ANALYSIS OF THE IMPACT OF COLLABORATIVE GROUND DELAY PROGRAMS IN AIR TRAFFIC CONTROL

by

John R. Jensen

Submitted to the Department of Civil and Environmental Engineering
in Partial Fulfillment of the Requirements for the
Degree of

Master of Science in Transportation
at the

Massachusetts Institute of Technology
February 1999

Author...
Department of Civil and Environmental Engineering
January 26, 1999

Certified by...
Amedeo R. Odoni
T. Wilson Professor
Department of Aeronautics and Astronautics
and Civil and Environmental Engineering

Accepted by ....
Andrew J. Whittle
Chairman, Departmental Committee on Graduate Studies

ANALYSIS OF THE IMPACT OF COLLABORATIVE GROUND DELAY PROGRAMS IN AIR TRAFFIC CONTROL

by

John R. Jensen

Submitted to the Department of Civil and Environmental Engineering on January 26, 1999, in partial fulfillment of the requirements for the Degree of Master of Science in Transportation

Abstract

Ground delay programs (GDPs) have been a pervasive feature in U.S. air traffic management for many years. Recently, an enhanced version of the ground delay programs has been instituted as part of collaborative decision making, a new approach to air traffic management. This enhanced versions holds out the promise of improving the system by making better use of the scarce resources during capacity-constrained operations, and by giving the airlines greater flexibility in managing their flights. Prototype operations of the enhanced program went into effect on January 23rd at two airports, and was subsequently extended to two additional airports on April 28th. This thesis analyzes and evaluates the information collected from the first nine months of prototype operations.

A computational analysis model has been developed to analyze the effects of the enhanced GDPs. The model uses information about flights arriving at airports under prototype operation and about the GDP programs themselves. The model extracts the critical pieces of information, and organizes this information in a database. This database is then used as the basis for all subsequent analysis. The model has been used to analyze the incidence of GDPs, ground hold delay savings gained by airline flight substitutions and GDP compression, capacity utilization during restricted operations, unexpected airborne holding, airborne delays during GDPs, and the increase in flight cancellations due to GDPs.

Thesis Advisor: Dr. Amedeo R. Odoni
Title: T. Wilson Professor of Aeronautics and Astronautics
and Civil and Environmental Engineering

Dedication

This study is dedicated to MJ, my wife, and to Mark and Todd, our two sons. Without their love and support over the years, I would never have been able to engage in my studies at M.I.T. I owe them a debt of gratitude, especially for their endless patience and understanding over the last ten months, when I would frequently absent myself to work on this study.

Acknowledgements

I am grateful for the support of the Federal Aviation Administration, which funded the research for this thesis.

I am especially indebted to professor Amedeo Odoni, who suggested I work on this study. His constant prodding and probing questions often induced me to reevaluate my line of thinking and explore new avenues, gaining a new understanding of and a better appreciation for the complexities of the subject along the way.

I would also like to thank Joe Sussman, Cynthia Barnhart and Carl Martland in the Center for Transportation Studies, and John Hansman and Peter Belobaba in the Department of Aeronautics and Astronautics. I have enjoyed their teachings and their wisdom over the last six years more than I can say in words.

I have had the opportunity to meet many gifted people and make many new friends during my stay at M.I.T. They are a big part of why I have found this time to be such an enjoyable and learning experience.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1 Introduction

1.1 Background

Ground delay programs (GDPs) have long been a feature in the management of the U.S. airspace. They are designed to manage operations at an airport, when arrival demand is projected to exceed the available capacity for a substantial amount of time. GDPs work by pre-assigning a specific, controlled time of arrival (CTA) to a flight, and then reserving the arrival slot for that flight. Since the flight time between the origin and destination airport can be determined with reasonable precision, this also determines a controlled time of departure (CTD). The flight is then held on the ground at the originating airport until the CTD time. At that time, it can depart, with the assurance of being able to proceed to the arrival airport and being able to land with a minimal amount of unexpected airborne holding (UABH).[1]

When working properly, GDPs have two major benefits. First, they increase the overall level of safety in the system by reducing the number of airplanes present in the terminal airspace of an airport. Since each airplane has a specific arrival time assigned to it, it will spend less time in airborne holding and hence in the terminal airspace. Second, there is a substantial gain in economic efficiency to the airlines operating the flights. Since delays are incurred on the ground rather than while aloft, fuel is not used. Fuel consumption is a major expense item for all airlines, so reducing the airborne holding directly reduces the amount of fuel spent to complete the trip.

The GDP process starts when the FAA determines the arrival acceptance rate (AAR) in response to current or foreseen reductions in airport capacity. This reduction may be caused by inclement weather, airport construction, or special runway operations. The AAR specifies how many flights the airport is projected to be able to handle for each given hour. AARs for several consecutive hours are typically defined at the same time.

A critical element in GDPs is the allocation and distribution of arrival slots. In the past, an algorithm known as the Grover-Jack algorithm has been used to perform this task. It allocates the available arrival slots at an airport according to the estimated time of arrival (ETA) of the

1 This section is heavily indebted to the discussion of this subject found in [1]. For a more detailed discussion, please refer to this paper.

incoming flights. The algorithm preserves the order of the original set of flights, but stretches out the arrivals in time to make the number of arrivals within each controlled hour stay within the available capacity at the airport, as defined by the AARs. It also has to factor in slots for flights that are exempt from the GDPs. Flights can be exempt for a number of reasons. International flights, general aviation flights and flights already airborne are exempt and cannot be issued a ground delay. Other categories may also be classified as exempt, such as flights originating at airports where de-icing is in effect.

Since Grover-Jack uses the flight ETAs as its means of rank ordering the set of flights to be controlled, any delay caused within an airline for a particular flight may end up resulting in a double penalty for that flight. This happens if the flight, after being delayed by the airline itself, is subsequently subjected to the Grover-Jack algorithm. This has caused some concern within the airline industry, and has acted as a disincentive for submitting updated ETAs to the FAA. This issue has been one of the factors driving the exploration of new ways to address the scheduling of flights into capacity-constrained airports.

Collaborative decision making (CDM) is a key component in the larger concept of the National Airspace System plan known as Free Flight. Its key goal it to seek a new way of ensuring that the airspace is used in a safe manner, while at the same time turning over more decision making responsibilities to the users of the system.

A CDM working group has been formed to steer the activities underway in this area. New algorithms for air traffic management have been developed, and a communications infrastructure has been established to ensure that the FAA and participating airlines have a common view of the arrival demand picture at every U.S. airport. Prototype operations of GDP-E (Ground Delay Program - Enhanced) under the CDM Program commenced at San Francisco (SFO) and Newark (EWR) on January 23rd, and were extended to La Guardia airport, New York (LGA) and St. Louis, Missouri (STL) on April 28th. GDP-E was subsequently extended to all major U.S. airports on September 8th.

A new computer network was set up as part of CDM to facilitate the flow on information between the airlines and the FAA. Named the AOCnet (Airline Operations Center network), it connects the participating airlines' operations centers with the FAA Air Traffic Control

System Command Center (ATCSCC) and the CDM hub-site at the Volpe National Transportation System Center (NTSC). Every five minutes aggregate demand lists (ADLs) are distributed across this network to the airlines. These lists contain the current status information of all flights in the system. The intent is for this information to be used by the airlines to facilitate them in their scheduling decisions. A computer application, the Flight Schedule Monitor (FSM), has been developed as part of CDM. FSM makes direct use of the ADL information, and enables its users to analyze the impact of alternative scheduling strategies before deciding upon which course of action to pursue.

A new arrival slot allocation algorithm named Ration by Schedule (RBS) has been developed as part of the CDM effort. As the name of the algorithm implies, arrival slot allocation is now being governed by the *scheduled* time of arrival, as opposed to the *estimated* time of arrival. Any delays internal to the airline companies do not influence the GDP arrival slot allocation process. This means that there no longer is an incentive for the airlines to withhold flight ETA updates, leading to a more accurate picture of the overall state of the system.

The airlines still maintain the ability to substitute flights within the pool of arrival slots allocated to them. This enables the airlines to swap an arrival slot allocation of one flight for another, moving up one flight in time and another back in time. It furthermore allows the airlines to cancel a flight and move another one up to take the place of the cancelled flight.

The RBS algorithm helps allocate the available arrival slot resources among the airlines. It does not, however, address the issue of what to do when an airline cannot make use of an allocated slot. This situation occurs when a flight is cancelled or moved up in time to an earlier slot, and the airline has no other flight that can be moved to take over the abandoned arrival slot. The *compression* algorithm has been developed to address the issue of allocated slots going unused. In simple terms, it does this by filling up any unusable slots with later-arriving flights, moving up flights to fill up the available slots as much as possible. The compression will respect the original rank ordering of the flights, while at the same time ensuring that a flight is not scheduled any earlier than its earliest possible arrival time.

An example may help illustrate the three interrelated concepts of ration by schedule, airline substitutions and compression. The example is borrowed from [1], but its graphical form given in figure 1-1 is new. In the example, 11 flights (01 through 11) from three different airlines (A,B,C) are scheduled to arrive between 7:00 and 8:30. Due to capacity restrictions at the airport, only one flight may land every 10 minutes. This arrival rate is admittedly unrealistic, and is used here only to make the exposition clearer. The initial situation is depicted to the left in Figure 1-1. Time of actions move from left to right in the figure.

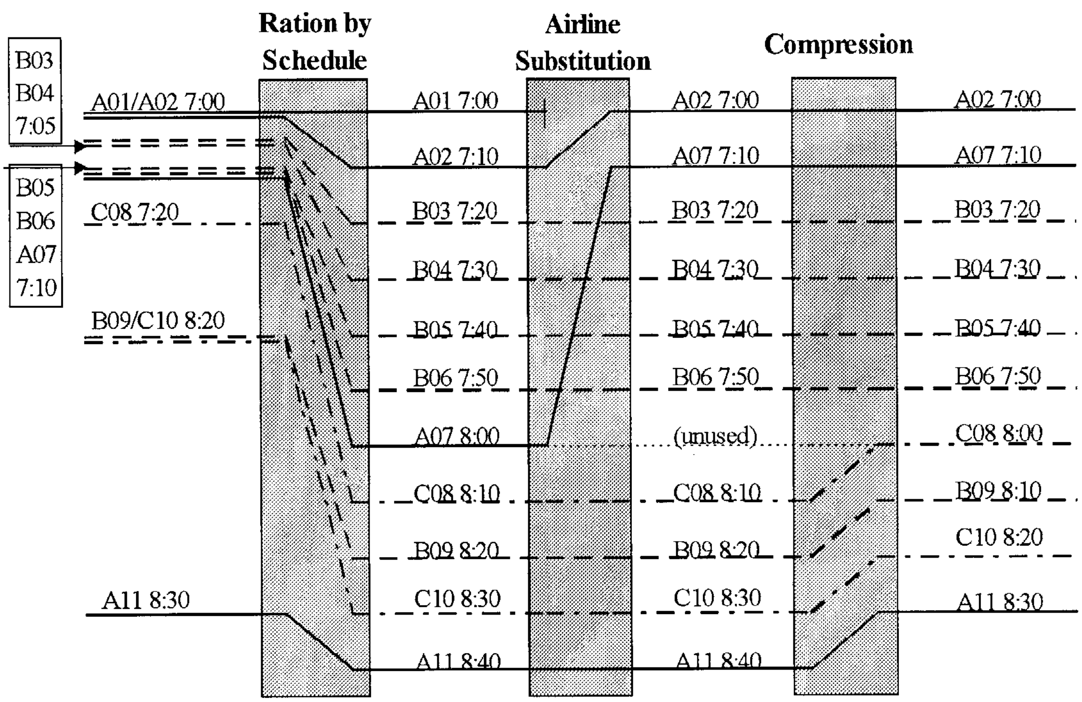

Figure 1-1: Compression example

Flights A01 through A11 are scheduled to arrive at the airport between 7:00 and 8:30. Due to an inability to handle more than six flights per hour, the RBS algorithm is run to reschedule the arrivals. This results in the arrival line-up shown to the right of the *Ration by Schedule* semi-transparent box. Overall, a total delay of 320 minutes is allocated, spread out over the airlines by 70 minutes to A, 150 minutes to B and 100 minutes to C. These delays are computed by calculating the delay for each individual flight as the difference between the scheduled and RBS-allocated arrival times.

After the initial RBS allocation has been completed, and the resulting arrival slot allocations broadcast to the airlines, they have a chance to substitute flights in the slots that have been assigned to them. These actions are shown in the *airline substitutions* box. In the example, flight

A01 is cancelled, flight A02 is moved up to take over A01's slot, and flight A07 is moved up in the now-vacated slot of flight A02. This leaves the slot at 8:00 unused. Airline A does not have any flights that can be moved into this slot, since its only other arriving flight has an earliest arrival time of 8:30, well after the 8:00 opening. Airlines B and C have flights that can move into this slot, but they cannot do so, since the slot is owned by A. Airline A is the only one to perform any substitutions, so it is the only one to see any change in its overall delay. Its delay drops from 70 minutes to 10 minutes, for overall savings of 60 minutes, but at the cost of having one flight cancelled. System-wide there is the same 60 minutes of savings, for a revised total delay of 260 minutes.

Finally, compression is run. The effects of this are shown in the *compression* box. In the example, each of the four last flights is moved up by one slot, in the process clearing out the unused slot. This results in an additional 40 minutes saved system-wide, with 10 minutes going to A, 10 minutes to B and 20 minutes to C. Overall, the delay is now 0 minutes for A (but with one cancelled flight), 140 minutes for B and 80 minutes for C. The total of 220 minutes of delay is the best that can be achieved on a system-wide basis.

1.2 Overview of content

An early assessment of CDM is described in the report issued by NEXTOR entitled *Collaborative Decision Making in Air Traffic Management: A Preliminary Assessment* [1]. This report dealt with several different topics related to CDM, analyzing the changes in the quality of the information, the impact of the distribution of information and the increased situational awareness that this has brought about, the ability of the airlines to make economic resource allocation decisions, and some of the delay reductions and other benefits associated with CDM-based ground delay programs.

The objective of this study is to continue the investigation started in [1], and to build a computational framework that will facilitate future work in this area. The study evaluates some of the effects of the first nine months of prototype GDP-E operations. A more detailed assessment has been made of the ability of CDM to reduce the GDP delays through airline substitutions and GDP compression. Capacity utilization has been reviewed to determine how fully the stated AAR capacity is being utilized, and if there are any signs of systematic or periodic under-utilization. The analysis on airborne holding started in [1] has been extended to cover additional airports and more dates. An investigation has been carried out to evaluate the impact of GDPs on airborne delays. Finally, cancellations have been investigated to determine if the overall number of cancellations increases on GDP days, and whether or not there is a correlation between the FAA arrival slot allocation delay and the number of flights being cancelled.

The data used for this study come from two sources. The first source is the ΔADL files. These files are organized by destination airport, with one file per airport. Each file contains information about all flights arriving at a given airport on any particular day. To limit the size of these files, only flight records that show one or more changes in their content fields from the previous record have been retained. Datafiles for the entire nine months and the four prototype airports have been made available for use within this study. The second source of data covers the GDP control information. This information has been retrieved from a Metron web-site. These two databases have been processed through a series of computational steps to create a database of flight-level information. This database has then served as the data-source for all the analysis work conducted as part of this study.

Chapter 0 describes the results of the analyses that have been performed on the data. The specific areas covered are ground delay programs, compression results, arrival slot utilization, unexpected airborne holding, airborne delay, and cancellations. Chapter 3 summarizes a set of conclusions, and offers many suggestions for further analysis that may be warranted and useful. A series of appendices provide additional details on the computational model. Appendix A contains a list of acronyms used throughout. Appendix B describes the computational model that has been built to help evaluate the large amounts of data being collected. It deals with the issues of collecting and reducing the flight data to a manageable level, collecting the GDP control information, setting up a database to manage this information, and finally extracting the information from this database to provide a basis for analysis. Appendix C provides reference information for the model, i.e., definitions of the database tables, file formats and computer programs used.

2 Results

This study looks at various indicators of performance during the first nine months of prototype operations of GDP-E. In the sections that follow, a number of aspects are examined: ground delay program statistics; effects of GDP compressions; slot utilization; unexpected airborne holding; en-route delays; and cancellations.

The data analyzed cover the first 9 months of 1998. During this period prototype GDP-E operations commenced in San Francisco, California (SFO) and Newark, New Jersey (EWR) on January 23rd, and were extended to La Guardia, New York (LGA) and St. Louis, Missouri (STL) on April 28th. GDP-E was subsequently extended to all major U.S. airports on September 8th. In this report, only data from the four prototype operations are analyzed due to the limited amount of data available for the additional airports that came on-line on September 8th.

2.1 Terminology used

2.1.1 GDP states

Flights can be categorized in a number of ways, e.g., by destination airport, by arrival hour, or by equipment used. One such category is the *GDP state* of a flight. Being able to assign such a state to each flight allows us to investigate, whether or not there are any material differences for flights in the different GDP states. There are three mutually exclusive GDP states.

Active-GDP flight: The flight had a current, active CTA allocated to it at the time it arrived at the destination airport.

Cancelled-GDP flight: The flight did not have an active CTA allocated to it at the time it arrived at the destination airport. It did, however, at one point in time have a specific CTA allocated to it prior to its arrival, and so was a part of a GDP program.

No-GDP flight: The flight never had an CTA assigned to it prior to its arrival at the destination airport.

In addition, the term *GDP flights* will be used as a composite for active-GDP flights and cancelled-GDP flights.

The GDP states will also on occasion be used in conjunction with other subjects. E.g., an active-GDP hour refers to an hour when a GDP is active, and a no-GDP day is a day that does not see any incidence of GDPs.

2.1.2 Hours

All references to hours in this report are given in local hours unless explicitly stated otherwise. To avoid any confusion, they are described by a range of time, e.g., 10:00 to 11:00. However, in tables and figures they are shown as a single value, e.g., the hour between 10:00 to 11:00 would be shown as 10. Daylight savings time is factored in to all applicable local hours.

2.2 Incidence of Ground Delay Programs

2.2.1 Methodology

There are two basic steps in the analysis. First, for each day and airport we determine the number of GDP programs that were run. For each program, we further determine on an hourly basis how many programs were initiated, how many revisions, extensions and compressions were applied, and whether the program was cancelled early or expired at the designated end-time. Second, we aggregate the results according to the needs of the analysis.

Each hour can be in one of three GDP states: (1) The no-GDP state, if a GDP was never defined for that hour; (2) the cancelled-GDP state, if a GDP was at one point defined for the hour, but was subsequently cancelled; and (3) the active-GDP state, if a GDP was in effect for that hour. Since the GDP state does not necessarily change on even-hour boundaries, any given hour may have two or even three different states associated with it. For simplicity, hours containing any amount of active-GDP time are classified as active-GDP hours. Hours containing any amount of cancelled-GDP time but no active-GDP time are classified as cancelled-GDP hours. All remaining hours are classified as no-GDP hours.

All information about the GDP programs has been obtained from the Metron web-site. The web-site lists events on a program by program basis. This information has been transcribed into a fixed-record format and then loaded into the database, where it is now available for use.

As of 7/22, the GDP program information has also become part of the content of the ΔADL files. This has the potential of improving the acquisition of the information, since it can now be obtained through a suitably written program, as opposed to the current, manual process in use. Ultimately this should lead to a faster and less tedious access to the GDP program information.

2.2.2 Monthly summary

Table 2-1 summarizes the number of GDP-E programs that were run at the four prototype operations airports during the first nine months of 1998. The table shows the number of days with GDPs (*days*), as well as the actual number of programs (*progs*), since a given day may have more than one program associated with it. An entry of '-' indicates that GDP-E did not go into effect for LGA and STL until the month of April. The same data is plotted in Figure 2-1.

Month	EWR Days	EWR Progs	LGA Days	LGA Progs	SFO Days	SFO Progs	STL Days	STL Progs
Jan	2	4	-	-	6	10	-	-
Feb	4	5	-	-	15	21	-	-
Mar	7	7	-	-	13	19	-	-
Apr	3	5	0	0	6	8	2	3
May	8	9	1	1	14	16	2	2
Jun	6	6	1	1	15	16	1	1
Jul	1	1	0	0	13	17	2	3
Aug	3	4	4	4	4	11	0	0
Sep	0	0	2	3	10	19	3	3
Total	34	41	8	10	96	137	10	12

Table 2-1: Ground delay programs

SFO stands out clearly in the overall number of programs run, with an average of 10.6 GDP days per month and 15.2 programs per month. Another way to look at this is that SFO had a GDP every 2.8 days. The other three airports being analyzed have a much lower overall incidence of GDP programs.

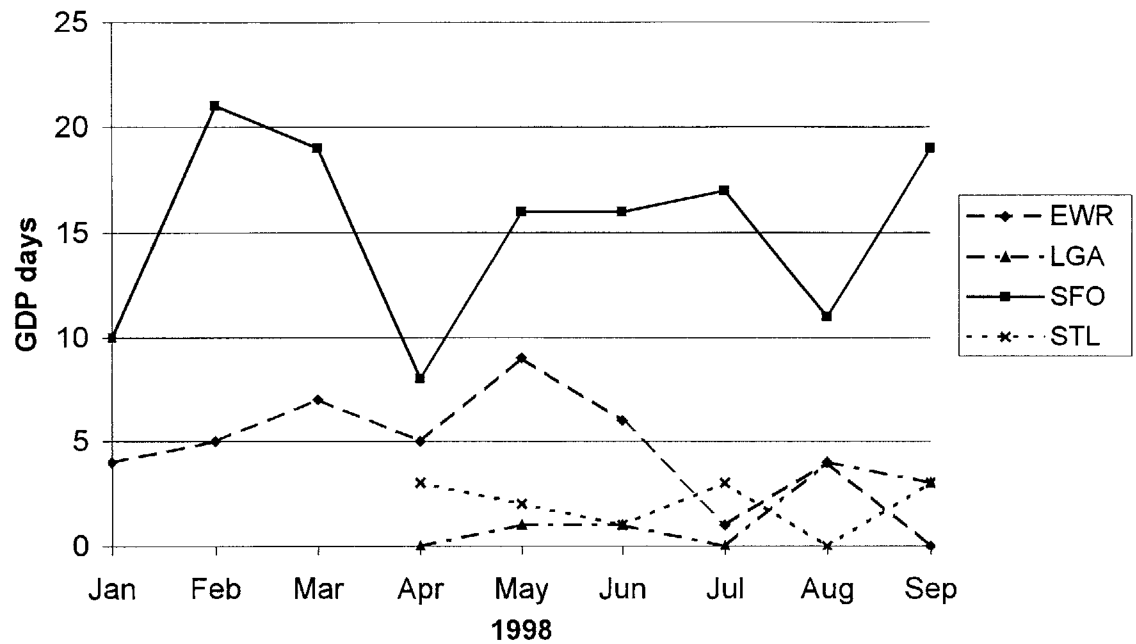

Figure 2-1: Ground delay programs

2.2.3 SFO hourly distribution of ground delay programs

Additional insight into the incidence of GDP programs can be obtained by examining the frequency with which GDPs were in effect, as a function of time of the day. SFO was used as the base for this analysis due to its large volume of programs actually run. Figure 5 shows, for each hour, the number of days that SFO had an active-GDP or a cancelled-GDP.

