Robust Crew Pairing: Delays Analysis and Implementation of Optimization Approaches

by

Pierre Leroyer

Ingénieur des Arts et Manufactures
Ecole Centrale Paris, 2004

Submitted to the Department of Aeronautics and Astronautics
in Partial Fulfillment of the Requirements for the Degrees of

Master of Science in Aeronautics and Astronautics

at the
Massachusetts Institute of Technology

September 2005
[February 2006]

Author..
Department of Aeronautics and Astronautics

Certified by..
Cynthia Barnhart, Thesis Supervisor
Professor, Civil and Environmental Engineering and Engineering Systems

Accepted by..
Jaime Peraire, Professor of Aeronautics and Astronautics
Chair, Committee on Graduate Students

Robust Crew Pairing: Delays Analysis and Implementation of Optimization Approaches

by

Pierre Leroyer

Submitted to the Department of Aeronautics and Astronautics
on August 19, 2005, in Partial Fulfillment of the Requirements for the Degree of
Master of Science in Aeronautics and Astronautics

Abstract

With increasing delays and airport congestion that disturb airline operations, the development of robust schedules is becoming crucial. Increased traffic and poor weather are a few of the causes of airport congestion, rising delays and lengthening passenger trips. In this thesis, we identify the latest trend in the flight arrival and departure delays, differentiating major U.S. airports from other smaller airports. We also quantify the types of delays airlines should work to mitigate. We then analyze the effects of schedules changes that were implemented by a major U.S. airline at their largest hub. We measure the effects of these schedule changes on on-time performance, taxi time, plane utilization, and passenger connection and total travel time. We also analyze how extensive is the practice of adding buffer time to flight times to improve schedule reliability. Finally, we propose and implement a new model to achieve robust crew schedules, that is, crew schedules that are less likely to be inoperable due to disruptions during operations. We show that with an increase in crew costs of 0.2%, we can decrease the number of times crews must connect between different aircraft by 32%.

Thesis Supervisor: Cynthia Barnhart
Title: Professor, Civil and Environmental Engineering and Engineering Systems

Acknowledgements

First, I would like to thank my thesis advisor Prof. Cynthia Barnhart. Our discussions were a great help for me to improve my work and develop my ideas. It has been a great joy to work with her. I would also like to address many thanks to Prof. Richard Dudley for his help in the statistical analysis, his passion for statistics and his very kind support along the year.

Many thanks to Curt Horjn, Stefan Karish and Martin Sorensen from Carmen Systems. Their help and enthusiasm has been great support to discover and utilize the Carmen Crew pairing software.

Then, I would like to thank my Rotary host counselor Ted Van Nahl and all the Rotary people I met in the US through my ambassadorial scholarship that supported me through the first semester.

Many thanks to all my friends at MIT and Harvard. I spent a wonderful year in the US and made so many great experiences with you. I thank my parents, my sisters and friends left in France. Their mental support was a great help.

Last but not least, special thanks to Katharina for a kindness and love that helped me to succeed.

Table of Contents

List of Figures

List of Tables

Chapter 1

Introduction

The crew scheduling problem is the last sub-problem to be solved in the *Airline Schedule Planning Process*. After defining the flight schedule, assigning aircraft types to the flight legs in the schedule, and routing individual aircraft to routinely visit maintenance stations, the final task in the Airline Schedule Planning Process is to ensure that every flight leg is covered by a crew and that total crew cost is minimized. The crew scheduling problem consists of two sub-problems, the Crew Pairing Problem to generate (partial) crew schedules, or pairings, that minimize crew costs and assign each flight leg to a single crew, and the Crew Rostering or Bidline Problem to assemble the selected crew pairings into month-long schedules.

1.1 The importance of delays

Recent indicators show that delays are rising again to the previous record levels experienced in 2000, so it is crucial that we to better understand their origins and how they propagate through the schedule so that we can reduce them. The majority of delays that disturb the airline schedule are created by external factors, such as, weather or airport and airspace congestion). However, some of the delays are created by hot spots in the airline process, or ineffective scheduling of aircraft, crew members or passengers. By identifying patterns of these types of delays that occur repeatedly, we believe that capturing their effects at the planning stage can lead to potentially significant reductions in delay during operations.

Our motivation stems from the fact that even a small reduction in delays can have an important impact on the airlines, in terms of their operations, image and financial results. Indeed, on-time performance has become one of the most important key performance indicators in the airline industry and an important service differentiator for customers, especially for valuable, high-yield customers. Booz Allen & Hamilton [6] estimate that airlines lose 0.6% to up to as much as 2.9% of their operating revenues as a result of delays Therefore, achieving even a small improvement in on-time performance through consideration of the delays during the planning process can potentially produce significant results.

1.2 Responses of the Airlines

Unlike Europe, where each airport is allocated a finite number of slots based on Instrument Flight Rules (IFR) conditions, all but 4 US airports allow airlines to schedule their flights, without any restrictions. As a result, sometimes the number of scheduled flights per unit of time exceeds the capacity of the airport. This inevitably results in airport congestion and delays, even in good weather conditions.
In their quest for on-time performance, airlines have adopted different means to reduce the delays they experience. At major airports that are often used as a hub by an airline, the dominant carrier can improve the situation for itself and all other carriers by de-banking its schedule. It requires giving up the common practice of operating a group of arrivals followed shortly by group of departures, in order that customers experience short connection times. By spreading its flights over the day, the dominant carrier can significantly improve operations at the airport. We'll study in this thesis the benefits resulting from the hub de-banking of a legacy carrier.

A second response of airlines to increases in delays is to add time into the schedule and incorporate into gate-to-gate time both air and taxi delays. This practice improves on-time performance directly. We'll study the policies of airlines and the quantities of buffer time they add to guarantee robustness of their operations.

1.3 Crew scheduling robustness

1. 3. 1. Definitions

A monthly crew schedule is composed of multiple *pairings*. A pairing is a sequence of *duties* that start and end at the same crew base. A duty is a set of flight legs covered by a crew in a day. Solving the crew scheduling problem involves finding a set of feasible crew pairings that cover all of the flights and minimizes crew costs while respecting the many rules imposed by the FAA or by regulatory and collective bargaining agreements. Some of the common rules include restrictions on the minimum and maximum connection time between two consecutive flights in a duty, the minimum and maximum rest time between duties in a pairing, and the number of duties in a pairing, etc. The cost of a pairing is usually the maximum of three quantities: the sum of the duty costs in the pairing, a fraction of the time away from base and a minimum guaranteed pay times the number of duties (Barnhart et al [3]).

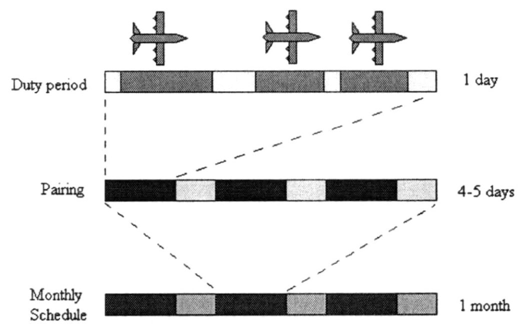

Figure 1.1: Decomposition of the monthly schedule in pairings and duties

1. 3. 2. Robustness

During operations, the assumption that every flight will be flown as planned and that every aircraft will arrive and depart on-time is erroneous. As a consequence, the realized cost of plans are greater, sometimes much greater, than those planned. Klabjan and Cherbalov [11] estimate that "the crew cost at the end of the month can be up to five times larger than the planned crew cost obtained by the optimal crew schedule". To decrease these additional crew costs, airlines can either use better recovery procedures or develop more robust crew scheduling solutions. In this thesis, we consider the second approach. By adding more slack to the crew pairing solution and allocating it wisely, we conjecture that solutions can be obtained that perform better during operations and achieve lower realized crew costs.

1.4 Thesis Objectives and Contributions

The objective of this thesis is to enhance crew pairing optimization models to capture the causes and effects of delays. We evaluate recent trends in delays, and airline responses to mitigate them, including de-banking of major hubs and adding buffer times to flight times. Lastly, we implement two robust crew scheduling models and discuss, for each model, the trade-off between robustness and crew costs.

1.5 Thesis Outline

This thesis is organized as follows. In chapter 2, we quantify the extent of flight delays in the US at the end of the first quarter of 2005. We use key performance indicators to measure the changes in delays and identify patterns of delay at major airports. In chapter 3, we analyze the effects of de-banking Delta's Atlanta hub, measuring delays and levels of congestion. In chapter 4, we analyze the effects of another common airline practice, that of adding buffer time to the schedule. We compare the policies of different airlines and their accomplishments in improving their on-time performance.

Finally, after reviewing robust scheduling models that are designed specifically to decrease delay propagation, we implement two models using the RAVE optimizer developed by Carmen Systems [7]. We then conclude with a short discussion of the trade-off between robustness and crew costs, and we suggest some new paths for research.

Chapter 2

Analysis of airline delays

2.1 Introduction

Flight delays and cancellations occur daily during airline operations. They have a direct impact on aircraft routes and crew schedules that might be disrupted or broken. Delays result from a broad range of causes. Some of them can be controlled by the airline, whereas others cannot.

The aim of this chapter is to draw a general picture of delays and to extract information that can be employed to improve the reliability of schedules. First, we identify causes of the delays from the point of view of airlines and from the point of view of the US Department of Transportation, and we compare their viewpoints. Then we look at the evolution of delays from different perspectives: yearly, seasonally and daily. Finally we address the pattern of delays for 10% of the major airports that receive 65% of the traffic.

2.2 Definition of the Causes of delays

2. 2. 1. The airlines' point of view

From the airlines' point of view, it is very important to identify delays and their causes. Front line people are responsible for reporting all delays that disturb operations. Delays are coded depending on their origin. The airlines use about 70 different codes to refer to all types of delays.

2. 2. 1. 1. Categories of delay

We studied reports of a major American airline company. The delay codes are aggregated into 9 different categories. We present a quick overview of them:

- *Airport services*: late loading of customers and/or bags, holding for connecting customers and bags, seat assignment duplication, soliciting over sale volunteers, inadequate resources to support the operation (skycaps, ramp services, etc…)
- *Technical services*: aircraft mechanical problem, adjusting, repairing or inspecting an aircraft, maintenance irregularity
- *Flight operations*: late release from system operations, crew disruption, unassigned crew member, holding for a connecting crew
- *Aircraft servicing*: late cleaning or supplying of aircraft
- *Catering/provisioning*: missing items, late provisioning,
- *In-flight service*: late crew to aircraft, late request for additional cabin service supplies, problem related to aircraft cabin where no maintenance is required.
- *System*: delay due to origin, enroute or destination weather, awaiting ATC clearance, substitution of an aircraft
- *Facilities*: failure of baggage system, ramps constructions interfering aircraft ground handling.
- *Miscellaneous* : damage to the aircraft discovered during a turn, failure of normal data processing or communications systems

2. 2. 1. 2. Categorization of delays

For the first 6 months of 2005 we plotted the relative importance of these categories in 2 ways: by their importance in minutes and in the number of reports they generated.

These front line data are very important in identifying the hot spots of the process that create delays. The following page presents our findings.

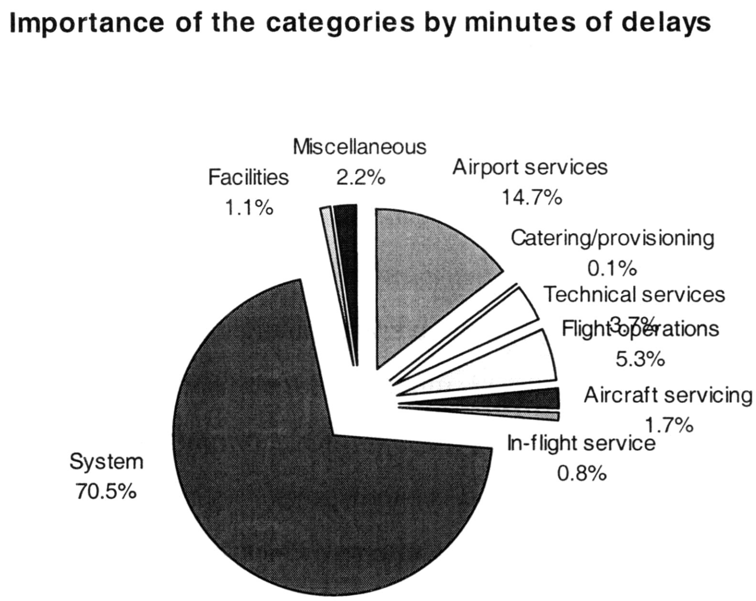

Figure 2.1: Delay categories by minutes (Major Airline, Jan – June05)

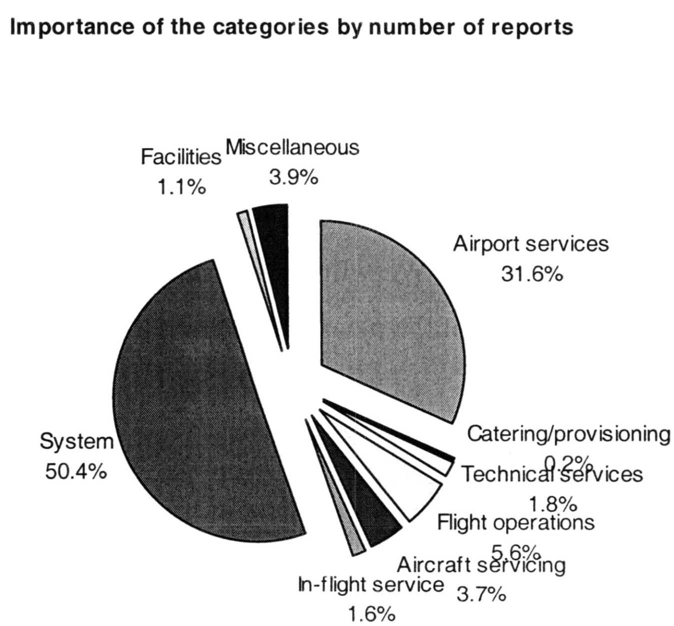

Figure 2.2: Delay categories by reports (Major Airline, Jan. – June 2005)

Figure 2.1 shows the very important role of "system" related delays (e.g., weather, heavy traffic, congestion at airports, lack of airspace, etc.). For this company, 70.5% of the minutes of delays result from the operations environment. These delays are responsible for 65% to 74% of the total delay minutes each month, with an average of 29 minutes of delay per report, as shown in Figure 2.3.

The cause of delay with the highest average number of delay minutes per report is the technical service category, with 42 minutes of delay per report. It refers to last-minute maintenance problems. Representing 5.3% of the total minutes of delay, it turns out to be the 4^{th} most important category behind system, airport services and flight operations. Yet, according to the opinion of a senior United Airlines pilot, the occurrence of maintenance delays is rising with the current attitude of the airline to increase outsourcing of maintenance. United Airlines decreased its maintenance staff from 12,000 mechanics to only 5,000 within a few years, and decreased the number of its maintenance stations by the same ratio. Outsourcing maintenance tends to make repair times longer, with potentially costly increases in delays. Indeed, according to La Mont [12], outsourcing heightens the risk of delays, because outsourced parts often have problems that need to be fixed before they can be put in service. Moreover, the logistics of getting the equipment to and from vendors are more complicated than having work done in-house.

Figure 2.2 shows that an important number of reports of delays involves airport services, with an average of 9.7 minutes per report. We will discuss later how flight operations delays are strongly connected with the quality of the planning function.

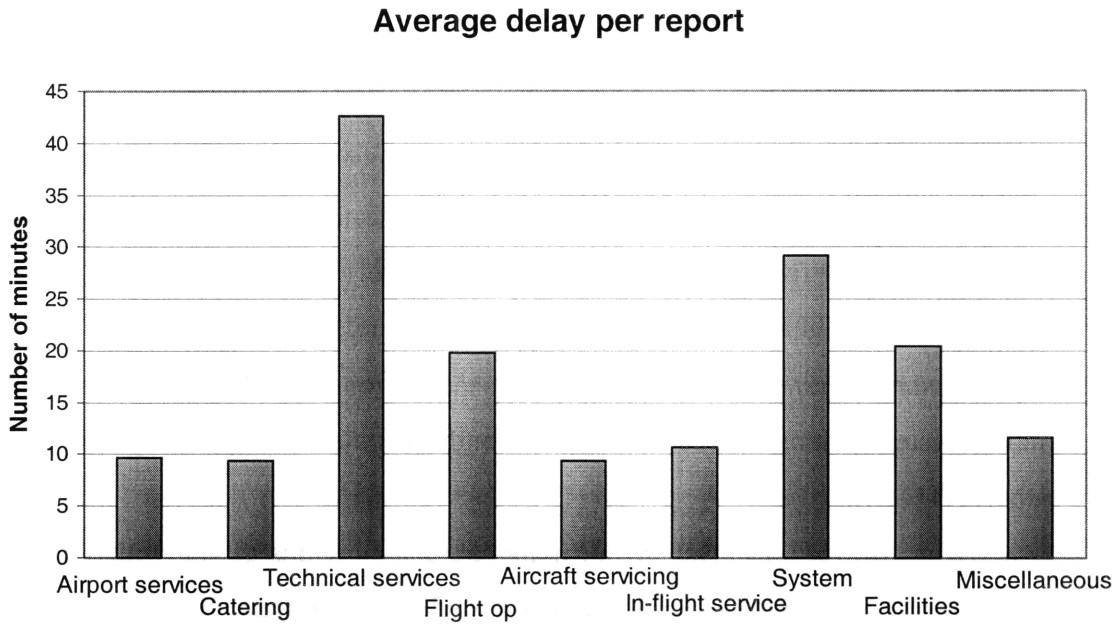

Figure 2.3: Average delay per report among the different categories

2. 2. 2. The Bureau of Transportation's point of view

2. 2. 2. 1. Source of the data

Within the US Department of Transportation, the Bureau of Transportation Statistics (BTS) collects and tracks flight data of the major domestic airlines (with more than 1% of total domestic scheduled passenger revenue). The number of reporting carriers varies between 10 (1997) and 20 (2005). The 9 major carriers present since the beginning of data collection are Alaska Airlines, America West Airlines, American Airlines, Continental Airlines, Delta Air Lines, Northwest Airlines, Southwest Airlines, United Airlines, and US Airways. Reporting airlines receive more than 90% of the total domestic operating revenues each year. Hence, we'll consider that the figures of these reports well represent airline industry trends. Later, we will compare this general trend to that at specific specifics airports to determine if a change in performance is due to better operations at the airport or just correlated with an overall improved performance of the entire air transportation system. The data we will use in this thesis is accessible at the following address:

http://www.bts.gov/programs/airline_information/

2. 2. 2. 2. The BTS categories

The BTS identifies five broad categories of delays:

- *Air Carrier*: Delays or cancellations attributable to the airline's operations (maintenance or crew problems, aircraft cleaning, baggage loading, fuelling, etc.).
- *Extreme Weather*: Delays or cancellations attributable to significant meteorological conditions (actual or forecasted) that, in the judgment of the carrier delays or prevents the operation of a flight (e.g. tornado, blizzard, hurricane, thunderstorm, etc.).
- *National Aviation System (NAS)*: Delays and cancellations attributable to the National Aviation System that refers to a broad set of conditions that we will detail further (e. g. weather, ATC …).
- *Late-arriving aircraft*: Delays are attributable to a previous flight with the same aircraft arriving late and causing the following flight to depart late.
- *Security*: Delays or cancellations caused by evacuation of a terminal or concourse, re-boarding of aircraft because of security breach, inoperative screening equipment and/or long lines in excess of 29 minutes at screening areas.

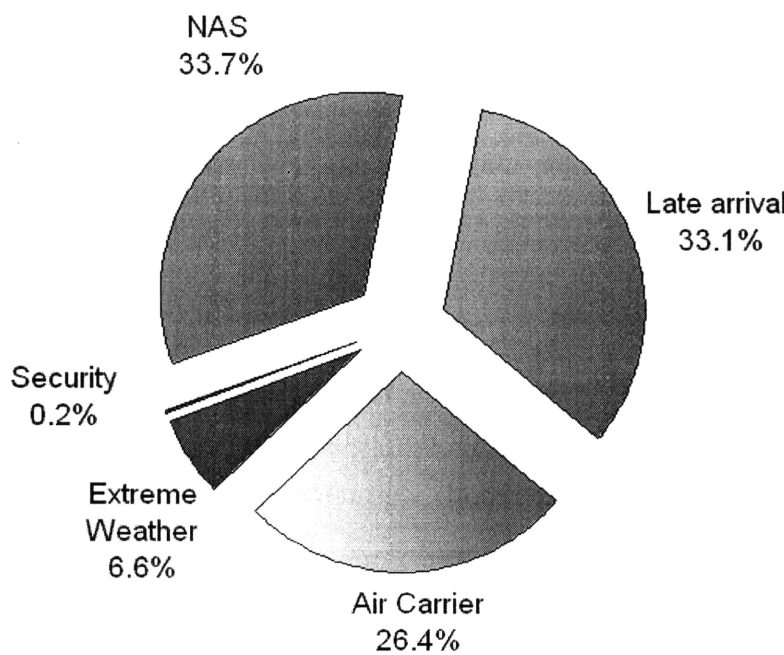

Figure 2.4: The causes of the disrupted flights (June 2003-June 2005)

Figure 2.4 represents the categorizations of the total number of delay minutes for the period (June 2003-June 2005). It identifies 3 main causes of delays: NAS delays, propagated delays (aircraft arriving late), and air carrier delay.

The main cause of delay with 33.7% of the total minutes of delay is the National Aviation System (NAS). The NAS is responsible for imposing some limits on airline traffic due to congestion of airspace or congestion of airports. Weather plays a very important role in these congestion problems.

Propagated delays are nearly as important with 33.1% of the total delay minutes. These delays get more and more important as the day progresses. However, different scheduling practices could probably avoid some of the delay propagation and amplification throughout the day.

The last important category, with 26.4% of the delay minutes, is carrier delays. In this category, the airlines can again change their schedules to decrease delays. We will review and propose some scheduling models in the last part of this thesis that target this objective, and focus on the scheduling of crews.

