}_{\text{dest}_{f_i}, t_{i_1}} \geq \tilde{o}_{j_1}$ and $\tilde{w}^{f_i}_{\text{dest}_{f_i}, t_{i_1}} \geq \tilde{o}_{j_2}$. Finally, let $\tilde{w}^{f_i}_{\text{dest}_{f_i}, t_{i_0-1}} = 0$ and $\tilde{w}^{f_i}_{\text{dest}_{f_i}, t_{i_1}} = 0.7$. We will show that under this setting $\nexists \tilde{u} \in P_{VB}$ such that $\pi_v(\tilde{u}) = \tilde{v}$.

From Constraint (3.1b), we have:

$$\begin{aligned}
\tilde{z}_{i1} &= \tilde{w}^{f_i}_{\text{dest}_{f_i},t_{i_1}} - \tilde{w}^{f_i}_{\text{dest}_{f_i},t_{i_0-1}} - \tilde{x}_{i1} \\
\tilde{z}_{i1} &= \tilde{w}^{f_i}_{\text{dest}_{f_i},t_{i_1}} - \tilde{w}^{f_i}_{\text{dest}_{f_i},t_{i_0-1}} - (\tilde{o}_{j_1} + \tilde{o}_{j_2}) \\
\tilde{z}_{i1} &= -0.3 < 0
\end{aligned}$$

The above implies that we cannot construct a vector $\tilde{u} \in P_{VB}$ such that $\pi_v(\tilde{u}) = \tilde{v}$.

Claim 2. $\pi_v(P_{VB}) \not\subseteq P_{BG}$.

Let $\tilde{u} = (\tilde{w}^f_{j,t},\ \tilde{o}_{dd'uu'},\ \tilde{x}_{ik},\ \tilde{z}_{ik}) \in P_{VB}$ such that $\tilde{z}_{ik} < \tilde{z}_{ik-1}$ and $\tilde{w}^{f_i}_{\text{dest}_{f_i},t_{i_{k-1}}} = 0$. Moreover, suppose there is only one offer corresponding to class k for flight f_i which implies $\tilde{x}_{ik} = \tilde{o}_{ikuu'}$ (from constraints (3.1d) and (3.1e)).

From constraint (3.1c), we have:

$$\begin{aligned}
\tilde{w}^{f_i}_{\text{dest}_{f_i},t_{i_k}} &= \tilde{w}^{f_i}_{\text{dest}_{f_i},t_{i_{k-1}}} + \tilde{x}_{ik} + \tilde{z}_{ik} - \tilde{z}_{i(k-1)} \\
\tilde{w}^{f_i}_{\text{dest}_{f_i},t_{i_k}} &= \tilde{x}_{ik} + \tilde{z}_{ik} - \tilde{z}_{i(k-1)} \\
\tilde{w}^{f_i}_{\text{dest}_{f_i},t_{i_k}} &= \tilde{o}_{ikuu'} + \tilde{z}_{ik} - \tilde{z}_{i(k-1)} \\
\tilde{w}^{f_i}_{\text{dest}_{f_i},t_{i_k}} &< \tilde{o}_{ikuu'}
\end{aligned}$$

The above is not satisfied by constraint (3.2a) and hence, $\pi_v(\tilde{u}) \notin P_{BG}$. □

Appendix C

Robust and Adaptive Optimization

C.1 Optimization Under Discrete Uncertainty Sets

Proof of Proposition 1. Let $\hat{\mathbf{x}}$ be feasible for OptDU, i.e., $\mathbf{A}\hat{\mathbf{x}} \leq \mathbf{b}, \forall \mathbf{b} \in \mathcal{U}_{\mathbf{0}}$. We will show that $\hat{\mathbf{x}}$ is feasible for OptDUConv as well. Let $\mathbf{b}^* \in \text{conv}(\mathcal{U}_{\mathbf{0}})$. Since, $\text{conv}(\mathcal{U}_0)$ is a polytope, therefore, by the Resolution Theorem, $\exists \boldsymbol{\lambda} \geq 0, \mathbf{e}'\boldsymbol{\lambda} = 1$, such that, $\mathbf{b}^* = \sum_{i=1}^{K} \lambda_i \cdot \mathbf{b}_i$.