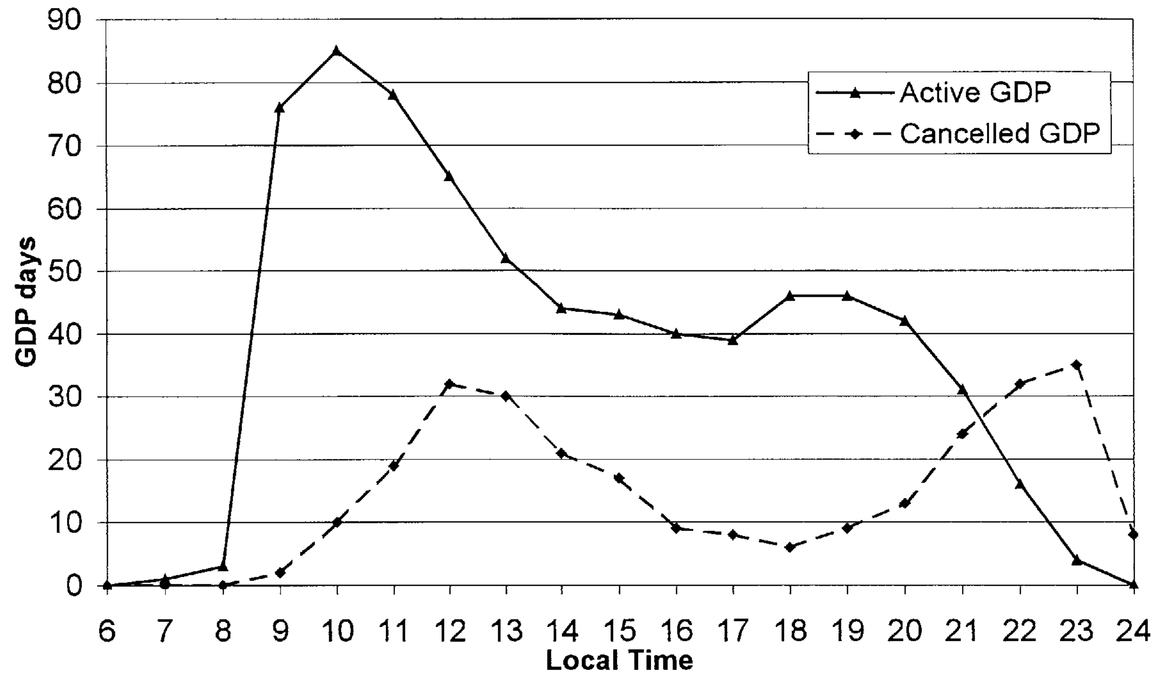

Figure 2-2: SFO incidences of GDPs, by hour

Since GDPs are instituted in response to anticipated shortfalls in capacity, GDPs have a close affinity to the scheduled demand profile. Table 2-2 and Figure 2-3 explores this further. In the table, *sched* is the average scheduled number of arrivals per hour and *AAR* is the average AAR.

Hour	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
Sched	29.3	36.2	42.5	45.5	39.0	36.8	29.6	36.3	35.5	32.5	38.2	40.5	46.9	36.4	25.7	12.4	2.7
AAR	30.0	30.3	29.6	29.6	29.6	30.1	29.3	30.2	30.1	28.9	30.1	30.7	30.5	30.1	29.9	29.9	28.8

Table 2-2: SFO average scheduled arrivals and active-GDP AARs

Table 2-2 and Figure 2-3 indicate a uniformly higher demand than the average AAR in the hours between 8:00 and 21:00. The morning demand peaks in the hour between 10:00 and 11:00, and the afternoon peak is between 19:00 and 20:00. These peaks are much more pronounced that the ones found in Figure 2-2, where the incidences of GDPs were plotted. The data shows that the airport activity level during a GDP remains almost constant throughout the day, which is due to the near-uniform use of an SFO AAR of 30. Note that the average active-GDP AAR remains at this level beyond 21:00. This can be attributed to accommodating flights that have been delayed by the GDP programs.

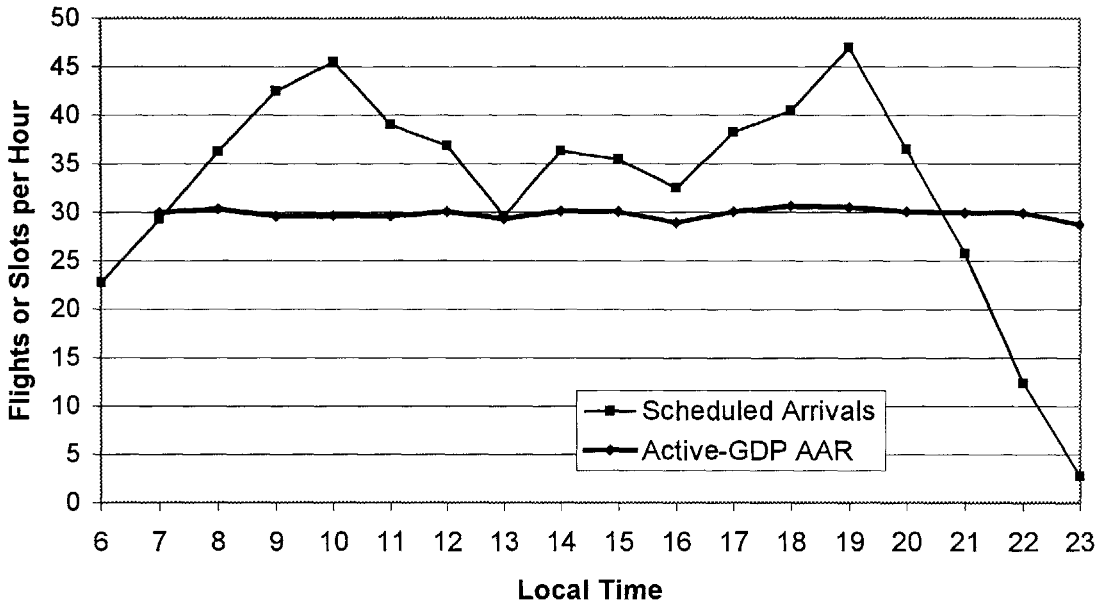

Figure 2-3: SFO average scheduled arrivals and active-GDP AARs, by hour

Figure 2-2 shows that the cancelled-GDPs also has a bimodal distribution, albeit shifted 3 to 5 hours later in the day from the active-GDP peaks. They thus follow the morning and evening post-peak arrival times as shown in when the arrival rates start to drop and capacity is less constrained.

2.2.4 *Seasonal variation of ground delay programs*

The seasonal variation can be explored by analyzing the data over the available quarters. Figure 2-4 shows the seasonal variation in the incidence of the SFO GDPs, and Figure 2-5 shows the similar variation of the EWR GDPs.

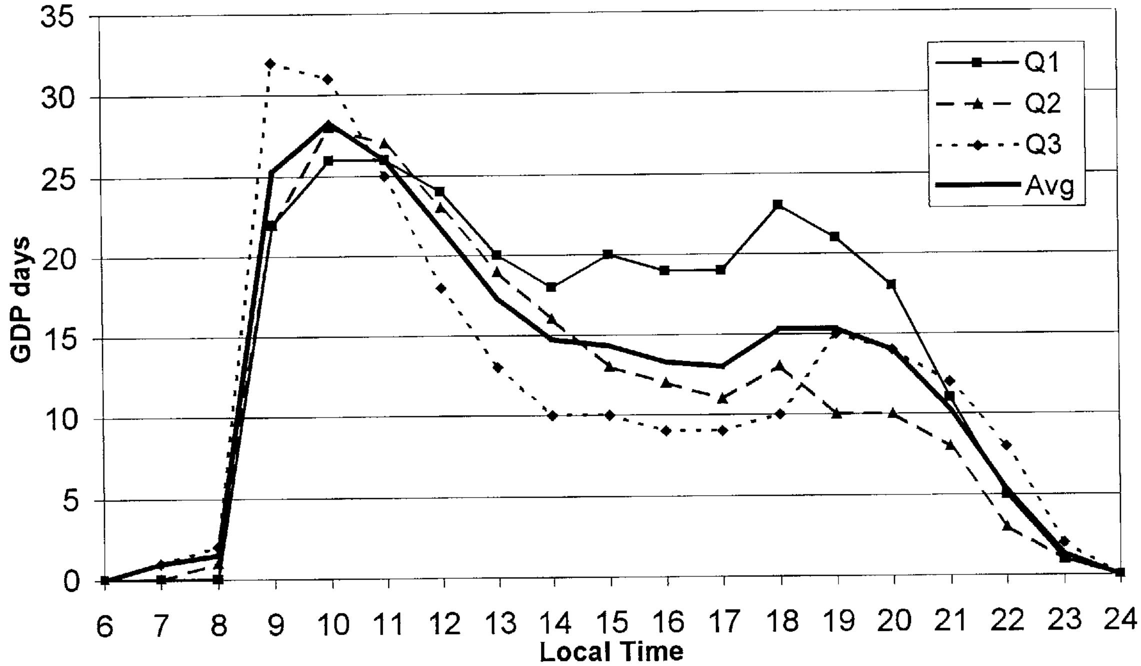

Figure 2-4: SFO quarterly incidences of GDPs, by hour

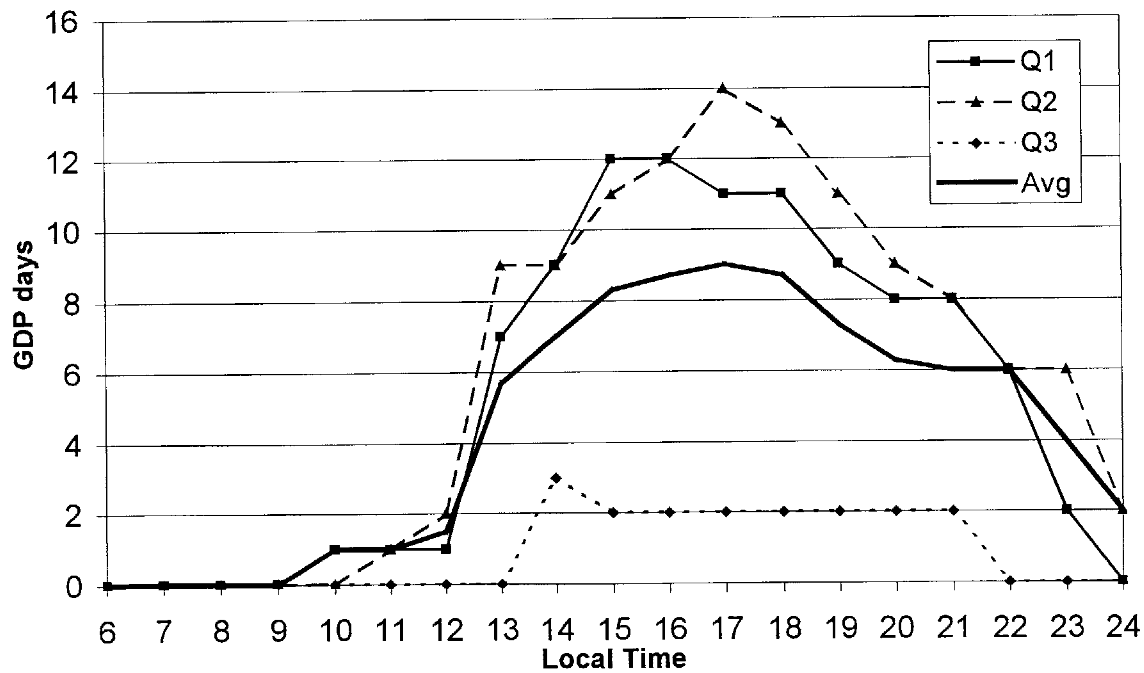

Figure 2-5: EWR quarterly incidences of GDPs, by hour

The seasonal variability of SFO GDP incidences seems quite low. As can be seen from Figure 2-4, each quarter exhibits the characteristic shape of the average, with some variations. The biggest difference is found in the early afternoon number, where there has been a substantial decrease of almost 50% from quarter 1 to quarter 3. What this decrease is due to is impossible to say based on the data at hand, but possible candidates are seasonal weather variations, better execution of the GDPs, the effects of compressions, or a combination of all of these and others.

Figure 2-5 shows that the EWR incidence of GDPs is quite different from the one at SFO. First, the averages are much lower overall. At EWR, the quarterly peak of 14 GDP days is reached in the 2nd quarter in the hour between 17:00 and 18:00. At SFO, the quarterly peak of 32 GDP days is reached in the 3rd quarter between 9:00 and 10:00. A similar pattern holds for the statistics of the full nine months studied. The average GDP days per quarter rises to a peak of 9 GDP days per quarter at EWR, and 28 GDP days per quarter at SFO.

Second, there is a much larger variation from quarter to quarter found at EWR. SFO basically maintains the same pattern for all three quarters, with some variations found especially during the secondary peak in the later afternoon hours. EWR on the other hand sees a dramatic lowering in the third quarter. This results from just three GDP days in this quarter.

Third, there is a fundamental difference in the shape of the two curves. SFO has a pronounced and sharp peak in the morning, probably as a result of the morning fog being a very consistent feature at this airport, followed by a substantial drop to a level that is then sustained from the noon hour through about 20:00 in the evening. EWR has a much smoother shape. GDPs typically do not take effect until after 12:00, but then last a bit longer, extending through to 23:00.

2.3 Substitution and Compression Analysis

2.3.1 Methodology

The effects of substitutions and compressions under CDM are analyzed by initially determining the effect on individual flights, and then aggregating the results as appropriate.

The ΔADL files contain information about each flight's change in estimated arrival time. Each of these changes are classified and aggregated into one of four categories.

FAA arrival slot allocation delays: Any change made during a 20 minute window[2], starting at the time of a GDP initiation, revision, extension or revision/extension, is deemed to belong to the FAA arrival slot allocation delay category. The incremental delay (calculated as the difference between the previous and current setting of the ETA field) is added to cumulative *FAA arrival slot allocation delay* field for that flight.

Airline substitution delays: If a flight moves into a controlled time of arrival (CTA) slot previously occupied by another flight from the same airline, and this happens outside the 20 minute window timeframe of any FAA-initiated GDP program action, then that change is deemed to belong to the airline substitution category. The incremental delay (calculated as the difference between the previous and current setting of the ETA field) is added to the cumulative *airline substitution delay* field for that flight. An airline and all of its subsidiaries are looked upon as a single entity when deciding which slot swaps belong in this category.

Compression delays: Any change made during a 20 minute window, starting at the time of a GDP compression, is deemed to belong to the compression category. The incremental delay (calculated as the difference between the previous and current setting of the ETA field) is added to the cumulative *compression delay* field for that flight. These delays are in actuality negative numbers, representing a time saving.

Other delays: This is a catch-all category established to collect those changes that cannot be ascribed to any of the above three reasons. It is typically used infrequently, i.e., typically less than 3 times per day, and the changes it makes are generally the result of some errors in the

---

2 The 20 minute window used during the determination of the delays is a historical artifact. Initially the GDP program files did not contain information about exactly when a given change took effect in the system (and hence in the ΔADL files), only when they were applied. There typically was a 10 minute gap between the time of applying the changes and the time when they took effect, but this gap could and did vary somewhat. Employing the 20 minute window ensured that the changes got attributed to the correct cause of the change.

This situation has already been corrected as far as the availability of the data is concerned. Any analysis of data after 7/22 can dispense with the time window and use the exact time of the changes taking effect, since this data is now contained in the ΔADL files. Before this can be put into use, though, some computer program changes will be needed.

underlying data. The incremental delay (calculated as the difference between the previous and current setting of the ETA field) is added to the cumulative *other delay* field for that flight.

These calculations take place during the data extraction phase, and only the net result per flight for each of the four categories is brought forward and stored into the database.

2.3.2 *Savings*

The benefits of substitutions and compressions were measured by calculating the savings directly attributable to each. Table 2-3 summarizes the results of the first 9 months of GDP-E prototype operations.

Airport	EWR	LGA	SFO	STL	Overall
All flights	57,190	122,540	145,278	172,740	597,748
GDP flights	9,241	1,571	28,713	3,263	42,788
GDP flights %	5.90%	1.30%	19.80%	1.90%	7.2%
Ground delay allocated (hrs)	14,681	1,692	59,841	6,936	83,151
Substitution savings (hrs)	1,533	176	5,341	670	7,720
Compression savings (hrs)	934	57	3,108	542	4,641
Net delay (hrs)	12,214	1,459	51,392	5,724	70,789
Substitution savings %	10.4%	10.4%	8.9%	9.7%	9.3%
Compression savings %	6.4%	3.4%	5.2%	7.8%	5.6%
Subst savings per Flight (min)	10.0	6.7	11.2	12.3	10.8
Comp savings per Flight (min)	6.1	2.2	6.5	10.0	6.5

Table 2-3: Airline substitution and GDP compression summary results

In this table, *All flights* is the total count of flights that arrived at the airport. *GDP flights* is the subset of flights that at one point or another had a GDP arrival slot allocated, and *GDP flights %* is the percentage of flights that were affected by a GDP.

Ground delay allocated is the net sum of all delays allocated by the GDPs. Each flight might be subject to multiple delay changes. The initial delay assigned will always be a positive delay, but subsequent changes may be either positive or negative, reflecting worsening or improving conditions. *Substitution savings* are the savings that are recouped by the airlines substituting one flight for another (and canceling some flights), and *compression savings* are the savings from running the GDP compression algorithm. The *net delay* is the final delay after airline and compression savings have been applied.

Substitution savings % measures how effective substitution is in reducing the delay allocated by the GDPs, and *compression savings %* similarly measures the effect of compressions. *Subst savings per flight* is calculated by averaging the substitution savings over every GDP flight, and *comp savings per flight* similarly measures the savings due to compression per GDP flight.

It is evident that substitutions and compressions both contribute substantial benefits to the system. The average compression savings is 5.6%, with a low of 3.4% at LGA and a high at 7.8% at STL. SFO is at 5.2% and EWR at 6.4%. Similarly, the average substitution savings are 9.3%, with a high of 10.4% in both EWR and LGA, and a low of 8.9% in SFO. It is also clear that GDPs affect SFO far more extensively than any of the other three airports studied. This is primarily due to the local weather conditions and overall arrival capacity at SFO.

2.3.3 Monthly results

The results can be analyzed in more detail by examining the data on a monthly basis. This is done in Table 2-4 and Figure 2-6. The table and the figure show the number of GDP flights for each airport (*count*), as well as the monthly percentage of flights affected by GDPs (*pct*).

Month	EWR (count)	EWR (pct)	LGA (count)	LGA (pct)	SFO (count)	SFO (pct)	STL (count)	STL (pct)
Jan	689	4.3%	-	-	1807	12.3%	-	-
Feb	1195	7.3%	-	-	5124	35.1%	-	-
Mar	2127	11.5%	-	-	3656	21.3%	-	-
Apr	659	3.7%	0	0.0%	1574	9.6%	707	3.6%
May	2286	12.6%	221	1.6%	3642	22.2%	710	3.6%
Jun	1217	8.0%	117	1.0%	3358	23.5%	283	1.7%
Jul	212	1.2%	0	0.0%	3559	20.8%	734	3.7%
Aug	856	4.5%	796	5.6%	2325	13.0%	0	0.0%
Sep	0	0.0%	437	3.1%	3668	21.9%	829	4.3%
Total	9241	5.9%	1571	1.9%	28713	19.8%	3263	2.8%

Table 2-4: Flown flights that were at some point affected by GDP

SFO again stands out clearly as being much more heavily subjected to GDP delays than the other three airports. February is an especially heavily impacted month, maybe due to the effects of El Niño.

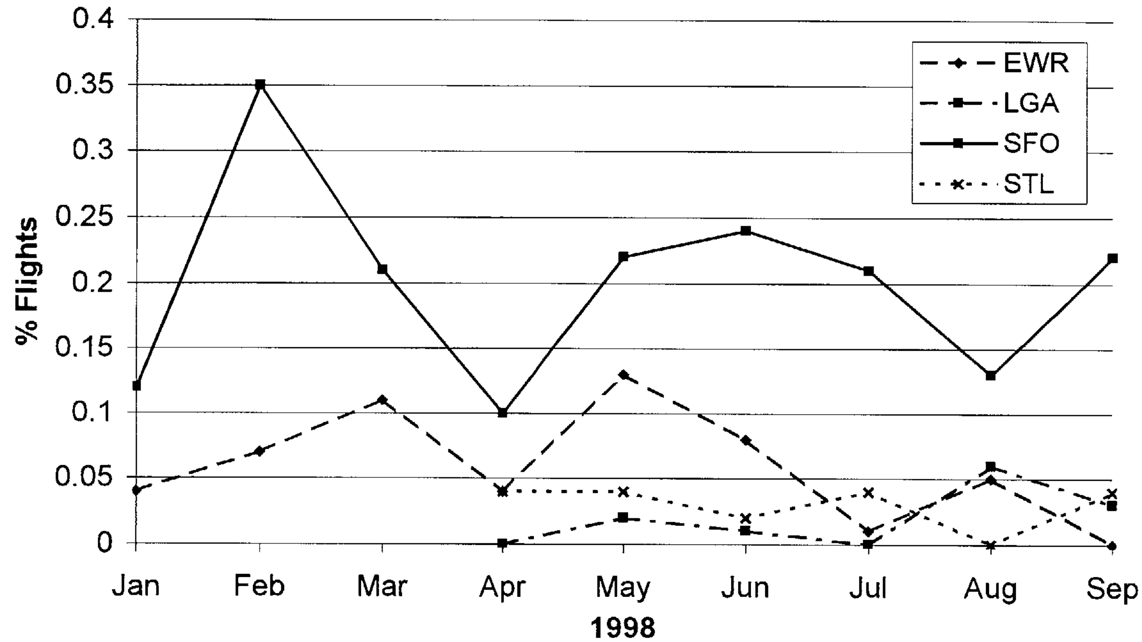

Figure 2-6: Flown flights affected at some point by GDP

2.3.4 *SFO monthly delay classification*

The net GDP delays can be determined by looking at the core components affecting these delays: first, the initial FAA arrival slot allocation delay, along with any subsequent revisions; second, savings associated with airline substitutions; and third, savings that come about as a result of using the compression algorithm.