2. 2. 3. Composition of NAS delays

The National Aviation System accounts for the most number of minutes of delays. NAS delays refer to a large number of "system" causes: non-extreme weather, heavy traffic volume, airport operations (equipment, closed runway), etc.

Figure 2.5 shows the importance of weather (responsible for 77% of the NAS delay minutes). Overall, we estimate the importance of weather. Extreme Weather accounts for 6.6% of NAS delays and "normal" bad weather accounts for 77%*33.7% of NAS delays. We add the weather delays resulting from propagated delays (33.1%), and the weather delays resulting from the propagated delays of the propagated delays and so on ...

In total, bad weather is responsible for 48.7% of the total NAS delays:

$$WeatherDelay = (\%WeatherInNas \times \%NASdelay + \%EtremeWeather) \times \sum_{i=0}^{\infty} (\%PropagatedDelay)^i$$

$$WeatherDelay = (\%WeatherInNas \times \%NASdelay + \%EtremeWeather) \times \frac{1}{1 - \%PropagatedDelay}$$

$$WeatherDelay = (77\% \times 33.7\% + 6.6\%) \times \frac{1}{1 - 33.1\%} = 48.7\%$$

The difference between Extreme Weather and Weather delay is explained by the BTS [5]. Weather delay corresponds to non-extreme weather delays that could be reduced with corrective action by the airports or the FAA. In the previous plot, extreme weather refers to delays that cannot be reduced by corrective action because of significant meteorological conditions, actual or forecasted, at the point of departure, en route, or point of arrival.

The second interesting element that we find in the composition of the NAS delay is the heavy volume delays. They represent 13% of total delays. These delays are related to the congestion problems and are likely to increase with increases in traffic. Already, airplanes can be held on the ground because there is no airspace available to fly to their destination. On the Figure 2.5, we can see that an en-route severe weather reroute some traffic over Atlanta so that some of the airplane in Atlanta are held on the ground because they have to space to insert themselves in the air traffic toward the North East Cost.

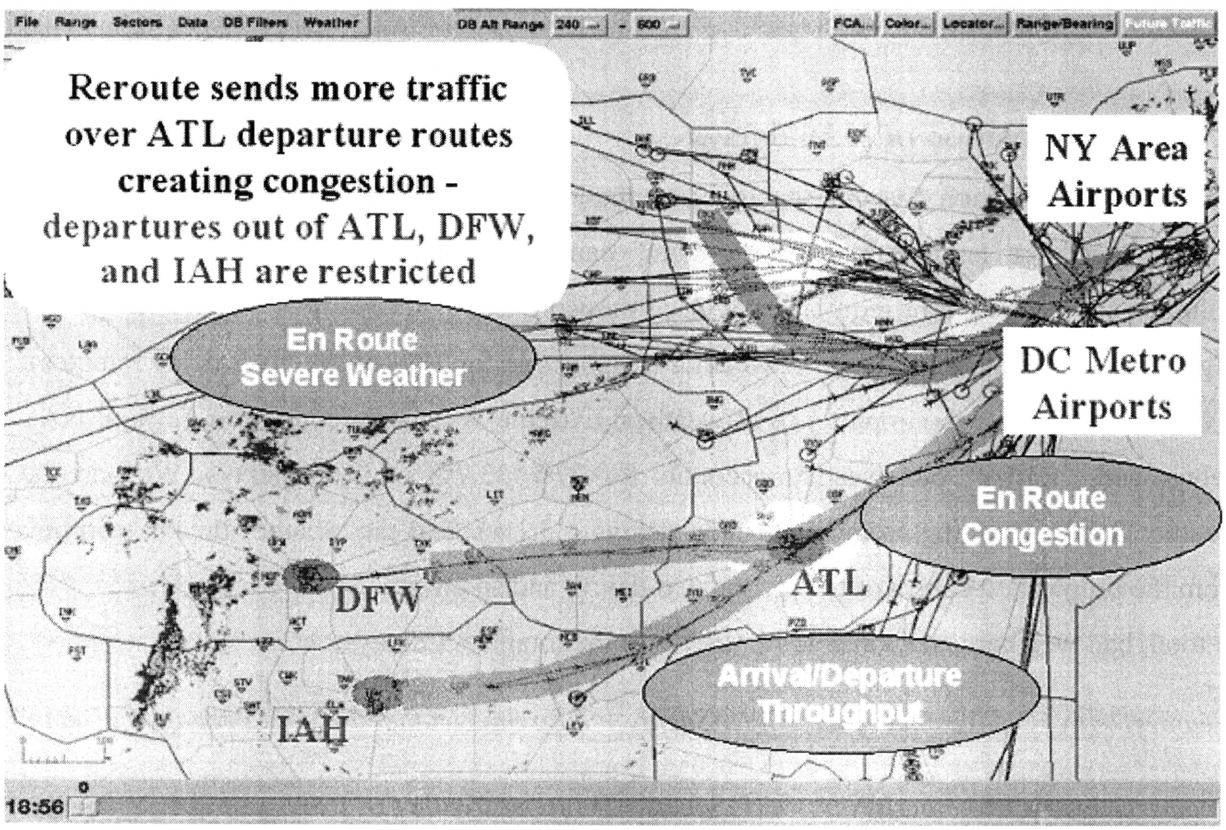

Figure 2.5: The causes of the disrupted flights (June 2003-June 2005)

Source : [15] The MITRE Corporation, Anatomy of Air Travel Delays - The Scenarios

2. 3. Evolutions of delays

Looking at the delay in a static way and taking the mean of the delay gives an idea of its importance, however it doesn't capture well all characteristics. Indeed, delays are not normally distributed (the median and the mean are very different because the mean is influenced by all the delay outliers). Moreover, delays evolve with time and have different shapes from a yearly, seasonal or daily perspective.

2. 3. 1. Industry trend

2. 3. 1. 1. Yearly evolution of on-time performance

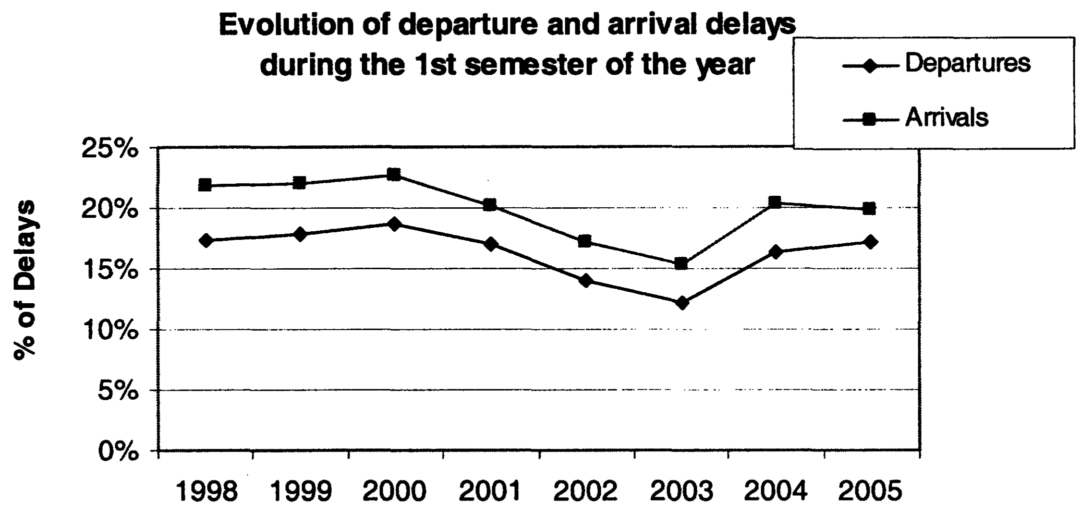

Figure 2.6: Departure and arrival delays for the first semester of the year (1998-2005)

Figure 2.6 shows that arrival delays are on average 3.6 percent higher than departure delays. A part of this difference comes from the fact that flexibility at the airport can absorb a part of the delays (flexibility takes the form of scheduled slack time, aircraft swapping, crew swapping, etc.). Another part of this difference is that some of the flights that depart on-time from the gate will be held on the ground because of taxi-out or NAS delays, and this will have a consequence on the on-time arrival performance if not enough time was scheduled for the flight.

2. 3. 1. 2. Taxi-out time

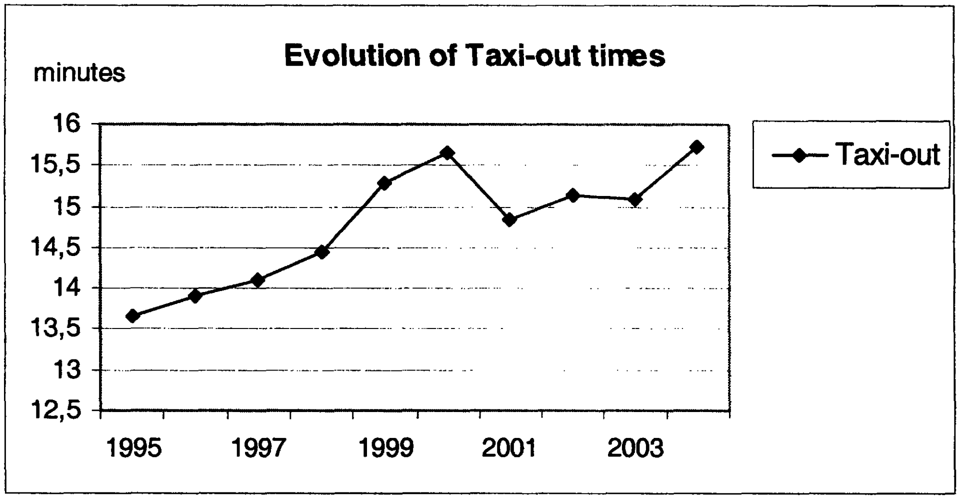

Figure 2.7: Taxi-out time for US domestic flights

Taxi-out is the time between the departure of an aircraft from the gate and its take-off time. It's a better indicator than the on-time performance of the real congestion of the airport because the on-time performance can be improved by an increase in the block time, whereas the taxi-out time can't be increased so easily.

We see that the taxi-out time went up until 2000, when it decreased with the decrease in traffic. Since then, it has been growing again so that in 2004, even though the percentage of delayed flight is lower than in 2000 due to block time adjustments (discussed in chapter 4), the average taxi-out time, and as a consequence, real congestion at the US airports is greater than in 2000.

2. 3. 2. Seasonal variations

When planning, it is important to take into account seasonal variations of delays. Figure 2.8 and Figure 2.9 show that delays and cancellations are much more prevalent in June, December and January. Weather and heavy summer traffic are the main causes of these increases in delays.

In the summer period, there is a 3-4% increase in the number of scheduled flights compared to the rest of the year.

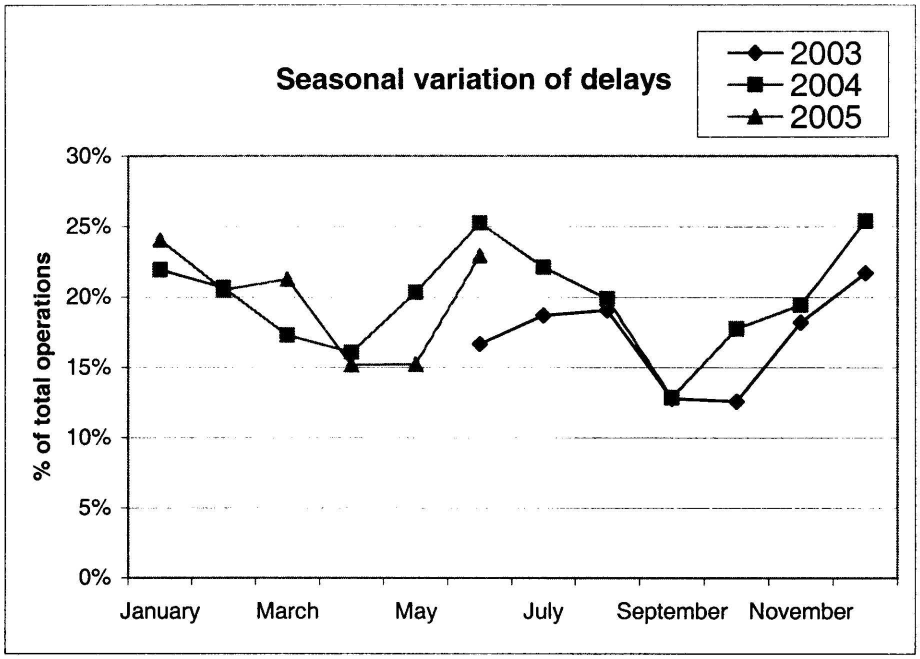

Figure 2.8: Seasonal arrival delays variations

We also notice a small increase from one year to another in the percentage of cancelled flights, shown in Figure 2.8, which impact delays by reducing the amount of propagated delay.

From 2002 to 2005, the percentage of cancelled flights increased from 1.31% to 1.90%.

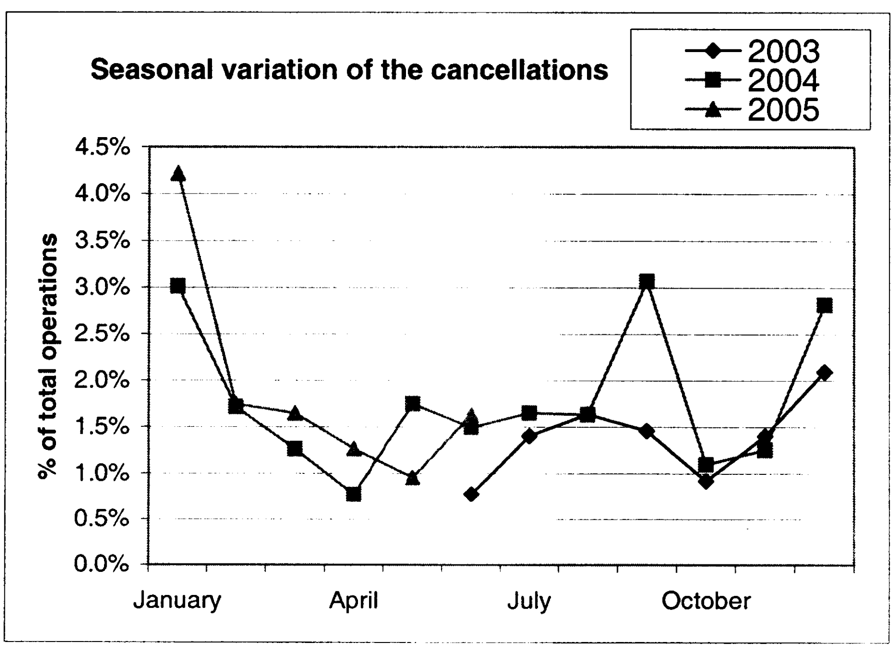

Figure 2.9: Seasonal cancellation variation

2. 3. 3. Variations during the day

Figure 2.10 shows the evolution throughout the day of flight delays, by category. Propagated delays grow trough the day and become the major contributor to average delay after 5pm. Earlier in the day, NAS delays are the major contributor, except at the very beginning of the day, when carrier delays are the ,major contributor.

This amplification of propagated delay is mainly due to the fact that the airlines don't have enough slack in their schedule to absorb the delay generated. In fact, these delays are often propagated by the airlines' schedules, with tight connections for crews and short turn times for aircraft. We will show in a later chapter the delay propagation resulting from crews connecting between different aircraft, and we will present ideas for reducing these connections with limited costs.

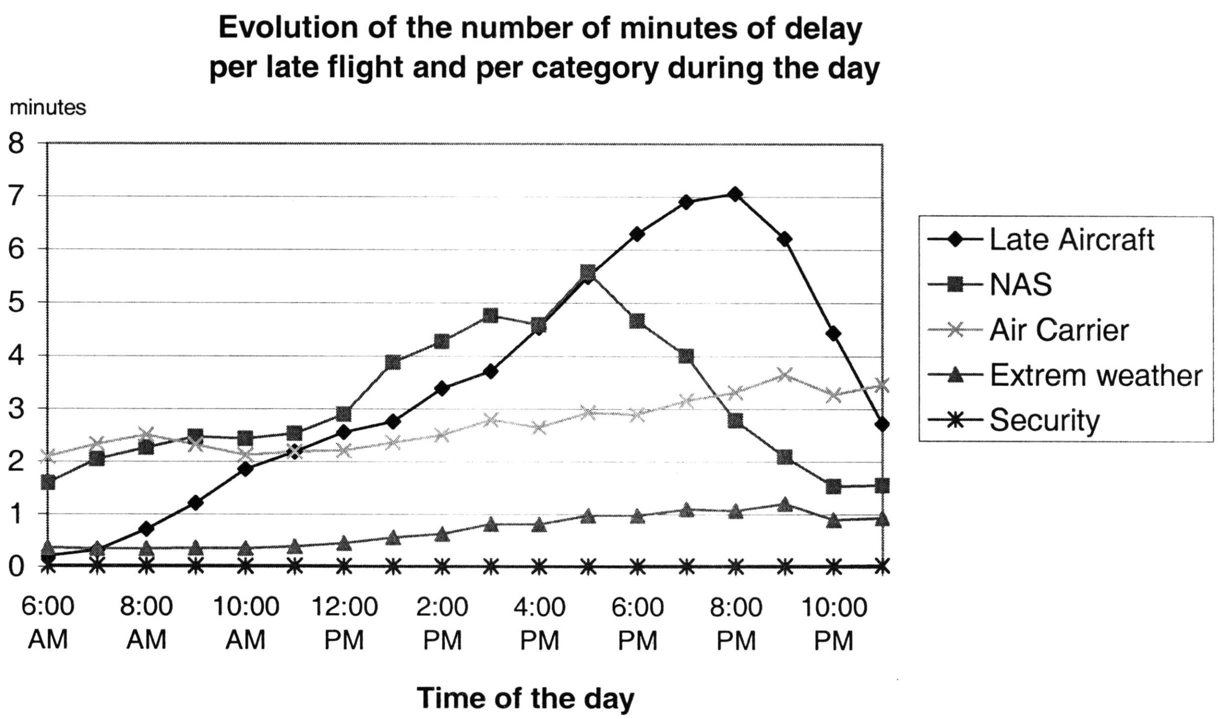

Figure 2.10: Summary of the evolution throughout the day of flight delays, by category

2. 4. Delays at Major Airports

Among the 286 US airports that serve more than 10,000 passengers per year, 10% of them serve 65% of the air traffic. Many of 33 major airports (listed by BTS [4]) are used as a hub for one or more airlines. Therefore, traffic conditions at these airports impact significantly on the rest of the system.

2. 4. 1. On-time performance

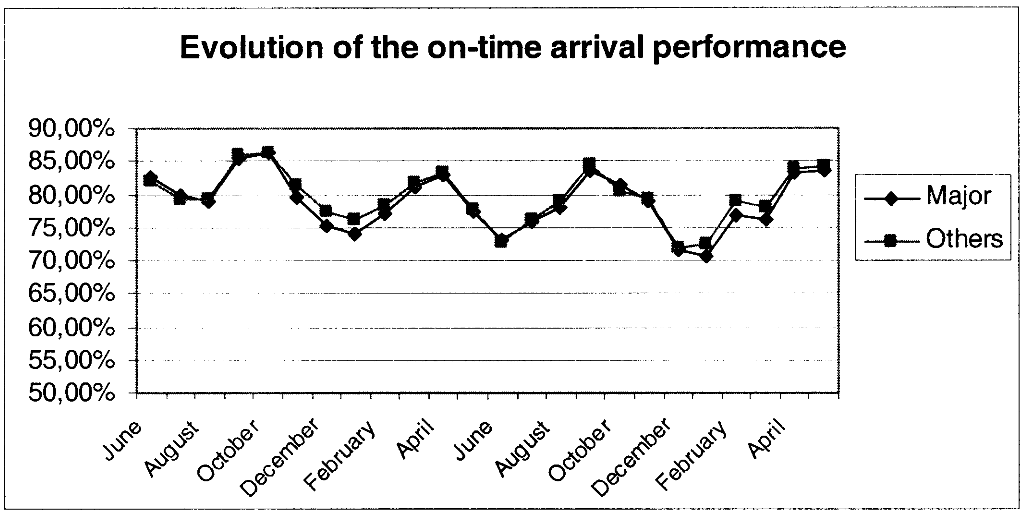

Figure 2.11: Comparison of the on-time performance (June 2003-May 2005)

On-time arrival performance of major airports and secondary airports (Figure 2.11) are very close: 78.8% for the major airports and 79.5% for the secondary airports over the 2 year period investigated. This is due to the fact that operations at the other airports depend substantially on the operations at major airports. Secondary airports have many flight legs to major airports, and those to secondary airports are often delayed by late arriving aircraft or crews, or other propagated delays from the major airports.

2. 4. 2. Taxi-time

We do not conclude, however, that major and secondary airports operate the same. By looking at taxi time, we can better identify the differences in congestion levels at these 2 types of airports. The taxi-time is 6 minutes longer on average at the 33 major airports (see Figure 2.12).

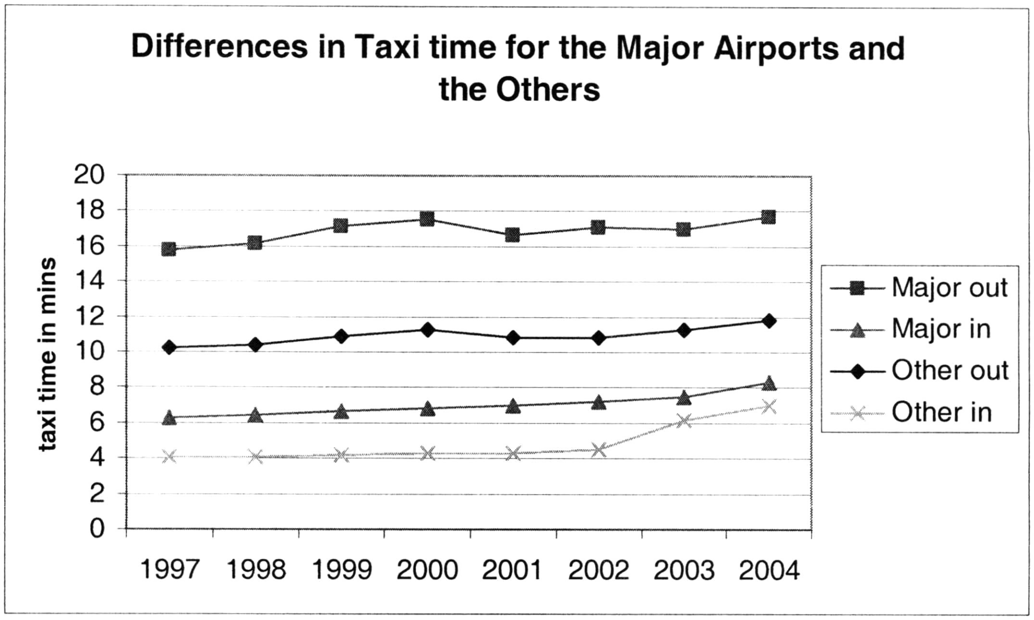

Figure 2.12: Difference in taxi times

Taxi-in time is also, on average, 2 minutes higher for major airports. This observation might be linked to the fact that the distances between the runway system and the terminal gates are higher in the case of the major airport.