$$
\begin{aligned}
& \mathbf{A}\hat{\mathbf{x}} \leq \mathbf{b}, \qquad \forall \mathbf{b} \in \mathcal{U}_{\mathbf{0}}, \\
\implies & \lambda_i \cdot \mathbf{A}\hat{\mathbf{x}} \leq \lambda_i \cdot \mathbf{b}_i, \qquad i = 1, ..., K, \\
\implies & \Big(\sum_{i=1}^{K} \lambda_i\Big) \cdot \mathbf{A}\hat{\mathbf{x}} \leq \sum_{i=1}^{K} \lambda_i \cdot \mathbf{b}_i, \\
\implies & \mathbf{A}\hat{\mathbf{x}} \leq \mathbf{b}^*.
\end{aligned}
$$

The above implies that $\hat{\mathbf{x}}$ is feasible for $\mathbf{b}^*$. Since, $\mathbf{b}^* \in \text{conv}(\mathcal{U}_{\mathbf{0}})$ was arbitrary (and hence spans $\text{conv}(\mathcal{U}_0)$), therefore $\hat{\mathbf{x}}$ is feasible for OptDUConv. Hence,

$$
\hat{\mathbf{x}} \in \text{OptDU} \implies \hat{\mathbf{x}} \in \text{OptDUConv} \tag{C.1}
$$

The opposite direction is obvious as $\mathcal{U}_0 \subseteq \text{conv}(\mathcal{U}_0)$. Therefore,

$$\hat{\mathbf{x}} \in \text{OptDUConv} \implies \hat{\mathbf{x}} \in \text{OptDU} \tag{C.2}$$

The proposition follows from (C.1) and (C.2). □

C.2 Power of Affine Policies in Multistage Adaptive Optimization

We closely mimic the proof technique presented in Bertsimas and Goyal [17] for the optimality of affine policies in two-stage adaptive problem under simplex uncertainty set.

Proof of Theorem 7. Let $\mathbf{w}_i^*(\mathbf{b})$ be an optimal solution of $\Pi^T_{\text{Adapt}}(\mathcal{U})$. We will construct an alternative solution $\hat{\mathbf{w}}_i(\mathbf{b})$ such that $\hat{\mathbf{w}}_i(\mathbf{b})$ are affine functions of $\mathbf{b}$ for $i \geq 2$ and the worst case cost of this solution is equal to Z_{Adapt}. Let

$$\mathbf{Q} = \Big[\big(\mathbf{b}^1 - \mathbf{b}^{m+1}\big) \ldots \big(\mathbf{b}^m - \mathbf{b}^{m+1}\big)\Big] \tag{C.3}$$

Since $\mathbf{b}^1, \ldots, \mathbf{b}^{m+1}$ are affinely independent, $(\mathbf{b}^1 - \mathbf{b}^{m+1}), \ldots, (\mathbf{b}^m - \mathbf{b}^{m+1})$ are linearly independent and $\mathbf{Q}$ is an invertible full-rank matrix. For any $\mathbf{b} \in \mathcal{U}$, $\exists \boldsymbol{\gamma},\ 0 \leq \boldsymbol{\gamma} \leq 1,\ \mathbf{e}'\boldsymbol{\gamma} = 1$,

$$\begin{aligned}
\mathbf{b} &= \sum_{i=1}^{m+1} \gamma_i \mathbf{b}^i, \\
\mathbf{b} &= \sum_{i=1}^{m} \gamma_i (\mathbf{b}^i - \mathbf{b}^{m+1}) + \mathbf{b}^{m+1}, \\
\mathbf{b} &= \mathbf{Q} \cdot \boldsymbol{\gamma} + \mathbf{b}^{m+1},\ \boldsymbol{\gamma} = (\gamma_1, \ldots, \gamma_m)^T.
\end{aligned} \tag{C.4}$$

Since $\mathbf{Q}$ is invertible, we have,

$$\mathbf{Q}^{-1}\big(\mathbf{b} - \mathbf{b}^{m+1}\big) = \boldsymbol{\gamma} \tag{C.5}$$