Each of these are measured on a per-month basis in Table 2-5. In this table, *FAA* tallies the FAA arrival slot allocation delays, *subst* is the savings resulting from airline substitutions., and *comp* is the savings resulting from compression. *S+C* is the combined savings from substitution and compression, and *net* is the net sum of all these measures. *S+C%* shows the monthly savings as a percentage of total delay handed out, and *S %* and *C %* shows the percentage savings resulting from substitution and compression respectively.

Month	FAA	Subst	Comp	S+C	Net	S+C %	S %	C %
Jan	143.2	-5.6	-4.7	-10.3	133.0	7.2%	3.9%	3.2%
Feb	168.4	-11.5	-8.6	-20.1	148.3	12.0%	6.8%	5.1%
Mar	140.8	-12.8	-3.3	-16.1	124.7	11.4%	9.1%	2.4%
Apr	121.7	-11.8	-2.4	-14.1	107.6	11.6%	9.7%	1.9%
May	101.6	-16.7	-2.1	-18.8	82.8	18.5%	16.5%	2.1%

Month	FAA	Subst	Comp	S+C	Net	S+C %	S %	C %
Jun	93.9	-10.1	-3.3	-13.4	80.5	14.3%	10.8%	3.5%
Jul	109.7	-6.2	-8.0	-14.2	95.6	12.9%	5.6%	7.3%
Aug	135.8	-11.1	-10.9	-22.1	113.7	16.2%	8.2%	8.0%
Sep	100.9	-11.8	-12.4	-24.2	76.7	24.0%	11.7%	12.3%
Average	125.0	-11.2	-6.5	-17.7	107.4	14.1%	8.9%	5.2%

Table 2-5: SFO average delay and delay savings per GDP flight, by month

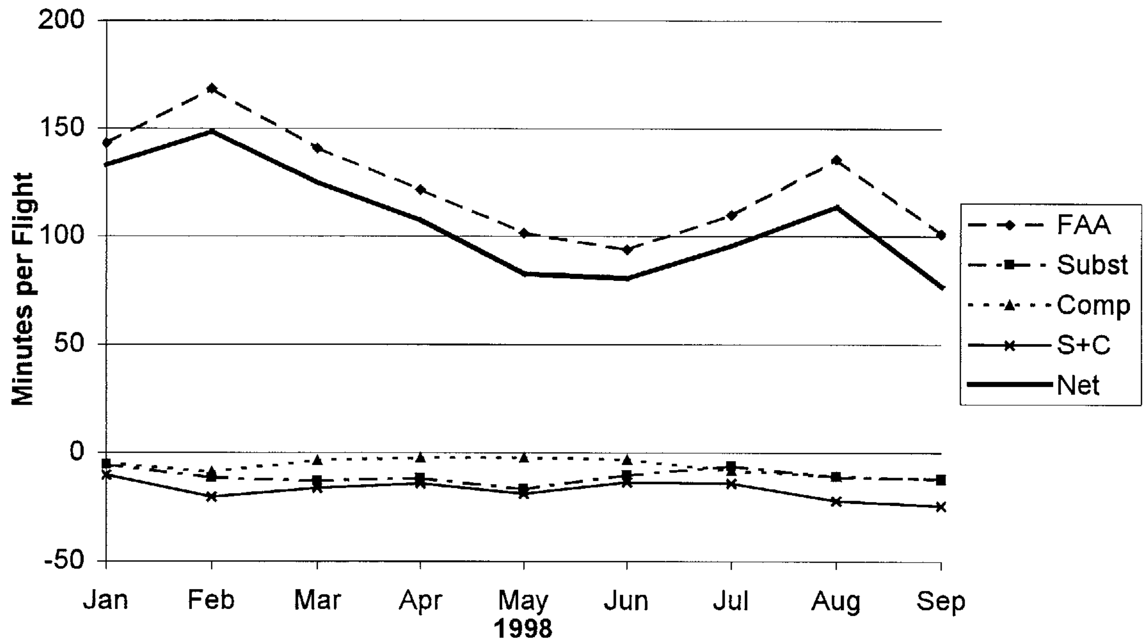

Figure 2-7: SFO substitution and compression savings per GDP flight, by month

The data suggests a possible improving trend in the delays, dropping from a net delay of 148.3 minutes per flight in February to a low of 76.7 minutes in September. There are, however, substantial month-to-month variations within this time-period. The drop in average may be due to seasonal factors, but this will become clearer only when data for several years become available. February clearly was a difficult month. Not only were there more flights delayed during that month (see Table 2-4), but each delayed flight was also delayed for a longer period of time. This may have been due to the effects of El Niño.

The delay savings resulting from the combination of airline substitutions and GDP compression are also quite noticeable, especially in the figure. Averaged over all GDP flights to arrive in SFO, the delay is reduced by 17.7 minutes per flight, from 125.0 minutes on average to 107.4 minutes or average. This is equivalent to a savings of 14.1%, of which 5.2% resulted from compression. The composition of the combined savings changed starting in July. Until then, airline substitutions clearly accounted for most of the savings, but from July on the savings from substitutions and compressions are almost identical. One possible explanation is that it reflects a maturing of the compression process. Another possibility is that the airlines are becoming more comfortable with the compression process and the results that it is producing, and are showing this by cutting back on the number of substitutions they are initiating. It does not, however, appear to be a consequence of fewer cancellations. *Table 2-24: Average daily cancellations for GDP and non-GDP days* shows that the number of cancellations per GDP day has remained fairly constant. This is important, since the benefit from a substitution, as far as a time saving is concerned, only arises when a flight gets cancelled and another flight assigned by the GDP to a later arrival slot can be moved up to take the former's CTA in the stream of arrivals. (An airline may obtain benefits from swapping two active flights with one another, but these benefits do not include any aggregate time saving, and so do not affect the results being measured here).

Note that the savings from airline substitutions and from GDP compression both increase over the nine month period studied. This is an indication that the CDM benefits are not derived solely from the introduction of compression, but that a portion of the airline substitution savings also needs to be attributed to CDM. However, it is also reasonable to assume that some of the benefits derived from GDP compression could have been obtained by additional airline substitutions in the absence of GDP compression. Since data has not been available for the period prior to the startup of GDP-E, it is not possible to state what a lower bound on the benefits of introducing CDM have been as SFO. The combined savings resulting from substitution and compression savings form an upper bound of 14.1% on the derived benefits. The 5.2% resulting from compression is an indication of this lower bound, but further analysis and data is required to determine this with greater certainty.

2.3.5 SFO hourly results

Analyzing the same data on an hourly basis yields further insight into the use of the SFO GDPs. The two following figures illustrate this. Figure 2-8 depicts the number of changes made to the CTA as a function of the time of day of the originally scheduled time of arrival.

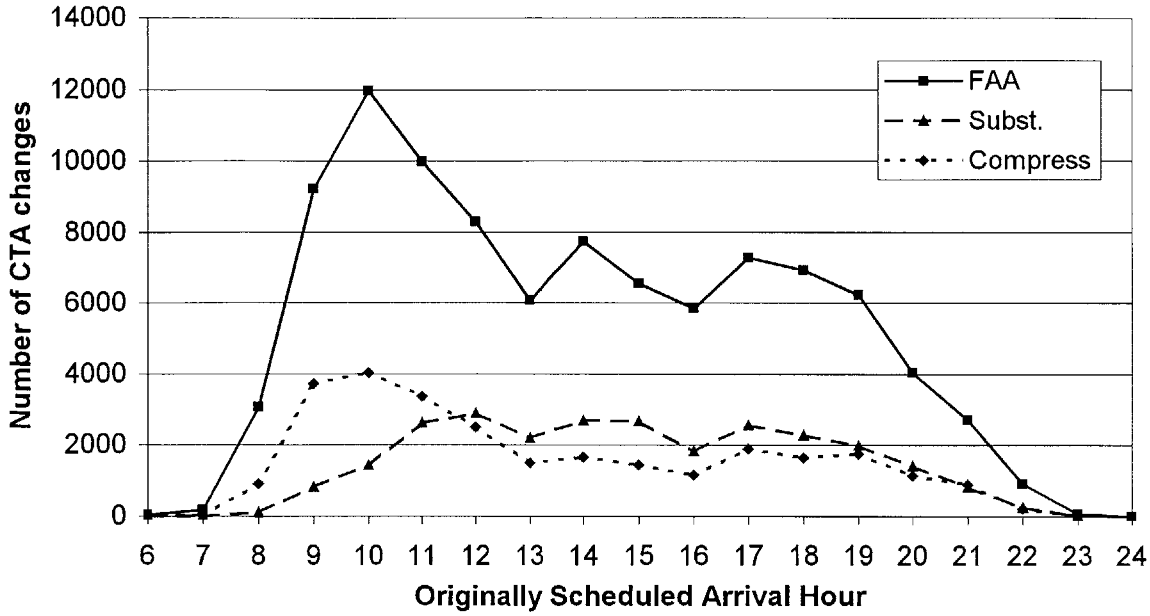

Figure 2-8: SFO number of CTA changes, by hour

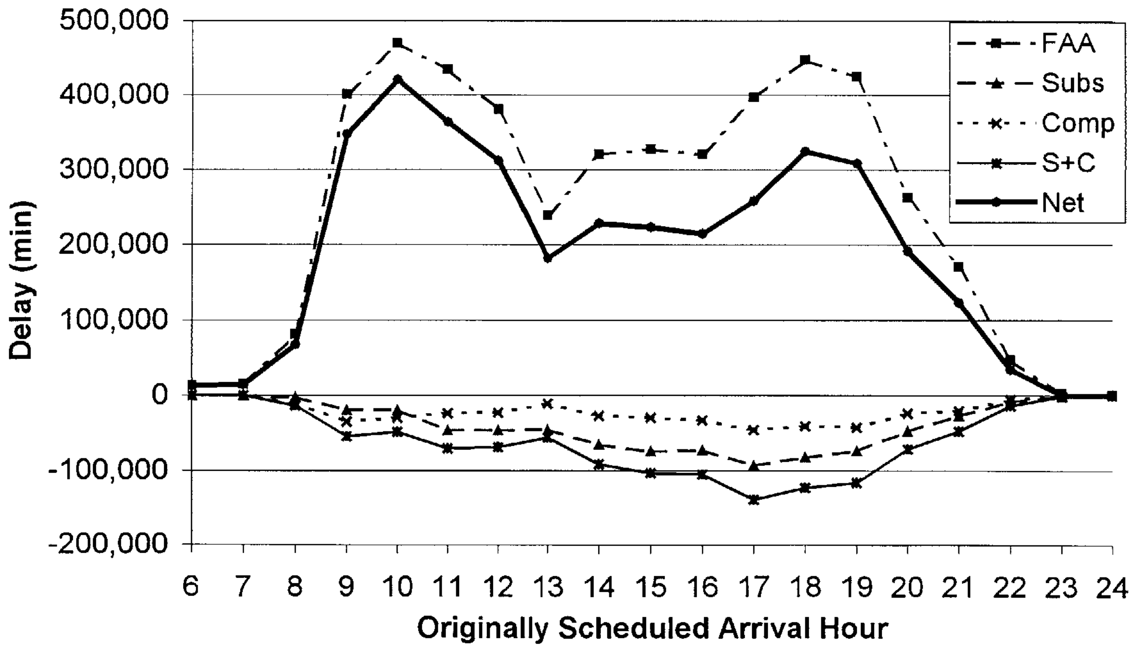

Figure 2-9: SFO ground delay, by hour

Figure 2-8 shows the large number of delay assignments made to the morning-hour flights into SFO, caused by the large number of scheduled flights during these hours and the prevalence of GDP programs. Two smaller secondary peaks are also in evidence later in the day, probably for the same reason.

It is also evident from this figure that compression picks up earlier than substitutions. The absolute number of compression-related changes is four times as high as the number of airline substitutions for arrivals between 9:00 and 10:00. However, from 11:00 on the number of changes attributable to these two factors are quite close to one another.

Figure 2-9 shows the ground delay allocated by the FAA arrival slot allocations, and the portion of this delay recovered by airline substitutions and compressions, again as a function of originally scheduled time of arrival. This shows the impact that the changes have on the total arrival delay. First, two peaks stand out, matching the time period of the morning and afternoon peak arrival demand (see also *Figure 2-3: SFO average scheduled arrivals and active-GDP AARs, by hour*). Second, the effects of both substitution and compression improve over the time of day, with the combined effect being twice as large on the afternoon peak as it is on the morning peak. Third, airline substitutions play a larger role than compression in reducing delays. However, as was shown earlier in the monthly analysis for SFO, savings attributable to airline substitutions and to compressions were almost equal in the last three month of the study period. The larger airline substitution savings shown in Figure 2-9 probably reflect in part the initial prototype operation of compression.

2.3.6 EWR hourly results

A similar analysis follows for EWR. Figure 2-10 shows the number of changes in the allocated delay, the airline substitutions and the compressions. Figure 2-11 shows the effect that these changes have on the various GDP delay measures.

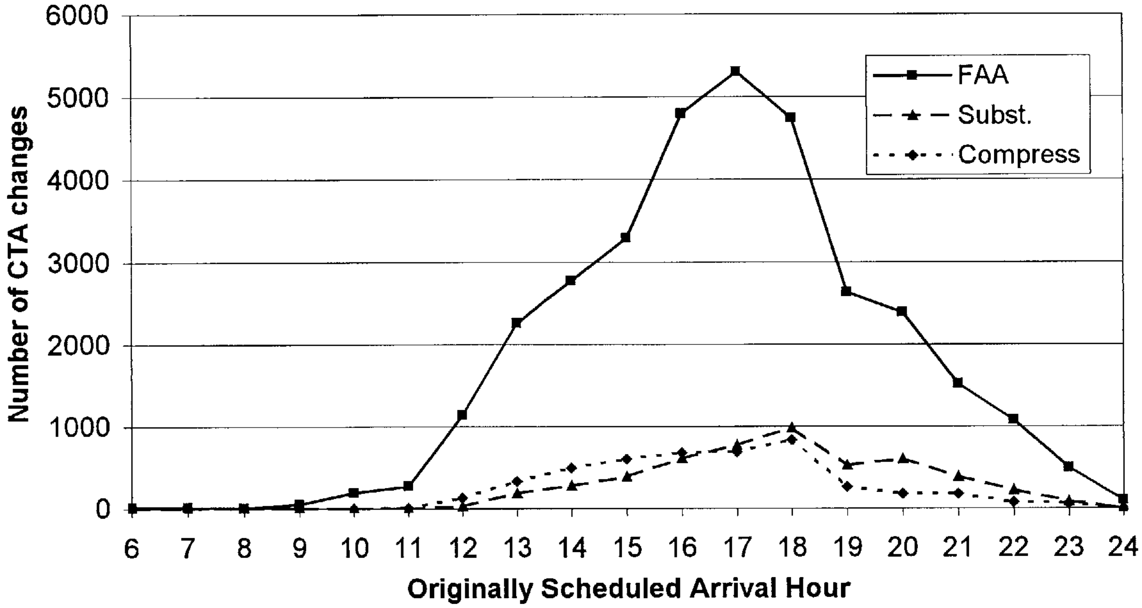

Figure 2-10: EWR number of CTA changes, by hour

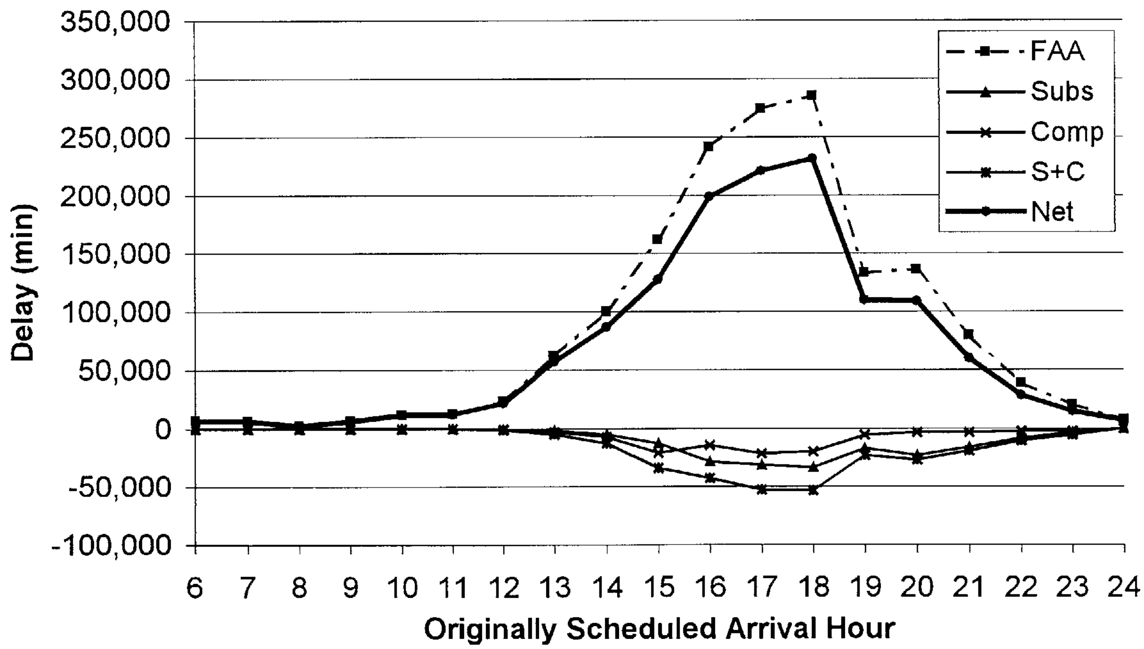

Figure 2-11: EWR ground delay, by hour

Figure 2-10 shows the situation at EWR to be considerably different from the one found at SFO. There are a lot fewer changes overall. The number of changes made in response to the FAA arrival slot allocations are also a lot fewer, both for airline substitutions and especially for GDP compressions. The impact of compression at EWR is much smaller than what was seen at SFO. This is also confirmed by the data in Figure 2-11, which depicts the delays. The delay savings attributable to compression after 20:00 are very small when compared to both the arrival slot allocation delay and even the substitution savings.

2.4 Arrival Slot Utilization

Compression and airline substitutions provide one avenue for improving the overall effectiveness of the system. They do so in two ways. First, they lower the average delay imposed on flights. Second, they may decrease the variability of the inter-arrival times. The use of compression increases the likelihood that a steady stream of flights arrive at the destination airport. This could ultimately lead to a lowering of the Managed Arrival Reservoir (MAR), which in turn should result in a lowered airborne holding.

A second avenue for improving the overall system is to ensure that the available arrival capacity at the GDP airport is utilized to its fullest. Measuring GDP arrival slot utilization helps answer this latter question, and is the subject of this section.

2.4.1 Methodology

There are two basic steps in the analysis. First, we determine the supply of and demand for capacity for every hour. Second, we aggregate the hourly results according to needs of the analysis.

Actual hourly demand is determined by counting how many flights actually arrive in a given hour at a given airport. In addition to this, scheduled demand is also needed, for reasons that will become clear shortly . Both types of information are readily available from the database.

Hourly capacity is provided as part of the GDP program parameters. Every time a GDP is initiated or revised, it is accompanied by a specification of the AARs and GA factors for the hours affected by the GDP. As such, capacity information is readily available for the GDP hours. Unfortunately, there is no similar source of capacity information available for the non-

GDP hours.[3] One impact this has on the analysis possibilities is to make it difficult to judge how quickly operations return to normal following the cancellation of a GDP program, since the arrival capacity and hence the upper bound on the arrivals is unknown.

To fully explain capacity utilization, two additional aspects need to be addressed. The first issue deals with the situation when multiple AARs have been successively specified for the same hour. Currently, only the last value specified is used, on the assumption that it contains the latest and therefore best estimate of the actual conditions for that hour. However, any earlier settings of an AAR will have governed the arrival slot allocation for some amount of time. As such, it is entirely possible that there is a relationship between when an AAR can be changed, and how quickly these changes can have a real effect on the system. This remains an open research topic to be investigated.

The second capacity aspect to be covered concerns the issue of what to do when demand is too low to fully utilize a given capacity. This is either caused by not allocating enough flights for the given hour, or if simply there are not enough scheduled and delayed flights to fill the available capacity for that hour. To deal with this an *adjusted capacity* AAR' is used in place of the stated AAR capacity. The adjusted capacity is defined as: (1) AAR' = AAR, when scheduled arrivals exceed or equal the available capacity, and (2) AAR' = actual arrivals, when scheduled arrivals is less than the available AAR capacity. The primary benefit of using AAR' in place of AAR is that it avoids denoting unused capacity as under-utilized.

GDPs are typically neither instituted nor cancelled on even-hour boundaries. This raises the issue of how to classify the hours that contain such a change. The approach adopted here is to treat these hours as active-GDP hours. This may register as a slight overuse of the restricted capacity, since the period prior to or following an active-GDP typically will have a larger capacity, but it errs on the side of caution. However, the overall effect of this is judged to be minimal.

Table 2-6 illustrates these concepts. The table contains a subset of the demand and capacity information for SFO on 9/30/1998, when a GDP was in effect.

---

[3] This is one of the gaps in the data needed for the analysis. Hopefully this can be remedied in the future, possibly by inclusion of airport configuration information in the AADL files that already provide the bulk of the core analysis data.