2. 4. 3. NAS delays

Another difference in delays at major and secondary airports is NAS delays. Indeed, 80% of the total NAS delay in 2004 occurred at major airports, whereas they represented only 65% of the total aircraft movements. For major airports, NAS delays cause, on average, 2% more of the delays than at secondary airports, as shown in Figure 2.13.

If weather is the same on average at major airports as secondary airports, this difference can be explained by the heavier traffic volumes at major airports.

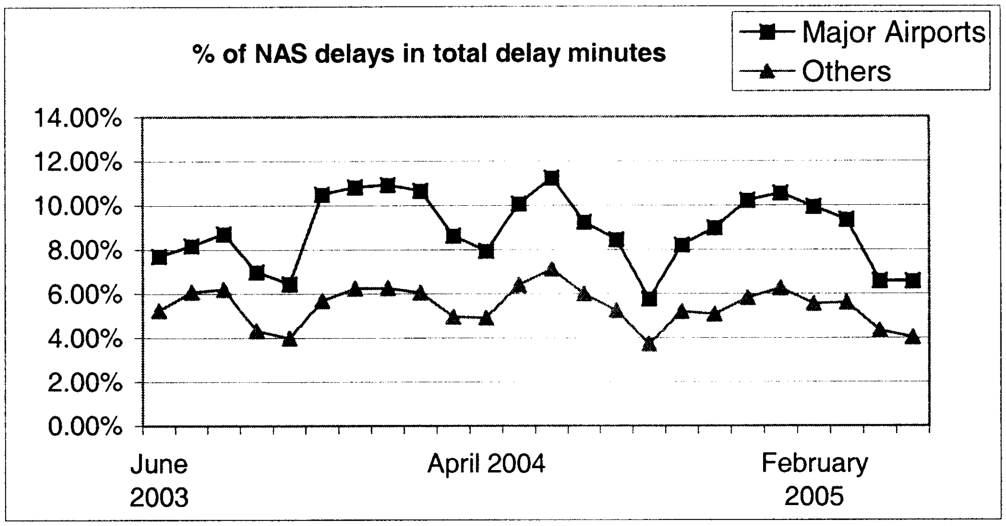

Figure 2.13: Difference of NAS delays between Major Airport and the others

2. 5. Conclusion

In this chapter, we identify the causes of delays and study their static and dynamic importance. Most of the causes of delay are independent of the airline, with weather accounting for 43% of total delay minutes. Where relevant, we point out opportunities for airlines to decrease delay propagation through scheduling and different operating procedures.

In this chapter, we identify the diverse set of parameters that airline schedulers should take into consideration, including industry trends, seasonal variations, time of the day variations, taxi-out times, and airport type.

In the next chapter, we describe one airline's attempt to mitigate delays through major schedule changes, especially at its hub airports. We study the operating benefits accrued and the impacts on delays.

Chapter 3

Delta Airline's De-banking of Atlanta Hartsfield Airport

3.1 Introduction

3.1.1. Banking and de-banking the major hubs

According to Bogusch [4], after deregulation in 1978, established carriers decided to compete not on fare but primarily on service (more itineraries and more frequency). The establishment of hub-and-spoke networks enabled the airlines to serve small communities, and to offer more frequency and more destinations. At these hubs, Bogusch explains that the priority was then to create short connections for passengers to minimize their total travel time and make the itinerary more likely to be selected by the travel agency booking system. Therefore, the airlines scheduled a bank of flights arriving in a short period of time followed by a bank of departing flights about 30-45 minutes later to enable the passengers to change airplanes. These were called "banked" hubs.

However, Bogusch believes that the recent changes in the airline industry (including Internet booking that gives better price information to the customers, competition from low cost carriers, etc.) and the costs of the banked structured raises some questions about the viability of banked operations, which create delays at the airport and are expensive to operate.

Continental was the first airline, with Newark in 1999, to de-peak one of its hubs. Continental spokesman David Messing claims that "it has been the real key to the improvements … seen at Newark."

In April 2002, American Airlines de-bank its hub in Chicago, O'Hare International Airport (ORD). It decided to spread the flights throughout the day. Bogusch shows that without losing market share, the operation was neutral or favorable from an operations perspective and likely favorable from a cost perspective.

The same year, American Airlines de-banked its other hub in Dallas-Fort Worth. Agbokou [1] argues that the benefits brought by this transformation include increased aircraft utilization; decreased operating costs, less congestion at the airport, on the taxi-ways, on the runways and at the gates. All of these effects occur without much increase in the average passenger connection time.

Our study will analyze the benefits of de-banking Delta Airline's hub in Atlanta, measuring effects on operations, on-time performance, and congestion levels at the airport. To conclude, we compare the level of progress of Delta after de-banking with that of American Airlines after de-banking.

3.1.2. Characteristics of the Atlanta Airport

Hartsfield-Jackson Atlanta International Airport is the world's busiest passenger airport (83,606,583 passengers in 2004) with 964,858 aircraft operations. Figure 3.1 shows the recent trend in the yearly passengers' traffic in Atlanta. The number of passengers is increasing again after a drop in 2001. This year has experienced a 5.35% increase so far in the number of passengers compared to last year. Figure 3.2 shows the dominance of Delta at the airport compared to the rest of the carriers. Atlanta is Delta's main connecting hub, serving numerous destinations around the globe. Because of Delta's dominance of the Atlanta airport, we'll show that on-time performance of the airport is correlated with the on-time performance of Delta.

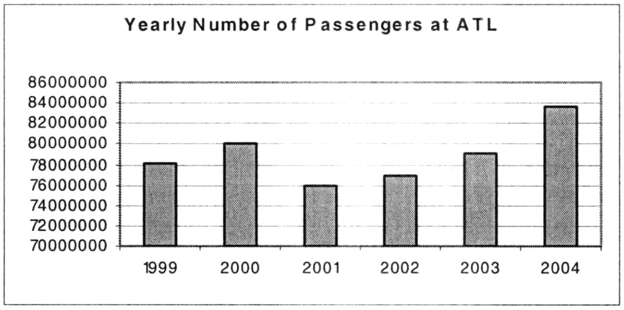

Figure 3.1: Traffic in terms of number of pax

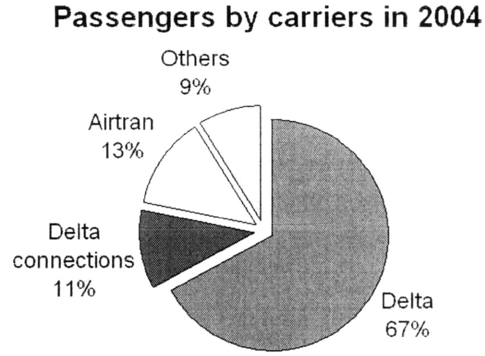

Figure 3.2 Repartition of the pax

3.1.3.: Airport runway capacity

Atlanta has a runway capacity of 180-188 movement per hours in optimum weather and 158-162 in IFR conditions, according to the Airport Capacity Benchmark [2]. Some more capacity will soon be added with the addition of a new independent runway in 2006. The airport capacity will increase to 237 movements per hours in optimum weather and 202 in IFR conditions.

3.2 Overview of schedule changes

3. 2. 1. Characteristics of schedule changes

On January 31 2005, Delta implemented Operation Clockwork, the "single largest schedule transformation in aviation history" according to Gerald Grinstein, Delta's CEO. Indeed, the airline restructured 51% of its network. A major point of this transformation was to redesign the hub in Atlanta: de-peaking the schedule and spreading flights over the day time while adding more flights to surpass every other airline in history in the number of flights operated from any one city.

Figures 3.3 and 3.4 enable us the compare the schedule change of Delta mainline flights at Atlanta at the end of January 2005. In the former schedule, there are 12 peaks of arrivals followed by 12 peaks of departures, each separated by about 45 minutes.

For the February schedule, Delta spread flights throughout the day so that there is no 15-minute interval between 6:30 am and 11pm when there is not at least one departure and arrival scheduled. The number of periods with more than 15 departures scheduled per 15 minute interval is decreased from 14 to 3 in the new schedule.

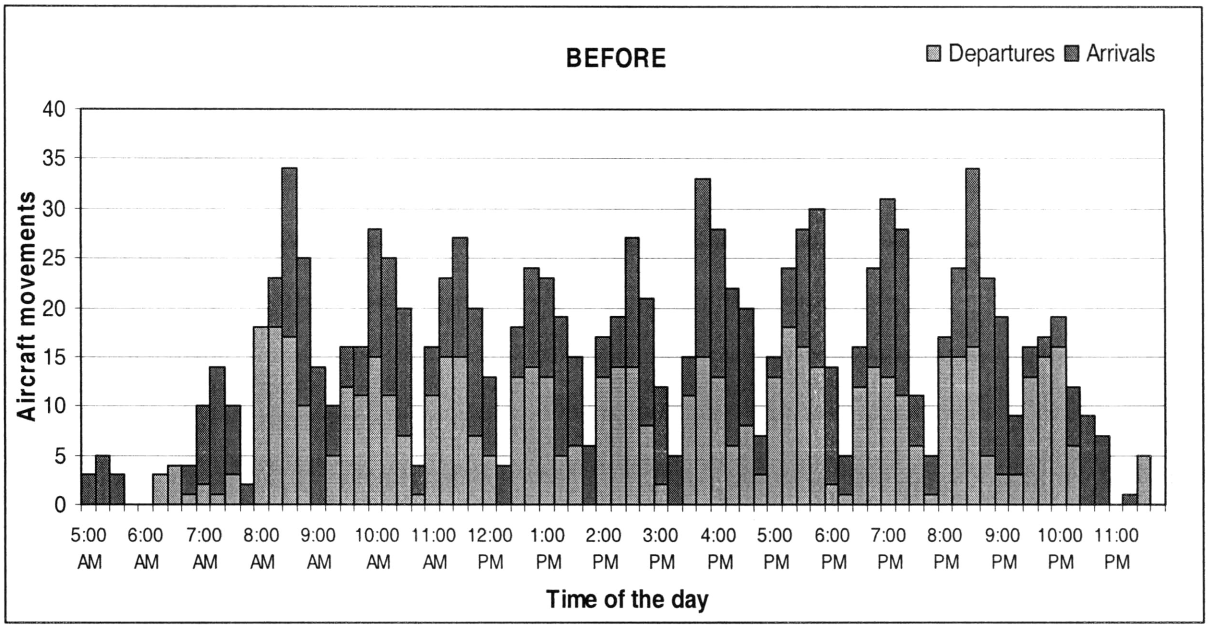

Figure 3.3: Aircraft movements of Delta mainline in January 2005 at Atlanta

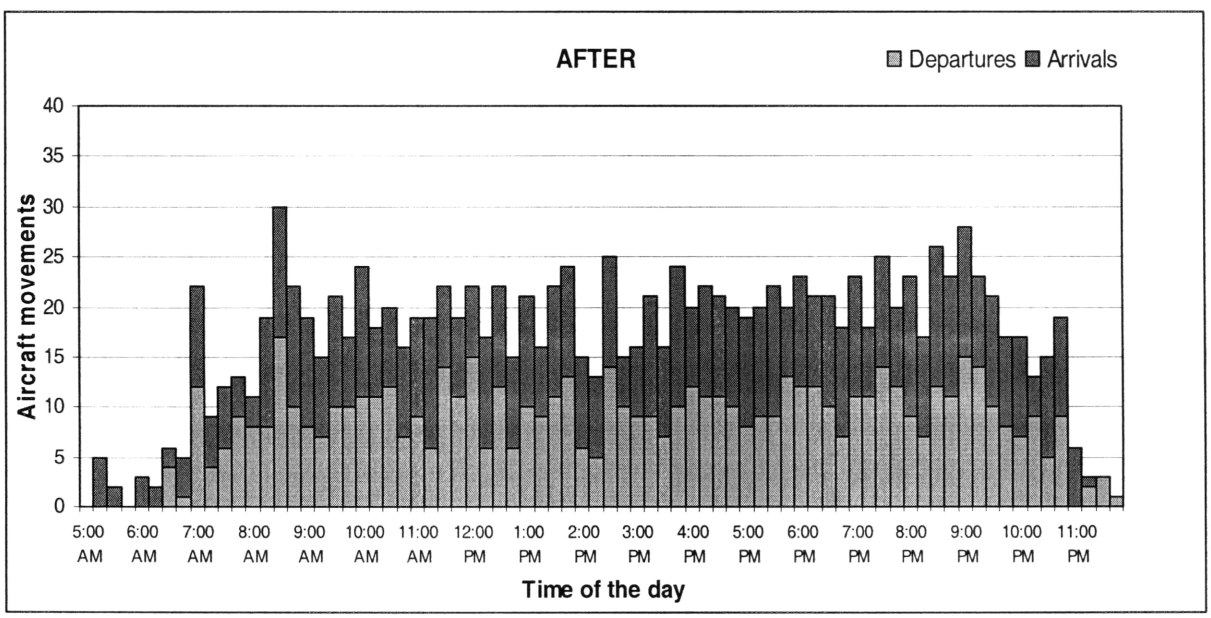

Figure 3.4: Aircraft movements of Delta mainline in February 2005 at Atlanta

What is remarkable is that at the same time, Delta added 63 new flights to the schedule, enabled by their increase in aircraft utilization and their de-hubbing of operations at Dallas. The result is an increase of 10% in mainline departures from Atlanta (Table 3.1), without an increase in fleet size.

ATL Schedule	Total Flights	Number of Nonstop Destinations	Mainline Flights	Delta Connection Flights	Seats per Departure (DL &DCI)
Jan. 2005	970	186	625	345	126
Feb. 2005	1,051	193	688	363	126

Table 3.1: Overview of the flights change in ATL

Source: Delta news

3. 2. 2. Comparison with other Delta hubs

The same day that they de-peaked the Atlanta hub, Delta stopped using its Dallas hub, decreasing the number of daily departures from Dallas from 258 to 21. In Cincinnati and Salt Lake City, Delta preserved its peaked schedule (Figures 3.5 and 3.6)

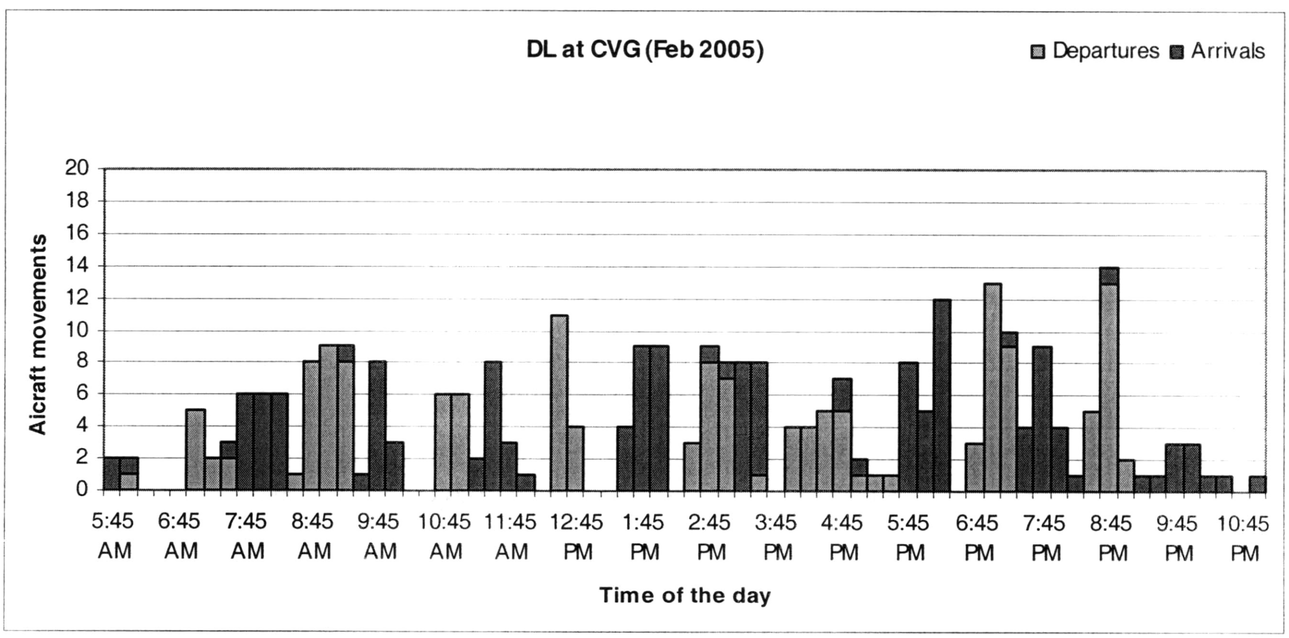

Figure 3.5: Delta mainline aircraft movements in February 2005 at Cincinnati

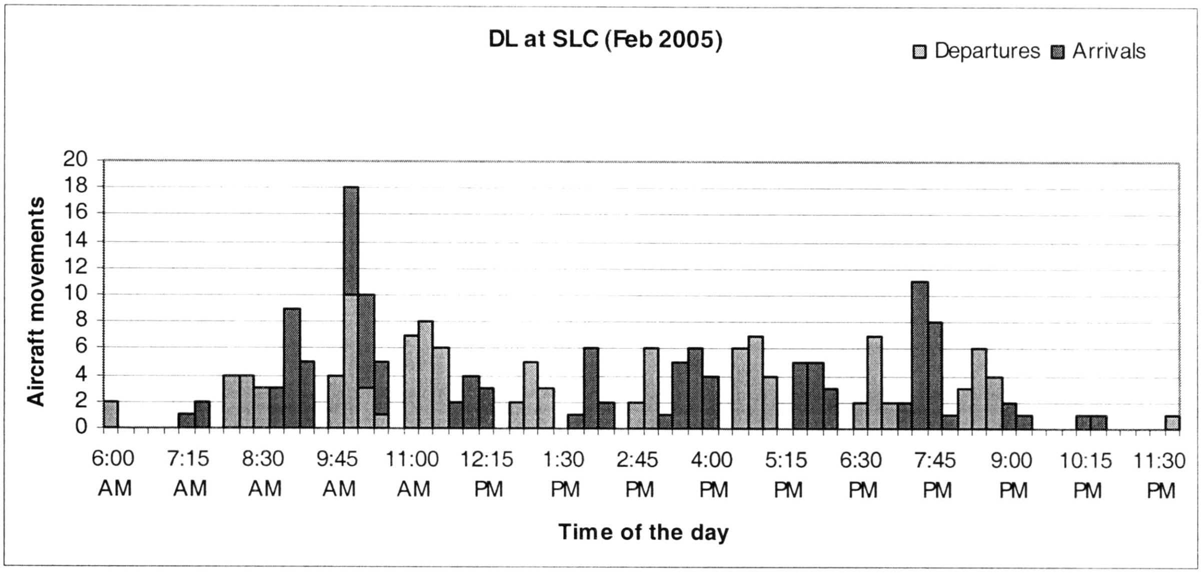

Figure 3.6: Delta mainline aircraft movements in February 2005 at Salt Lake City

We define the *peaking degree* of a departure schedule as the average number of departures scheduled per 15 minute interval divided by the standard deviation of the group. (Peaking degree =Average/ StDev).

These results show that the peaked scheduled in Atlanta during January had a smaller peaking degree than the other 2 hubs. This is mainly due to the fact that Atlanta had 12 peaks between 6:30 am and 11pm that were very tightly scheduled, sometimes without breaks between them. Hence, the schedules at Cincinnati and Salt Lake City are more peaked than in Atlanta during January due to the gaps between groups of departing flights in Cincinnati and Salt Lake City.

	ATL Jan 2005	ATL Feb 2005	CVG	SLC
Average number of departures per 15 minute intervals	9.25	9.95	2.41	1.76
St Dev	5.82	2.72	3.57	2.63
Peaking degree	0.63	0.27	1.48	1.49

Table 3.2: Degree of peaking of Delta's hub airports

3.3 Analysis of the effects of de-peaking on airport delays

3. 3. 1. Methodology

When analyzing the effect of a schedule modification, we must be aware that other factors can also influence the performance indicators that we compare. In her analysis, Agbokou [1] defines 4 factors that can disturb the performance analysis: seasonality of traffic, industry trends, one-time shocks and incremental changes (block time adjustments, boarding procedures, etc.).

In our analysis, we compare the operations during the spring of 2004 (March to May) with the spring of 2005 (same months) in order to normalize the effects of weather on operations. (We assume that the spring weather in Atlanta was similar in 2004 and 2005).

We will use the U.S. mean, our industry trend indicator, enables us to compare the evolution of performances of overall airline traffic with the evolution of performances at Atlanta with the de-banking of Delta.

Because air traffic is somewhat reduced during weekends (see Figure 3.7) and airlines perform slightly different schedules on these days, we limit our analysis to the weekdays (Monday through Friday).