Let, $\mathbf{W}_i = \Big[\big(\mathbf{w}_i^*(\mathbf{b}^1) - \mathbf{w}_i^*(\mathbf{b}^{m+1})\big) \ldots \big(\mathbf{w}_i^*(\mathbf{b}^m) - \mathbf{w}_i^*(\mathbf{b}^{m+1})\big)\Big]$

For all $\mathbf{b} \in \mathcal{U}$, where $\mathbf{b} = \sum_{j=1}^{m+1} \gamma_j \mathbf{b}^j$ for $0 \leq \boldsymbol{\gamma} \leq 1$, consider the following solution:

$$\begin{aligned}\hat{\mathbf{w}}_i(\mathbf{b}) &= \mathbf{W}_i \mathbf{Q}^{-1}(\mathbf{b} - \mathbf{b}^{m+1}) + \mathbf{w}_i^*(\mathbf{b}^{m+1}) \\ &= \mathbf{W}_i \boldsymbol{\gamma} + \mathbf{w}_i^*(\mathbf{b}^{m+1}) \\ &= \sum_{j=1}^{m+1} \gamma_j \mathbf{w}_i^*(\mathbf{b}^j)\end{aligned}$$

The worst case cost of the solution $\hat{\mathbf{w}}_i(\mathbf{b})$ can be bounded as follows:

$$\begin{aligned}\max_{\mathbf{b}\in\mathcal{U}} \mathbf{c}_i' \hat{\mathbf{w}}_i(\mathbf{b}) &= \max_{\mathbf{b}\in\mathcal{U}} \mathbf{c}_i' \sum_{j=1}^{m+1} \gamma_j \mathbf{w}_i^*(\mathbf{b}^j) \\ &= \max_{\mathbf{b}\in\mathcal{U}} \sum_{j=1}^{m+1} \gamma_j \mathbf{c}_i' \mathbf{w}_i^*(\mathbf{b}^j) \\ &\leq \max_{j=1,\ldots,m+1} \mathbf{c}_i' \mathbf{w}_i^*(\mathbf{b}^j) \qquad \text{(C.6)}\end{aligned}$$

Consider the worst-case objective cost for the last two stages:

$$\begin{aligned}\max_{\mathbf{b}\in\mathcal{U}} \Big[\mathbf{c}_{T-1}' \hat{\mathbf{w}}_{T-1}(\mathbf{b}) + \max_{\mathbf{b}\in\mathcal{U}} \mathbf{c}_T' \hat{\mathbf{w}}_T(\mathbf{b})\Big] &\leq \max_{\mathbf{b}\in\mathcal{U}} \Big[\mathbf{c}_{T-1}' \hat{\mathbf{w}}_{T-1}(\mathbf{b}) \\ &\qquad + \max_{j=1,\ldots,m+1} \mathbf{c}_T' \mathbf{w}_T^*(\mathbf{b}^j)\Big] \\ &\leq \max_{j=1,\ldots,m+1} \mathbf{c}_{T-1}' \mathbf{w}_{T-1}^*(\mathbf{b}^j) + \max_{j=1,\ldots,m+1} \mathbf{c}_T' \mathbf{w}_T^*(\mathbf{b}^j)\end{aligned}$$

This indicates that the worst-case objective cost becomes separable. Using the same argument one step at a time in the backward direction (i.e., for $i = T-2, \ldots, 1$), we have,

$$\begin{aligned}\mathbf{c}_1' \hat{\mathbf{w}}_1 + \max_{\mathbf{b}\in\mathcal{U}} \Big[\mathbf{c}_2' \hat{\mathbf{w}}_2(\mathbf{b}) + \cdots + \max_{\mathbf{b}\in\mathcal{U}} \big[\mathbf{c}_{T-1}' \hat{\mathbf{w}}_{T-1}(\mathbf{b}) + \max_{\mathbf{b}\in\mathcal{U}} \mathbf{c}_T' \hat{\mathbf{w}}_T(\mathbf{b})\big]\Big] \\ \leq \mathbf{c}_1' \mathbf{w}^* + \sum_{i=2}^{T} \Big[\max_{j=1,\ldots,m+1} \mathbf{c}_i' \mathbf{w}_i^*(\mathbf{b}^j)\Big] \leq Z_{\text{Adapt}}^T\end{aligned}$$