Dest	Hour	GDPs	Sched	Flown	Cnx	Land	AAR	AAR'	Unused	Pct
SFO	20	1	48	36	12	31	30	30	-1	-3%
SFO	21	1	36	25	11	32	30	30	-2	-6%
SFO	22	1	24	18	6	28	30	28	0	0%

Table 2-6: Adjusted arrival capacity (AAR') example

The hour starting at 20:00 (*hour*) covers 1 hour of GDP (*GDPs*). It originally had 48 flights scheduled to arrive at that hour (*sched*), out of which 36 flights were flown (*flown*) and 12 were cancelled (*cnx*). 31 flights actually arrived during this hour (*land*). The data provide no information about when these particular 31 flights originally were scheduled to land, but it is unlikely that they all came from the original pool of 48 scheduled flights for that hour. The stated AAR capacity was 30 flights per hour (*AAR*), which is the last set value for the AAR capacity for this hour. The adjusted AAR capacity for this hour is also 30 (*AAR'*). The number of unused slots are -1 (*unused*), calculated as *AAR' - land.* Finally, the slot capacity under-utilization is -3% (*pct*), calculated as *unused/AAR*. In this instance, the last two columns show that one more flight landed than was planned for.

The hour starting at 22:00 shows the effect of the adjusted capacity calculation. 24 flights were originally scheduled to arrive at this hour, which is less than the available capacity of 30. As a consequence, the adjusted capacity is set to 28, the number of flights that actually arrived at this hour.

2.4.2 SFO hourly arrival slot utilization

The hourly arrival slot utilization at SFO has been analyzed to answer the question of whether or not the available capacity is being used to its fullest extent. Table 2-7 summarizes this analysis. In this table, *Local Hour* is the local hour of day, i.e., in PST or PDT. *Active-GDP days* counts the number of days when a GDP was active for that particular hour, and *cancelled-GDP days* similarly counts the number of days when a GDP had initially been planned for that hour, but was subsequently cancelled. *Unused active slots* is the sum total of active-GDP slots that went unused in that hour, and similarly for the *unused cancelled slots.* In the latter case this number is going to be negative, representing an overuse which is to be expected when the capacity constraints are lifted. *Net unused slots* is the difference between the previous two columns. Finally, the last three columns, *unused active slots/day, unused cancelled slots/day*, and *net unused slots/day* show a daily average for each slot measurement.

Local Hour	Active GDP Days	Cancel. GDP days	Unused Active Slots	Unused Cancel Slots	Net Unused Slots	Unused Active Slot/day	Unused Cancel Slot/dy	Net Unused Slot/day
6	0	0	0	0	0	0.00	0.00	0.00
7	1	0	0	0	0	0.00	0.00	0.00
8	3	0	19	0	19	6.33	0.00	6.33
9	76	2	212	-4	216	2.79	-2.00	2.77
10	85	10	141	80	61	1.66	8.00	0.64
11	78	19	67	138	-71	0.86	7.26	-0.73
12	65	32	81	170	-89	1.25	5.31	-0.92
13	52	30	27	6	21	0.52	0.20	0.26
14	44	21	48	-9	57	1.09	-0.43	0.88
15	43	17	69	62	7	1.60	3.65	0.12
16	40	9	-3	20	-23	-0.08	2.22	-0.47
17	39	8	-1	-2	1	-0.03	-0.25	0.02
18	46	6	-33	31	-64	-0.72	5.17	-1.23
19	46	9	21	42	-21	0.46	4.67	-0.38
20	42	13	-21	83	-104	-0.50	6.38	-1.89
21	31	24	30	8	22	0.97	0.33	0.40
22	16	32	4	0	4	0.25	0.00	0.08
23	4	35	0	0	0	0.00	0.00	0.00
24	0	8	0	0	0	0.00	0.00	0.00
All	711	275	661	625	36	0.93	2.27	0.04

Table 2-7: SFO slot utilization, by hour

The rightmost three items in the last line (*All*) are calculated as averages of the total line values, not as sums of the columns. The interpretation of the information in this table is discussed below.

2.4.2.1 *Net matching of slots to demand*

Plotting the average net slots per day from Table 2-7, one obtains Figure 2-12. The plot depicts the average difference between the AAR and the number of flights that actually landed. All slot allocations are taken into account in this figure, whether active at the time of arrival, or if they had been cancelled.

The results indicate a marked under-utilization during the early hours of the day, with an average number of slots of 6.33 going unmatched between 8:00 and 9:00, and 2.77 between 9:00 and 10:00. The only other major deviation is the over-utilization between 20:00 and 21:00, where the average number of net slots is -1.89. Apart from these three hours, every other reading falls with a range of plus or minus one slot.

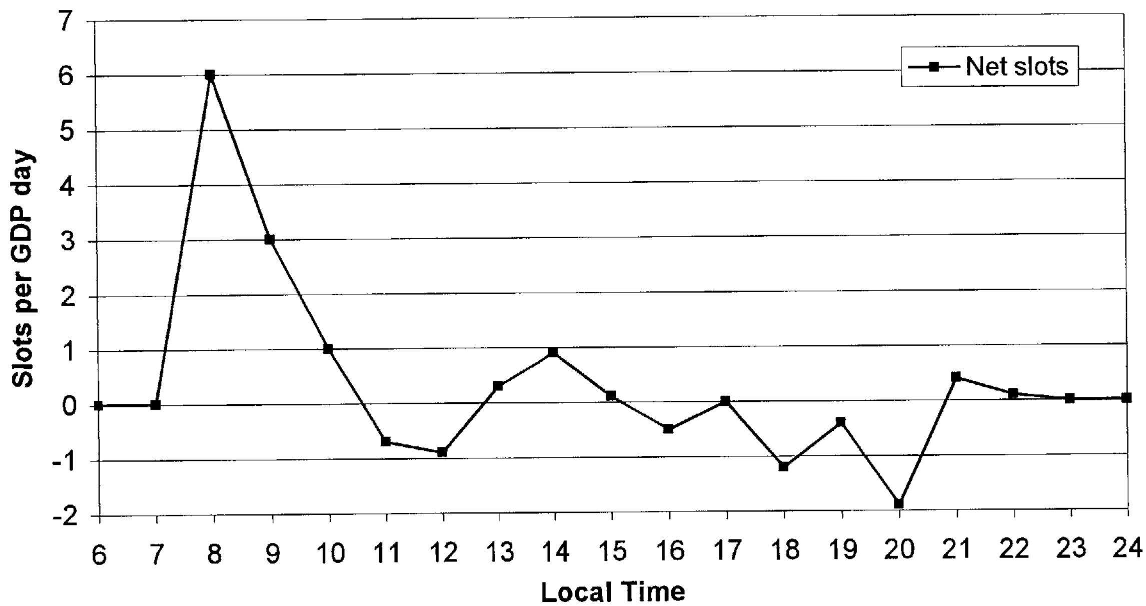

Figure 2-12: SFO average net matching of slots to demand, by hour

Further analysis of this requires separating the active-GDP hours from the cancelled-GDP hours, which is done in the following two sub-sections.

2.4.2.2 *Matching slots with demand during active ground delay programs*

Focusing on the active-GDP hours exclusively, slots are matched up with demand as shown in Figure 2-13. The initial spike remains, which lends further credence to the possibility that there is an under-utilization of capacity during the morning hours. This time period corresponds directly with the time when most GDP programs go into effect at SFO. This suggest the possibility that the initial GDP allocation may have been set too tight and could be relaxed to accommodate additional flights in each of the two hours between 8:00 and 10:00. Before a more definitive statement can be made, however, it is necessary to look at the demand and variability of demand for these hours to ensure that this under-utilization is not caused by a lack of demand. Furthermore, the observed pattern must also be based on a reasonable number of daily observations, since a few days may be subject to aberrant behavior.

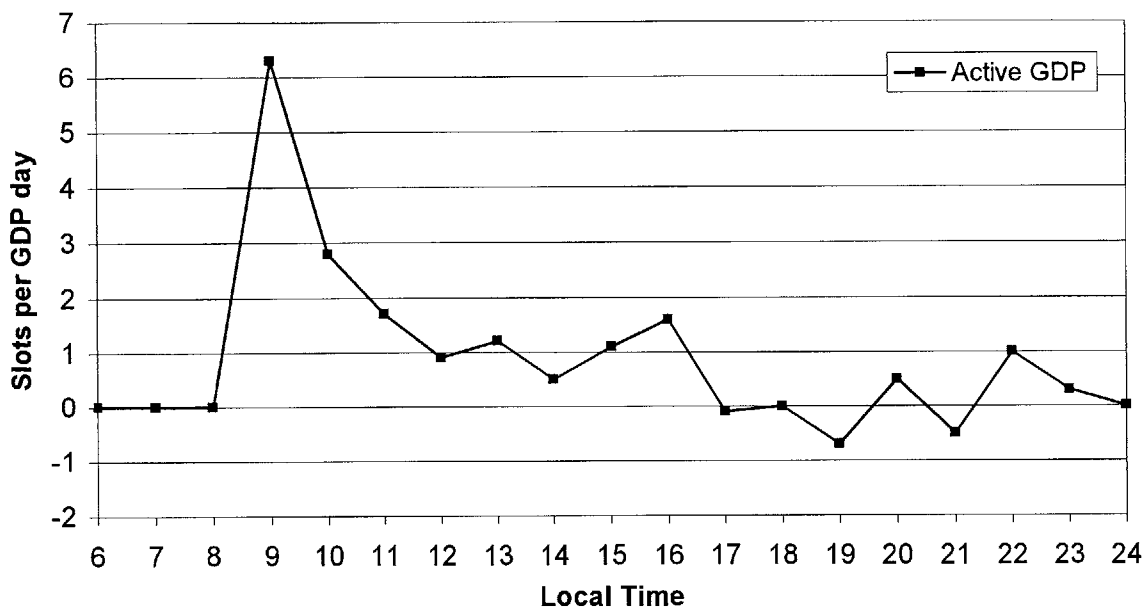

Figure 2-13: SFO average active-GDP matching of slots with demand

The graph also indicates the potential of a systematic under-utilization of one slot in the mid-day period from 10:00 and 15:59. For this to be true, though, the variability of the utilization must be quite low. The variability is an indication of how much the utilization will vary from day to day. A high variability would mean that there would be days where more flights would arrive than the capacity could handle expeditiously, leading directly to unexpected airborne holding. If this is the case, then increasing the stated AAR would only further worsen the situation by increasing the number of days that would experience some form of unexpected airborne holding, and increasing the unexpected airborne holding on days that already were impacted. If, however, the variability is quite low, then it might be feasible to add another arrival slot to the stated capacity.

Table 2-8 provides the additional information needed to explore these issues further. The table contains a number of different items measured on an hourly basis. *Sched* is the average daily demand for the nine months, and *GDP* is the number of days that were impacted by a GDP, further broken down into active (*act*) and cancelled (*cnx*) days. *AAR'* is the average adjusted AAR, and *slots* is the average number of unused slots. *Match* lists how well the available capacity is matched to the controlled demand on active-GDP days, counting the number of active-GDP days where the demand has been less than the AAR' capacity (*under*), exactly matching the capacity (*exact*), and greater than the capacity (*over*). Finally, *cap ok* shows the fraction of days where the stated AAR capacity could handle all arriving flights, and *cap+1* shows the fraction of days where the stated AAR capacity could handle at least one additional flight without exceeding the capacity constraints. For instance, during the 9:00-10:00 hour, *cap ok* = 75% and *cap+1* = 67%, meaning that 75% of the days handled all arriving flights within the stated capacity, and 67% of the days could have handled at least one more flight without exceeding the limits. From this can be inferred that 8% of the days had a number of flights that exactly matched the stated AAR capacity.

Hour	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
Sched	29	36	43	46	39	37	30	36	35	32	38	40	47	36	26	12	3
GDP	1	3	78	95	97	97	82	65	60	49	47	52	55	55	55	48	39
- Act	1	3	76	85	78	65	52	44	43	40	39	46	46	42	31	16	4
- Cnx	0	0	2	10	19	32	30	21	17	9	8	6	9	13	24	32	35
AAR	25	30	30	30	30	31	30	29	30	29	29	30	30	30	29	26	17
- Act	25	30	30	30	30	30	29	28	30	29	29	30	30	30	29	29	20
- Cnx	0	0	30	30	33	33	33	30	30	30	31	30	29	30	30	25	16
Slots	0.0	6.3	2.8	0.6	-0.7	-0.9	0.3	0.9	0.1	-0.5	0.0	-1.2	-0.4	-1.9	0.4	0.1	0.0
- Act	0.0	6.3	2.8	1.7	0.9	1.2	0.5	1.1	1.6	-0.1	0.0	-0.7	0.5	-0.5	1.0	0.3	0.0
- Cnx			2.0	-8.0	-7.3	-5.3	-0.2	0.4	-3.6	-2.2	0.3	-5.2	-4.7	-6.4	-0.3	0.0	0.0
Match																	
- under	1	3	51	49	34	27	22	30	24	16	22	17	25	15	21	10	3
- exact			6	8	9	6	9	3	4	3	3	6	6	10	5	2	1
- over			19	28	35	32	21	11	15	21	14	23	15	17	5	4	
Cap ok	100%	100%	75%	67%	55%	51%	60%	75%	65%	48%	64%	50%	67%	60%	84%	75%	100%
Cap +1	100%	100%	67%	58%	44%	42%	42%	68%	56%	40%	56%	37%	54%	36%	68%	63%	75%

Table 2-8: SFO capacity analysis

Analyzing the situation between 8:00 and 9:00 first, Table 2-8 shows that the spike of 6.3 average unused slots is generated by just three active-GDP days. Further examination of the underlying data reveals that the three days in question are 4/30, 8/20 and 8/31. On 4/30, the AAR' was 30 between 8:00 and 9:00, 31 flights were scheduled to arrive, but only 26 actually arrived in the allotted hour, and four arrived later. Similarly, for 8/20 the numbers were AAR' of 31, 34 scheduled, 23 actual and 8 late arrivals, and for 8/31, AAR' of 30, 33 scheduled, 23 actual and 5 late arrivals. In summary, the three days had deficits of 4, 8 and 7 on-time flight arrivals, for a total of 19 as listed in the table. Based on this limited number of samples, it is not possible to suggest an under-utilization of the available capacity.

The situation between 9:00 and 10:00 is a bit different. Table 2-8 shows that this hour sees a lot more GDP days, 78 in total, of which 76 are active, so there clearly is enough data on which to base some tentative conclusions on. With an average of 43 scheduled flights for this hour, there are also more than enough flights to fill the stated AAR capacity of 30 flights per hour. The average number of unused slots is 2.8 on active-GDP days and 2.0 on cancelled-GDP days. This is an indication that the problem may not be exclusively capacity related, since any limitations in capacity should feature much less prominently for cancelled-GDP hours. Another factor also pushing in the same direction is the large variability in the average number of arriving flights. Out of the 76 active-GDP days, only 51 days show under-utilized capacity. Of the remaining 25 days, 6 days matched the stated capacity exactly, and 19 days showed that more flights landed than was planned for with the stated AAR. In percentage terms this means that SFO had sufficient capacity to handle all arriving flights on 75% of the active-GDP days. This percentage would drop to 67%, if the stated capacity had been increased across the board by one flight per day for the hour between 9:00 and 10:00. Figure 2-14 plots the fraction of days, where the stated capacity is sufficient to handle the flights actually arriving. It also shows the fraction of days where the stated AAR capacity could handle at least one more flight.

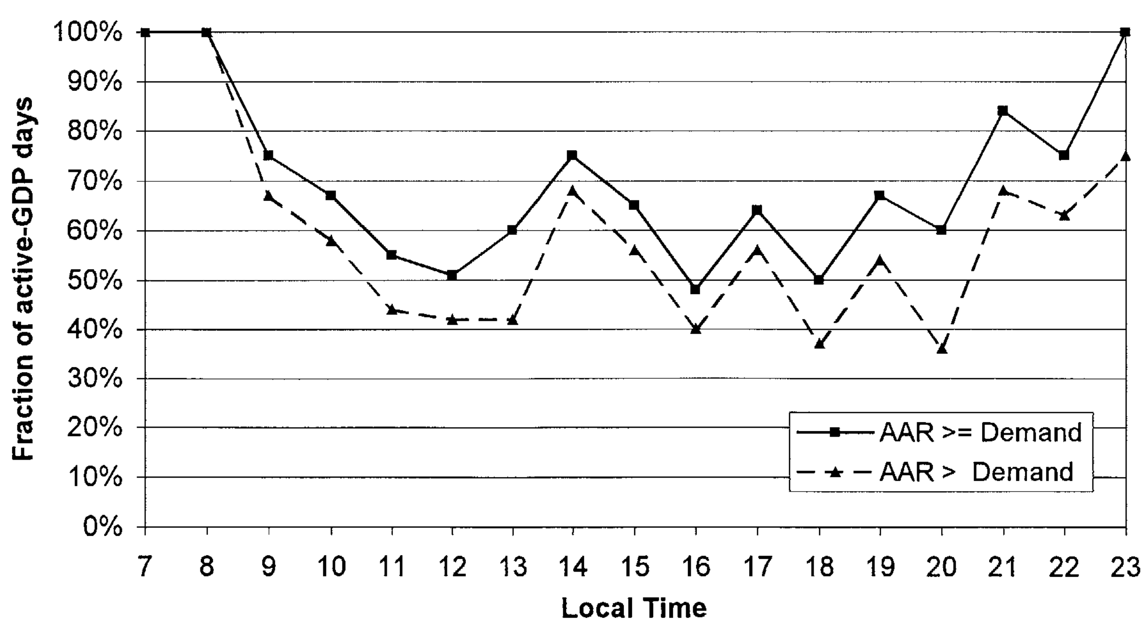

Figure 2-14: SFO fraction of days where AAR capacity is sufficient to handle actual arrivals

Figure 2-14 indicate that the capacity during the afternoon hours is fully utilized. Additional flights can only be added at the risk of more frequent occurrences of flights finding the capacity being fully utilized upon arrival, and thus being forced into unexpected airborne holding. The hours between 16:00-17:00 and 18:00-19:00 are currently the most constrained. However, there is a large variability from hour to hour.

Finally, Figure 2-15 shows the seasonal variation in the utilization of the slots. The large variability shown in this figure reinforces the conclusion reached earlier about the impossibility of increasing the stated capacity by one additional slot during the afternoon period.

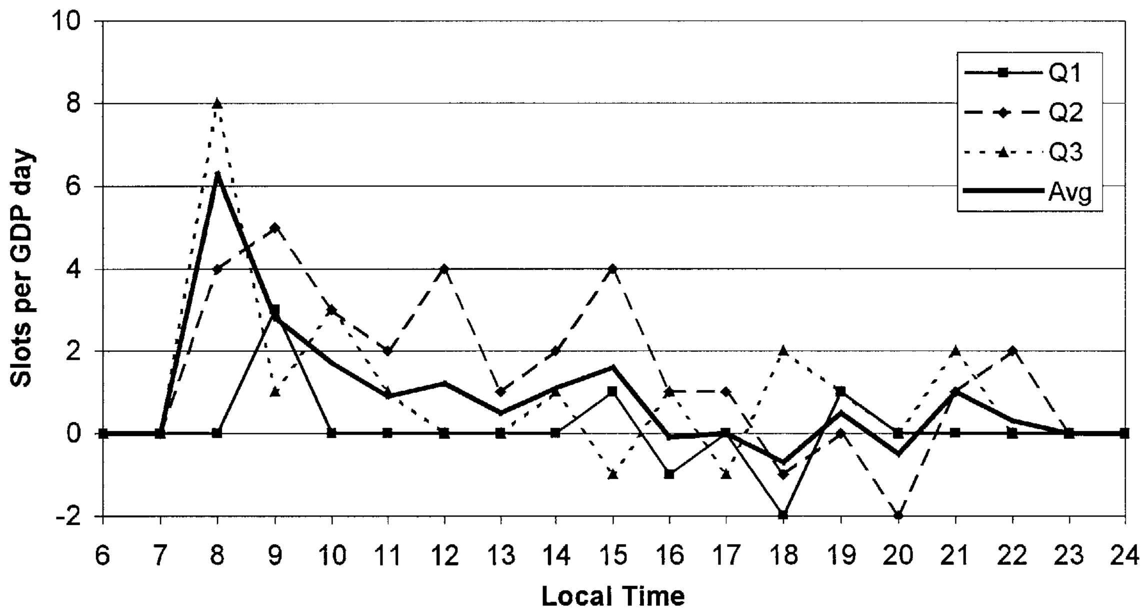

Figure 2-15: SFO average quarterly matching of slots with demand, by hour

2.4.3 Recovering slots for cancelled GDPs

Slots may be recovered when a GDP is cancelled or the acceptance rate is increased. A partial indication of how well and how quickly this can be done answer to this question can be found by analyzing what happens during cancelled-GDP hours. These are the hours of the day that originally had an AAR, but where the AAR was subsequently lifted due to the cancellation of the GDP. Since the GDP was cancelled for these hours, additional capacity is now available.

By comparing with the AAR specified for these hours, it is possible to determine how many more flights have been handled beyond this stated capacity limitation. Figure 2-16 shows the results of this analysis for SFO.

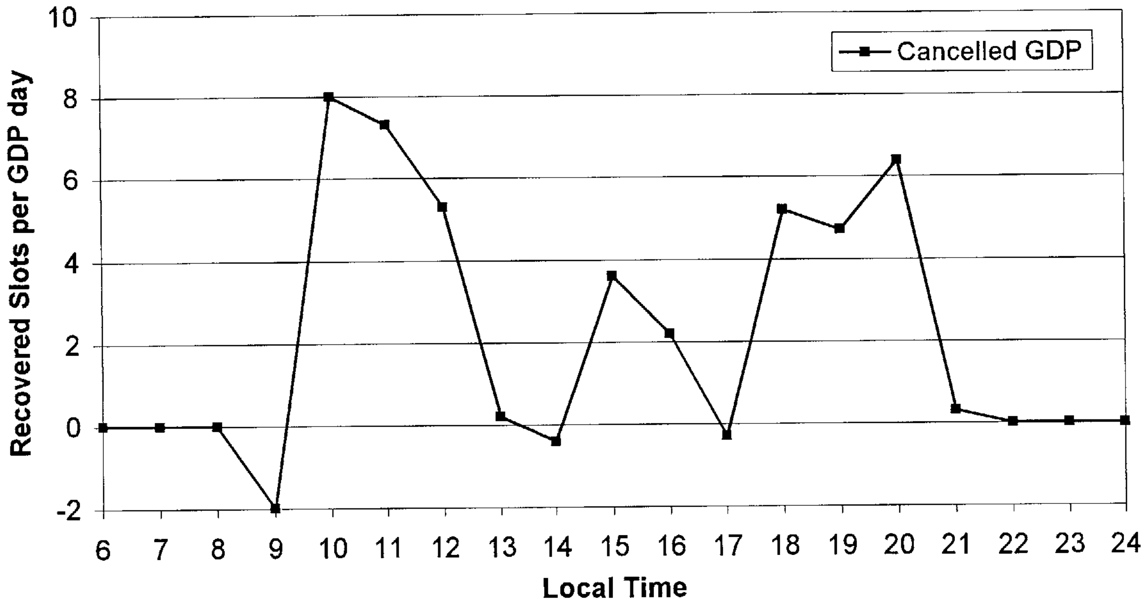

Figure 2-16: SFO average recovered slots from a cancelled GDP

Figure 2-16 shows a quite large number of recovered slots. This is especially noticeable in the 10:00-12:59 and 18:00-20:59 timeframes. This is an indication that a greater number of flights are arriving than originally planned for by the (now-defunct) AAR settings.