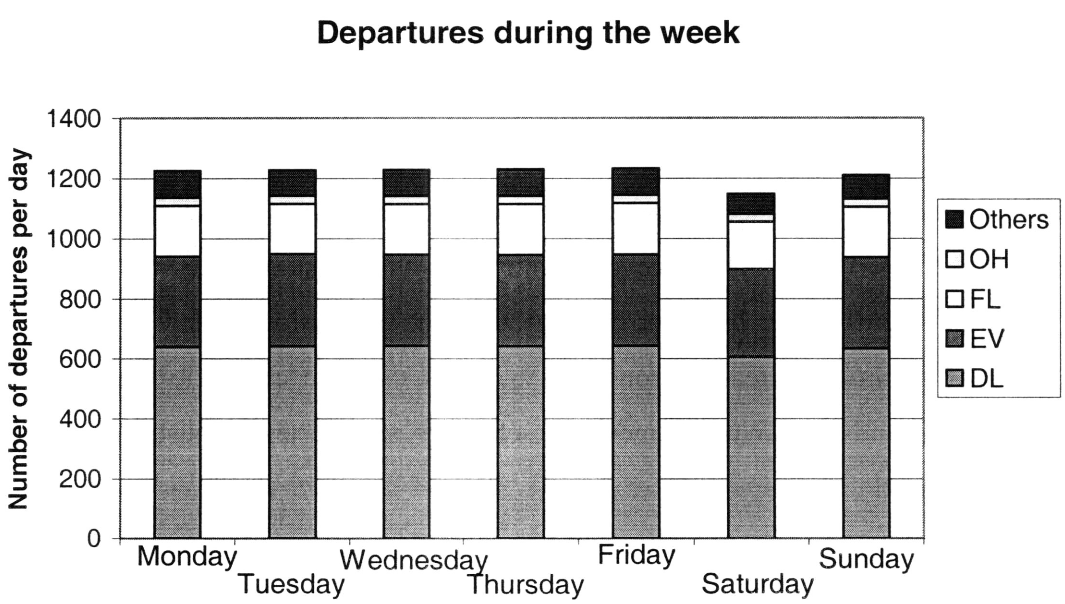

Figure 3.7: Number of departures per day

3. 3. 2. On-time departures

We start by looking at on-time departures. On-time departures are influenced by many external factors, such as weather and congestion. However, it is a good indicator to assess the quality of an airline's operations because it represents an airline's ability to get aircraft, crew, aircraft service and passengers aboard on-time. At first glance, the benefits of the schedule change appear a bit disappointing. Indeed, the on-time departure percentage of Delta decreased from 84.7% in spring 2004 to 82.5% in spring 2005. However, we notice that the on-time departure percentage of the other airlines operating in Atlanta was reduced 5.9% from 81.4% to 75.6%. (Table 3.3).

	2004	2005	Variation
Delta	84,7%	82,5%	-2,1%
Others	81,4%	75,6%	-5,9%

Table 3.3: Variation of on-time departure percentages

The decrease in the on-time departure percentage is mainly due to the month of March 2005, during which less than 72% of the flights departed on-time. During that month, Atlanta experienced a lot of bad weather, establishing, in particular, a new daily record of 2.87 inches of rainfall on March 27th.

The relatively good performance of Delta can be attributed to its schedule changes. We will study in further detail at which stage of the process the most benefits occur. We note immediately that from Spring 2004 and Spring 2005, the proportion of delay minutes due to "late arrival of aircraft" increased from 27.1% to 39.7%. We conclude that with higher aircraft utilization and shorter turn times, the schedule became less robust and more delay propagation occurred.

3. 3. 3. Taxi-out at Atlanta

Taxi-out is the time between the departure of an aircraft from the gate and its take-off time. For a single flight, it depends on the distance from the gate to the runway, the runway configuration, the rate of arrivals and of course, the congestion levels at the airport. By considering Spring 2004 and Spring 2005 aggregated data, we assume that the different runway configurations and the weather are the same on average, so that the difference of taxi-time is related to a difference in the queuing time of the aircraft before take-off.

A decrease in taxi-out time appears to be one of the major contributions of de-banking. Indeed, between 2004 and 2005, we have a decrease of 20% in the taxi-out time for Delta (Table 3.4), translating to 106,155 minutes of savings compared to the same period in the previous year.

In spring 2004, Delta had average taxi-out times greater by 2.5 minutes than the other airlines, primarily because Delta's banked operations resulted in flights queuing one behind another before departing. For this reason, the benefits of the new schedule and the spreading of the flights were disproportionately reaped by Delta.

During Spring 2005, average taxi-out times for the airlines are close to one another because when there is no special pattern (no banking of one airline), flights go through the system at a "random" time and, on average, they experience the same delay. In this case, the average taxi time at the airport is a good estimation of the taxi-time for any specific airline.

	Spring 2004	Spring 2005	Variation
Delta	22.8	18.3	-19.9%
Others	20.3	18.0	-11.3%

Table 3.4 Average Taxi-out Times (in minutes)

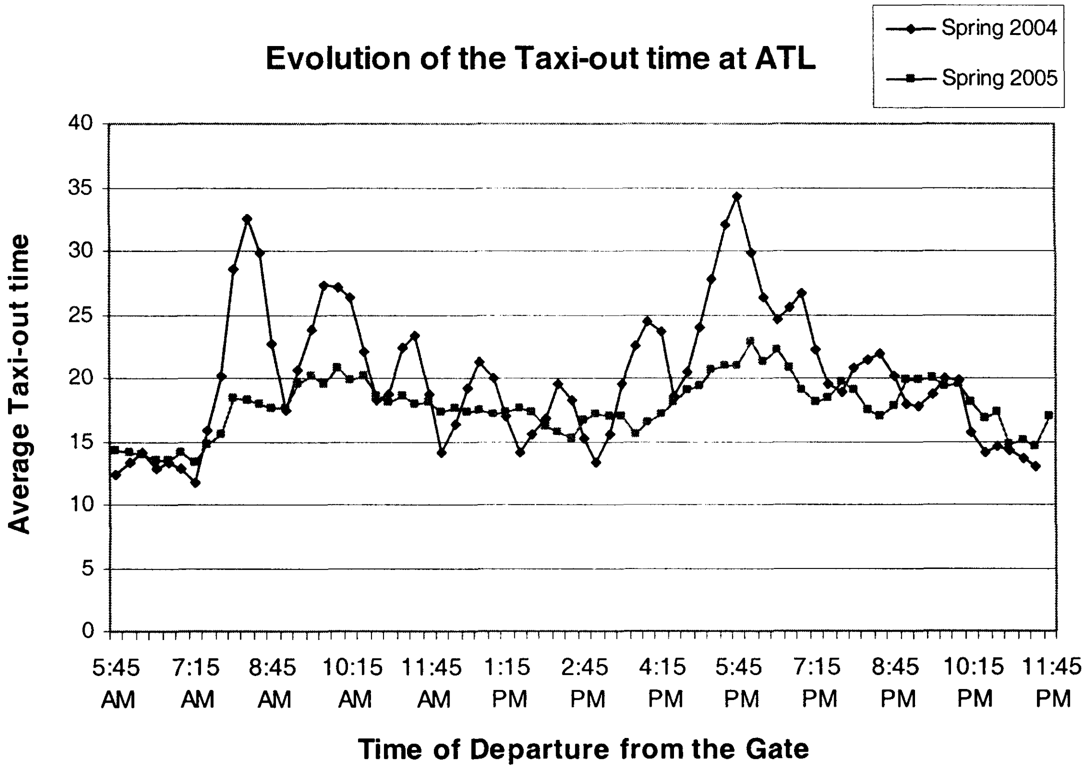

Figure 3.8: Taxi-out time as a function of departure from the Gate

Figure 3.8 shows the change in the taxi-out time at Atlanta for all airlines throughout the day. Note that the peaks during Spring 2004 we easily identifiable, but they are shaved in the new schedule, avoiding many of the delays.

This plot allows us to visualize the reduction in delay achieved with the new schedule as delay produced during the day is represented by the area under the curve.

The schedule changes allowed Delta planners to reduce the variation of taxi-out time, an important contributor to arrival delay. Interestingly, Delta's schedule changes also benefit its competitors by reducing their taxi-out times.

3. 3. 4. Importance of meeting airport capacity

When designing a de-banked schedule, we argue that an airline should constrain its schedules to adhere to expected airport capacity.

BEFORE

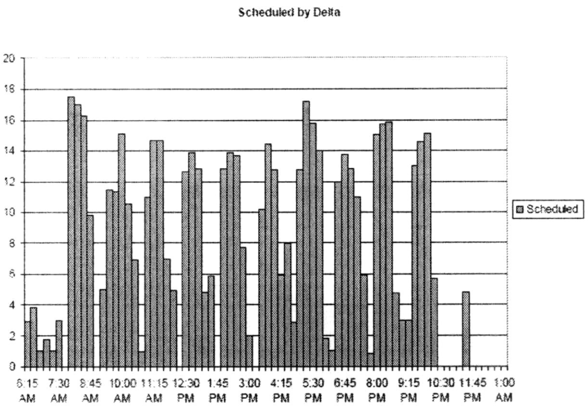

Figure 3.9 Schedule before

BEFORE

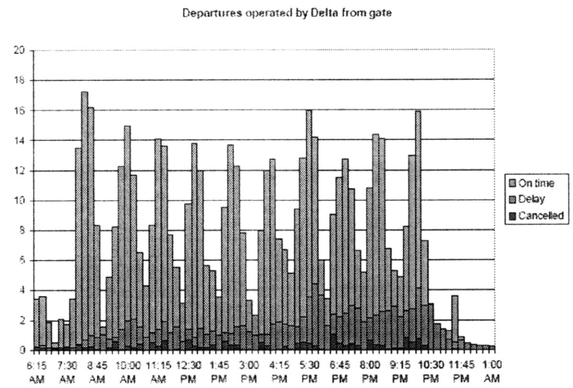

Figure 3.10: Operated at gate before

Figures 3.9 and 3.10 depict the planned and actual gate departure schedules, respectively for January 2005. In the actual schedule, peaks are decreased a bit compared to the plan, with flights spread out more evenly. In Figure 3.11 we present the average number of departures from the Atlanta airport throughout the month of January. Note that compared to gate departures, take-off times further reduce the peaks and spread out the departures of aircraft. In Figure 3.12, we display Delta's February 2005 schedule. Note the similarities between it and Figure 3.12. We conjecture that Delta designed its new schedule, recognizing the constraints of airport capacity, and adding new flights at times when excess capacity existed.

.

BEFORE

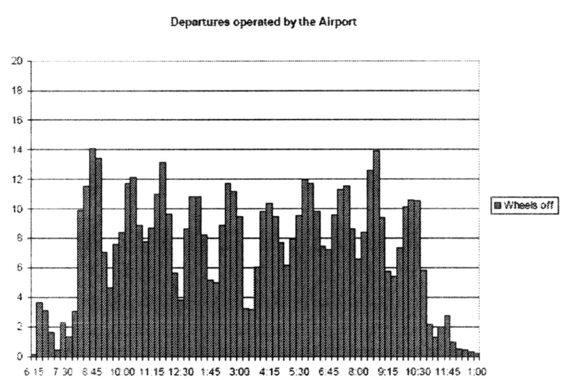

Figure 3.11 Take-off time before

AFTER

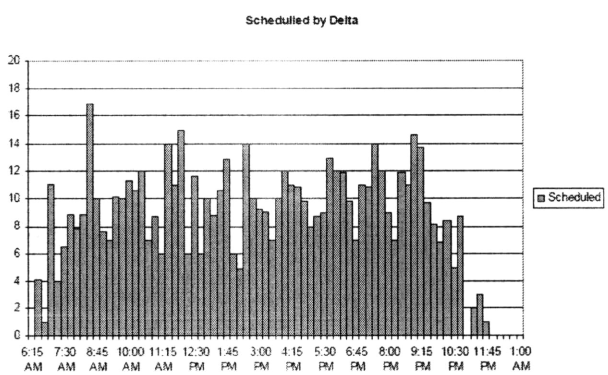

Figure 3.12: Schedule after

3. 3. 5. On-time arrivals at ATL

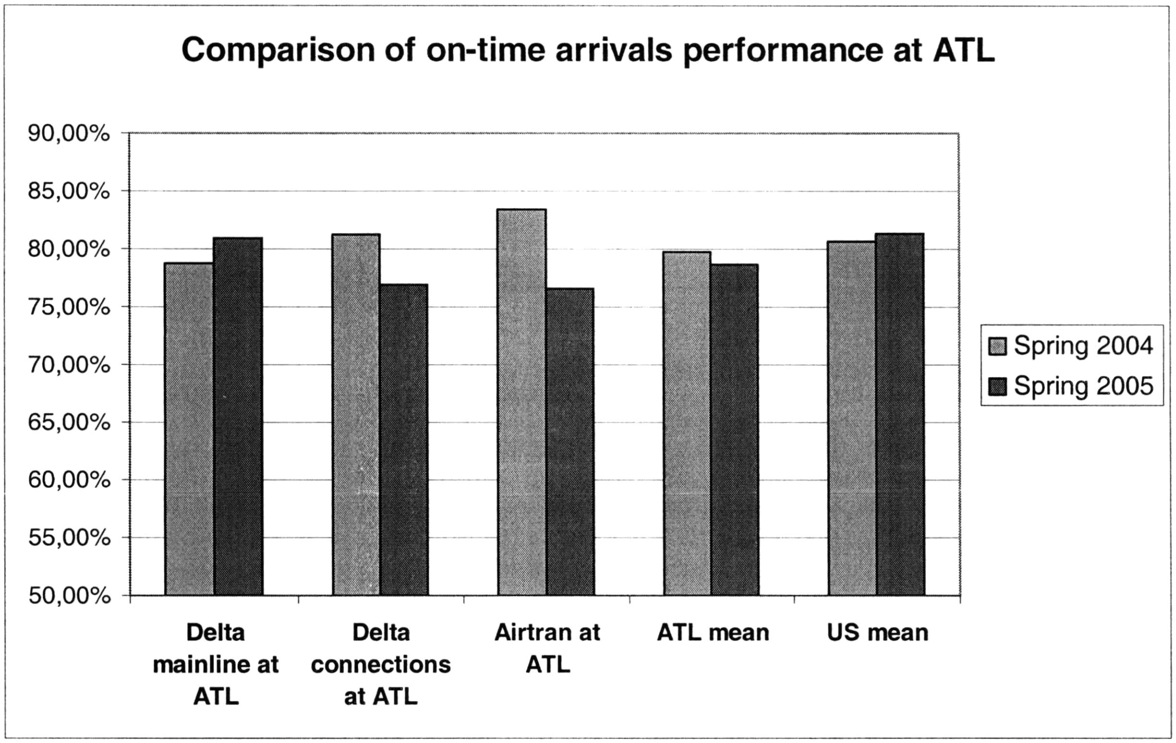

Figure 3.13 On-time arrivals in Atlanta

Figure 3.13 shows the improvement in the on-time arrival rate at Atlanta for Delta. Improved schedule reliability in Atlanta leads to improved downstream operations, and overall reductions in delays. The term "Delta Connections" corresponds to the regional partners of Delta operating in Atlanta (Atlantic Southeast Airlines and Comair). Interestingly, while Delta's mainline operations became more reliable in Atlanta, the arrival on-time performance of Airtran, Delta Connections and the Atlanta airport in general, worsened. This perhaps can be explained by the fact that arrival performance depends heavily on operations at the departing airports.

The US mean shows that the percentage of on-time arrivals increased for the whole system.

3.4 Comparison with other de-banking experiments

3. 4. 1. Comparison

We report in the table below the key performance indicators calculated by Bogusch and Agbokou to measure the benefits of the de-banking of American Airlines operations in O'Hare and Dallas.

Airport	New schedule	Period compared	Average Daily Departures	Average Taxi-out (minutes)	Average Taxi-in (minutes)	Average departure delay (minutes)
ORD	April 2002	July 2001	312	22.4	8.8	
ORD		July 2002	313	21	8.5	
DFW	November 2002	March 2002	414	18.1		11.5
DFW		March 2003	404	14.9		5.5
ATL	February 2005	Spring 2004	625	22.8	6.4	7.0
ATL		Spring 2005	688	18.3	6.3	5.0

Table 3.5: Comparison of performance indicators

Table 3.5 shows that de-banking has been successful and has improved the operations of each participating airline. The main benefit for the de-banking airline is the decrease in taxi-out times, helping to stem down-stream delay propagation

3. 4. 2. Who is next?

Given the positive results of de-banking, we ask if this trend will continue or not.

In June 2005, United Airlines implemented a de-peaking effort at LAX with the aim of achieving better utilization of aircraft, staff, and infrastructure. According to Yu [17], the results of the de-banking for an airline are considerable cost savings for the airline and a small increase in transfer times for passengers.

		Hub	Aircraft movements
Chicago	(ORD)	UA, AA (de-banked)	992,427
Atlanta	(ATL)	Delta (de-banked)	964,858
Dallas/Ft.	(DFW)	AA (de-banked)	804,865
Los Angeles	(LAX)	UA (de-banked)	655,097
Denver	(DEN)	UA	560,198
Phoenix	(PHX)	America West	546,763
Las Vegas	(LAS)	Southwest	544,679
Minneapolis	(MSP)	Norhtwest	541,093
Detroit	(DTW)	Norhtwest	522,538
Cincinnati	(CVG)	Delta	517,520

Table 3.6: Airport ranking by number of operations

Currently, the 5 top US airports in terms of aircraft movements have been successfully de-banked by the airlines using them as a hub. Because United is reported by the Airport operation fact sheet [16] to have made an effort to de-peak its schedule, we consider it as de-banked.

3. 5. Conclusion

De-banking benefited Delta operationally in that it allowed the company to decrease its taxi-out delays by 20% and increase its on-time performance for departing and arrival flights. However, what is striking in this study is that the dominant carrier is in a position of weakness at its own hub. Indeed, the established airline cannot operate its banked schedule as planned . When taxi-time delays in particular increase to a certain level, the dominant carrier is driven to adopt this strategy because delays in banked operations affect its flights more than it affects those of its competitors.

Figure 3.14 shows that Airtran maintains a banked schedule at peak hours. This suggests that it is their strategy to provide shorter connections than on-time reliability of its flights.

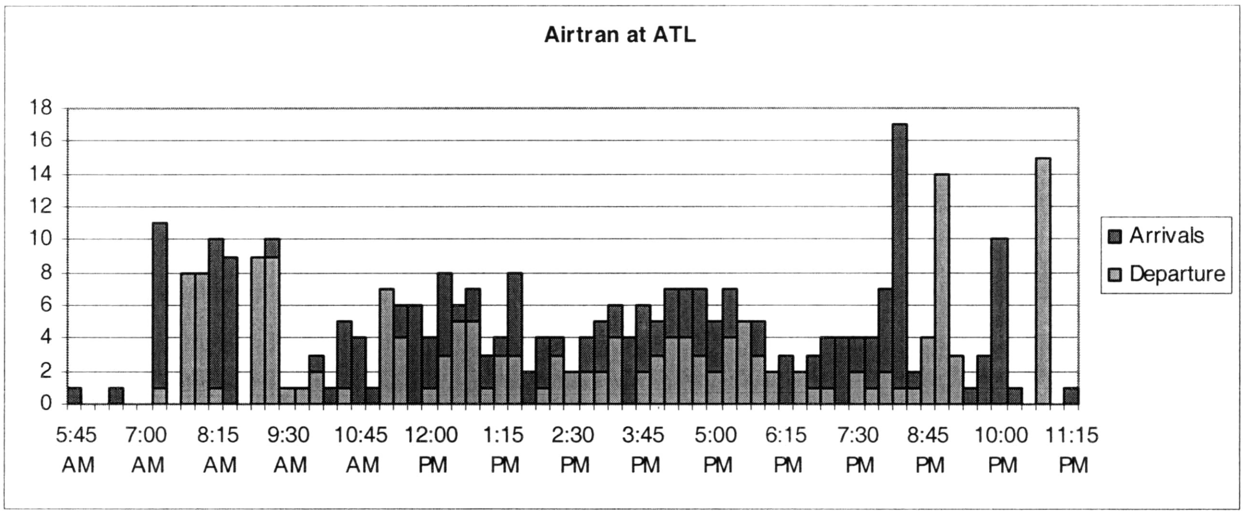

Figure 3.14 Airtran schedule at Atlanta

Chapter 4

Addition of buffer time

4.1 Introduction

The elapsed time from schedule flight departure time to flight arrival time represents the planned *block time*, including flying time between the flight's origin and destination, taxi-out time at the flight's origin, taxi-in time at the flight's destination, and expected delays due to congestion and other disruptive effects. In fact, we found that the actual flying time represents on average only 62% of the total scheduled block time for the flights departing Atlanta in February 2005. Actual flying time as a percent of planned block time measures between 50% for flights covering short distances and 97% for flights covering the longest distances.

In this chapter, we analyze and compare airlines' buffering strategies, that is, their approach to estimating block times given information about flying times, taxi times, congestion levels and their desired on-time performance.

4.2 Block Time Comparisons

4. 2. 1. Example of Atlanta-Dallas

With 30 departures per day, Dallas Forth Worth is the most served destination from Atlanta. Delta schedules 13 departures per day, American Airline 11 and Airtran 6.

The average flight time to Dallas, 732 miles away from Atlanta, is 117 minutes. However, because of taxi times and delays, the average gate to gate time is 145.4 min. Figure 4.1 is a plot of the density function of the gate-to-gate time of the flights from Atlanta to Dallas, obtained with S-plus. We find that 10% of the flights need more than 160 minutes to complete the trip. We will compare the performance of the 3 carriers to this destination during the month of February 2005.

4. 2. 2. Analysis

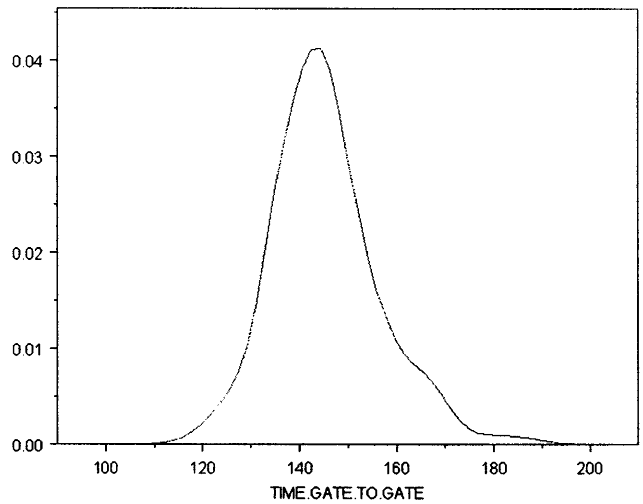

Figure 4.1: Density function of actual time gate to gate (ATL-DFW) in February 2005

In Figure 4.2, we compare the block times announced by the 3 airlines for times throughout the day. First of all, we can see that the airlines don't propose the same block time through out the day. Indeed, they anticipate taxi delays at rush hour in the morning and the end of the afternoon. The second reason is that they use different types of aircraft with different cruise speed to cover the flight.