Therefore, the worst case cost of the solution $\hat{\mathbf{w}}_i(\mathbf{b})$ is equal to the optimal cost of $\Pi^T_{\text{Adapt}}(\mathcal{U})$, which implies that the affine policy for each stage (after deterministic first stage variables) is optimal. □

To get a more qualitative understanding of the optimality of affine policies for uncertainty set with less than $m+1$ extreme points (say K), note that the solution in the i^{th} stage, $\hat{\mathbf{w}}_i(\mathbf{b})$ is described completely by describing the solution at all the extreme points. Thus, we need to find $\mathbf{P}_i \in \mathbb{R}^{n_i \times m}$ and $\mathbf{q}_i \in \mathbb{R}^{n_i}$ such that:

$$\mathbf{P}_i \mathbf{b}^j + \mathbf{q}_i = \mathbf{w}_i^*(\mathbf{b}^j), \qquad \forall j = 1, \ldots, K,$$

This system has a solution for $\mathbf{P}_i$ and $\mathbf{q}_i$ as long as $K \leq m+1$. Subsequently, the remainder of the proof consists of showing the feasibility of $\hat{\mathbf{w}}_i(\mathbf{b})$ for an arbitrary $\mathbf{b}$ followed by bounding the worst-case cost of $\hat{\mathbf{w}}_i(\mathbf{b})$ by Z^T_{Adapt}.

Equivalent LP for Two-stage Adaptive Problem with Affine Recourse

Bertsimas, Maes [11] show that a two-stage adaptive optimization problem with affine recourse can be converted to a single deterministic linear program. The two-stage adaptive optimization problem

$$\begin{aligned} \min_{\mathbf{x}, y(\mathbf{b})} \ & \mathbf{c}'\mathbf{x} + \max_{\mathbf{b} \in \mathcal{U}} \mathbf{d}' y(\mathbf{b}) \\ \text{s.t. } & \mathbf{A}\mathbf{x} + \mathbf{B} y(\mathbf{b}) \geq \mathbf{b}, \qquad \forall \mathbf{b} \in \mathcal{U} \\ & y(\mathbf{b}) \geq 0, \qquad \forall \mathbf{b} \in \mathcal{U} \\ & \mathbf{x} \geq 0. \end{aligned} \tag{Adapt}$$

with polyhedral uncertainty set of the form: $\mathcal{P} = \{\mathbf{b} | \ \mathbf{G}\mathbf{b} \leq \mathbf{f}\}$ where $\mathbf{G} \in \mathbb{R}^{l \times m}$ and $\mathbf{f} \in \mathbb{R}^l$ can be converted to a single deterministic linear program:

$$\begin{aligned} \min_{\mathbf{x},\mathbf{F},\mathbf{q},\eta,\mathbf{W},\mathbf{V},\mathbf{w}} \quad & \mathbf{c}'\mathbf{x} + \eta \\ \text{s.t.} \quad & \mathbf{f}'\mathbf{w} \leq \eta - \mathbf{d}'\mathbf{q}, \\ & \mathbf{G}'\mathbf{w} = \mathbf{F}'\mathbf{d} \\ & \mathbf{W}'\mathbf{f} \leq \mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{q}, \\ & \mathbf{G}'\mathbf{W} = \mathbf{I} - \mathbf{F}'\mathbf{B}', \\ & \mathbf{V}'\mathbf{f} \leq \mathbf{q}, \\ & \mathbf{G}'\mathbf{V} = -\mathbf{F}', \\ & \mathbf{x}, \mathbf{w}, \mathbf{V}, \mathbf{W} \geq 0. \end{aligned}$$

The problem above contains $n+(m+1)p+(m+p+1)l$ variables and $1+2m(1+l)+p$ constraints.