This is only a partial answer to the underlying question of how quickly capacity can be increased and fully utilized. To obtain a more complete answer, it is necessary to look at what happens in all the hours immediately after a cancelled GDP, and not just look at the cancelled-GDP hours. The after-effects of the cancelled-GDPs may well stretch beyond the end-time of the GDP itself, but have not been measured in this study. They could possible be measured by basing the measurements of the number of arriving flights on the length of time since the capacity constraints were lifted at the airport. Another factor to take into account would be the lead-time being given between the time of the cancellation, and the hours whose capacity restrictions are being lifted. One should expect to see less and less impact of these former capacity restrictions, as the length of time increases and the longer the lead-time of the cancellation notice. This type of analysis has not been performed yet.

2.5 Unexpected Airborne Holding

One objective of ground delay programs is to keep airborne delays to a minimum, under the assumption that it is cheaper and safer for a flight to incur the delay while sitting on the ground than while aloft. One major component of airborne delay is unexpected airborne holding. This section addresses the issue of whether or not unexpected airborne holding increases with the introduction of GDPs, and if so, by how much.

Note that the term used here is "unexpected airborne holding", not merely "airborne holding". Airlines may already plan for a certain amount of airborne holding and factor that into the flight plans they file, but the focus in this section is to measure the airborne holding that is not planned for and therefore unexpected.

2.5.1 Methodology

There are two basic steps in the analysis. First, we determine the unexpected airborne holding experienced by each flight. Second, we aggregate the results according to the specific needs of the analysis.

Unexpected airborne holding is not directly available as a data element in the ΔADL files, but it can be estimated fairly accurately from the estimated time of arrival information that is available. The method used to calculate unexpected airborne holding is described fully in [1], appendix B. The following description borrows heavily from this appendix.

(1) Let T_d be the actual time of departure. At T_d + 15 minutes, record the ETA (estimated time of arrival) field, and call this ETA_d. The 15 minute interval after departure before recording the ETA makes it possible for the flight to sustain any departure-related delays and update its ETA accordingly.

(2) Let T_t be the estimated time of arrival near to or at the terminal airspace. T_t can be estimated as $T_t = ETA_d - 30$. At this time, the flight may have experienced (most of) any en-route delays that it will sustain, and so will have an accurate prediction of its ETA. Record the ETA at T_t, and call it ETA_a.

(3) Finally, record the actual time of arrival at the runway, ARTA. Note that for reasons described in the [1], appendix B, it is better to use the last value of ETA prior to arrival than the actual ARTA value found in the ΔADL files. The unexpected airborne holding UABH can then be calculated as:

$$UABH = ARTA - ETA_a$$

The unexpected airborne holding is calculated for every flight using this algorithm. Due to what appears to be data corruption in the ΔADL data, flights occasionally show up with very large UABH delays, both positive and negative. Since these flights might conceivably skew the results, they are filtered out before any aggregation takes place. The valid range of UABH is set to be [-60,200], i.e., any flights registering an unexpected airborne holding of less than -60 minutes or greater than 200 minutes are automatically disregarded in all aggregate values.[4]

2.5.2 Unexpected airborne holding by month

Table 2-9 shows the average unexpected airborne holding by month. The results are given for each of the four airports being studied. *All flights* measures the average over all flights in the particular sample, whereas *AGDP flights* only uses active-GDP flights in its calculations, i.e., only those flights whose arrival was actively being managed by a GDP. In the latter category a given month's reading is listed as 'n/a' if there were no active-GDP flights into that airport on that month.

4 Using a filtering value of 200 minutes may seem excessive; after all, this represents an airborne holding in excess of three hours. However, lowering the value to 120 minutes does not change the results in any substantial fashion. For instance the average active-GDP SFO UABH for January would change from 2.9 minutes per flight to 2.7 minutes per flight, and for September would remain unchanged at 2.4 minutes per flight. All other changes fall between these two extremes.

Month	EWR All Flights	EWR AGDP Flights	LGA All Flights	LGA AGDP Flights	SFO All Flights	SFO AGDP Flights	STL All Flights	STL AGDP Flights
Jan	1.2	1.9	1.1		2.9	2.9	1.4	
Feb	1.9	4.3	1.8		2.3	4.3	1.6	
Mar	1.1	2.7	0.7		1.3	3.2	2.1	
Apr	1.5	3.7	1.6	n/a	1.0	2.8	2.1	16.3
May	1.5	3.5	1.1	1.3	1.6	2.9	1.6	8.3
Jun	1.9	2.8	1.2	4.3	1.7	2.6	1.0	5.1
Jul	1.3	0.3	0.8	0.0	1.9	6.3	1.0	3.3
Aug	1.1	1.1	1.1	1.7	1.0	6.9	1.1	n/a
Sep	1.4	n/a	0.5	0.4	0.8	2.4	0.7	1.8
Average	1.4	3.1	1.1	1.8	1.6	3.8	1.4	8.9
Low	1.1	4.3	0.5	0.0	0.8	2.4	0.7	1.8
High	1.9	0.3	1.8	4.3	2.9	6.9	2.1	16.3

Table 2-9: Average unexpected airborne holding, by month

Figure 2-17 shows the results when including all flights in the averages. Figure 2-18 only includes the data for the active-GDP flights.

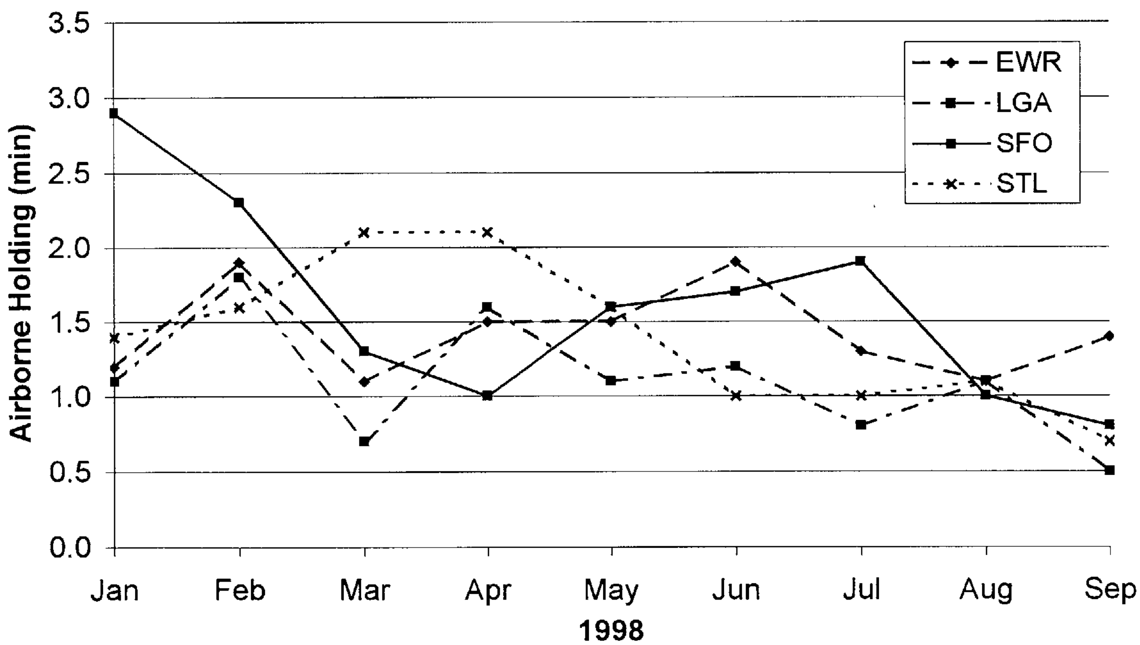

Figure 2-17: Average unexpected airborne holding for GDP and non-GDP flights

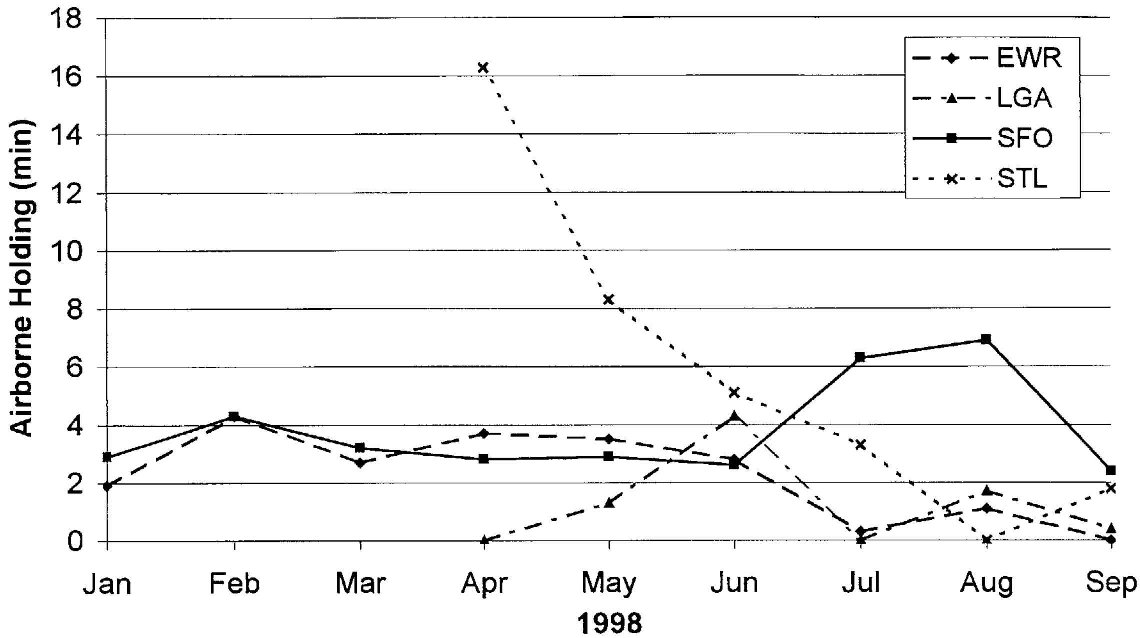

Figure 2-18: Average unexpected airborne holding for active-GDP flights

Figure 2-17 shows that the average unexpected airborne holding for all flights varies between 1.1 minutes at LGA, 1.4 minutes at EWR and STL, and 1.6 minutes at SFO. It has a relatively tight distribution around those averages. There is no discernible trend over time suggesting either a decrease or an increase in the average unexpected airborne holding.

Figure 2-18 shows that the average unexpected airborne holding increases substantially when considering just the active-GDP flights. The unexpected airborne holding increases by a factor of 2.2 at EWR, 1.6 at LGA, 2.4 at SFO and 6.4 at STL. STL shows particularly high active-GDP averages for April (16.3 minutes) and May (8.3 minutes). Each of these months had just two days with active-GDP operations. Table 2-10 shows the data for the four GDP days. The flights on these four days have been separated into three different categories, *active-GDP, cancelled-GDP* and *no-GDP,* and two columns for each of these categories: *flights* is the number of flights, and *ABH* is the average unexpected airborne holding.

Date	Active GDP Flights	Active GDP UABH	Cancelled GDP Flights	Cancelled GDP UABH	No GDP Flights	No GDP UABH
4/28/98	159	19.3	33	7.7	375	6.2
4/29/98	113	12.1	339	6.5	162	14.6
5/15/98	292	10.7	123	5.3	97	8.9
5/22/98	128	2.8	35	0.6	333	17.3

Table 2-10: STL Unexpected airborne holding for 4/28, 4/29, 5/15 and 5/22

The high average for no-GDP flights is quite surprising, especially on 4/29 and 5/22, when they actually exceed the averages for the active-GDP flights. Examination of the underlying data for 4/29 shows that there were two separate GDPs instituted, the first from 7:00 to 12:30, and the second from 18:05 to 20:40. Table 2-11 summarizes the state of the GDPs, the AAR, the flights actually landed and the unexpected airborne holding (*UABH*).

Hour	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
GDP	No	Act	Act	Act	Act	Act	Act	Cnx	Cnx	No	No	No	Act	Act	Act	Cnx	Cnx	Cnx
AAR		32	32	32	32	32	32	32	45				32	32	32	42	42	52
Landed	27	29	28	31	27	40	48	41	41	34	36	42	34	22	33	47	19	6
UABH	3	6	5	12	24	6	1	0	3	25	22	22	16	32	5	3	0	0

Table 2-11: STL hourly AAR and unexpected airborne holding on 4/29

The information shows a circumstance arising around 15:00 local time that dramatically reduced the number of flights being able to land, and also caused the average unexpected airborne holding to increase substantially. A second GDP was instituted at 18:05, probably in response to these conditions, and after two hours the unexpected airborne holding was reduced to 5 minutes per flight, despite landing no more than 33 flights in the hour of 20:00-21:00.

The situation on 5/22 differs from the one on 4/29 in several respects. Table 2-12 summarizes the state for that date.

Hour	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
GDP	No	No	No	No	No	No	No	No	No	No	No	No	Act	Act	Act	Cnx	Cnx	Cnx
AAR													32	36	56	56	56	56
Landed	17	29	35	24	26	28	62	30	35	26	9	8	31	31	48	38	40	36
UABH	5	13	13	20	25	27	15	17	33	3	6	18	14	5	0	0	3	0

Table 2-12: STL hourly AAR and unexpected airborne holding on 5/22

On this date, the GDP was instituted very late in the day, despite there having been several periods during the day when the capacity was clearly affected. The reduced capacity is reflected in the low number of flights that actually landed and in the excessive buildup of unexpected airborne holding. For instance, during the hour of 11:00 to 12:00, only 28 flights landed, and the unexpected airborne holding was 27 minutes on average. No information is available on the events of this day that caused the buildup of delays.

2.5.3 Unexpected airborne holding in relation to GDP

Figure 13 provides a comparison of the unexpected airborne holding at SFO for flights that were not subjected to a GDP program, for active-GDP flights, and for cancelled-GDP flights.

The unexpected airborne holding data show the expected relative ordering. No-GDP flights have the smallest average airborne delay, and active-GDP flights the longest. Cancelled-GDP flights fall between these two, which again is as expected. They operate by definition on days where there are some capacity constraints (or the GDP would not have been instituted), yet the conditions are not judged severe enough to carry through with the GDP for those flights.

There appears to be no structural reason for the increase in unexpected airborne holding at SFO during July and August. The number of flights impacted are in line with those for other months, unlike the situation described above for STL, where the discrepancies were the result of four particularly bad days.

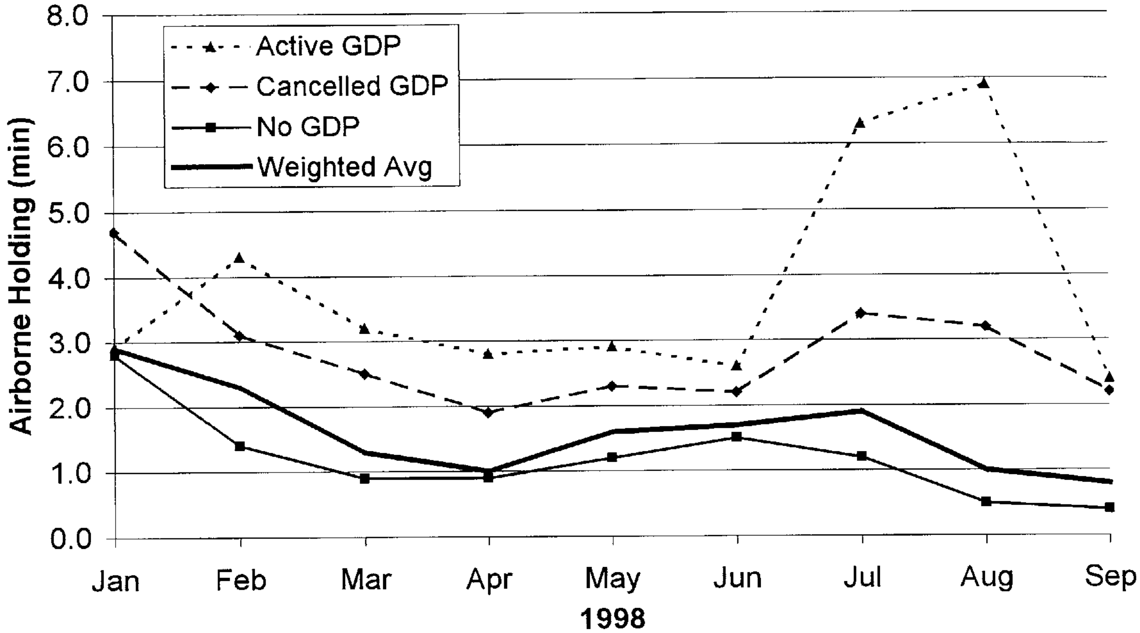

Figure 2-19: SFO average monthly UABH, by type of GDP flight

Table 2-13 examines the flights that experienced unexpected airborne holding. *Dest* is the airport (EWR or SFO), *GDP* is the state of the GDP (Active, Cancelled or None), *Count* is the number of flights that experienced unexpected airborne holding, *mean* is the average unexpected airborne holding, and *stddev* is the standard deviation of the unexpected airborne holding.

Dest	GDP	Count	Mean	Stddev
EWR	Active	5866	3.0	8.4
EWR	Cnx	2351	4.1	19.0
EWR	No	124290	1.2	6.8
SFO	Active	11517	3.7	8.9
SFO	Cnx	10890	2.7	9.8
SFO	No	86078	1.1	6.2

Table 2-13: Unexpected airborne holding variance at EWR and SFO

The mean values can be compared pair-wise by testing the statistical hypothesis that one mean indeed differs from the other. For instance, to test if the EWR mean found for active-GDP days (3.0) differs from the one found for cancelled-GDP days (4.1), the statistical null hypothesis to be tested is: *Is the average EWR UABH on active-GDP days the same as the EWR UABH on cancelled-GDP days.* Using a two-sample, two-tailed t-test to compare the two means, one finds that the null hypothesis can be rejected at the 0.69% significance level. This means that the two means being compared are indeed significantly different. Conducting a similar analysis for all pairs of means for a given airport shows that each mean differs from each other mean. The actual values of the t-test are given in Table 2-14.

Comparing UABH for:	Critical value
EWR active-GDP days to EWR cancelled-GDP days	$6.91\ 10^{-3}$
EWR active-GDP days to EWR no-GDP days	$1.56\ 10^{-57}$
EWR cancelled-GDP days to EWR no-GDP days	$2.00\ 10^{-13}$
SFO active-GDP days to SFO cancelled-GDP days	$1.59\ 10^{-15}$
SFO active-GDP days to SFO no-GDP days	$4.39\ 10^{-195}$
SFO cancelled-GDP days to SFO no-GDP days	$2.00\ 10^{-13}$

Table 2-14: Comparing the unexpected airborne holding means statistically

All the test values are highly significant. The closest call is between the active-GDP and cancelled-GDP means at EWR, which are statistically significant at the 0.69% level. All other pairs show higher levels of significance than this. The weighted averages of the active-GDP and cancelled-GDP averages are 3.3 minutes for EWR and 3.2 minutes for SFO. Based on these weighted averages, both airports show an average increase of 2.1 minutes in unexpected airborne holding during GDPs. The overall conclusion is that the imposition of a GDP increases the average unexpected airborne holding by about two minutes per flight.

2.5.4 SFO unexpected airborne holding by hour

Additional insight into the nature of unexpected airborne holding can be gained by determining how it varies over the course of a day. Intuitively one would expect to see unexpected airborne holding increase during times when traffic volumes approach or exceed the capacity of the airport. This situation occurs during the morning and evening peak hours, and during times of restricted capacity, as manifested by the GDP programs. Figure 2-20 explores this hypothesis on the available data for SFO, showing the actual unexpected airborne holding experienced at SFO during active-GDP, cancelled-GDP and no-GDP hours.

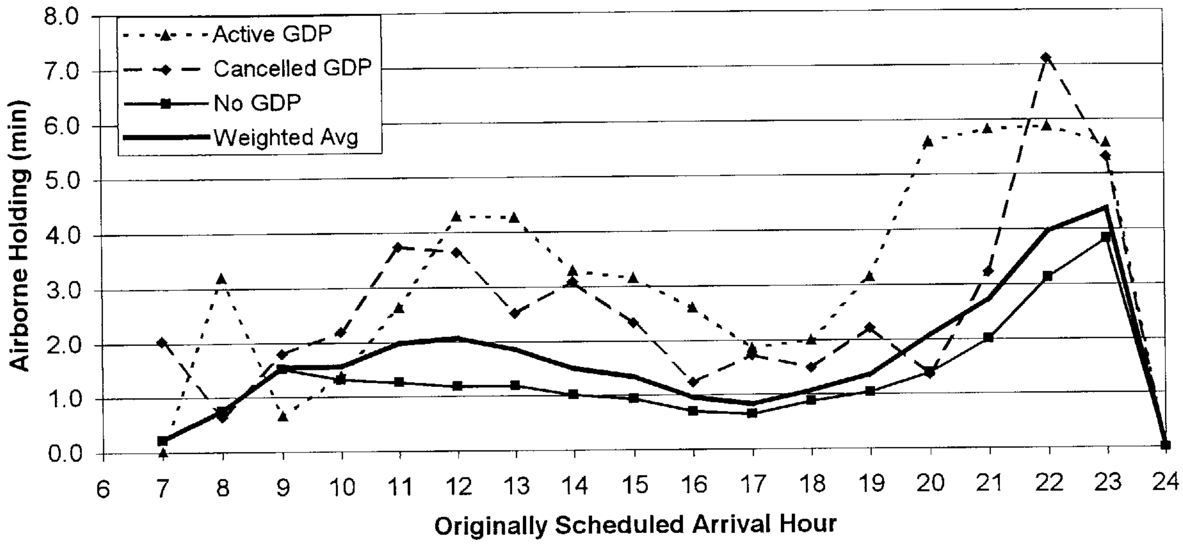

Figure 2-20: SFO average monthly UABH, by type of GDP flight

The data confirms the intuition. First, there is a definite increase in the late morning and evening hours as the morning and evening flights arrive. These increases match the times of the incidences of the GDPs, albeit shifted out in time by two hours. For reference, Figure 2-2 shows the incidences of GDPs, and Figure 2-3 shows the scheduled demand. Obviously, the weather patterns factor into this, since the incidences of fog are more prevalent in the morning than in the afternoon. Second, the evening increase is larger than the morning one. One possible explanation for this is that not enough GDPs are being declared for or extended into these evening hours, or that they are cancelled prematurely. Figure 2-2 and Figure 2-3 indicate that this may be so. The scheduled demand is as high in the evening peak hour as the morning peak, yet the incidence of GDPs during the evening hours is only about half of the morning peak. Also, the number of late-hour cancelled-GDP hours correspond to the large increase in unexpected airborne holding for cancelled-GDP flights. Third, the situation found in the monthly data, where No-GDP flights have the smallest average unexpected airborne holding and Active-GDP flights the longest, is again generally the case for the hourly data.