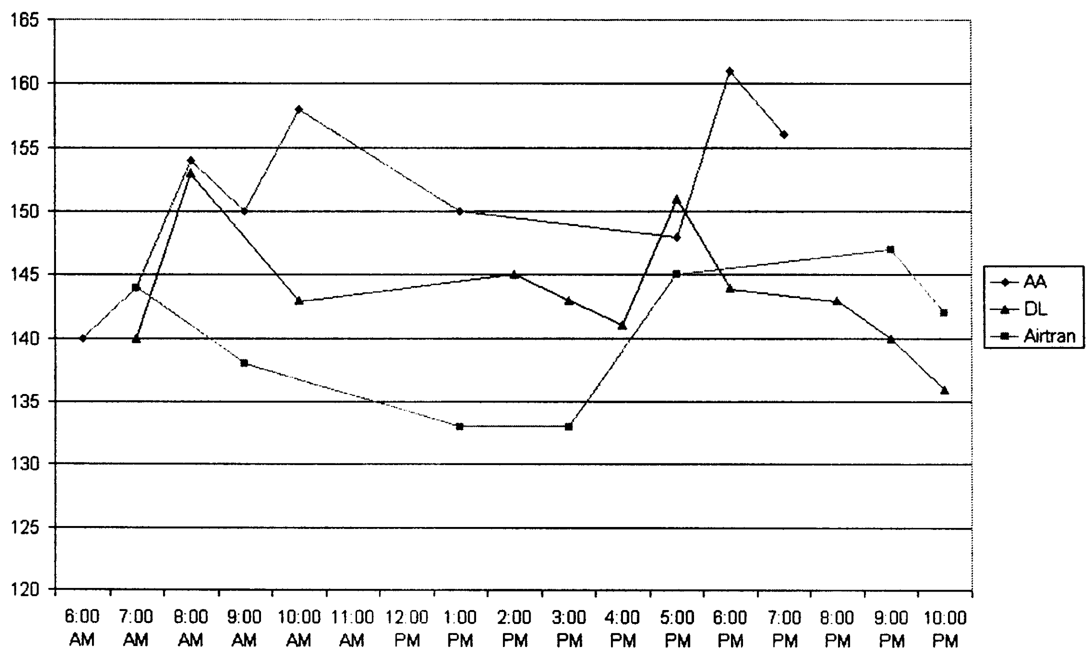

Figure 4.2: Block Time of the flights departing from Atlanta to Dallas

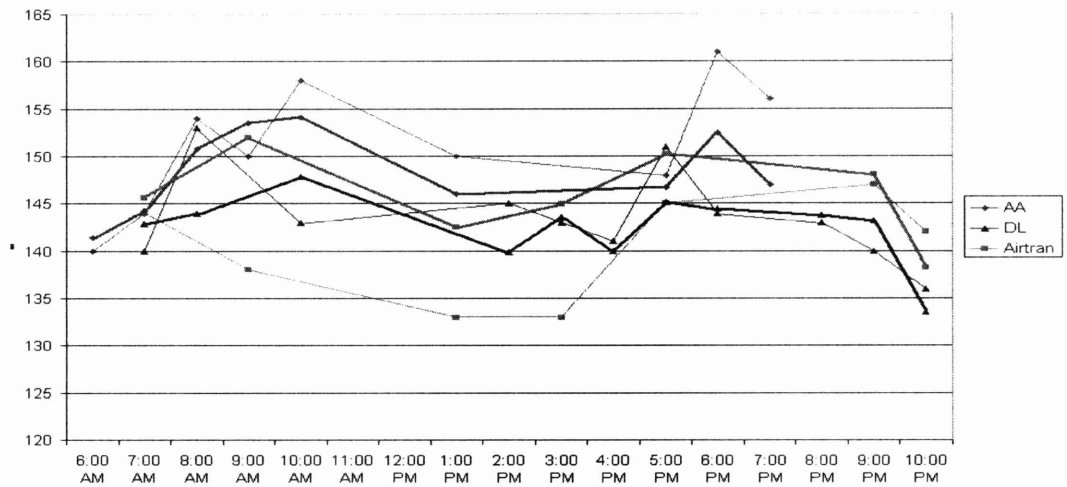

Figure 4.3: Time gate to gate (bold) versus scheduled time

In Figure 4.3, we compare their planned schedule with the realized gate-to-gate times. We can see that Delta has the shortest gate-to-gate times (on average, nearly 6 minutes shorter than American and 4 minutes shorter than Airtran). The obvious explanation is that Delta uses faster airplanes than the others. This is partially true. Indeed Delta has the shortest flying time and this alone explains the time difference with Airtran, but not with American Airline. The taxi- out time in Atlanta is on average the same for all departing flights.

Interestingly, the difference in gate-to-gate time between American and the others stems from the fact that American's average taxi-in time is 57% higher than Delta's and 37.5% higher than Airtran flights, as shown in Table 4.1.

	Taxi out	Flight time	Taxi in	Gate to Gate
AA	19.2	117.9	11.0	148.1
Delta	19.5	116.2	7.0	142.7
Airtran	19.5	119.0	8.0	146.5

Table 4.1: Mean value in minutes for ATL-DFW flights in February 2005

The reason is not that the terminal of American is further from the runways than the terminals of the others. To understand the difference, consider the histogram in Figures 4.4 – 4.6 showing the density line of the taxi-in times for the 3 airlines.

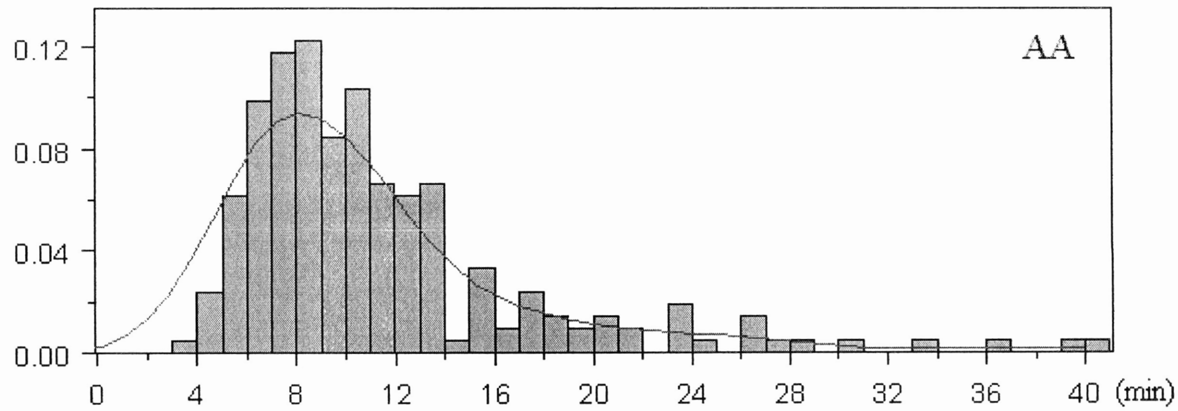

Figure 4.4: Taxi-in time of American Airline Flights ATL-DFW in February 2005

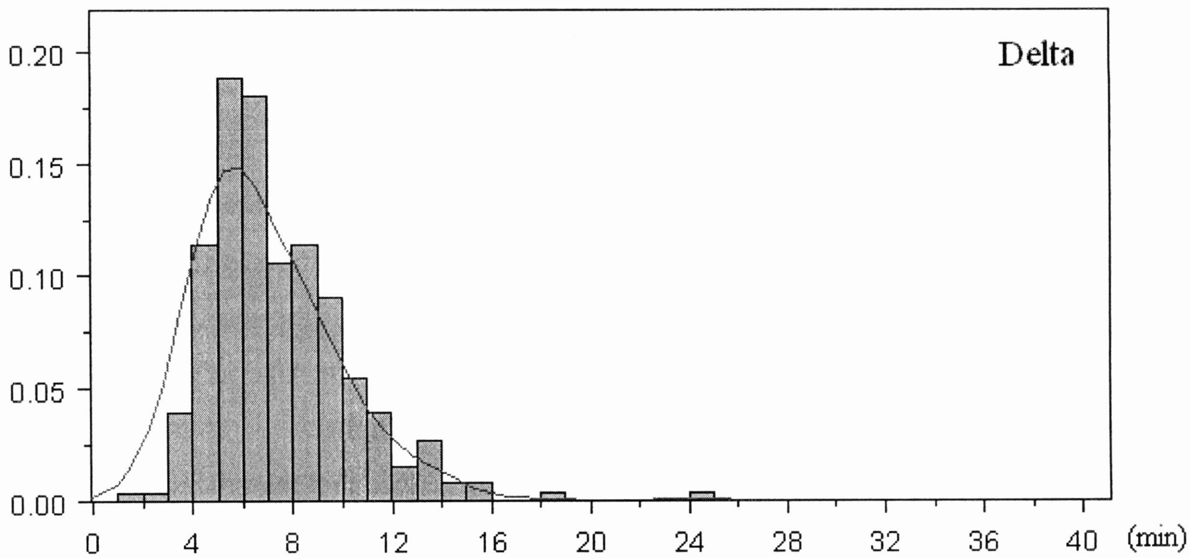

Figure 4.5: Taxi-in time of Delta Flights ATL-DFW in February 2005

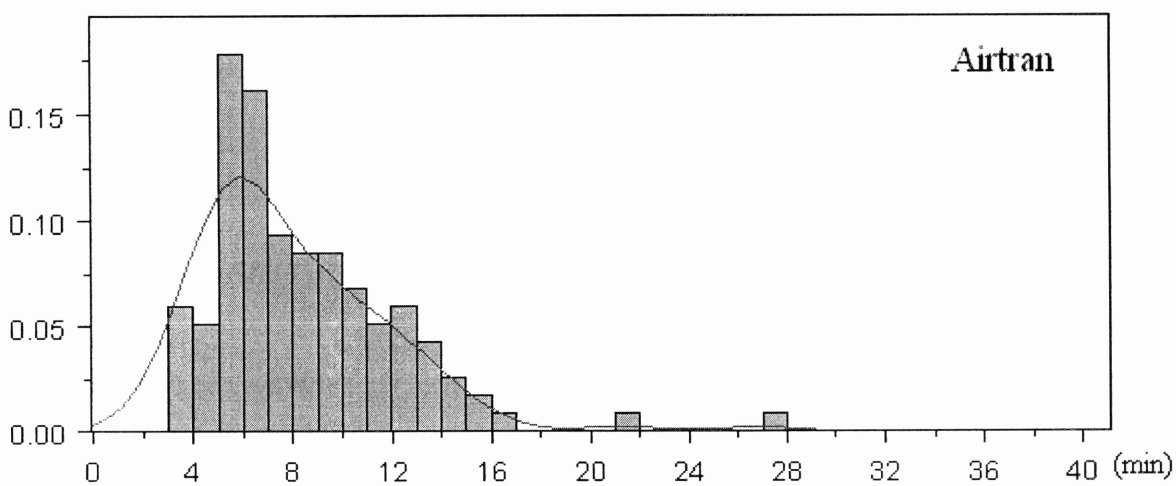

Figure 4.6: Taxi-in time of Airtran Flights ATL-DFW in February 2005

The histograms with density line demonstrate that the taxi-in delays don't come from the fact that the terminal of American is further to reach than the terminal of Delta or Airtran. Indeed, all airlines have nearly the same median time of 6-7 minutes.

However, in the case of American Airline we see that 10% of the flights were experiencing long delays greater than 20 minutes after landing and before reaching their gate. The main explanation for this is that the gate was occupied and unable to receive the incoming aircraft.

Consider the on-time performance results shown on Figure 4.7 and table 4.2 These results are quite poor even though the airline estimates of the gate-to-gate times needed are quite close to the realized times.

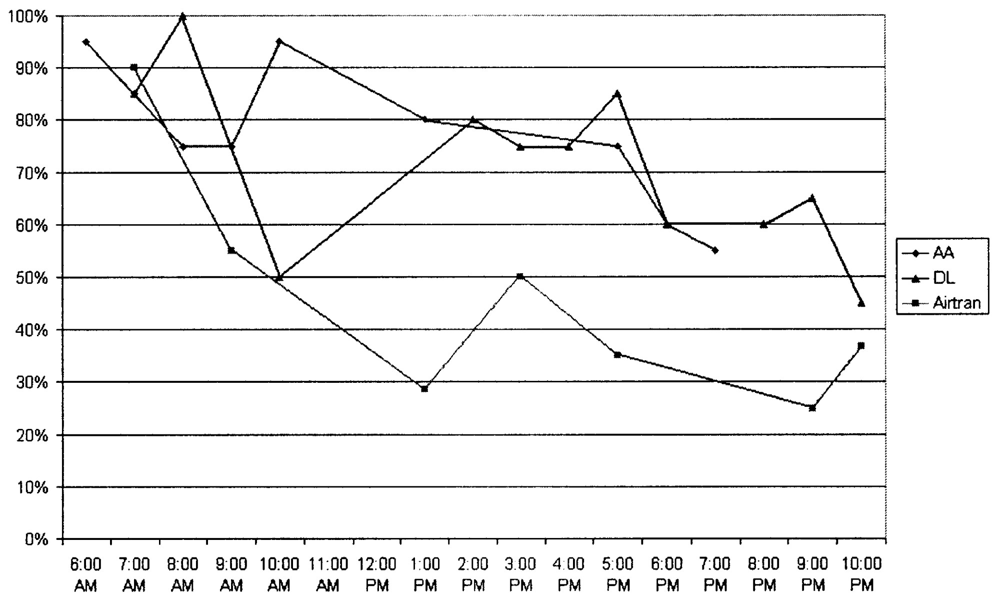

Figure 4.7: On-time arrival in Dallas performance

Table 4.2 shows that American Airlines achieves the highest on-time arrival rate with its strategy of longer block-times. Delta is second, due to its short gate-to-gate times.

	On-time	Delayed	Cancelled	Diverted
AA	81,3%	16,0%	2,8%	0,0%
Delta	76,4%	22,2%	1,4%	0,0%
Airtran	57,8%	41,7%	0,0%	0,5%

Table 4.2: Carrier performances at Dallas

The major contributors to on-time performance degradation are late departures of aircraft from the gates at Atlanta. These late departures are mainly attributable to propagated delays, as shown in Table 4.3.

	Average Departure Delay (in minutes)	Main causes of the delays		
		Carrier delay	NAS delay	Late aircraft
AA	9.7	27.7%	14.8%	56.8%
Delta	9.7	24.8%	21.0%	54.0%
Airtran	15.9	25.6%	24,0%	50.1%

Table 4.3: Causes of Delays at Dallas

4. 2. 3. Calculation of the gate-to-gate time

By knowing the distribution of the flight time needed (Figure 4.8), and the average taxi-out time at the scheduled departure time (Figure 4.9), we can better estimate the gate-to-gate times that should be planned. In the case of American Airlines, taxi-in times can not be considered constant due to their gate availability problems.

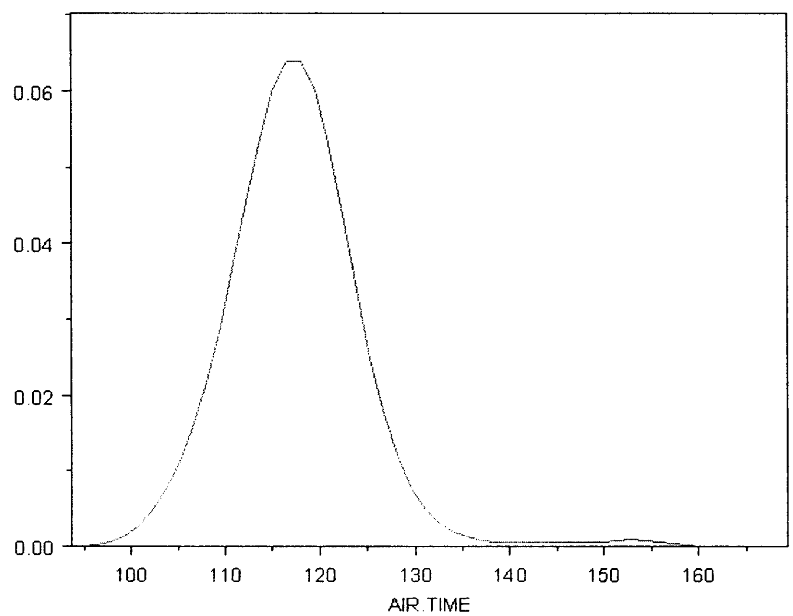

Figure 4.8: Airtime needed between ATL and DFW

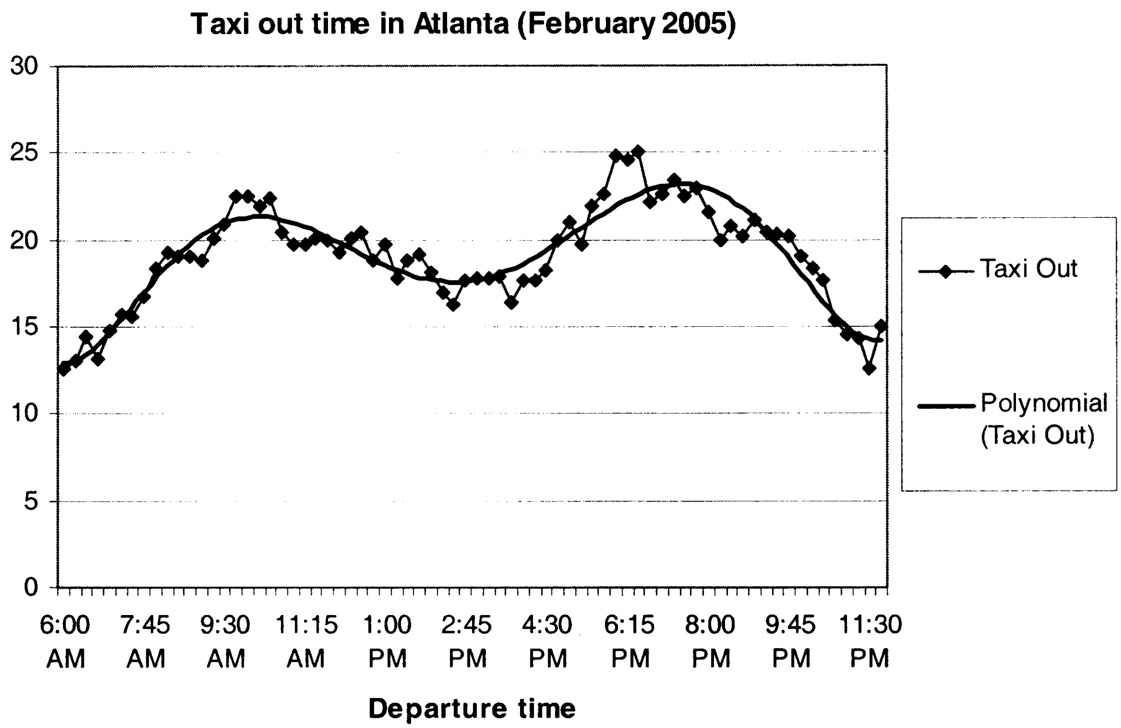

Figure 4.9: Taxi-out time in February for all the flights departing from Atlanta

4.3 Practices of timetable modification

4. 3. 1. Introduction

To achieve schedule reliability, airlines add time to their flying times and thereby increase their block-time estimates. This is a reasonable approach, given current levels of delays in taxi-time and flight time. In addition, some airlines add *buffer time* to their estimations of block-time, that is, they add additional time beyond the necessary estimated gate-to-gate time in order to gain robustness during operations. We compare airline practices concerning the use of *buffer time* in their planned schedules. We focus our analysis on a specific airport, namely Boston Logan Airport, which is not a hub of a major airline. We gather data of all departures and arrivals at Logan during January 2005. To estimate buffer times introduced by airlines, we select one airline at random, specifically, JetBlue, to determine among different types of regression, which one best fits this kind of data. Then, we utilize the selected regression approach to analyze the buffer time practices of other airlines and we compare it with their on-time performances.

4. 3. 2. Linear regression (base case: Least Square Fit)

The data we use contains all the departure and arrival delays for the flights operated by jetBlue at Boston during January 2005. We begin by test how well a linear regression works with the data. Then we give an interpretation of its intercept.

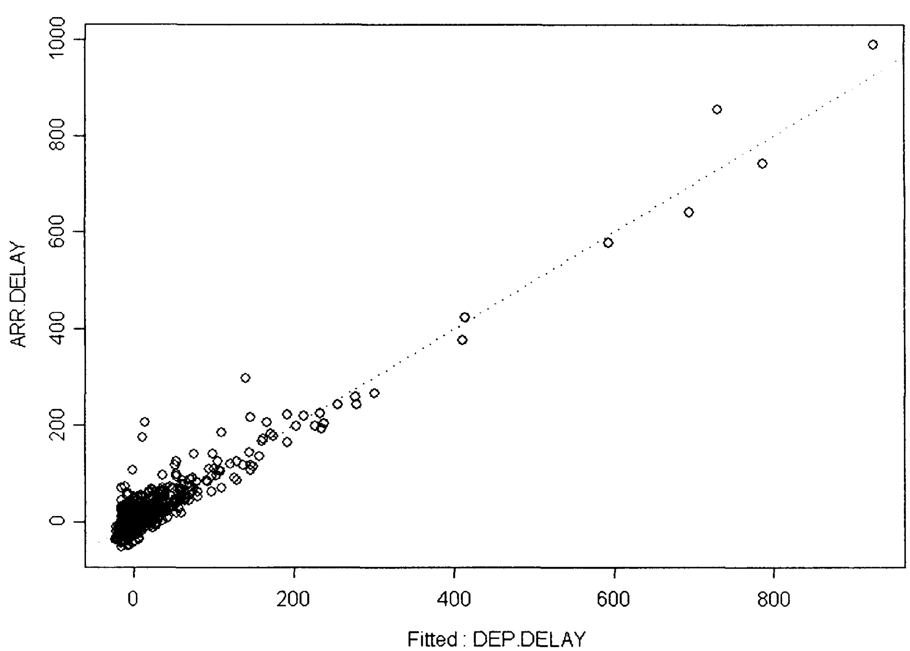

Figure 4.10 Linear regressions between arrivals and departures delays

Figure 4.10 shows the linear regression of the arrival delays in minutes versus the departure delays of the jetBlue flights that ended or began in Boston.. The formula used is ARR.DELAY=α*DEP.DELAY +β.