Relating Robust and Adaptive Problems

The optimal objective values of the robust and adaptive problems are related by a specific property of the uncertainty set – whether the uncertainty set contains the component-wise minimum element ($\mathbf{b}_{\min}$) or not.

Proposition 6. *Consider the multi-stage adaptive optimization problem* Π^T_{Adapt} *and its robust counterpart* Π_{Rob}. *If* $\mathbf{b}_{\min} \in \mathcal{U}$, *then,*

$$IZ_{\text{Adapt}} = IZ_{\text{Rob}}.$$

Proof. Let $\mathbf{w}_1^*, \mathbf{w}_i^*, \ i \geq 2$ be an optimal solution of Π_{Rob} and $\hat{\mathbf{w}}_1, \hat{\mathbf{w}}_i(\mathbf{b}), \forall \mathbf{b} \in \mathcal{U}$ be an optimal solution of Π^T_{Adapt}. Since $\mathbf{w}_1^*, \mathbf{w}_i^*, \ i \geq 2$ is the optimal solution, it is clearly feasible for Π_{Rob}. Therefore,

$$\mathbf{A}_1\mathbf{w}_1^* + \sum_{i=2}^{T} \mathbf{A}_i\mathbf{w}_i^* \leq \mathbf{b}, \ \ \forall \mathbf{b} \in \mathcal{U}$$

This implies that the solution $\mathbf{w}_1 = \mathbf{w}_1^*;\ \mathbf{w}_i(\mathbf{b}) = \mathbf{w}_i^*,\ \forall \mathbf{b} \in \mathcal{U}$ is feasible for Π^T_{Adapt}.

$$\begin{aligned} IZ_{\text{Adapt}} &\leq \mathbf{c}_1'\mathbf{w}_1^* + \sum_{i=2}^{T} \mathbf{c}_i'\mathbf{w}_i^* \\ &= IZ_{\text{Rob}} \end{aligned} \tag{C.7}$$

Since $\mathbf{b}_{\min} \in \mathcal{U}$,

$$\begin{aligned} IZ_{\text{Adapt}} = \mathbf{c}_1'\hat{\mathbf{w}}_1 + \max_{\mathbf{b}\in\mathcal{U}} \Big[\mathbf{c}_2'\hat{\mathbf{w}}_2(\mathbf{b}_2) + \cdots + \\ \max_{\mathbf{b}\in\mathcal{U}} \left[\mathbf{c}_{T-1}'\hat{\mathbf{w}}_{T-1}(\mathbf{b}_{T-1}) + \max_{\mathbf{b}\in\mathcal{U}} \mathbf{c}_T'\hat{\mathbf{w}}_T(\mathbf{b}_T)\right]\Big] \end{aligned} \tag{C.8}$$

$$\geq \mathbf{c}_1'\hat{\mathbf{w}}_1 + \sum_{i=2}^{T} \mathbf{c}_i'\hat{\mathbf{w}}_i(\mathbf{b}_{\min}) \tag{C.9}$$

Now, $(\hat{\mathbf{w}}_1, \hat{\mathbf{w}}_i(\mathbf{b}_{\min}))$ is a feasible solution of Π_{Rob}, therefore,

$$\begin{aligned} IZ_{\text{Rob}} &\leq \mathbf{c}_1'\hat{\mathbf{w}}_1 + \sum_{i=2}^{T} \mathbf{c}_i'\hat{\mathbf{w}}_i(\mathbf{b}_{\min}) \\ &\leq IZ_{\text{Adapt}} \end{aligned} \tag{C.10}$$

From (C.7) and (C.10),

$$IZ_{\text{Rob}} = IZ_{\text{Adapt}} \tag{C.11}$$

□

Figure C-1 depicts two examples of uncertainty sets. For the inverted triangle on the left, $\mathbf{b}_{\min} \notin \mathcal{U}$. In contrast, for the triangle on the right, $\mathbf{b}_{\min} \in \mathcal{U}$, and thus, $IZ_{\text{Rob}} = IZ_{\text{Adapt}}$.