There is a large increase in average unexpected airborne holding seen during the late hours of the day, i.e., 22:00 to 24:00. This result, however, is based on a substantially smaller number of flights than found during earlier hours, say., between 20:00 and 21:00. There is an approximate drop-off in the number of flights with unexpected airborne holding of 50%

between 21:00-21:59 and 22:00-22:59, and a further 50% between 22:00-22:59 and 23:00-23:59. The situation differs in one crucial respect during the evening hours from the one found earlier in the day: delaying a flight may no longer be an option for the airlines, since there is a limit to how late a flight can arrive at its destination. This may induce airlines to try to squeeze in the late-evening flights to get the equipment positioned correctly for the next day, even at the cost of experiencing long airborne holding delays.

2.6 Airborne delay

A flight can be delayed for a number of reasons. One such reason is *unexpected airborne holding* (UABH), as discussed in the section 2.5. Another set of reasons can broadly be termed *other airborne delays (OABD)*. Weather-related route stretching, speed reductions and miles-in-trail restrictions are examples of en-route delays. The sum of UABH and OABD is termed *airborne delay (ABD)*. The airborne delay of a particular flight can be calculated as the amount of time beyond the scheduled wheels-off to wheels-on time that the flight remains airborne for.

2.6.1 Methodology

The methodology employed in the previous sections remains valid for this section as well. It consists of two basic steps. First, we determine the airborne delay for each flight. Second, we aggregate these results in a manner dependent on the analysis required.

The overall airborne delay is defined as the excess airborne time above and beyond the scheduled flight time, i.e., (actual airborne time) - (scheduled airborne time). The scheduled airborne time can be determined from the original estimated time of departure (OETD) and original estimated time of arrival (OETA) fields that the airlines submit to the FAA. Both of these are available in the database, and so can be used with ease in the computational model. They are used to create the original estimated time en-route (OETE) variable, defined as:

$$\text{OETE} = \text{OETA} - \text{OETD}$$

The actual airborne time is determined in similar fashion. Again, two data items submitted by the airlines to the FAA are used, estimated time of departure (ETD) and estimated time of arrival (ETA). However, unlike the OETD and OETA variables, these variables will typically change during the course of the day, as the airlines adjust to the day-to-day conditions.

Therefore it is the *last* instance of the variables that are used, i.e., the *last* known ETD prior to departure, and the *last* known ETA prior to arrival. These values are also available in the database, and their difference forms the estimated time en-route (ETE) variable, defined as:

$$ETE = ETD - ETA$$

A question to be answered is why the last ETD and last ETA is used instead of the actual runway time of departure (ARTD) and actual runway time of arrival (ARTA) variables, since the latter pair also exists in the ADL data. The reason for this choice has to do with apparent discrepancies between the two sets of variables. During the development of the unexpected airborne holding algorithm it was noted that the last ETA and the ARTA could differ from each other, sometimes quite substantially, although they should have been essentially identical. The ETD/ETA pair provided the better choice of the two, in the sense of consistency and fewer outliers and missing data, and so was adopted for the unexpected airborne holding analysis. This choice has been carried over to this area of the analysis. See [1], appendix B, for a full discussion of this subject.

A potential source of difficulty lies in the definition of the scheduled airborne time. As currently defined, the scheduled airborne time is determined by using the OETD and OETA variables. The value of these variables are provided to the FAA as part of the routinely filed flight plans that are submitted automatically by the airlines. As such, they may reflect general seasonal variations, but they do not take into account specific, late-arising day-to-day events and conditions. If it is deemed advisable to employ more up-to-date information, then an alternative definition of the scheduled time can be constructed by using the ETD and ETA variables at the time of departure, changing the definition of OETE to:

$$OETE' = (\text{ETA at departure}) - (\text{ETD at departure}).$$

2.6.2 *Average airborne delays*

Table 2-15 and Figure 2-21 show the average airborne delay per flight for the four airports studied. The delay is calculated for each airport on a monthly basis by aggregating the individual delays incurred by all incoming flights, and then dividing by the number of flights. Results are also given for all four airports viewed as a whole, and for all nine months.

Month	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	All
EWR	2.1	4.1	2.4	2.1	0.7	-0.7	-2.0	-1.1	-2.1	0.6
LGA	0.4	1.8	0.1	1.3	0.4	0.3	-0.4	-0.1	-1.1	0.3
SFO	10.1	9.7	6.6	1.2	0.1	-1.0	0.1	-1.4	-1.7	2.5
STL	-1.6	-1.6	-0.8	-0.5	-0.7	-1.2	-1.3	-2.0	-2.7	-1.4
All	2.4	3.2	1.9	1.0	0.1	-0.7	-1.0	-1.2	-2.0	0.4

Table 2-15: Average airborne delay per flight, by month

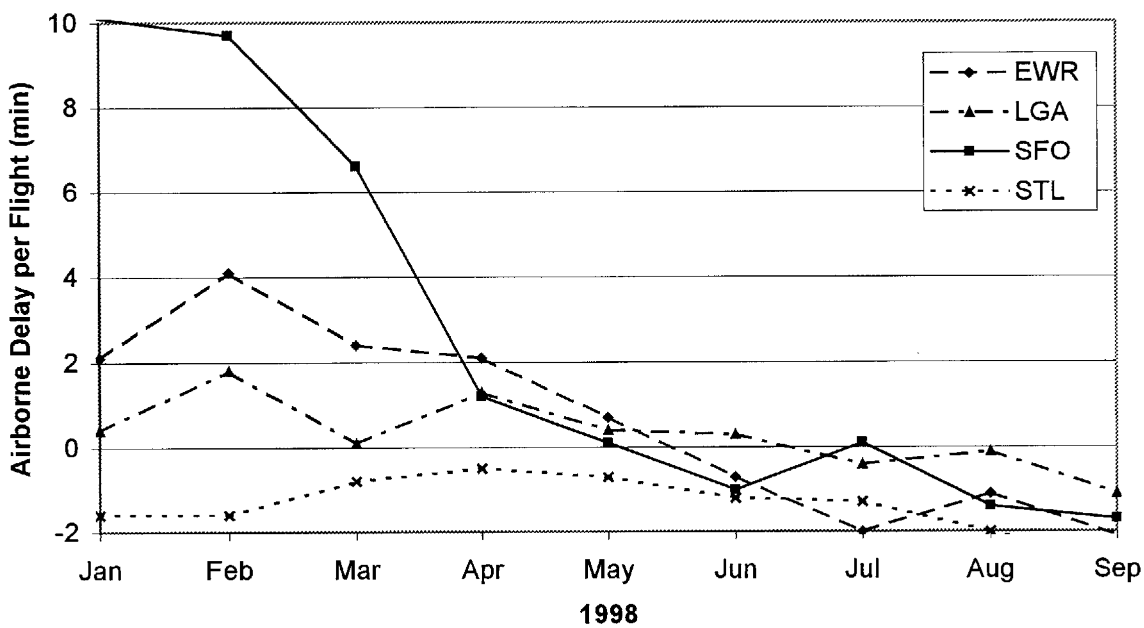

Figure 2-21: Average airborne delay per flight, by month

The first three months of 1998 at SFO stands out in Table 2-15 and Figure 2-21 as having been particularly bad months. Since there is no corresponding increase in the unexpected airborne holding for these months (see Table 2-9 and Figure 2-17), the delay measured can be attributed to en-route delays. The underlying reason for this delay cannot be determined from the available information, but may have been due to the effects of El Niño. As stated earlier in section *2.3, Substitution and Compression Analysis,* February clearly was a difficult month at SFO.

SFO, EWR and LGA all show indications of a possible improving trend. There is a clear decline in the average airborne delay for each of these three airports over the nine months studied. However, not enough data exist to form any conclusions on whether this is part of a general trend, or if it is caused by seasonal effects. The answer to this question will have to await the availability of several years worth of data.

Further insight can be gained by classifying the data according to the GDP state of each flight. This is done for SFO in Table 2-16 and Figure 2-22, and for EWR in Table 2-17 and Figure 2-23. These figures also include the averages for EWR and SFO from Table 2-15, for ease of comparison.

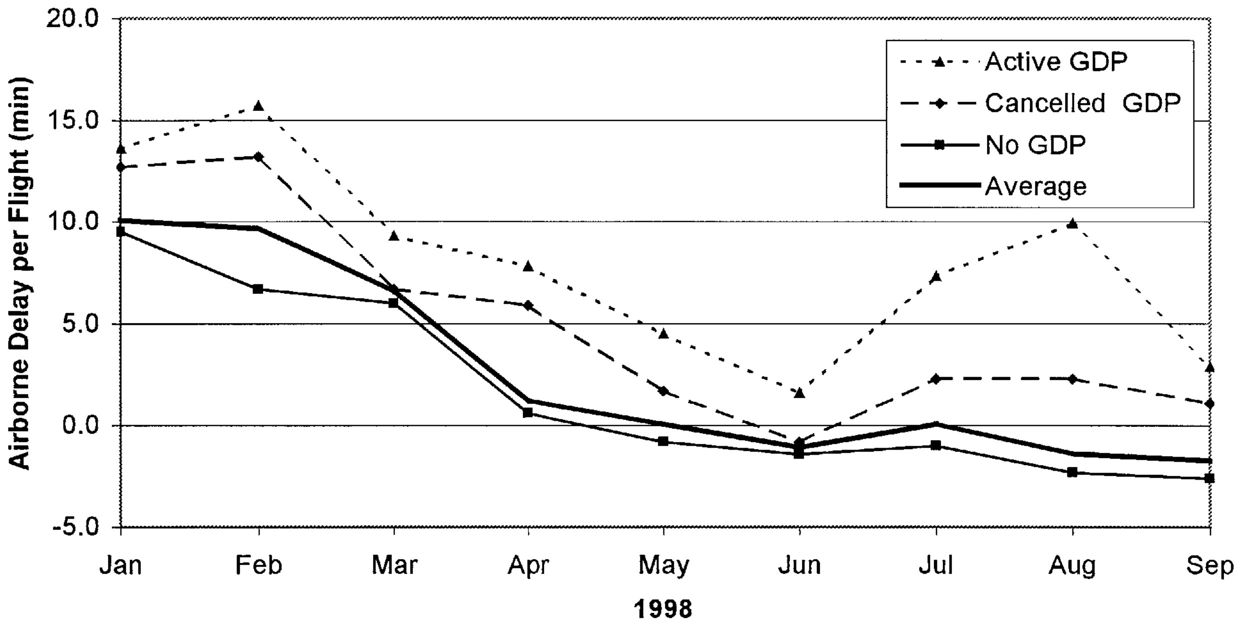

Figure 2-22: SFO average airborne delay per flight, by month

	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	All
Active	13.6	15.7	9.3	7.8	4.5	1.6	7.3	9.9	2.9	9.0
Cancelled	12.7	13.2	6.7	5.9	1.7	-0.8	2.3	2.3	1.1	4.1
None	9.5	6.7	6.0	0.6	-0.8	-1.4	-1.0	-2.3	-2.6	1.4
All	10.1	9.7	6.6	1.2	0.1	-1.0	0.1	-1.4	-1.7	2.5

Table 2-16: SFO average airborne delay per flight, by month

Table 2-16 shows that the SFO average airborne delay is 1.4 minutes for non-GDP flights. This increases to 4.1 minutes for cancelled-GDP flights and 9.0 minutes for active-GDP flights, representing increases of 2.7 minutes and 7.4 minutes respectively over and above the delay for non-GDP flights.

Table 2-16 and Figure 2-22 both show that the airborne delay for non-GDP flights tracks very closely to the average airborne delay. Only a single month, February, shows a significant difference between the two measures. On the other hand, active-GDP and cancelled-GDP flights both show substantial additional delay. The biggest difference is found in August, when active-GDP flights are delayed by 9.9 minutes on average, and non-GDP flights record a savings of 2.3 minutes, for a net difference of 12.2 minutes. There is a distinct increase in the active-GDP delays in the third quarter. This increase is similar to the increase found for unexpected airborne holding, as shown in Table 2-9 and Figure 2-18. The increase from June to August is 8.3 minutes of additional airborne delay, and 4.3 minutes of unexpected airborne holding. Thus, one can conclude that slightly more than half of the increase in active-GDP airborne delays is caused by unexpected airborne holding. This differs from the situation early in year, where the en-route delay dominated the unexpected airborne holding delay. For instance, February saw an average active-GDP airborne delay of 15.7 minutes, of which 4.3 minutes can be attributed to unexpected airborne holding.

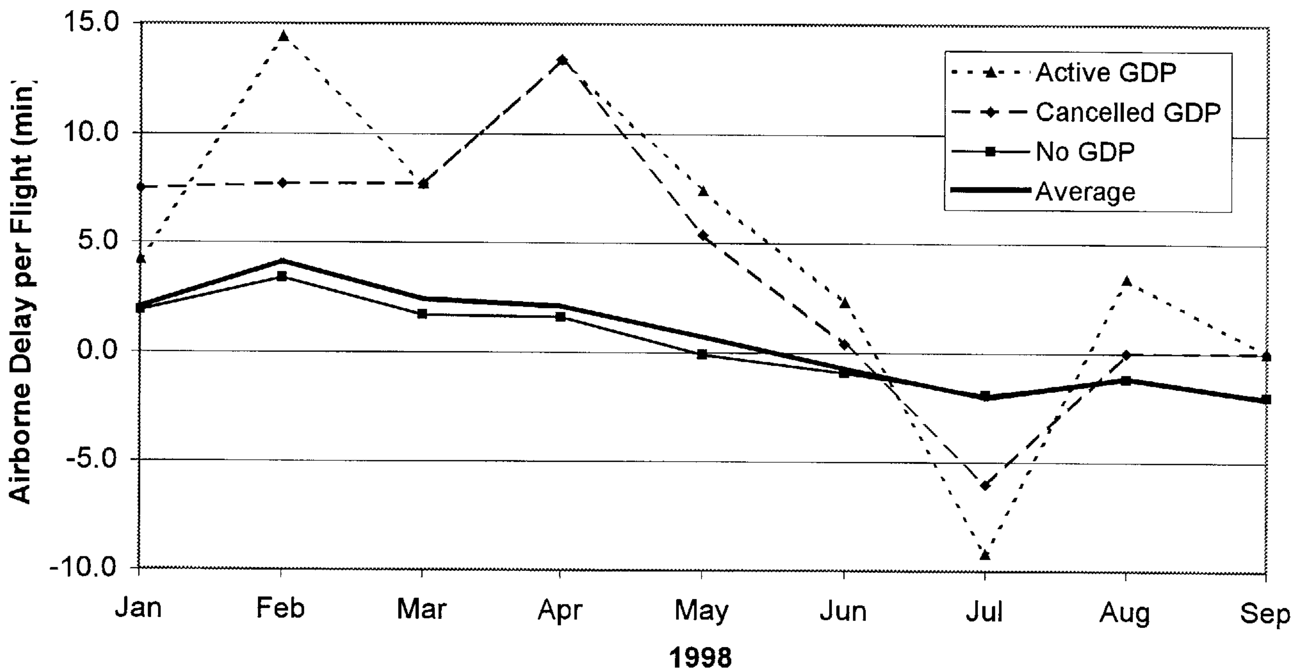

Figure 2-23: EWR average airborne delay per flight, by month

	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	All
Active	4.2	14.4	7.7	13.4	7.4	2.3	-9.2	3.4	n/a	7.8
Cancelled	7.5	7.7	7.7	13.4	5.4	0.4	-6.0	0.0	n/a	4.0
None	1.9	3.4	1.7	1.6	-0.1	-0.9	-1.9	-1.2	-2.0	0.2
All	2.1	4.1	2.4	2.1	0.7	-0.7	-2.0	-1.1	-2.1	0.6

Table 2-17: EWR average airborne delay per flight, by month

Table 2-17 shows that the EWR average airborne delay is 0.2 minutes for non-GDP flights. This increases to 4.0 minutes for cancelled-GDP flights and 7.8 minutes for active-GDP flights, representing increases of 3.8 minutes and 7.6 minutes respectively over and above the delay for non-GDP flights. These increases are nearly identical to the ones found for SFO.

Table 2-17 and Figure 2-23 shows the situation at EWR. Non-GDP flights again track very close to the average airborne delay, as was the case for SFO. Two months, February and April, show a substantial difference between non-GDP flights and active-GDP flights. The difference per flight is 11.0 minutes in February and 11.8 minutes in April. The difference in unexpected airborne holding for the same months are 2.4 minutes and 2.2 minutes respectively, showing that the bulk of the increase can be attributed to wind or en-route delays. July shows an unusual situation occurring, with active-GDP days showing *less* airborne holding than on non-GDP days. However, this result stems from a single incidence of GDP, so no particular conclusions can be drawn from the occurrence.

Figure 2-24 offers a different perspective on the airborne delays. The figure shows the cumulative percentage of flights arriving at a particular destination with less than a given amount of delay. For instance, SFO saw 63% of all of its flights arrive in the scheduled amount of airborne time or less. This increases to 86% when including all flights with delay less than 15 minutes.[5]

[5] This should not be compared to the FAA on-time arrival statistics, for several reasons: first, what is measured in the figure is airborne time, not arrival time; second, cancellations are not factored into the data in the plot -- if they were, the final percentage reached would be less than 100%; and third, each category on the horizontal (X) axis includes data for 5 minutes centered on the value shown, e.g., the 15 minutes value includes all flights arriving with delays between 13 and 17 minutes.

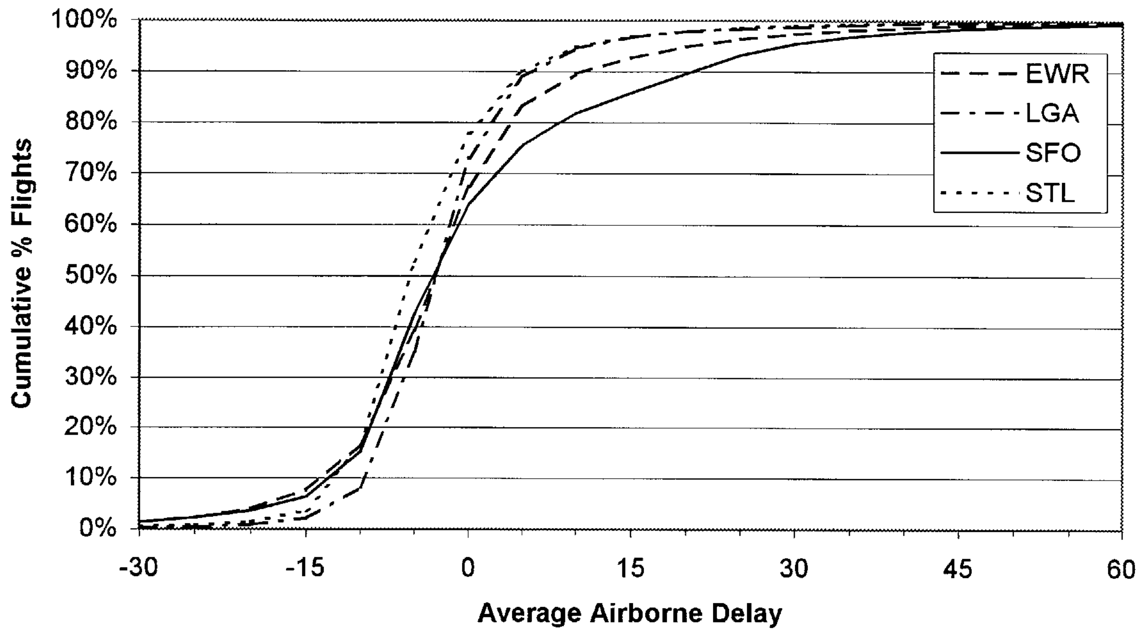

Figure 2-24: Average airborne time delay, by destination

Figure 2-24 shows very similar patterns for the four airports being studied. SFO shows the largest deviations from the norm, with a larger percentage of its flights arriving with a substantial delay. This reiterates the observation made previously in connection with the discussion of Table 2-15 and Figure 2-21. The figure also shows that the airborne portion of a substantial number of flights is completed in less time than scheduled. LGA not surprisingly has the tightest fit to the schedule, given its emphasis on relatively short flight.

2.6.3 Examining SFO airborne delays in more detail

Figure 2-25 compares the scheduled with the actual airborne times for flights arriving at SFO from LAX (Los Angeles International), SEA (Seattle-Tacoma), DEN (Denver), DFW (Dallas-Forth Worth), ORD (Chicago O'Hare) and EWR (Newark). These six airports have been chosen to provide a sample of different travelling distances and high frequency routes. The figure shows data from just two days, 9/1 and 9/30, one without a GDP (9/1), and one with a GDP (9/30). This provides a direct, side-by-side comparison. The figure shows all flights arriving at SFO from the six origin airports. Every flight flown is indicated by a symbol at the intersection of its scheduled and actual en-route time. Different marker symbols are used depending on the GDP status of the flight, and linear regression lines are plotted through each set of markers. The respective linear regression equations are listed on the chart.