The linear regression fits the data fairly well, with an R-Squared of 0.9074. The coefficients of the regression in Table 4.1 are useful in understanding buffer times.

	Value	Std. Error
Intercept (β)	-4.9892	0.71129
DEP.DELAY(α)	1.0649	0.01079

Table 4.1: Coefficients for the linear regression of jetBlue data

The absolute value of the intercept in Table 4.1 indicates the average number of minutes that a flight departing on time will arrive early at its destination. We refer to it as the *actual buffer* time. Because the slope of the regression is very near to one, it means that on average a flight's gate to gate time is nearly 5 minutes faster than the scheduled block time.

With this average of nearly 5 minutes of slack time incorporated in its schedule, jetBlue appears to be making efforts to guarantee schedule robustness. For this low cost airline with remarkably high aircraft utilization, schedule robustness is critical to achieving planned productivity levels.

The residuals (Figure 4.11) can give us an indication of the goodness of fit of the regression.

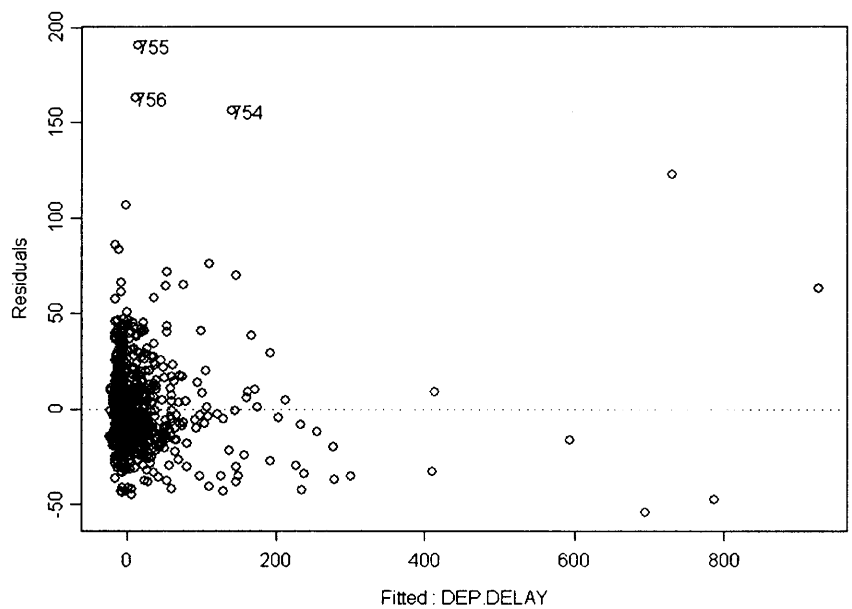

Figure 4.11 Residuals of the linear regression between arrival and departures delays

Figure 4.11 represents the residuals obtained for the linear regression of arrival delays versus departure delays for jetBlue flights. We can identify 3 outliers with small departure delays and large arrival delays. These are all flights destined to Boston Logan Airport that were delayed by the National Aviation System on the 24th of January. The disruptions were caused by a period of very bad weather over Boston and congestion at the airport resulting from Logan's closure for some hours during the previous day due to a blizzard.

4. 3. 3. Robust regressions (by erasing the outliers)

Because we don't want the outliers to have a greater influence on the results than the majority of the data, we eliminate the outliers from the jetBlue arrival and departure delays data. Then we draw the linear regression ARR.DELAY=α*DEP.DELAY +β for the rest of the data and we collect the Intercept and the Slope. Table 4.2 summarizes our findings.

	Linear regression	minus 3 outliers	minus 7 outliers	minus 10 outliers	Mean	Standard deviation of Mean
Intercept	-4.9895	-4.4073	-4.4812	-4.7737	**-4.6629**	0.2691
Slope	1.0638	1.0188	1.0244	1.0079	**1.0287**	0.0243

Table 4.2: Comparison of the Robust Regressions

The results for various numbers of outliers removed are quite similar. As more outliers are removed, the slope tends to decrease to 1. To quantify the correspondence of our mean results with the data, we define the quality of the pick around the mean value as Q = Mean/Stdev.
For the intercept, we find Q= 17.3 and for the slope Q=42.3. Both values are much greater than 10 and as a consequence, we can conclude that the means are good estimators for both the intercept and slope.

4. 3. 4. Least Trimmed Squares Robust Regression (LTS Regression)

Removing outliers is not easy to implement because the definition of outliers is not always very clear. Moreover, we would like to find a systematic method that enables us to characterize the properties of the data. We again use the data set of jetBlue flights and we study the results of another type of regression. The least trimmed squares regression approach minimizes the sum of the smallest "half" of the squared residuals. The regression has a high breakdown point (nearly 50%): by definition, it means that if nearly 50% of the data is corrupted, the regression will not be influenced by all the outliers and will still reveal the main trend of the data set. Its usual rate of convergence is higher than the least median of squares regression [10]. The objective of the least trimmed squares approach is to minimize the sum of the q smallest squared residuals. To determine q, the residuals are ordered in increasing value and q is set to be slightly larger than ½ of n, where n is the number of observations. q is thus set equal to floor(n/2) + floor((p + 1)/2), where p is the rank of x. The objective then is:

$$\min \sum_{i=1}^{q} | y_i - x_i b |^2$$

In this case, p=2 estimated parameters for the regression.

Using again the linear formula ARR.DELAY = α * DEP.DELAY + β, and the same data set as in the previous analysis, the robust approach gives us an absolute value for the intercept of 7.4 and a linear coefficient very close to 1, as presented in Table 4.3

Intercept (α)	-7.4094
DEP.DELAY (β)	1.0005

Table 4.3: Coefficients for the LTS regression of jetBlue data

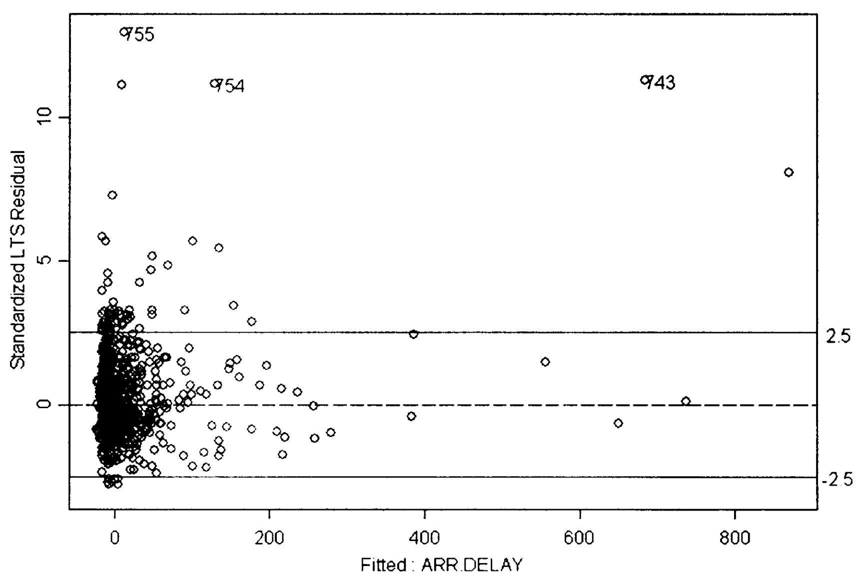

Figure 4.12 Standardized LTS Residuals versus Fit

S-plus indicates that this estimation considers 896 data points, representing 90% of the data. We find an intercept of -7.4 based on. 90% of the flights departing on-time, implying they arrived 7.4 minutes ahead of schedule on average. Thus, for 90% of the data, the buffer time added to the expected gate-to-gate time was 7.4 minutes on average.

However, as illustrated in Figure 4.12, there are too many standardized residuals greater than 2.5. This indicates that this kind of regression doesn't take into account a particular pattern and therefore doesn't capture the structure of the data. Therefore, we will try to apply another robust regression to estimate better the average amount of buffer time airlines add to their gate-to-gate time.

4. 3. 5. Robust MM Linear Regression

The Least squares method carries the assumption that observations are normally distributed. This is not the case in our dataset. Hence, the LTS regression returns inaccurate estimates. Therefore, because the dataset contains significant outliers, it is more accurate to use the Robust MM regression, a nonparametric technique that is very useful for fitting linear relationships. The Robust MM regression is also less sensitive to erratic observations than the nonparametric approach. [14].

In Robust MM Regression, robust initial regression coefficients are used as starting values. The robust regression coefficients are found by minimizing a scale parameter, S. χ is a bounded function chosen so that it will decrease the influence of outliers. Here, we use $\chi(u) = u^6 - 3u^4 + 3u^2$ for $|u| \leq 1$ $\chi(u) = 1$. χ is an integral of χ(u) in the formula :

$$\sum_{i=1}^{n} \chi(\frac{y_i - x_i b}{c_0 s}) = (n-p)\beta$$

$c_0 = 1.548$ (Tuning constant), $\beta = 0.5$

The M-estimate is derived according to the loss function from the S estimate and the fixed scale estimate produced. With S-PLUS, the procedure is generated with the *lmRobMM* function. [10]

Using again the linear formula ARR.DELAY = α * DEP.DELAY + β, we get the following coefficients (Table 4.4).

	Value	Std. Error
Intercept (α)	-9.209	0.8209
DEP.DELAY (β)	0.9889	0.0154

Table 4.4: Coefficients for the MM Robust regression of jetBlue data

The absolute value of the intercept is 9.2 minutes. This result obtained by the robust method giving less weight to the outliers is considered as the buffer time planners added to the expected gate-to-gate time in order to gain robustness. We refer to this as the *"planned buffer time"*.

Figure 4.13 shows that the regression fits the data points fairly well.

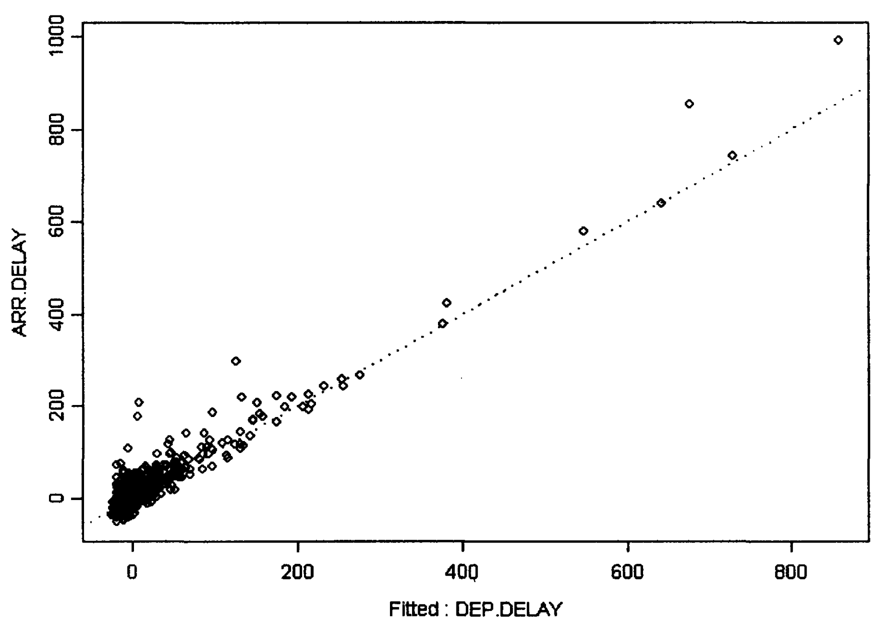

Figure 4.13 Robust MM regression

The absolute value of the intercept obtained by the Robust MM regression is thus employed to quantify what we call the *"planned buffer time"*, that is, the amount of time the airline added to the expected gate-to-gate time to guarantee more robustness during operations. The *"actual buffer time"* is estimated by the intercept of the linear regression.

4. 3. 6. Comparison Robust MM Linear Regression versus LS fit

The comparison of the intercept of these 2 regressions gives us some information about the plans and achievements of the airline. Our findings are presented in Table 4.5. The small standard deviations demonstrate the accuracy of the estimations.

	Intercept (α)		DEP.DELAY (β)	
	Value	Std. Error	Value	Std. Error
Linear Regression	-4.9892	0.71129	1.0649	0.01079
MM Robust Regression	-9.2098	0.82089	0.9889	0.01541

Table 4.5: Comparison of the coefficients of the regressions

Given the property of these two regressions, we will estimate the *"actual buffer time"* with the intercept of the linear regression because it takes into account the entire data set. The *"planned buffer time"* is estimated with the intercept of the Robust MM regression because it is not influenced by outliers.

4.4. Application to different airlines

4. 4. 1. Method

We apply the linear regression and the Robust MM regression to the departures and arrivals of different airlines operating in Boston for the month of January. We also report the arrival on-time performance of flights during this month.

4. 4. 2. Results

	Continental	JetBlue	Airtran	Comair	UA
On-time	67%	67%	63%	62%	59%
"actual buffer time"	4.7	-5.0	3.5	-3.0	2.7
"planned buffer time"	-0.77	-9.2	-1.4	-7.3	-3.0

Table 4.6: Buffer time (in minutes) practice for different airlines in January in Boston

In Table 4.6, we report the intercepts of the regressions. An *actual buffer time* of -3.0 (for Comair) means that aircraft were experiencing gate-to-gate time that were on average 3 minutes shorter than those scheduled. An *actual buffer time* of 2.7 (for United Airlines) implies that aircraft were arriving 2.7 minutes later than scheduled. From this result, we can see that jetBlue was the most conservative in planning its flight block time by adding 9.2 minutes to the average gate-to-gate time.

The on-time performance for this airport is very poor, for all the airlines, and it doesn't appear to be correlated with the quality of operations at the airport. Indeed, Continental achieved the same on-time percentage as jetBlue, but Continental was adding no buffer time to their expected gate-to-gate times, while jetBlue was adding significant buffer time.

This non-correlation comes from the fact that on-time performance of one flight is also influenced by the preceding flights, which arrive late or not.

Another factor that we believe should influence the amount of buffer time airlines allocate to a flight leg is the distance flown by the aircraft. In Table 4.7, we show the mean distance of flights flown by each airline, and their planned buffer times. Clearly, different airlines have different strategies for allocating buffer times.

	Continental	JetBlue	Airtran	Comair	UA
Average Distance (miles)	628	1541	511	346	1364
"planned buffer time"	-0.77	-9.2	-1.4	-7.3	-3.0

Table 4.7: Mean distance

4.5. Conclusion

Adding buffer time to account for congestion at airports and weather delays is necessary to achieve schedule reliability. Buffer time estimates should include the effects of seasonal variations. Some airlines, like United Airlines, Airtran and Continental, underestimated the need for buffer times at Boston Logan in January 2005, a period in which extreme weather severely disrupted operations at Logan. Buffer time estimates should also recognize that there are periods of the day when taxi-times are higher, namely, when departure queues form. In the case of the American Airlines' flights from Atlanta to Dallas, not capturing gate delays has serious delay effects. All of these factors should be considered in estimating buffer times.

Of course, there are some parameters that are specific to the particular day of operations and cannot be predicted by the planners. However, the task of the planners is to consider as many parameters as possible in order to decrease the risk and consequences of delay propagation while maintaining strict cost controls. This question will be discuss and evaluated in the last part of this thesis.

Chapter 5

Implementation of a robust crew scheduling model

5.1 Introduction

Our motivation for a different crew scheduling approach comes from the fact that analyzing delay data from a major U.S. airline, we find that 18 % of the non-system delays are related to crews (that is, waiting for a crew connection, insufficient turn time to complete a crew change, etc.). It suggests that improvements in how crew pairings are constructed can lead to decreased numbers of schedule disruptions and reduced delay propagation and amplification for subsequent flights.

After having presented two approaches used by the airlines to make their schedules more reliable, this part of the thesis will address scheduling optimization solutions for crews. In particular, we will focus on optimization models targeting reductions in delay propagation, as we have seen that it is the major source of delays starting at 5pm. Our review of relevant optimization models will be followed by an implementation of our selected models using the Carmen Crew Pairing software [7].We evaluate the trade-off between schedule robustness and cost.

5. 2. Review of crew pairing models

The crew scheduling problem is the last sub-problem to be solved in the *Airline Schedule Planning Process*. After defining the flight schedule, assigning aircraft types to the flight legs in the schedule, and routing individual aircraft to routinely visit maintenance stations, the final task in the Airline Schedule Planning Process is to develop crew schedules that ensure every flight leg is covered by a crew and that total crew cost is minimized. To deal with uncertainty and delays, different robust approaches have been proposed during recent years with the objectives to decrease the occurrence of disruptions and ease the recovery process. Depending on the definition of robustness taken, different robustness criteria have been defined. Our focus is to reduce delay propagation and amplification caused by crews.

5. 2. 1. Generic Crew Pairing Model

We focus on the crew pairing problem, that is, the problem of finding the minimum cost set of crew schedules that cover all flight legs. In crew scheduling, the crew pairing problem is solved and then the selected problems are then assembled into monthly crew work schedules.

The first crew pairing model we present is a "base" case that does not include any robustness criteria. Its objective is to find the set of pairings that cover all the flights and minimizes crew costs. Its solution provides a useful baseline from which we can measure the increased costs needed to achieve increases in robustness.

The parameters and variables in this model are defined as followed:

Parameters

- F is the set of flight legs i
- P is the set of pairings p
- c_p is the cost of pairing p
- δ_{ip} equals 1 if pairing p includes flight leg i, and 0 otherwise

Variables

- y_p equals 1 if pairing p is selected, and 0 otherwise

Given the above notation, the formulation of the generic crew pairing problem (Barnhart et al [3]) is:

$$\min \quad \sum_{p\in P} c_p y_p \qquad (1\text{-}1)$$

subject to

$$\sum_{p\in P} \delta_{ip} y_p = 1 \ , \quad \forall i \in F \qquad (1\text{-}2)$$

$$y_p \in \{0,1\} \qquad (1\text{-}3)$$

The objective (1-1) minimizes the cost of the chosen set of pairings. The cover constraints (1-2) and the binary constraints (1-3) ensure that each flight leg i is covered by exactly one pairing.

5. 2. 2. Bi-criteria approach

Ehrgott and Ryan [9] focus on developing a measure of non-robustness for each pairing based on the effect of potential delays within the pairing. If the crew stays with the aircraft between 2 flights, there will be no penalty. However, if the crew has to change aircraft within a duty period, the penalty will reflect the potential disruption effect of the possible delay caused by the aircraft change.

Given a pairing p consisting of f flights, they define for each flight $i=1\ldots f\text{-}1$, three different times related to crew connections (see Figure 5.1), namely:

- Ground duty time (GDT_i^{i+1}): the minimum connection time; usually 45 minutes or more if meal breaks are included.
- Measure of delay of incoming flight (DM_i): the mean delay plus its standard deviation for the incoming flight i
- Scheduled Ground time (SGT_i^{i+1}): the time between two consecutive flights on different aircraft in the duty.

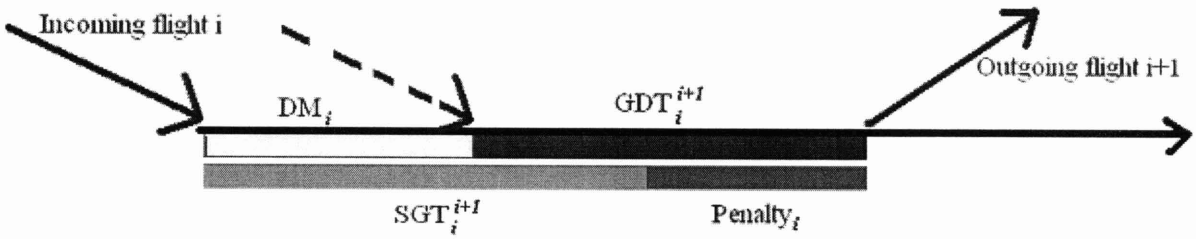

Figure 5.1: Definitions of the different times in Ehrgott's and Ryan's model

The penalty for each crew connection is defined as:

$p_i^{i+1} = Max[0, GDT_i^{i+1} + DM_i^{i+1} - SGT_i^{i+1}]$.

The non-robustness criterion is then defined for each pairing p, composed of f flights, as:

$$r_p = \sum_{i=1}^{f-1} p_i^{i+1}.$$

They use the non-robustness measure as a second objective. The parameters and variables in their model are defined as followed:

Parameters

- F is the set of flight legs i
- P is the set of pairings p
- c_p is the cost of pairing p
- r_p is the penalty define above of pairing p
- δ_{ip} equals 1 if pairing p includes flight leg i, and 0 otherwise

Variables

- y_p equals 1 if pairing p is selected, and 0 otherwise

Given the above notation, the formulation of the Ehrgott and Ryan's model is:

$$\text{Min} \quad \sum_{p \in P} c_p y_p \qquad (2\text{-}1)$$

$$\text{Min} \quad \sum_{p \in P} r_p y_p \qquad (2\text{-}2)$$

subject to

$$\sum_{p \in P} \delta_{ip} y_p = 1 \ , \ \forall i \in F \qquad (2\text{-}3)$$

$$y_p \in \{0,1\} \qquad (2\text{-}4)$$

The objectives (2-1) and (2-2) minimize the cost and non-robustness, respectively of the chosen set of pairings. The cover constraints (2-3) and the binary constraints (2-4) ensure that each flight leg i is covered by exactly one pairing.

To generate solutions, they use the ε-constraint method (Chankong and Haimes, 1983) based on the idea of only minimizing one of the objectives and transforming the other one into a constraint limited by an upper bound.