C.3 Number of Extreme Points

We now quantify the number of extreme points in the polyhedral description of our uncertainty set $\mathcal{U}$ for the TFMP problem (done in Table C.1) so as to classify the

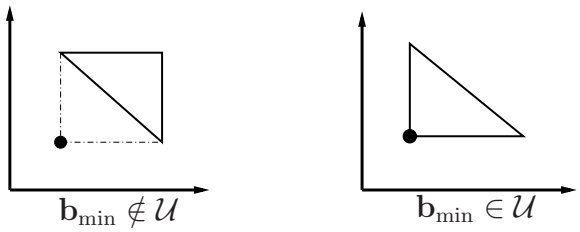


Figure C-1: Example uncertainty sets with and without $\mathbf{b}_{\min}$ (black filled circle denotes $\mathbf{b}_{\min}$).

cases in which affine recourse would be optimal. We introduce some more notation:

- $\mathcal{E}_i^j$: number of extreme points in the polytope corresponding to phase j of weather front i.
- $\mathcal{E}_i$: number of extreme points in the polytope corresponding to weather front i (across all phases).
- $\mathcal{E}$: number of extreme points in the polytope corresponding to all weather fronts.

Description (Denoted By)	**Uncertainty Set**	**# of Extreme Points**
1 WF, Single airspace element, one phase (EP_i)	$\mathcal{P}_{\underline{\alpha}_i}=\Big\{\mathbf{b}\in\mathbb{R}_+^m \mid b_t=C(1-y_t)+\underline{\alpha}_i C y_t$ $+(1-\underline{\alpha}_i)Cz_t,\ \forall t\in\{\underline{T}_{a_i},\ldots,\overline{T}_{b_i}\};$ $b_t=C,\ \forall t\in\mathcal{T}\setminus\{\underline{T}_{a_i},\ldots,\overline{T}_{b_i}\};$ $y_t\le y_{t+1};\ z_t\le z_{t+1};\ z_t\le y_t;\ 0\le y_t,z_t\le 1\Big\}$	$(\overline{T}_{a_i}-\underline{T}_{a_i}+1)\times(\overline{d}_i-\underline{d}_i+1)$
1 WF, Airspace, one phase ($\mathcal{E}_i^j$)	$\mathcal{P}_{\underline{\alpha}_i}^{\mathcal{AS},\mathcal{OP}}=\bigoplus_{k\in\mathcal{A}}\mathcal{P}_{\underline{\alpha}_i}^k$	$EP_i^{\lvert\mathcal{W}_i^j(\mathcal{S})\rvert+\lvert\mathcal{W}_i^j(\mathcal{K})\rvert}$
1 WF, Airspace, all phases ($\mathcal{E}_i$)	$\mathcal{P}_{\underline{\alpha}_i}^{\mathcal{AS},\mathcal{AP}}=\bigoplus_{j=1}^{p_i}\mathcal{P}_{\underline{\alpha}_i}^j$	$\Pi_{j=1}^{p_i}\mathcal{E}_i^j$
k WF, Airspace, all phases ($\mathcal{E}$)	$\mathcal{P}_{\underline{\alpha}}^{\mathcal{AS},\mathcal{AP}}=\bigoplus_{i=1}^{k}\mathcal{P}_{\underline{\alpha}_i}$	$\Pi_{i=1}^{k}\mathcal{E}_i$

Table C.1: Number of extreme points for the weather-front based polytope. $\mathcal{AS}$ denotes airspace, $\mathcal{OP}$ denotes one-phase and $\mathcal{AP}$ denotes all-phases. $\bigoplus$ denotes polyhedron concatenation.

Number of Extreme Points ($\mathcal{E}$) Relative to the Number of Constraints (m). Table C.1 implies that $\mathcal{E}$ increases exponentially with an increase in the number of airspace elements affected by weather and an increase in the number of time-periods. In contrast, $m \propto (2|\mathcal{K}|+|\mathcal{S}|)|\mathcal{T}|$ which implies that it increases linearly in the number of airspace elements and the total number of time-periods. Consequently, in the

asymptotic regime, $\mathcal{E}$ will dominate m. But, the more interesting (and practically relevant) case pertains to the number of extreme points when there is a reasonable upper bound on $|\mathcal{W}_i^j(\mathcal{K})|, |\mathcal{W}_i^j(\mathcal{S})|$ and the uncertainty set of T_a and d.