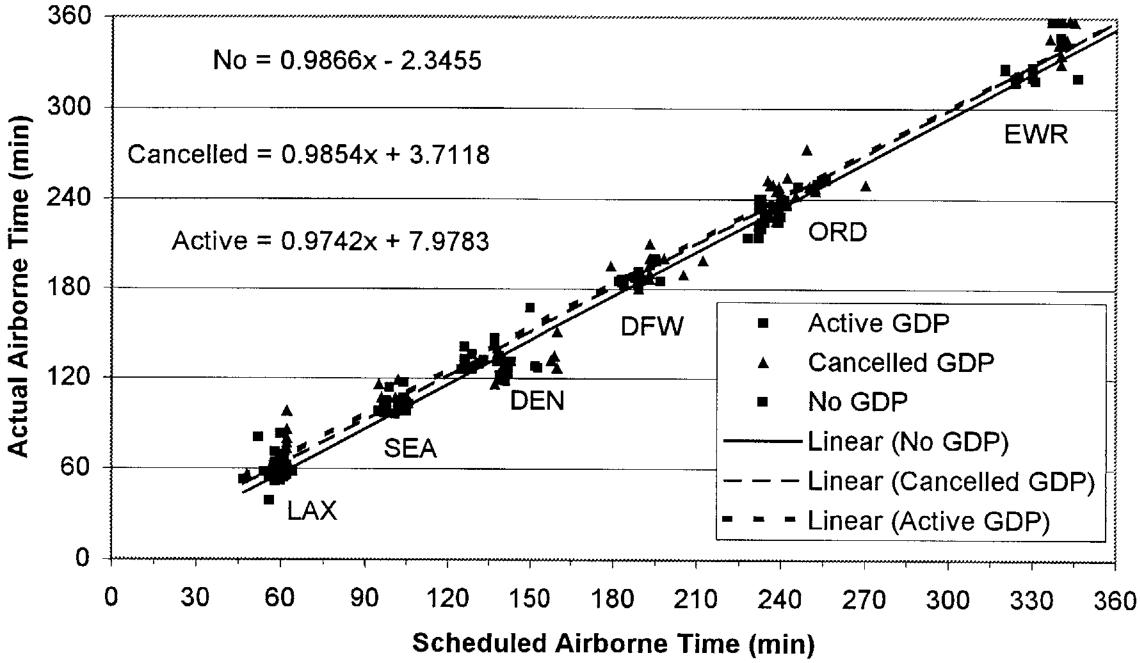

Figure 2-25: SFO scheduled vs actual airborne time on 9/1 (non-GDP) and 9/30 (GDP)

The regression lines highlight the differences both mathematically and, with a sharp eye, visually. Several tentative observations can be made based on the information contained in these charts. There seems to be very little difference between the three regression lines, suggesting little difference in flying time between non-GDP and GDP days. This is a clear indication that the ground delay program is helping to maintain a safe and efficient system by preventing excessive airborne delays.

The impact of the GDP programs on airborne time decreases as the travel distance increases. This can be seen graphically at the upper-right hand corner of the chart, and is brought about by the slight difference in the slope terms of the three linear regression equations. This behavior is consistent with the way that the GDP programs are conducted, since long-haul flights will only be affected by long-running programs, and in some cases (e.g., EWR - SFO flights) may be exempted from GDPs altogether. Furthermore, the effects of the additional airborne holding may be dominated by the effects of winds in a four or five hour flight.

There is a more noticeable impact of the GDP programs on short-haul flights especially, due to the difference in the three intercept terms. This is the reverse situation of that found for the long-haul flights. Since unexpected airborne holding on average increases during GDPs as shown in the previous section, most short-haul flights will experience an increase in their overall time airborne.

The three different states that a flight can be in vis-à-vis a GDP (active, cancelled or none) have the correct ordering, with Active-GDP flights having the biggest airborne delay and No-GDP flights the smallest. This matches the conclusion reached in section 2.5 for unexpected airborne holding.

However, since only two days are included in the data used in Figure 2-25, one must be careful not to carry this analysis too far. The remaining part of this section analyzes the situation in more detail, using the entire nine months' worth of data for flights from the six origin airports.

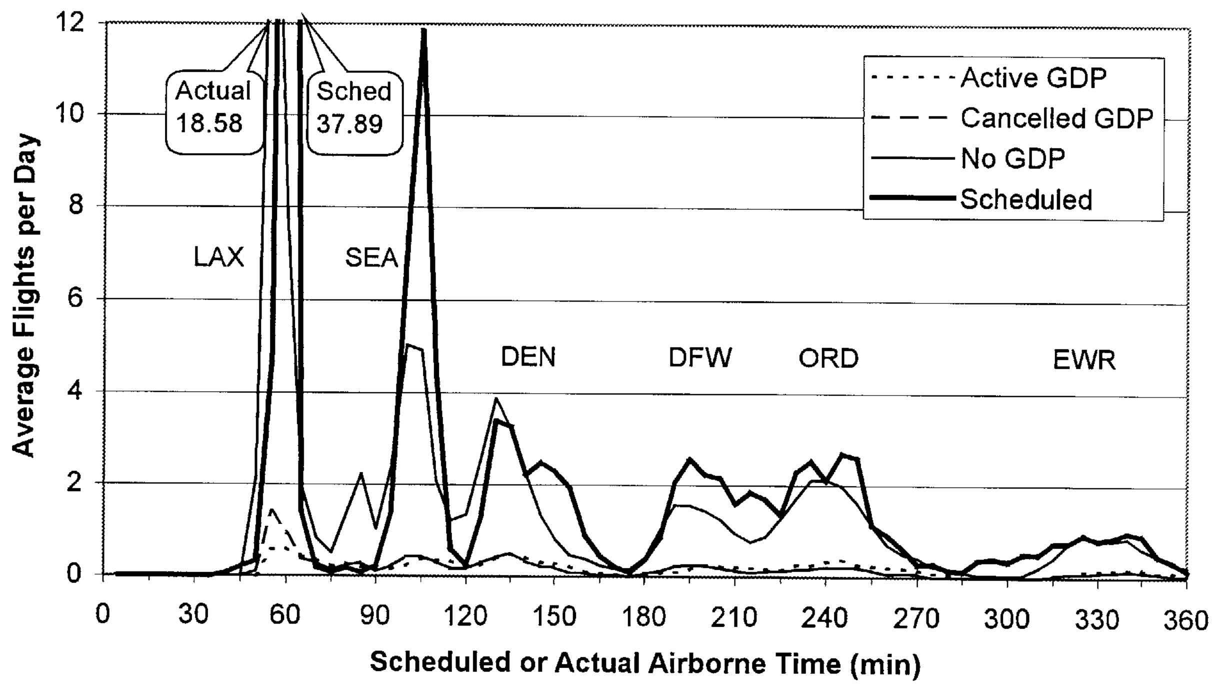

Figure 2-26: SFO number of scheduled flights versus actual flights for six origins

Figure 2-26 shows the scheduled and actual number of flights per day. The flight times are grouped in five-minute intervals on the horizontal axis by the scheduled or actual airborne time. The actual number of flights are divided according to their GDP state. The scheduled number of flights shown are a computed daily average number of flights, ignoring the fact that there may be both day-of-week and seasonal variations over the nine months study period.

The scheduled time airborne shows very little variation for the short-haul flights from LAX and SEA. Figure 2-26 indicates that a substantial number of flights arrive five minutes early from these two origins. For flights of longer duration the picture is more diffuse, with both scheduled and actual airborne times varying much more. This is especially true for EWR, where there is no discernible peak, and where both scheduled and actual airborne time vary by more than thirty minutes.

Figure 2-27 shows the individual cumulative distribution of the actual airborne time of flights from the six origin airports. Figure 2-28 and Figure 2-29 plots the same information in different ways. Figure 2-28 shows the cumulative number of flights as a function of the airborne delay, i.e., the difference between the scheduled and actual airborne time, enabling comparisons between the different flight origins. Finally, Figure 2-29 shows the cumulative number of flights as a function of the relative airborne delay. The relative airborne delay for a flight is calculated as the airborne delay experienced by that flight, divided by the scheduled airborne time for the flight.

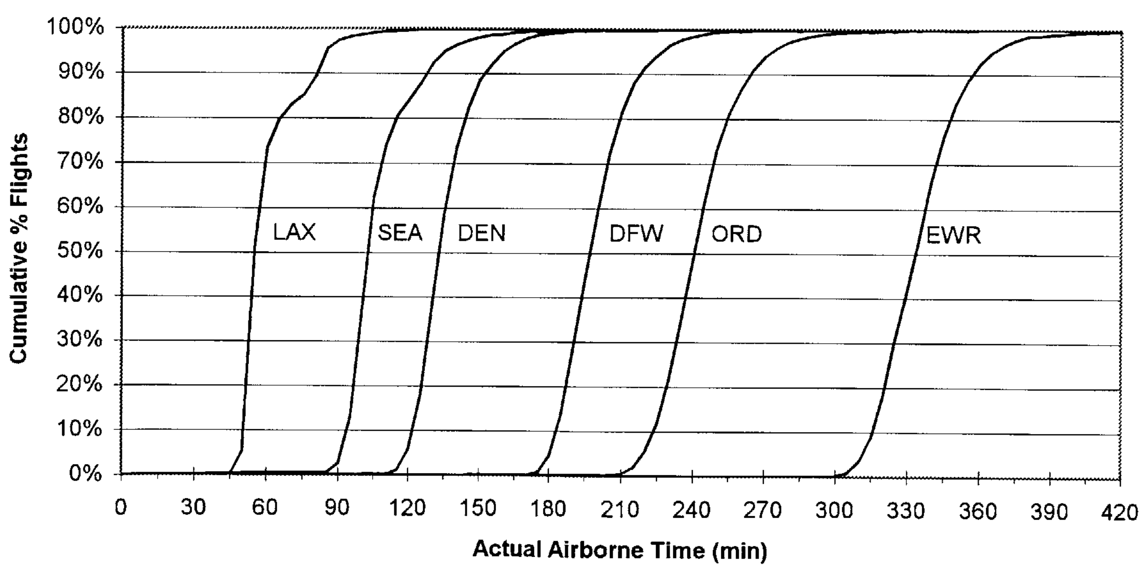

Figure 2-27: SFO airborne time distribution for six origins

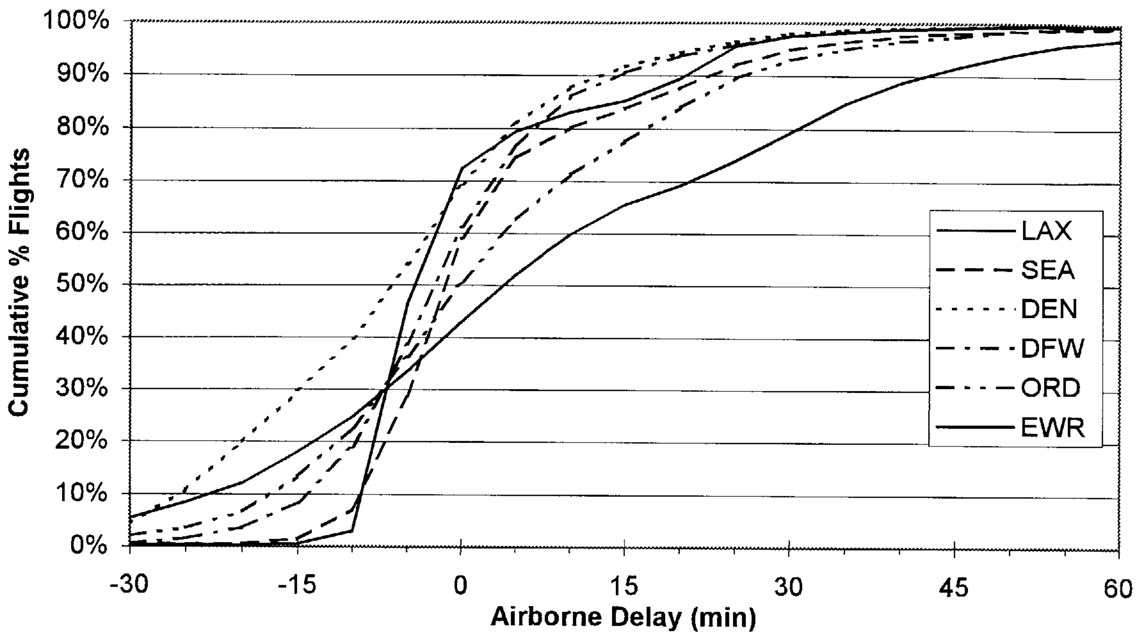

Figure 2-28: SFO airborne delay distribution for six origins

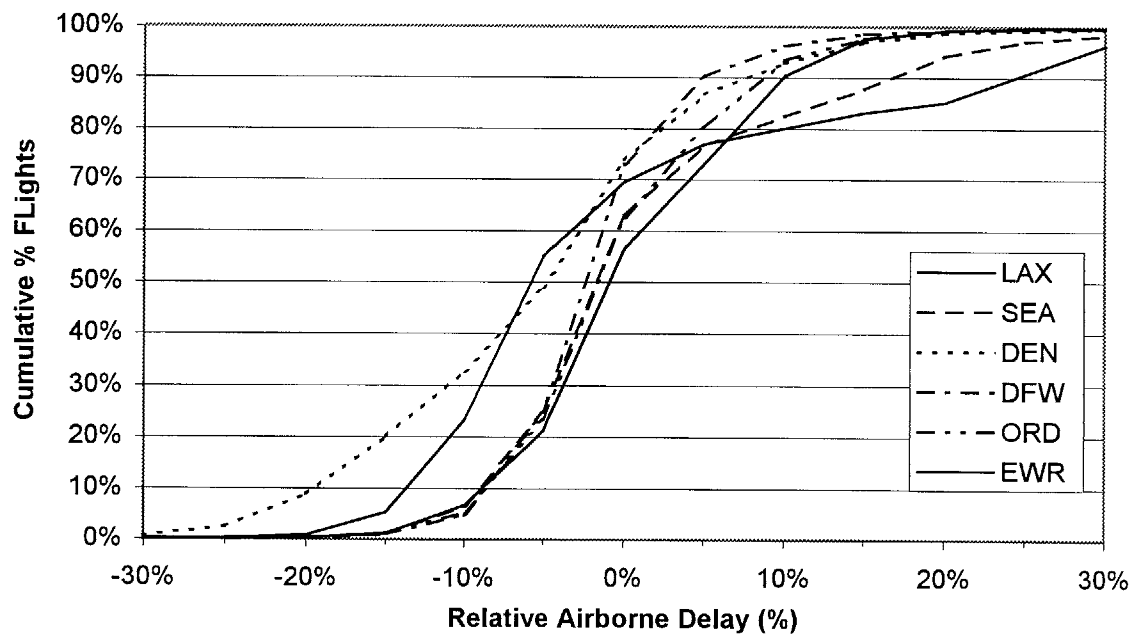

Figure 2-29: SFO relative airborne delay distribution for six airports

Figure 2-28 seems to indicate that the flights from ORD and especially EWR are subject to greater variation in the airborne delay they experience than what is found for the other origin airports examined. The source of this variation is a combination of winds, other en-route delays and unexpected airborne holding delays. If the main reason for the variation is found in winds and other en-route delays, then one would expect to see the variation increase as a function of the duration of the flights. Thus, the amount of en-route delay should then be about the same per hour of flying time. Figure 2-29 shows the results of applying this reasoning to the data, since the figure shows the percentage of delay per unit flying time. Plotting the information in this fashion indicates that the airborne delay as expected is indeed strongly influenced by the en-route delay component, since all six curves are grouped much more tightly together in Figure 2-29 than is the case in Figure 2-28. The effect of winds are shown most clearly on the not insignificant number of very early arrivals of EWR-SFO and ORD-SFO flights seen most clearly in Figure 2-28.

Table 2-18 shows the average airborne delay for flights from the six origin airports. The delay is given in both absolute and relative measures, and is furthermore broken down by GDP state.

	Active GDP (min)	Active GDP (%)	Cancelled GDP (min)	Cancelled GDP (%)	No GDP (min)	No GDP (%)	All (min)	All (%)
DEN	1.2	0.2%	-2.0	-1.8%	-4.7	-3.7%	-3.7	-3.0%
DFW	5.5	2.2%	3.6	1.6%	0.6	0.4%	1.6	0.7%
EWR	19.5	5.3%	11.0	2.8%	7.7	2.3%	9.9	2.8%
LAX	9.6	11.1%	5.0	5.4%	1.7	0.8%	2.5	1.9%
ORD	12.4	9.7%	5.5	5.6%	3.4	2.6%	4.9	3.6%
SEA	12.7	4.4%	7.3	2.3%	3.8	1.3%	5.0	1.9%
All	10.0	6.1%	4.7	3.2%	1.7	0.6%	2.9	1.4%

Table 2-18: Average airborne delay, by destination

The major stand-out in Figure 2-28, Figure 2-29, and Table 2-19 are the DEN-SFO flights. 30% of these flights arrive 15 minutes or more early, which is approximately 10% of the scheduled flying time. The average airborne delay saving is 3.7 minutes, representing a 3.0% saving in airborne time. No further analysis has been performed to obtain an explanation for this.

Table 2-19 shows the results of comparing the mean airborne delay of flights from each of the six origin airports against the average airborne delay, using a two-pair, two-tailed t-test. Each individual mean airborne delay for a particular airport (e.g., *ORD*) is compared to the average airborne delay to determine, if the two means differ statistically. The comparisons are further categorized by flight GDP state to determine what effect, if any, the GDP state has on the test results. The table contains the results of these tests. For example, the ORD-SFO mean airborne delay differ significantly at all levels from the average airborne delay for no-GDP flights, as well as when all ORD-SFO flights are grouped together irrespective of their individual GDP states. The active-GDP ORD-SFO flight airborne delay mean is statistically different from the average at a very strong level of 0.64%. However, the cancelled-GDP ORD-SFO flights do not differ significantly from the average, since the test statistic is 42.29%

GDP state	DEN	DFW	EWR	LAX	ORD	SEA
Active	0.00%	0.00%	0.00%	51.00%	0.64%	0.53%
Cancelled	0.00%	15.73%	0.01%	45.54%	42.29%	1.07%
No	0.00%	0.01%	0.00%	87.86%	0.00%	0.00%
All	0.00%	0.00%	0.00%	3.05%	0.00%	0.00%

Table 2-19: SFO mean airborne delay test results for six airports

The test results given in Table 2-19 indicate that the airborne delay for LAX-SFO flights do not differ significantly from the average airborne delay. This can be explained by noting that the average airborne delay is strongly influenced by the flights from LAX, due to the sheer number of LAX-SFO flights that are included in the calculation of this average. Cancelled-GDP flights from DFW and ORD to LAX also do not have airborne delays that differ significantly from the average airborne delay. However, all other combinations differ significantly from the average at the 5% test significance level.

Figure 2-30 and Figure 2-31 show the effect of the ground delay programs on airborne delays. Both these figures use data from flights flown between SEA and SFO, grouping the data by the state of flights' GDP. Figure 2-30 shows the cumulative percentage of flights as a function of the time airborne, and it also includes the distribution of the scheduled flights. Figure 2-31 shows the cumulative percentage of flights as a function of the delay experienced.

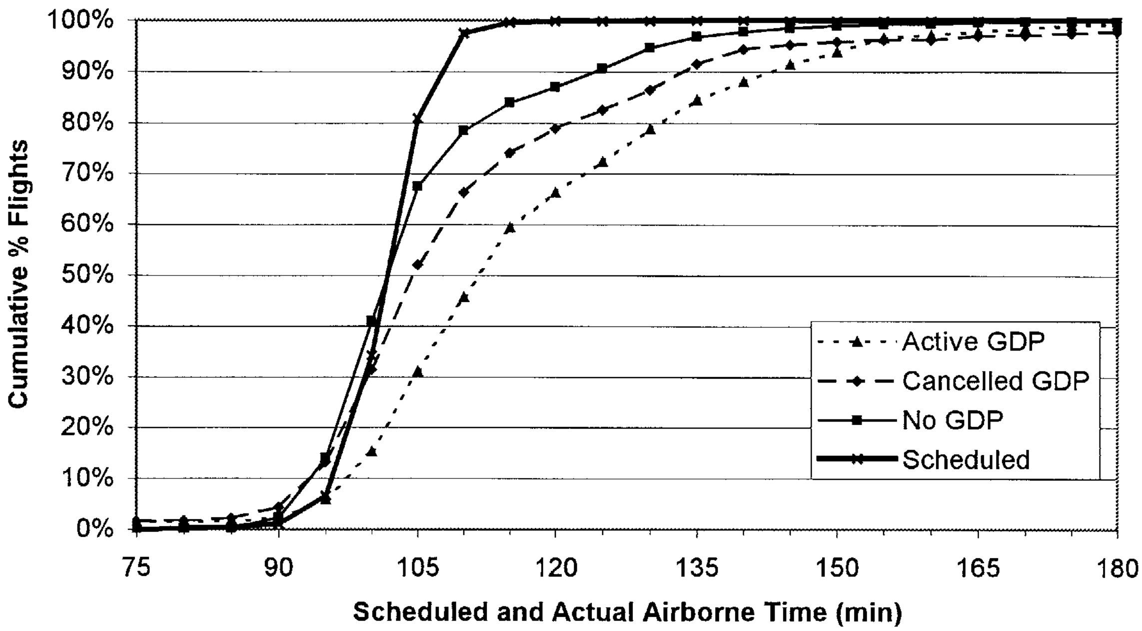

Figure 2-30: SEA-SFO scheduled and actual airborne time, by GDP state

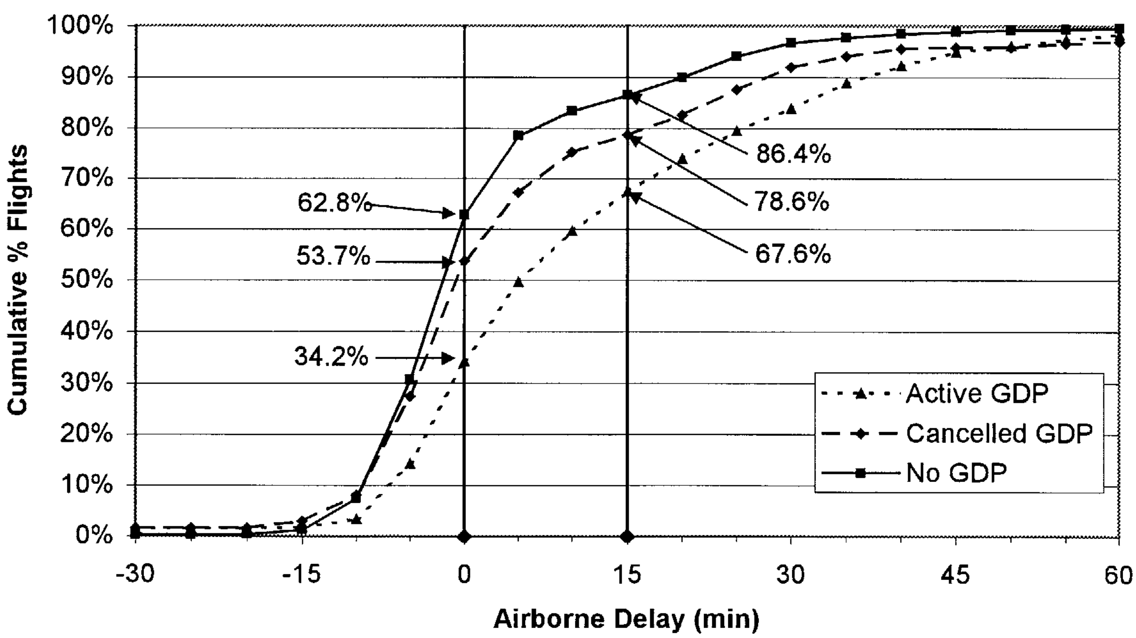

Figure 2-31: SEA-SFO airborne delay, by GDP state

Figure 2-30 indicates that the vast majority of SEA-SFO flights are scheduled for an airborne time between 95 minutes 110 minutes. This figure and Figure 2-31 both indicate that the introduction of a ground delay program affects the ability of flights to adhere to the scheduled airborne time. This can be seen by looking at the percentage of flights that arrive in less than or exactly at the scheduled airborne time. This percentage drops from 62.8% for non-GDP flights to 34.2% for active-GDP flights, for an absolute difference of 28.6%. Similarly, the percentage of flights that arrive with 15 minutes or less of delay are 86.4% for non-GDP flights and 67.6% for active-GDP flights, for a difference of 18.8%. (See footnote 5 for why this should *not* be compared to the FAA on-time arrival statistics.) Overall, the difference between the means of the no-GDP flights and the active-GDP flights is statistically significant at all levels, using a two-tailed t-test to perform the test.