5. 2. 3. Maximizing Short Connect Utilization

Our approach adopts a similar point of view. However, without using historical data, we guide our optimization approach to select solutions that will reduce delay propagation and amplification by making the crew follow the routing of the plane as much as possible.

A short connect is defined as a connection which is feasible for a crew only if two sequential flights comprising that connection have been assigned to a common aircraft.

Figure 5.2 illustrates how short connect utilization decreases the risk of delay propagation.

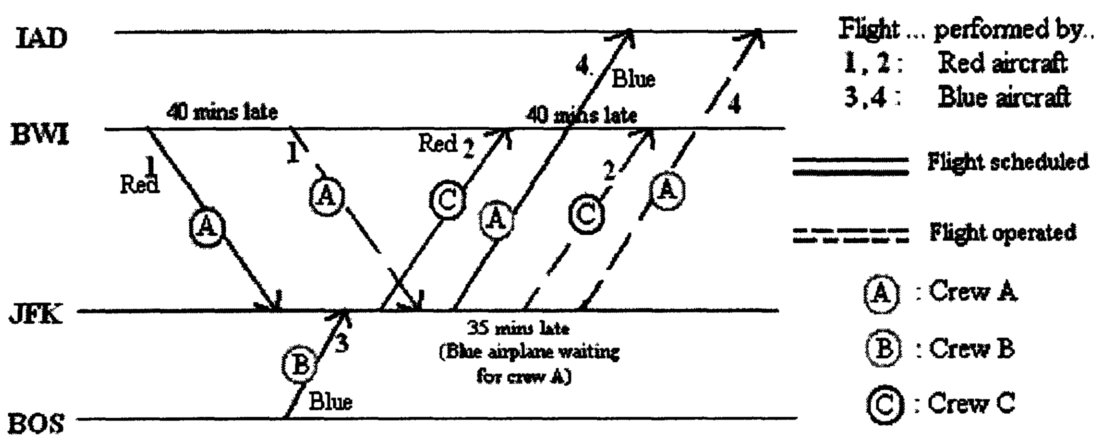

Figure 5.2: Illustration of maximizing short connection

Suppose that the Red aircraft is assigned to cover Flights 1 and 2 and the Blue aircraft is assigned to cover Flights 3 and 4. Further suppose that Crew A is assigned to cover Flights 1 and 4, Crew B is assigned to Flight 3 and Crew C is assigned to Flight 2 (see Figure 5.2).

Crew A needs 45 minutes to connect to flight 4 and has 50 minutes scheduled connection time. Assume that the red aircraft makes a quick turn of 30 minutes following flight 1. In this case, the short connection of the red aircraft from flight 1 to 2 is not utilized, that is no crew is assigned to it. Instead, a crew is assigned to the *critical connection* (that is, a crew connection between two different aircraft with duration less than some critical threshold value) between the red aircraft operating flight 1 and the blue aircraft operating flight 4.

Now consider that Flight 1 experiences 40 minutes of delay. Flight 2 will be delayed because its aircraft is delayed and Flight 4 is also cannot depart on-time because its crew is delayed. If, instead of assigning a crew to the critical connection between flights 1 and 4, a crew is assigned to the short connection between flights 1 and 2, delay propagation and amplification is reduced because only flight 2 is delayed, rather that both flights 2 and 4.

Given a solution to the maintenance routing problem, we can improve upon the operational flexibility of a crew schedule by maximizing the number of short connections used.

To formulate this problem, we introduce the following parameters and variables:

Parameters

- F is the set of flight legs i
- P is the set of pairings p
- SC is the set of short connects provided by the maintenance routing solution
- c_p is the operating cost of pairing p
- b_{jp} is equals 1 if pairing p includes short connect i, and 0 otherwise
- δ_{ip} equals 1 if pairing p includes flight leg i, and 0 otherwise

Variables

- y_p equals 1 if pairing p is selected, and 0 otherwise.

Given this notation, we formulate the crew pairing problem that maximizes the number of short connections used as:

$$\max \quad \sum_{i \in SC} \sum_{p \in P} b_{ip} y_p \qquad (3\text{-}1)$$

subject to

$$\sum_{p \in P} \delta_{ip} y_p = 1 \ , \quad \forall i \in F \qquad (3\text{-}2)$$

$$y_p \in \{0,1\} \qquad (3\text{-}3)$$

$$\sum_{p \in P} c_p y_p \leq (1+r).c_{OC} \qquad (3\text{-}4)$$

The objective (3-1) of this crew pairing model is to maximize the number of short connections used by the crew schedule. The cover constraints (3-2) and the binary constraints (3-3) ensure that each flight leg i is covered by exactly one pairing. Constraint (3-4) guarantees that planned crew costs are within a certain tolerance level above the minimum possible crew costs given by the generic crew pairing model.

The main limitation of our model is that the number of possible short connects is limited by the aircraft routing. Therefore it would be interesting to create an aircraft routing model where we would maximize the number of possible short connect before solving this robust crew pairing model.

5. 2. 4. Integrated Robust Routing and Crew Model

The "Integrated Robust Routing and Crew Pairing Model" presented by Agbokou [1] augments the basic crew pairing model with a set of feasible aircraft routings and then selects simultaneously the maintenance routing solution and crew pairing solution that provides a *robust,* yet near minimum-cost, crew solution.

While our model favors short connect utilization (and as a consequence disfavors any crew connections), her model distinguishes between critical crew connections and others, stating that if the crew connection time is longer than 1 hour and 15 minutes, the crew is much less likely to be disturbed and can thus be included in the solution without penalty.

Agbokou defines a critical crew connection (see Figure 5.3) defined as one in which a crew is required to change aircraft between successive flights, and the connection time is between 45 minutes and 1 hour 15 minutes.

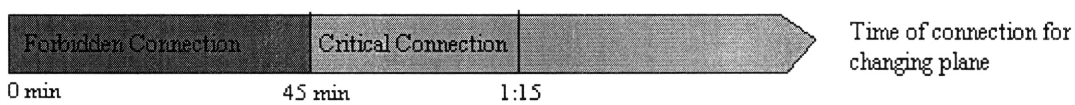

Figure 5.3: Definition of a critical connection

As a consequence, Agbokou's model achieves the same benefits that we identify regarding decreases in delay amplification. Indeed, it pushes crew connections out of the critical zone, where possible, if flight delays are less than 30 minutes (the time of the critical connection zone). In 2004, actual delays and cancellations, by length of delay, are displayed in Table 5.1.

On-time	Delayed between 15 and 30 minutes	More than 30 minutes	Cancelled	Diverted
78.08%	8.44%	11.50%	1.79%	0.19%

Table 5.1: Delays and cancellations for the US domestic flights in 2004

Table 5.1 shows that 11.5% of the flights in 2004 were delayed by more than 30 minutes. These flights will continue to create disruptions to crews using the critical connection thresholds set by Agbokou. To achieve a more robustness solution, we would need to extend the critical connection window.

The parameters and variables of the Agbokou model are:

Parameters

- F is the set of flight legs *f*
- P is the set of pairings *p*
- S is the set of maintenance solutions *s*. A maintenance solution s is a set of aircraft strings that satisfy the basic aircraft maintenance routing requirements. It determines the feasible short connects and the number of critical connects.
- R_s is the set of route strings included in maintenance solution *s*
- $\tilde{C}$ is the set of critical connections annulled by S
- $\tilde{S}$ is the set of short connections allowed by S
- b_{cr} equals 1 if route string *r* includes (that is, assigns the same aircraft to) critical connect *c*, and 0 otherwise
- β_{hr} equals 1 if route string *r* allows short connect *h*, and 0 otherwise
- δ_{fp} equals 1 if pairing *p* includes flight *f*, and 0 otherwise
- a_{cp} equals 1 if pairing *p* includes critical connect *c*, and 0 otherwise
- α_{hp} equals 1 if pairing *p* includes short connect *h*, and 0 otherwise
- r is the "robustness" factor

Variables

- x_s equals 1 if maintenance solution *s* is in the solution and 0 otherwise.
- y_p equals 1 if pairing *p* is picked, and 0 otherwise
- α_c , β_c equals (0, 0) if critical connect *c* is covered by one crew and one aircraft or if critical connect *c* is not included in the maintenance routing solution and is not in the crew pairing solution; (0, 1) if critical connect *c* is not in the crew pairing solution included and is included in the maintenance routing solution; (1, 0) if critical connect *c* is included in the crew pairing solution and not in the maintenance routing solution.

Given the above notation, here is the Agbokou formulation of the integrated robust aircraft routing and crew model:

$$\min \sum_{c \in \tilde{C}} \alpha_c \tag{4-1}$$

subject to

$$\sum_{s \in S} x_s = 1 \tag{4-2}$$

$$\sum_{p \in P} \delta_{fp} y_p = 1 \quad \forall \ f \text{ in F} \tag{4-3}$$

$$\sum_{s \in S} \sum_{r \in R_s} \beta_{hr} x_s - \sum_{p \in P} \alpha_{hp} y_p = 0 \ \ \forall \text{ short connections } h \in \tilde{S} \tag{4-4}$$

$$\sum_{s \in S} \sum_{r \in R_s} b_{cr} x_s - \sum_{p \in P} a_{cp} y_p - \beta_c + \alpha_c = 0 \ \ \forall \text{ critical connections } c \in \tilde{C} \tag{4-5}$$

$$\sum_{p \in P} c_p y_p \leq (1 + r).c_{OC} \tag{4-6}$$

The objective (4-1) is to minimize the number of critical connects in the selected crew pairings. Constraint (4-2) ensures that exactly one maintenance solution is selected. Constraints (4-3) guarantee that each flight leg is covered by exactly one crew. Constraints (4-4) ensure that only feasible short connects are included in the crew pairing solution. Constraints (4-5) count the number of critical connects in the pairing solution that are not in the aircraft maintenance routing solution. Constraint (4-6) ensures that the cost of the selected crew pairings is close to minimum crew pairing cost.

5.3 Implementation

5. 3. 1 Introduction

Our objective is to obtain a sense of the trade-off between schedule robustness and crew costs. Crew cost is the second largest expense of the airlines after fuel (Barnhart et al [3]), and hence, increases in robustness must be weighed carefully against the increases in crew costs necessary to achieve this added schedule reliability. We implement the model of Agbokou: "Integrated Robust Routing and Crew Model" and our model. We run these models on the Crew Pairing optimization software of Carmen System. The advantage of using an industry product is that it can solve important scheduling problems; the drawback is that it limits us in the model formulation.

5. 3. 1. 1. Fleet used

The fleet used in our analysis is the American Airlines Boeing 737 fleet. It comprises 77 aircraft that can transport up to 142 passengers with an average stage length of 1108 miles. In 2004, their average number of block hours per day was 9 hours and the aircraft performed an average of 3.2 departures per day with a load factor of 71.3 %. (source:Aviation Daily)

The Boeing 737 is the second smallest airplane in the fleet of American Airlines (see Figure 5.4). It transports more passengers (142 passengers compared to 129 passengers) and travels a longer distance on average (1010 miles on average compared to 891 miles on average) than the M80. This aircraft is appropriate for our analysis because average of 3.2 aircraft departures per day provides opportunities for crews to operate different aircraft. Therefore, this relatively small problem is useful for testing our ideas for adding robustness to crew schedules.

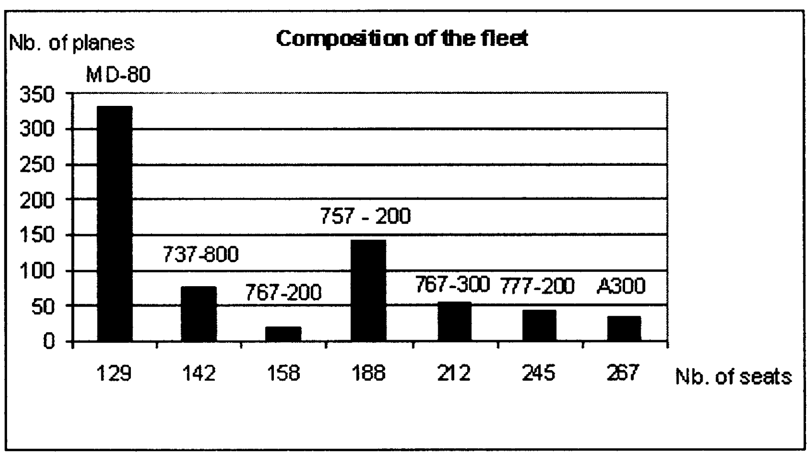

Figure 5.4: Composition of the American Airline fleet

5. 3. 1. 2. Destinations

Figure 5.5 was generated with the "geoplot" function of RAVE. It shows graphically the destinations covered by the Boeing 737 fleet. The fleet is used mainly to cover domestic destinations in the United States. In addition, it flies to Canada (Toronto and Montreal), Central America and the Caribbean. There are 233 flight legs for the daily problem and 1589 flights legs for the weekly problem.

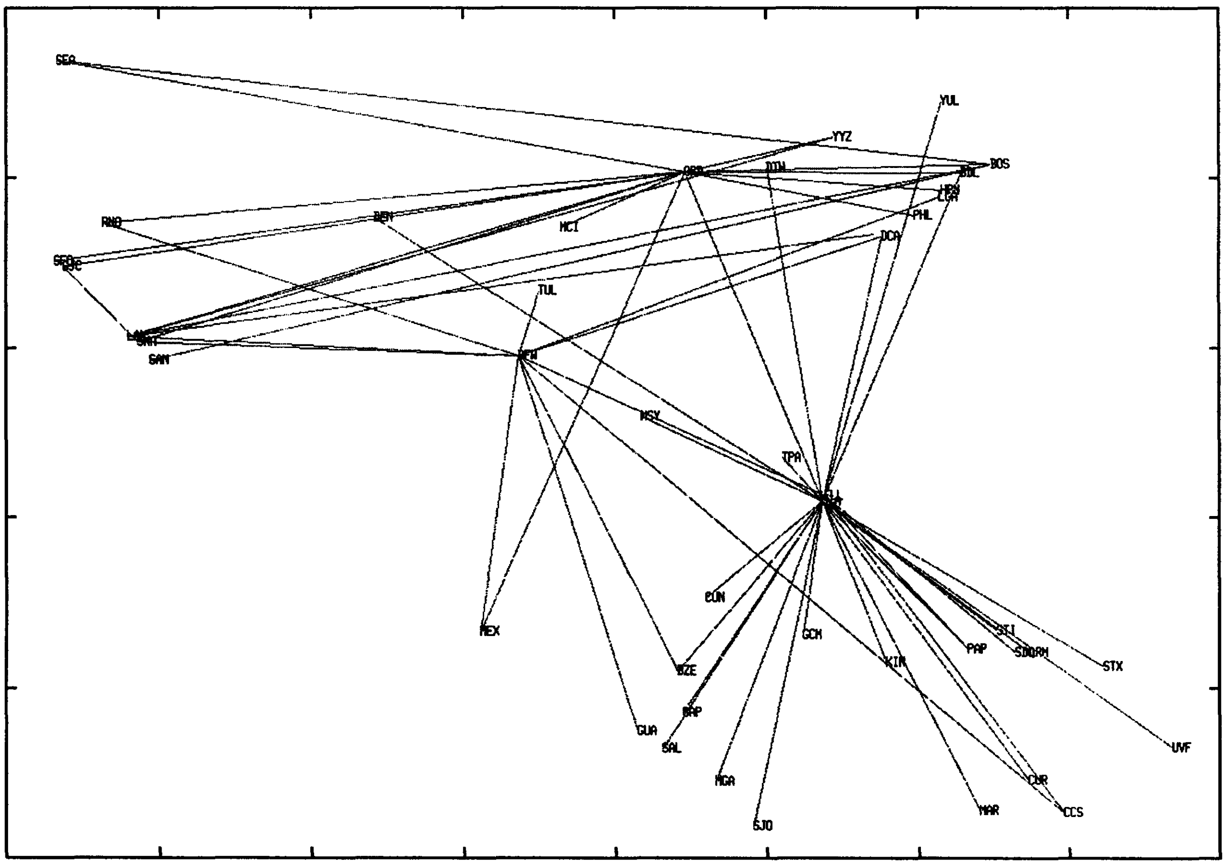

Figure 5.5: Geographic plot of the flights legs and destinations of the fleet

On geographic plot, Miami or Fort Lauderdale initially appears to be a hub for this fleet, because a significant number of flight legs are connected to these airports. However, this impression comes from the fact that the plot does not represent the frequency of each flight leg. If we look at the most active airport, it turns out to be Dallas.

5. 3. 1. 3. More about the RAVE Optimizer

To implement some of the robust crew scheduling ideas discussed, we used the Carmen RAVE optimizer. This software (RAVE stands for "Rule and VAlue Evaluator") is currently used by 20 airlines and three railway companies. It enables us to solve large scale problems; however, one downside is that the software does not allow us to implement the robust models exactly as presented above. Instead, we use the general ideas of these models and incorporate them into our approach by manipulating the basic crew pairing's cost function, as follows:

$$\min \sum_{p \in P} c_p y_p$$

subject to

$$\sum_{p \in P} \delta_{ip} y_p = 1 \ , \quad \forall i \in F$$

$$y_p \in \{0,1\}$$

The Rave optimizer minimizes the cost function and ensures that all flight legs are covered. All pairings formed respect FAA and airline rules that are coded in the optimization tool. The optimizer calculates the cost for all these pairings and selects the cost minimizing set.

All pairings are represented in the graphic environment (Figure 5.6), enabling planners to manipulate and modify them as desired. For example, planners might protect, that is, require some good pairings to be contained in the optimizer's solution.

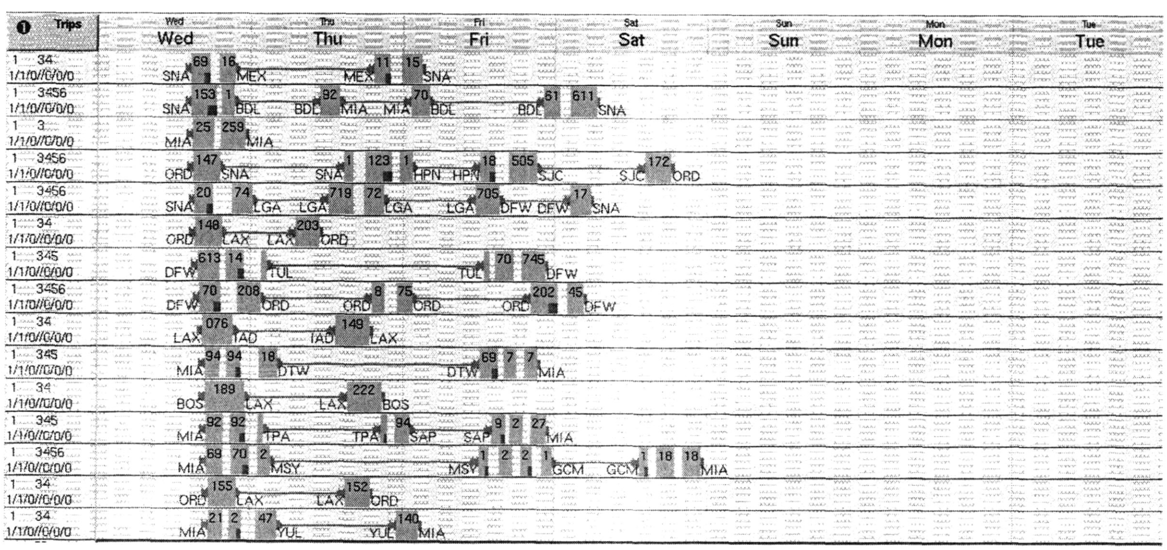

Figure 5.6: Screen shot of Carmen Crew Pairing optimizer

5. 3. 1. 4. Limitations

Our analysis is limited by two factors. First, aircraft routings are fixed and second, we cannot modify the cost function; we can only add penalties to the costs of the pairings in the model. As a consequence, we create penalties that capture our model's objective to maximize the use of short connects and the objective of Agbokou's model to select robust routings and crew pairings. In Table 5.2, we compare these models and present our modifications to these formulations.

	Robustness Criteria	Objective	Penalty of the pairing in RAVE	Advantages of the model
Our approach	Maximize the number of short connects	Crew to follow their plane the maximum number of times	We give a penalty for each aircraft change in the pairing	The criterion guarantees a strong level of robustness
Integrated Robust Routing and Crew Model (Agbokou)	Minimize the number of critical connections	Crew either to stay with the plane or to have *enough* time to connect between aircraft (more than 1hour 15)	We give a penalty for each critical connection in the pairing	Tight crew connections between different aircraft are replaced by short connects or long connections

Table 5.2: Comparison and Adaptation of the models

5. 3. 2. Implementation of the "Maximization of the short connects" model

5. 3. 2. 1. Adaptation

We need to adapt our model to the Carmen Crew Pairing software. Therefore, to maximize the number of short connects and have crews stay with their aircraft instead of changing aircraft between flights, we place a penalty whenever a crew makes an aircraft change between flights.

5. 3. 2. 2. Results

We solve this model for the weekly problem for the American Airlines Boeing 737 fleet, involving 1589 flight legs. By varying the penalty placed on an aircraft change, we were able to obtain solutions with different numbers of aircraft changes and hence different crew costs. In computing crew costs, we did not include the penalty costs placed on aircraft changes by crews, thereby ensuring that our cost comparisons were valid. We were therefore able to examine the effect of the number of aircraft changes on crew cost.

	Crew Costs	Number of Duty Days	1 leg Duty	Percentage Change in crew costs compared to the baseline	Aircraft Changes	Percentage Change
Baseline Model	$1,617,199	2018	555	--	336	--
Penalty - $0.5	$1,617,238	2072	592	0.002%	321	-4.464%
Penalty - $2.50	$1,617,404	2112	620	0.013%	304	-9.524%
Penalty - $5	$1,617,563	2070	603	0.023%	293	-12.798%
Penalty - $25	$1,618,353	2130	636	0.071%	239	-28.869%
Penalty - $50	$1,620,298	2138	643	0.192%	227	-32.440%

Table 5.3: Crew cost and Aircraft change in the weekly problem

In Table 5.3, we present our findings. Crew costs correspond to the weekly cost of the set of pairings selected to cover the 1589 flights leg of the schedule. This calculation is based on the hypothesis that the next week's schedule is exactly the same, which might not be the case at the end or at the beginning of the month.