Theorem 8. *Let k be the number of weather fronts and P be the maximum number of phases across all fronts. Moreover, let, $\tau = \max_i \left\{ \overline{T}_{a_i} - \underline{T}_{a_i} + 1,\ \overline{d}_i - \underline{d}_i + 1 \right\}$. Then, $\mathcal{E}$ can be upper bounded as follows:*

$$\mathcal{E} \leq \tau^{4k*\Delta*P} \tag{C.12}$$

Proof. Note that:

$$\begin{aligned}
\mathcal{E} &= \Pi_{i=1}^{k} \Pi_{j=1}^{p_i} EP_i^{|\mathcal{W}_i^j(\mathcal{S})|+|\mathcal{W}_i^j(\mathcal{K})|} \\
&\leq \Pi_{i=1}^{k} \Pi_{j=1}^{p_i} \tau^{2*(|\mathcal{W}_i^j(\mathcal{S})|+|\mathcal{W}_i^j(\mathcal{K})|)} \\
&\leq \Pi_{i=1}^{k} \Pi_{j=1}^{p_i} \tau^{4\Delta} \\
&\leq \tau^{4k*\Delta*P}
\end{aligned}$$

□

The left side of Figure C-2 lists an upper bound on $\mathcal{E}$ for various combinations of τ and Δ. The main observation is that if any of the two parameters τ or Δ is less than 3, then $\mathcal{E}$ is in the thousands and is thus, quite likely to be less than the number of constraints. This table suggests that in many real-world instances, a single affine policy is likely to be optimal. In addition, the right side of Figure C-2 depicts the surface plot of the upper bound on $\mathcal{E}$ highlighting the speedy increase with an increase in τ and Δ.

τ	Δ	Upper Bound on $\mathcal{E}$	
		$P = 1$	$P = 2$
1	1	**1**	**1**
2	2	**256**	**65536**
2	3	**4096**	16777216
3	2	**6561**	43046721
3	3	531441	$2.8*10^{11}$

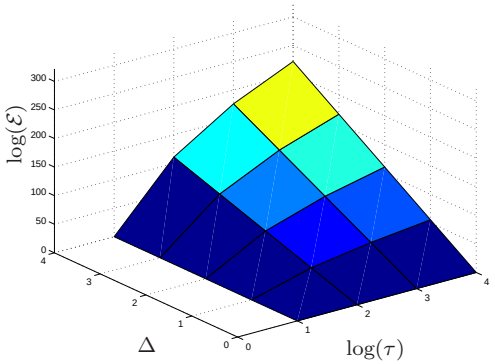


Figure C-2: **Left**: Upper bound on the number of extreme points in the weather-front based uncertainty set ($k = 1$). The bold numbers indicate the attractive cases from the point of view of obtaining a single affine policy; and **Right**: Plot of the upper bound on $\mathcal{E}$ ($k = 1, P = 2$).

Bibliography

[1] Bemporad A., Borrelli F., and Morari M. Min-max control of constrained uncertain discrete-time linear systems. *IEEE Transactions on Automatic Control*, 48(9):1600–1606, 2003.

[2] Ben-Tal A., Ghaaoui El L., and Nemirowski A. *Robust Optimization*. Princeton Univ. Press, 2009.

[3] Mukherjee A. and Hansen M. A dynamic stochastic model for the single airport ground holding problem. *Transportation Science*, 41:444–456, 2007.

[4] National severe weather playbook Air Traffic System Command Center, FAA.

[5] Kotnyek B. and Richetta O. Equitable models for the stochastic ground-holding problem under collaborative decision making. *Transportation Science*, 40:133–146, 2006.

[6] M. Ball, C-Y. Chen, R. Hoffman, and T. Vossen. Colloborative decision making in air traffic management: Current and future research directions. *ATM'99*, 1999.