The number of duty days represents the number of pilot days needed for covering all the legs. If this number increases, it means that each pilot flies less on average.

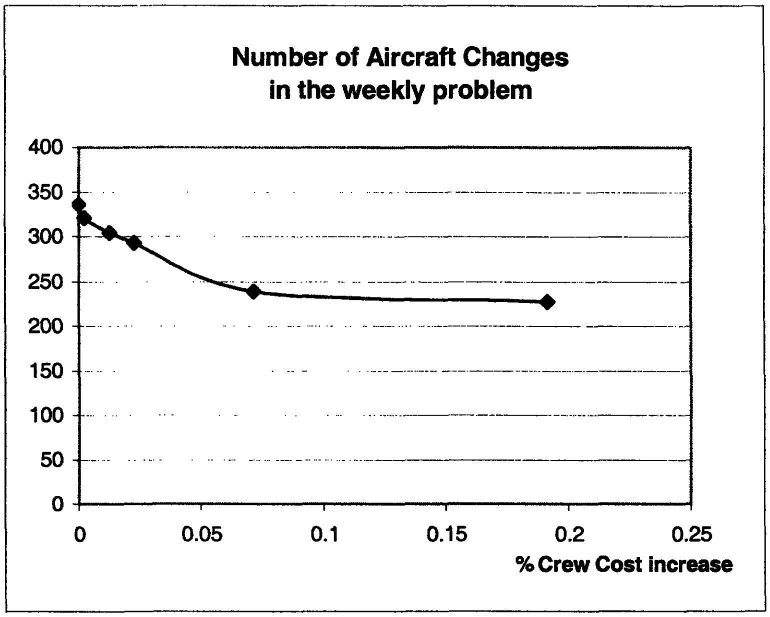

Figure 5.7: Trade off between aircraft changes and crew cost increase

At first glance, the results of Figure 5.7, that a 0.2% increase in the crew costs can generate a gain of 32.4% in robustness by decreasing the number of aircraft connections, are quite remarkable. However, by further examination of the data, this impressive result is tempered by the fact that we have at the same time an increase in the Duty Days (that is, the total number of pilots needed to cover the schedule), especially in the number of 1 leg crew duties. More details are provided in Table 6.2.

	Aircraft Changes	Short Connects	Number of Duty Days	1 leg duty	2 legs duty	3 legs duty	4 legs duty
Baseline Model	336	252	2018	555	334	106	14
Penalty - $0.5	321	250	2072	592	323	103	14
Penalty - $2.50	304	271	2112	620	320	105	15
Penalty - $5	293	269	2070	603	314	106	12
Penalty - $25	239	307	2130	636	297	105	13
Penalty - $50	227	316	2138	643	292	106	13

Table 5.4. : Aircraft changes, shorts connects and number of legs per duty

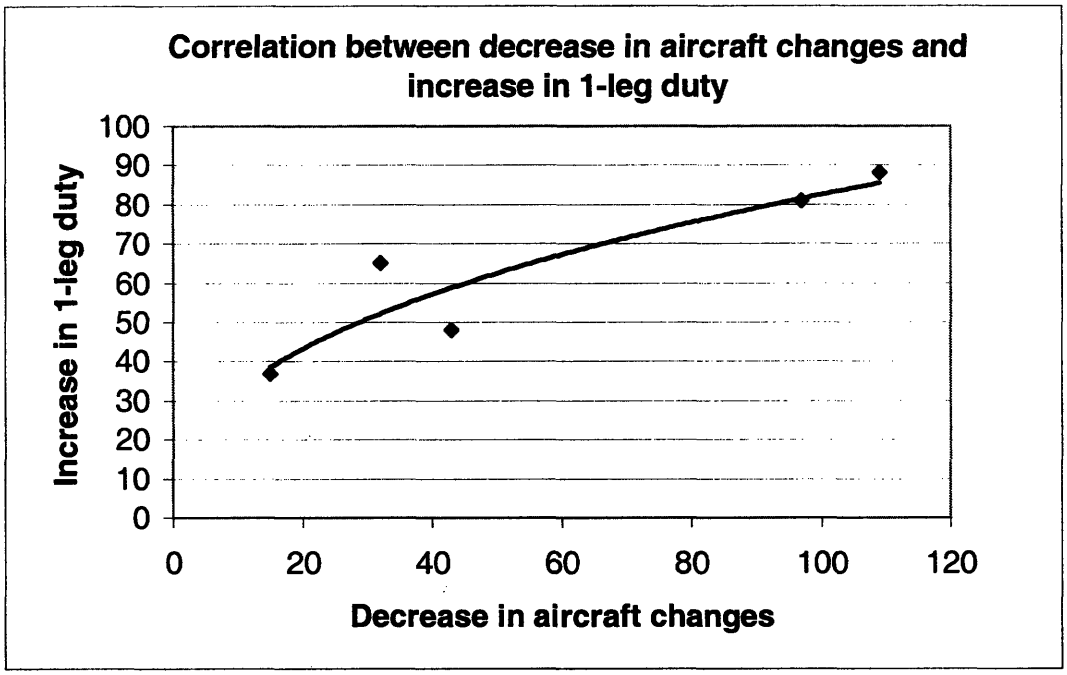

Figure 5.8: Correlation between decrease in aircraft changes and increase in 1-leg duty

In Figure 5.8, we show the correlation between the number of reductions in crew connections between different aircraft and the number of increases in 1-leg duties. We conclude that most of the connections between different aircraft were eliminated by splitting the duty at the connection into two and assigning two crews.

Replacing these connections between aircraft are short connects, with the number of aircraft connections decreasing by 32%, the number of short connects increased by: 25%.

The drawback of this solution is that it requires 6% more pilots to fly the same schedule. Therefore, the value of this approach depends on how crew are paid. We don't include in our model the additional compensation crews receive when they don't fly. Instead, we consider only crew costs related to total assigned block time. In our solution, this block time expense increases by only 0.2% .

The need to have more crew has associated costs that are difficult for us to quantify.

5. 3. 3. Implementation of the Agbokou model

5. 3. 3. 1. Adaptation

In our adaptation of Agbokou's model, we discourage the use of critical connections in the crew pairing solution by placing a penalty on each critical connection. We evaluated this model using data representing the daily problem for the American Airlines Boeing 737 fleet, involving 240 flight legs. .

All the critical crew connections are not equivalent. Indeed, a critical crew connection 45 minutes after the scheduled arrival time is more likely to be disrupted than a critical crew connection 1 hour and 15 minutes later. Therefore, we consider 5 cases, each of which has a different shape for the penalty placed on critical crew connections. Figure 5.9 shows a plot of the penalty for each case.

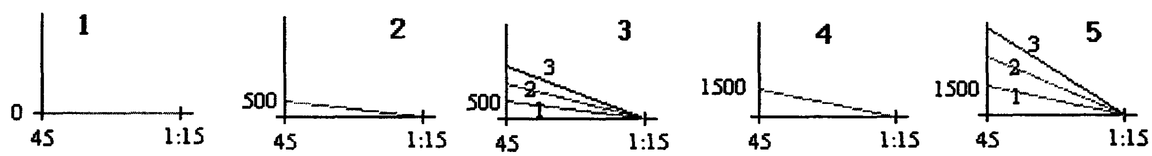

Figure 5.9: Five different Penalty shapes

<u>Case 1 (0 penalty)</u>: This is the baseline case. There is no penalty for the critical connection.

$$CritCost_{pairing} = 0$$

<u>Cases 2 and 4 (linear)</u>: The critical connection penalty is linear starting from \$500 and \$1500 respectively for Cases 2 and 4. The penalty is highest when the connection time is 45 minutes (minimum connection time) and it decreases linearly with time until the connection time is 1 hour 15 minutes, where the penalty is \$0. By definition, a connection beyond 1 hour 15 minutes is not a critical connection.

$$CritCost_{pairing} = \sum_{leg \in pairing} Penalty \times \frac{(1:15 - ConnectionTime_{leg})}{30}$$

Cases 3 and 5 (linear special): In addition to being a linear function of time from the minimum connection time (as per Cases 2 and 4), the penalty on a critical connection is also a function of the number of flight legs left in the duty after the critical connection. We have illustrated how critical connections are undesirable because a flight delay is likely to cause a crew to delay their next flight or worse, miss it. In addition, this effect is propagated to the additional flights the crew has remaining in its duty. Therefore, in order to improve the robustness of the crew schedule, it is desirable to limit critical connections to those with fewer remaining flight legs. The penalty cost function in Cases 3 and 5 accounts for this by multiplying the penalty by the number of flights remaining in the duty.

$$CritCost_{pairing} = \sum_{duty \in pairing} \sum_{leg \in duty} Penalty \times \frac{(1:15 - ConnectionTime_{leg})}{30} \times RemainingFlightLegs_{leg}$$

5. 3. 3. 2. Results

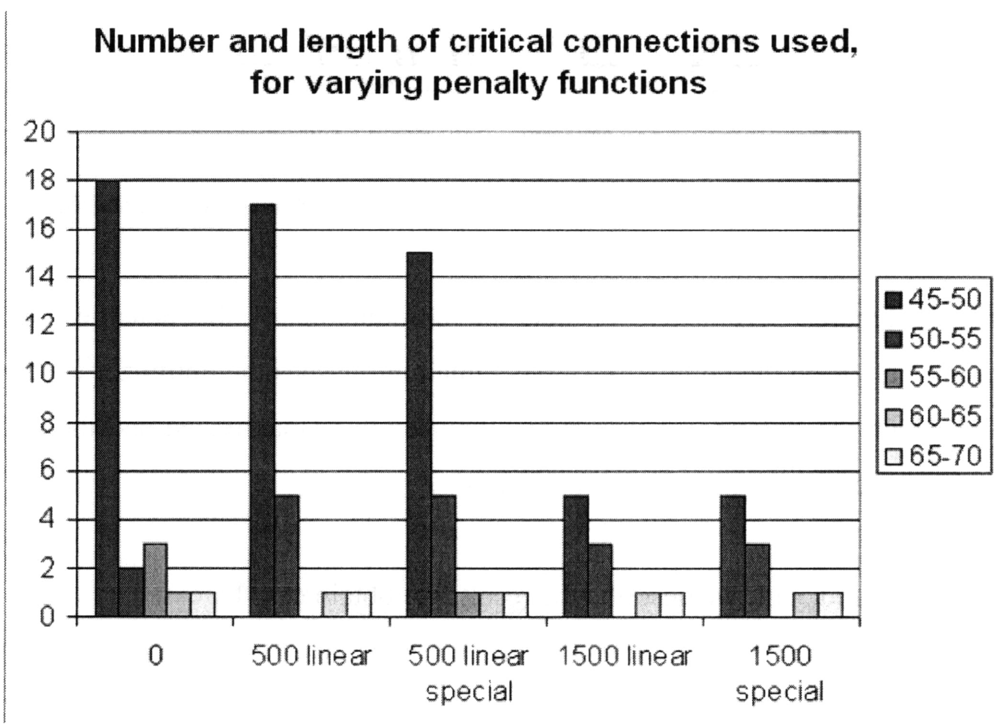

Figure 5.10: Number and length of critical connections used, for varying penalty functions

The Figure 5.10 shows the number and length of critical connections used in the solution, for the different penalty functions. The objective of the different penalties was to increasingly push out of the solution those critical connections with a small amount of connecting time. It works fairly well and we see that the optimizer was able to reduce the number of critical connections between 45 and 50 minutes from 18 to 5, which can represent a significant gain in robustness.

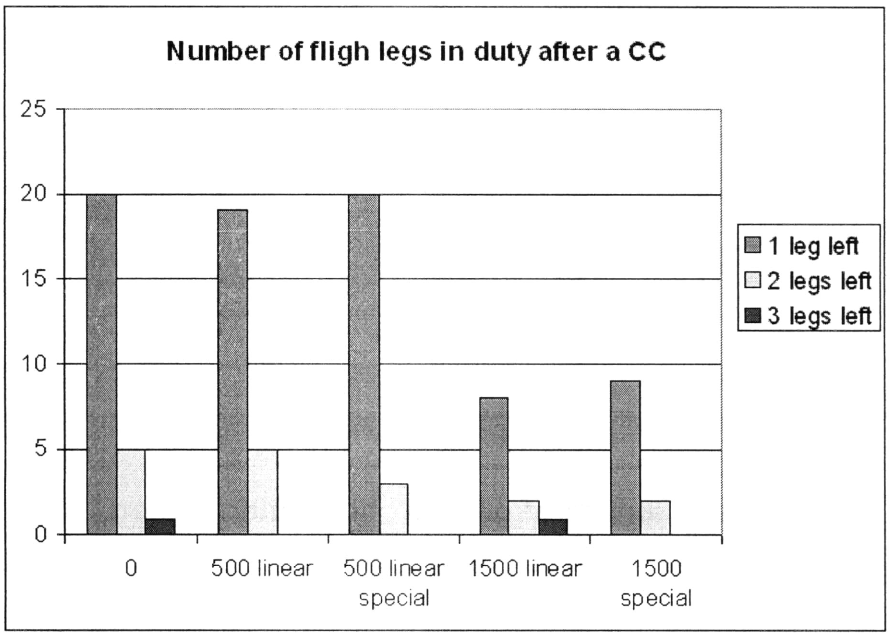

Figure 5.11: Number of Flight legs in duty after a Critical Connection

We observe in Figures 5.10 and 5.11 that the introduction of the penalties on critical connections as well as on the number of flights remaining pushes the optimizer to select crew schedules in which the duration of the critical connection is maximized and the number of flight legs after the critical connection is minimized.

Using Figures 5.10 and 5.11, we describe the results we obtain case by case.
Case 1 is the baseline case without any penalty. We have 26 critical connections among the 49 aircraft changes. In this crew schedule, we observe that there are a large number of connections between 45 and 50 minutes. Furthermore, there is one critical connection which has three flight legs remaining. Therefore, this crew schedule is not very robust.

Cases 2 and 4 are the cases in which the penalty on a critical connection takes a higher value with no penalty on the number of flights remaining in the duty after the critical connection. Referring to the plots, we observe that the number of critical connections between 45 and 50 minutes has decreased. Robustness is improved in these cases.

In Cases 3 and 5, in addition to a penalty on critical connections, there is also a penalty on the number of flights remaining in the duty after the critical connection. We observe that this penalty, as expected, pushed the optimizer to select a solution that seems more robust than Cases 1, 2, and 4.

We also note that the aim is not to get rid of all the critical connections. If you break certain critical connections, cost savings might not be achieved because they are unlikely to lead to delays. Therefore, we essentially seek to find a tradeoff between robustness of the solution (and savings in recovery) and savings in the planned operations.

Concerning the cost of the solution, the costs associated with all solutions are within 1% of the cost of $298,714.5 for the baseline case.

5. 3. 4. Conclusions

We adapted 2 optimization models, our model and that of Agbokou [1] to study their benefits in terms of reducing delays resulting from crews.

The advantage of the Agbokou model adapted to this problem is the possibility to give a shape to the penalty depending on the connection time and the number of flight legs left in the duty. It would also be possible to add some parameters like the expected delay of the flight, or airport congestion levels.

We conclude that many crew solutions exist within 1% of the baseline optimal cost. Hence, there are opportunities to find near-optimal solutions that are more robust than those being generated with conventional models that ignore robustness. Hence, generating robust crew schedules can potentially reduce the 18 % of the non-system delays that are related to crews, without excessive costs to the airlines.

However, our solution requires 6% more pilots to fly the schedule for a 25% increase in short connection utilization. Hence, our estimation of an associated 0.2% cost increase should be augmented to include the costs to compensate pilots for non-flying duty time. These costs represent an important part of crew costs that are airline dependent. We are unable, however, to compute these costs.

Chapter 6

Conclusions

6.1. Summary

Delays and congestion are certain to grow in the near future with the increasing trend of air traffic. This study analyzes delay trends and proposes and evaluates new models aimed at reducing delays caused by crews. We begin this thesis by conducting an analysis of current delays in the airline industry, followed by a discussion of some of the measures airlines are taking to maintain on-time performance, and we end with a review, implementation and evaluation of crew scheduling models aimed at achieving increased reliability.

In chapter 2, we present a broad picture of the delays in the US. We examine the causes as reported by airlines and by the US Department of Transportation. 90% of delays stem from the 3 following sources: the National Aviation System, the Air Carrier, delay propagation (aircraft arriving late). From the airline viewpoint, more than 70% of the delays are caused by the system, with airlines having little to no control over these delays. The causes of system delays include weather, heavy traffic volume, and closed runways. We compute that weather itself is directly responsible for 48.7% of the U.S. flight delays.

In our study, we show that delays don't appear totally at random in the system. Instead there are yearly variations in some performance indicators, including on-time arrivals and taxi-out times; seasonal variations in delays and cancellation; and daily variations in the causes of delays. The 10% of US airports that serve 65% of the air traffic display similar on-time performance to that of smaller airports, but experience much greater taxi-out times (6 minutes on average), reflecting high levels of congestion at these large airports.

In chapter 3, we study Delta Airline's de-banking of their Atlanta hub in response to increased delays and inability to execute the flight schedule as planned. With de-banking, Delta has removed their banks in Atlanta and spread-out flight departures and arrivals throughout the day. The key performance indicators show that de-peaking has had positive effects on the operations of Delta, and generally speaking, on all airport operations, even taking into account that the on-time arrival rate of the competitors decreased. The on-time percentage of Delta increased and taxi-out times decreased. Although de-banking theoretically reduces the number of opportunities to swap airplanes at peak hours, Delta compensated for this by adding more flights so that more aircraft are on the ground at the same time.

In chapter 4, we discuss an approach widely utilized by the airlines to gain on-time performance, namely: adding buffer time to scheduled operations to gain robustness and improve on-time performance. Our case study involving flights from Atlanta to Dallas show different practices and accomplishments among the 3 airlines that serve this market. American Airlines adds the most amount of buffer time and Airtran the least. However, the on-time performance of American Airlines is disappointing, not because of its tight bad schedule, but instead because long taxi-in times result from gate unavailability. We present least square regression and the Robust Majorize-Minimize approach to evaluate and compare the airlines' buffering practices. Some airlines, like United Airlines, Airtran and Continental, underestimated the need for buffer times at Boston Logan in January 2005, a period in which extreme weather severely disrupted operations at Logan.

In chapter 5, we review different robust scheduling models specifically targeted to decrease delay propagation. We propose a model, aimed at reducing delays caused by crews, that minimizes the number of times crews must transfer between different aircraft during their workday. From our implementation and evaluation of two different models, each with an objective to reduce delays resulting from crew unavailability, we conclude that an increase of 0.2% in crew costs could enable a decrease of 32% in the number of times crew must transfer between aircraft during their workday and an increase by 25% in the number of times a crew continues on the same aircraft.

6.2. Future Research

In this study we illustrate potential improvements in schedule reliability that are attainable with robust scheduling approaches, without incurring large increases in crew costs. A further study could address evaluate, using historical data, how much schedule non-robustness affects realized costs as compared to planned costs. This would shed light on how optimization models should be formulated to ensure that realized, and not planned, costs are minimized. This suggests another important direction of research: how to integrate robustness considerations into the schedule planning optimization process. There are many associated questions, including is it profitable to cater to time-sensitive passengers and if so, how should airline schedules be structured and how should resources by deployed to achieve on-time performance?

References

[1] Agbokou, B. (2004) "Robust Airline Schedule Planning: Review and Development of Optimization Approaches" *SM Thesis, Massachusetts Institute of Technology*

[2]: Airport Capacity Benchmark Report (2004)
http://www.faa.gov/events/benchmarks/2004download.htm

[3] Barnhart, Cohn, Johnson, Klabjan, Nemhauser and Vance (2002) "Airline Crew Scheduling" *Handbook of Transportation Science, Chapter 14*

[4] Bogusch, L. (2003) "Rethinking the Hub and Spoke Airline Strategy: An Analysis and Discussion of American Airlines's Decision to De-peak its Schedule at O'Hare International" *SM Thesis, Massachusetts Institute of Technology*

[5] BTS (2005) "Airline On-Time Performance and Causes of Flight Delays"
http://www.bts.gov/help/aviation.html

[6]: Booz·Allen & Hamilton (2001) "Punctuality: How Airlines Can Improve On-Time Performance?" *Airline consultants*

[7] Carmen Systems Crew Paring software
www.carmen.se/air_products/pdf/air_crew_pairing.pdf

[8] Credeur M.J. Operation Clockwork radical, risky (2005) *American City Business Journals Inc.*

[9] Ehrgott and Ryan (2002), Constructing robust crew schedules with bicriteria optimization. *Journal of Multi-Criteria Decision Analysis 11(3):139-150, 2002.*

[10] Insightful Corporation (2001). "S-Plus 6 for Windows User's Guide". *Insightful Corporation, Seattle, Washington, July 2001.*

[11] Klabjan, Chebalov (2002) "Robust Airline Crew Scheduling: Move-up crews" *Proceedings of the 2002 NSF Design, Service, and Manufacturing Grantees and Research Conference.*
https://netfiles.uiuc.edu/klabjan/www/proceedings/nsf02.pdf

[12] LaMont R (2004) "Airline Maintenance Outsourcing: Savings at Any Cost, Putting Safety At Risk While Exporting High-Paying American Jobs"
http://www.amfa4.org/Outsourcing2.htm

[13]: Maynard (2005) "Will the Flight Be on Time? It's Anybody's Guess", *New York Times, August 08 2005*

[14] Noether (1991). "Introduction to Statistics the Nonparametric Way" Springer-Verlag, New York, NY.

[15] The MITRE Corporation (2005) "Scenario B–Congestion Due to En Route Severe Weather"
http://www.mitrecaasd.org/work/project_details.cfm?item_id=139

[16] United Airlines (2005) "Airport Operation United Airlines fact sheet"
http://www.catravelsmart.com/UA%20Fact%20Sheet.htm

[17] Zhang, Y., M. Menendez, and M. Hansen, (2004) "Analysis of De-peaking Strategies Implemented by American Airlines: Causes and Effects" *TRB National Conference.*