[7] M. Ball, Hoffman R., Odoni A.R., and Rifkin R. A stochastic integer program with dual network structure and its application to the ground holding problem. *Operations Research*, 51:161–171, 2003.

[8] C. Barnhart, D. Bertsimas, C. Caramanis, and D. Fearing. Equitable and efficient coordination of traffic flow management programs. *Transportation Science*, 2012.

[9] D. Bertsimas and C. Caramanis. Finite adaptability in multistage linear optimization. *IEEE Transactions on Automatic Control*, 55(12):2751–2766, 2010.

[10] D. Bertsimas, G. Lulli, and A.R. Odoni. The air traffic flow management problem: An integer optimization approach. *Operations Research*, 59(1):211–227, 2011.

[11] D. Bertsimas and C. Maes. On the complexity of constructing adaptive policies in two-stage linear optimization. *Working Paper*, 2011.

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

[13] D. Bertsimas and S. Stock. The air traffic management problem with enroute capacities. *Operations Research*, 46:406–422, 1998.

[14] D. Bertsimas and S. Stock. The traffic flow management rerouting problem in air traffic control: A dynamic network flow approach. *Transportation Science*, 34:239–255, 2000.

[15] D. Bertsimas and R. Weismantel. *Optimization Over Integers*. Dynamic Ideas, 2005.

[16] Bertsimas D., Brown. D.B., and Caramanis C. Theory and applications of robust optimization. *SIAM Review*, 53(3):464–501, 2011.

[17] Bertsimas D. and Goyal V. On the power and limitations of affine policies in two-stage adaptive optimization. *Math Programming*, 2009.

[18] Bertsimas D. and Goyal V. On the power of robust solutions in two-stage stochastic and adaptive optimization problems. *Mathematics of Operations Research*, 35(2):284–305, 2010.

[19] G.B. Dantzig. Linear programming under uncertainty. *Management Science*, pages 197–206, 1955.

[20] FAA.

[21] D. Fearing. *The Case for Coordination: Equity, Efficiency and Passenger Impacts in Air Traffic Flow Management*. PhD thesis, MIT, 2011.

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

[23] R. Hoffman, A. Mukherjee, and Vossen T. Air traffic flow management. Technical report, 2007.

[24] Goh J. and Sim M. Robust optimization made easy with rome. *Operations Research*, 59(4):973–985, 2011.

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

[26] G. Lulli and A. Odoni. The european air traffic flow management problem. *Transportation Science*, 41(4):431–443, 2007.

[27] A.R. Odoni. The flow management problem in air traffic control. *Flow Control of Congested Networks*, pages 269–288, 1987.

[28] Air Transport Association of America Inc.

[29] Joint Planning and Development Office. Concept of operations for the next generation air transportation system, version 2.0. 2006.

[30] O. Richetta and A.R. Odoni. Solving optimally the static ground-holding policy problem in air traffic control. *Transportation Science*, 27:228–238, 1993.

[31] O. Richetta and A.R. Odoni. Dynamic solution to the ground-holding policy problem in air traffic control. *Transportation Research*, 28a:167–185, 1994.

[32] H. Sherali, J. Hill, M. McCrea, and A. Trani. Integrating slot exchange, safety, capacity, and equity mechanisms within an airspace flow program. *Transportation Science*, 45(2), 2011.

[33] M. Terrab and A.R. Odoni. Strategic flow control on an air traffic network. *Operations Research*, 41:138–152, 1993.

[34] M. Terrab and S. Paulose. Dynamic strategic and tactical air traffic flow control. Technical report, RPI, 1993.

[35] Bureau of Transportation Statistics U.S. Dept. of Transportation. Airline on-time statistics.

[36] T. Vossen and M. Ball. Optimization and mediated bartering models for ground delay programs. *Naval Research Logistics*, 53(1):75–90, 2006.

[37] T. Vossen and M. Ball. Slot trading opportunities in collaborative ground delay programs. *Transportation Science*, 40(1):29–43, 2006.

[38] P.B. Vranas, D. Bertsimas, and Odoni A.R. The multi-airport ground holding problem in air traffic control. *Operations Research*, 42:249–261, 1994.

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