The cancelled-GDP flights fall between the non-GDP flights and the active-GDP flights, as expected. The cancelled-GDP flights track quite closely to the non-GDP flights when looking at delays of less than 0 minutes. Beyond this point there appears to be more of an impact on the cancelled-GDP flights. This can be explained by noting the limited impact on cancelled-GDP flights, when the GDP is cancelled well in advance of departure time. These flights operate in an environment very similar to that found for non-GDP flights. Conversely, flights where the GDP was cancelled relatively close to the departure time are bound to encounter the after-effects of the GDP, both in terms of the pent-up traffic demand as well as weather conditions that, while not cause for continuing the GDP, may still be sub-par. Using a two-tailed t-test to compare cancelled-GDP flights against active-GDP flights and no-GDP flights respectively, one finds that the difference between the active-GDP and cancelled-GDP flights is statistically significant at the 0.01% level, and the difference between the cancelled-GDP and no-GDP flights is statistically significant at the 0.04% level, both very strong results.

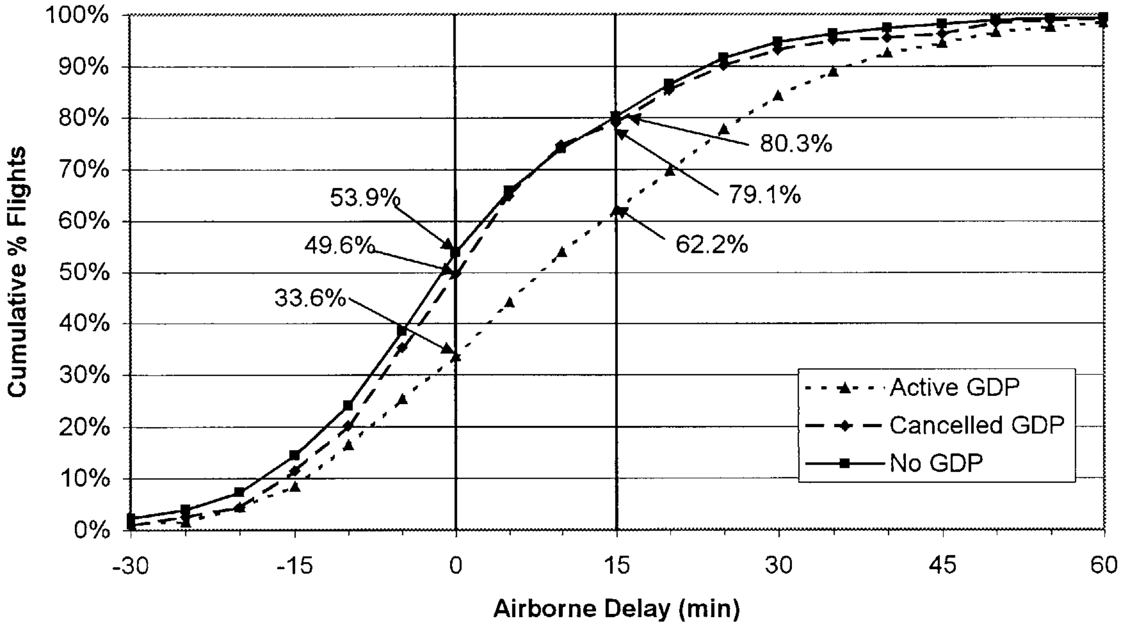

Figure 2-32: ORD-SFO airborne delay, by GDP state

Figure 2-32 shows the result of performing a similar analysis on the ORD-SFO data. The result differs from the SEA-SFO in that the cancelled-GDP flights track much closer to the no-GDP flights. This can be explained by noting that the airborne time between ORD and SFO is much greater than between SEA and SFO. This allows more time for any residual impact at SFO (e.g., weather or pent-up demand) to be cleared away, resulting in fewer delays being imposed on the cancelled-GDP flights. On the other hand, active-GDP flights differ from no-GDP flights in very similar fashion to that found for SEA-SFO flights. Statistically, comparing the means of the different GDP states show that active-GDP flights differs from both the cancelled-GDP and no-GDP flights at all levels, whereas the difference between the cancelled-GDP and no-GDP flights is significant at the 4.34% level.

Table 2-20 lists the results from comparing the airborne delay means from the different GDP states in a pair-wise fashion for each origin airport, using a two-tailed t-test. *Act-Cnx* shows the results of comparing active-GDP and cancelled-GDP flights. Similarly, *Cnx-No* compares cancelled-GDP to no-GDP flights, and *No-Act* compares no-GDP to active-GDP flights. Table 2-20 shows that all tests are statistically significant at the 5% level.

	Act-Cnx	Cnx-No	No-Act
LAX	0.00%	0.00%	0.00%
SEA	0.01%	0.04%	0.00%
DEN	0.40%	0.06%	0.00%
DFW	3.59%	0.00%	0.00%
ORD	0.00%	4.34%	0.00%
EWR	0.00%	4.83%	0.00%
All	0.00%	0.00%	0.00%

Table 2-20: SFO average airborne delay test results, by GDP state

2.6.4 Generalizing the results

The previous section dealt with the airborne delays suffered by flights flying into SFO from six specific origins. This section generalizes the results found by analyzing all flights arriving at all four airports under study. Table 2-21 shows the average airborne delay by GDP state and overall for the four destination airports being studied. Figure 2-33 plots the cumulative percentage of flight arrivals as a function of percentage airborne delay, while categorizing flights by their GDP state. Similarly, Figure 2-34 plots the cumulative percentage of flight arrivals as a function of percentage airborne delay, while categorizing flights by their destination. Both figures maintain the S-shaped curves found while examining the SFO data.

	Active GDP (min)	Active GDP (%)	Cancelled GDP (min)	Cancelled GDP (%)	No GDP (min)	No GDP (%)	All (min)	All (%)
EWR	7.8	4.7%	3.9	2.1%	0.2	-0.8%	0.6	-0.6%
LGA	1.9	2.1%	2.5	2.3%	0.3	-0.1%	0.3	0.0%
SFO	9.0	5.4%	4.1	2.0%	1.4	-0.4%	2.5	0.4%
STL	8.1	7.4%	2.0	0.6%	-1.5	-2.9%	-1.4	-2.8%
All	8.3	5.2%	3.8	1.9%	-0.1	-1.2%	0.4	-0.9%

Table 2-21: Average airborne delay, by destination

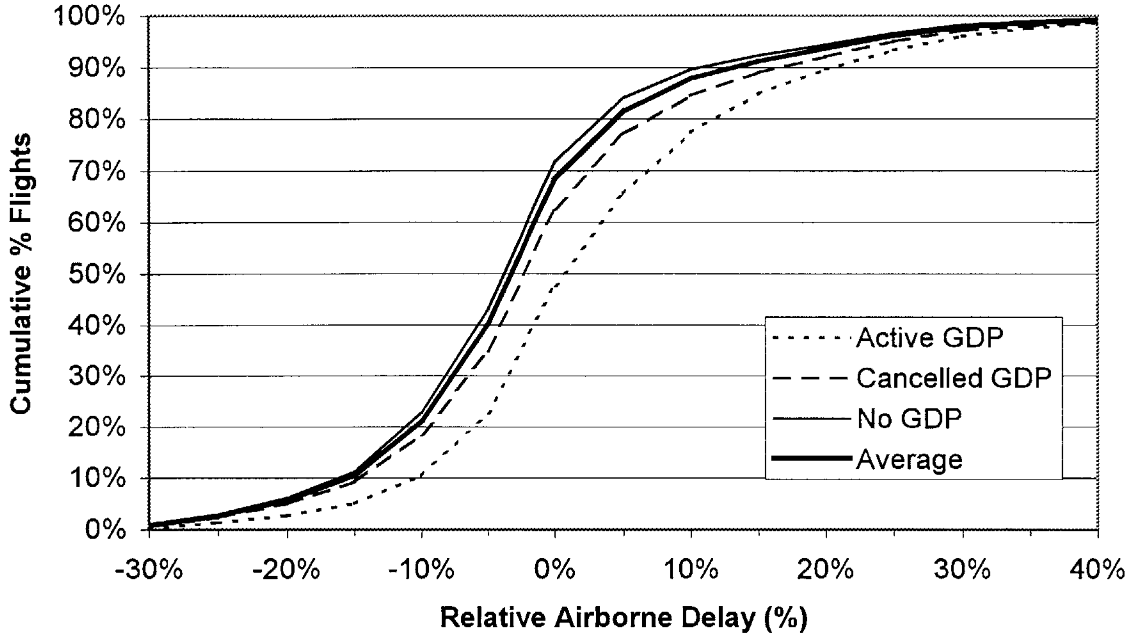

Figure 2-33: Cumulative arrival percent by relative airborne delay, by GDP state

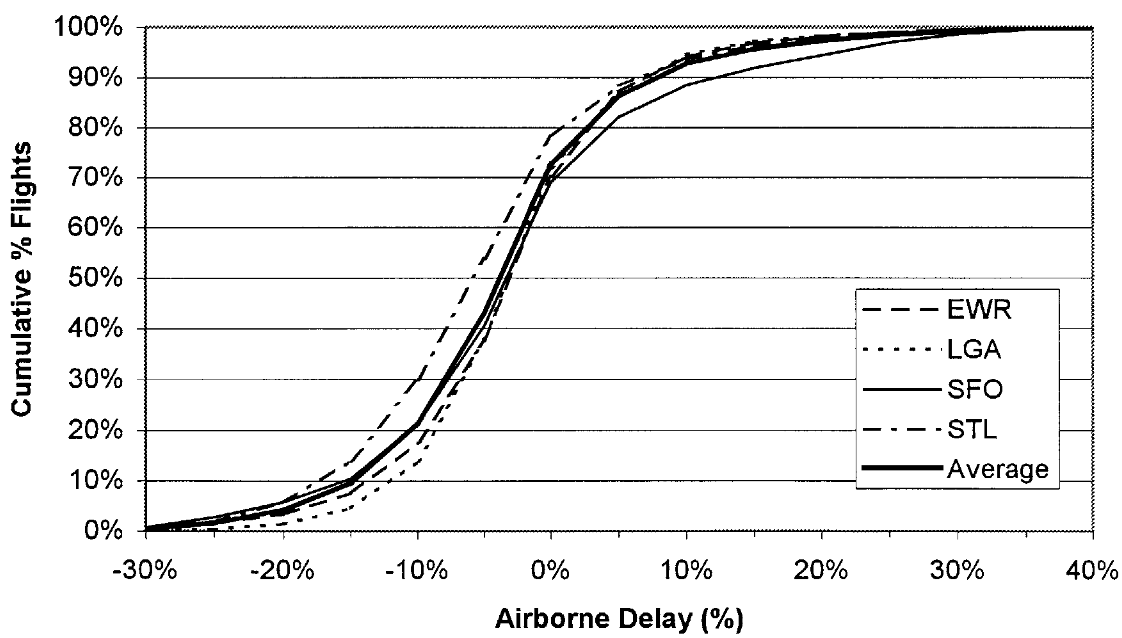

Figure 2-34: Cumulative arrival percent by airborne delay, by destination

Table 2-21 shows the average airborne delay for the four airports being studied. Much of the information presented in the table has already been covered in the discussion of Table 2-16 and Table 2-17. The percentage figures given are for the average relative airborne delay, as defined earlier on in this section. Since these percentages are not weighted according to the length of the flights, some apparent inconsistencies show up. For example, the average airborne delay for EWR for all flights irrespective of GDP state is 0.6 minutes and -0.6% of scheduled flight time. This is caused by the combination of different number of flights with different scheduled airborne times experiencing different airborne delays.

Table 2-22 tests the three airborne delay distributions shown in Figure 2-33 and Table 2-21 on a pair-wise basis for identical means, using a two-tailed t-test. *Act-Cnx* shows the results of comparing active-GDP and cancelled-GDP flights for identical means. Similarly, *Cnx-No* compares cancelled-GDP to no-GDP flights, and *No-Act* compares no-GDP to active-GDP flights. The comparisons are done on the composite level that includes all flights into each of the four airports (as plotted in Figure 2-33) as well as for each airport individually (not plotted).

	Act-Cnx	Cnx-No	No-Act
EWR	0.00%	0.00%	0.00%
LGA	71.27%	0.00%	0.00%
SFO	0.00%	0.00%	0.00%
STL	0.00%	0.00%	0.00%
All	0.00%	0.00%	0.00%

Table 2-22: Airborne delay statistical comparisons, by GDP state

The tests show, with one exception, that the mean airborne delays associated with the three GDP state are significantly different from one another, for each airport as well as overall. The one exception occurs at LGA, where the means associated with active-GDP flights and cancelled-GDP flights are not statistically different.

Table 2-23 tests the four mean airborne delay shown in Figure 2-34 and Table 2-21 against the overall average airborne delay, using a two-tailed t-test. The tests are done on the composite level for all flights (*All*, as plotted in Figure 2-34), as well as by GDP flight state (not plotted). For instance, the SFO mean airborne delay is tested against the average airborne delay to determine if they are statistically different. This can be confirmed at all statistical test levels when looking at all flights as a whole (i.e., irrespective of GDP state), as well as for no-GDP flights. It cannot be confirmed for active-GDP flights and cancelled-GDP flights, since the tests associated with these GDP states have significance levels of 15.17% and 35.26% respectively.

	EWR	LGA	SFO	STL
Active	0.72%	0.00%	15.17%	0.00%
Cancelled	48.46%	30.52%	45.26%	0.00%
No	0.00%	0.00%	0.00%	0.00%
All	0.00%	0.00%	0.00%	0.00%

Table 2-23: Airborne delay statistical comparisons, by destination

The tests show a variety of results. First, on the composite level, when the flights' GDP state is ignored, each individual airport's mean airborne delay is significantly different from the average airborne mean, at all statistical test levels. Second, for no-GDP flights, each airport's mean airborne delay is significantly different from the average airborne mean at all statistical test levels. Third, for active-GDP flights, mean airborne delay for EWR, LGA and STL are significantly different from the average airborne mean at a statistical test level of 1.0% or stronger. Fourth, for cancelled-GDP flights, only STL has a mean airborne delay that is statistically different from the average.

2.7 Cancellations

GDP programs increase the ground delay of the affected flights. Intuitively, one would expect to find a correlation between the length of the ground delay and the likelihood that a flight will be cancelled. However, this relationship works in both directions. The more flights that are cancelled, the better the opportunity for the remaining flights to recover some of the delay time imposed on them. This section addresses these two issues.

2.7.1 Methodology

An issue to be dealt with is how to classify the cancellations themselves vis-à-vis the GDP programs. There is no direct information in any of the basic data available that enables an unequivocal assignment of some cancellations to GDP related causes, and others to non-GDP related causes. All that can be observed is the cancellations themselves on a flight by flight basis, and the incidence of GDPs on a daily basis.

To overcome this problem, the days studied have been divided into GDP days and non-GDP days, and the data aggregated on a per-airport basis. This provides an initial basis for comparing GDP days with non-GDP days, and for analyzing the effects across all GDP days. An example of the former type of analysis is to determine if the average number of cancellations increase on GDP days. An example of the latter is to correlate the average ground delay per flight to the number of cancellations for that day. Both are described in the following two sections.

2.7.2 *Monthly overview*

Table 2-24 lists the average daily cancellations for the four airports being studied. In this table, *non-GDP* refers to the days without GDP programs, and *GDP* to days with a GDP program. For comparison, the overall daily average number of flights scheduled (*Avg Flts*) and the cancellation percentages of these flights (*Cnx%*) are also included in the table.

Month	EWR non-GDP	EWR GDP	LGA non-GDP	LGA GDP	SFO non-GDP	SFO GDP	STL non-GDP	STL GDP
Jan	131	129	101	-	152	102	151	-
Feb	91	141	63	-	77	127	114	-
Mar	88	128	63	-	61	88	104	-
Apr	96	201	72	n/a	76	104	100	115
May	114	136	76	209	98	126	103	163
Jun	174	200	133	105	117	146	160	177
Jul	113	116	78	n/a	100	116	92	249
Aug	106	114	72	142	77	113	82	n/a
Sep	109	n/a	68	152	74	109	87	97
Avg Cnx	113	149	80	148	94	117	110	161
Avg Flts	694	708	532	573	634	642	746	775
Cnx %	16%	21%	15%	26%	15%	18%	15%	21%

Table 2-24: Average daily cancellations for GDP and non-GDP days

The increase seen in cancelled flights per day is 36 at EWR (31%), 68 at LGA (85%), 23 at SFO (24%) and 51 at STL (46%). This suggests that there is a substantial increase in the number of cancellations on days with a GDP program. This hypothesis can be tested statistically using a one-tailed t-test. The result from this test is given in Table 2-25.

EWR	LGA	SFO	STL	All
0.20%	0.04%	0.14%	3.99%	0.00%

Table 2-25: Testing increases in cancelletion on days with GDPs

Based on the results in Table 2-25, all four airports show a statistically significant increase at the 5% level in the number of cancellations on GDP days. In fact, the results are significant at the 0.25% level for EWR, LGA and SFO.

2.7.3 Correlating cancellations to FAA arrival slot allocations

Another way of evaluating the effect that the GDP programs have on cancellations is to compare the average FAA arrival slot allocation ground delay on a given day with the number of cancellations. The expectation would be that the higher the average delay is, the larger the number of cancellations.

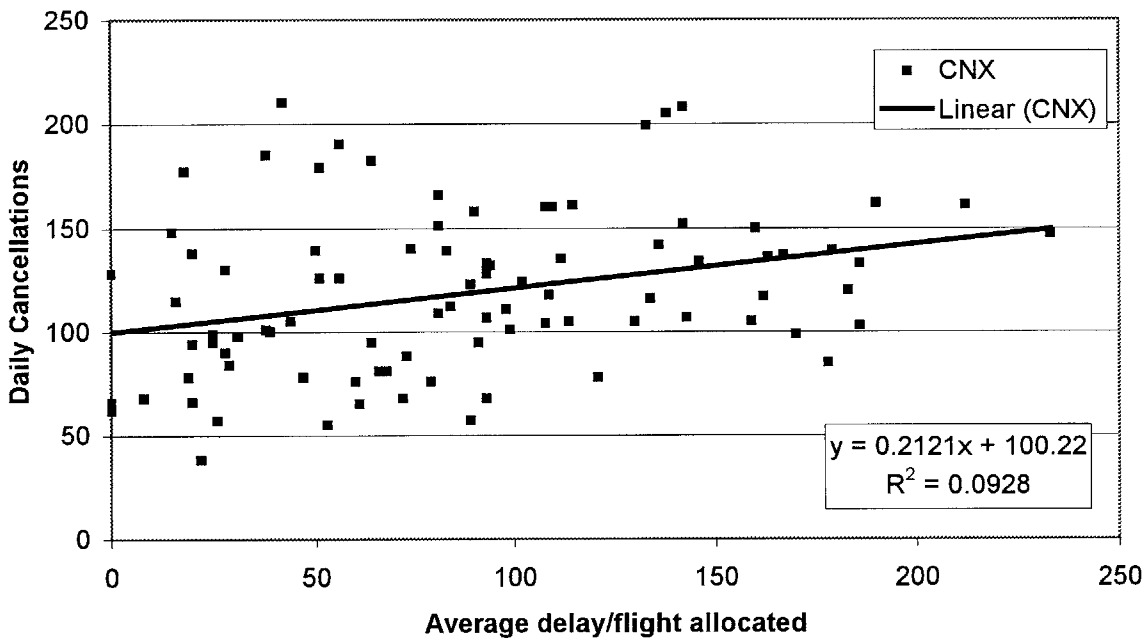

Figure 2-35: Correlating number of cancellations to average GDP delay

Figure 2-35 shows the results of conducting this analysis on the SFO data, using only the active-GDP data to calculate the average GDP delay imposed. The trend is as expected, but it is clear from both visual inspection of the chart and from the low correlation coefficient R^2 that the relationship, if any, is a weak one. Similar types of analysis have been performed using the average number of FAA arrival slot allocation changes and using the average net overall delay, but neither yielded any substantially stronger results than using the average FAA arrival slot allocation ground delay.

Using the regression formula indicated on the chart, one can estimate that a lowering of 30 minutes of the average ground delay incurred by incoming flights to SFO would have the effect of avoiding 6 cancellations. As seen in Table 2-24, there is an increase in the average number of cancellations at SFO from 94 on non-GDP days to 117 on GDP days, for a net difference of 23 cancellations. Thus, avoiding 6 cancellations would represent a 26% improvement in the number of GDP-related cancellations, but only a 6.3% improvement when compared to the average non-GDP number of cancellations.

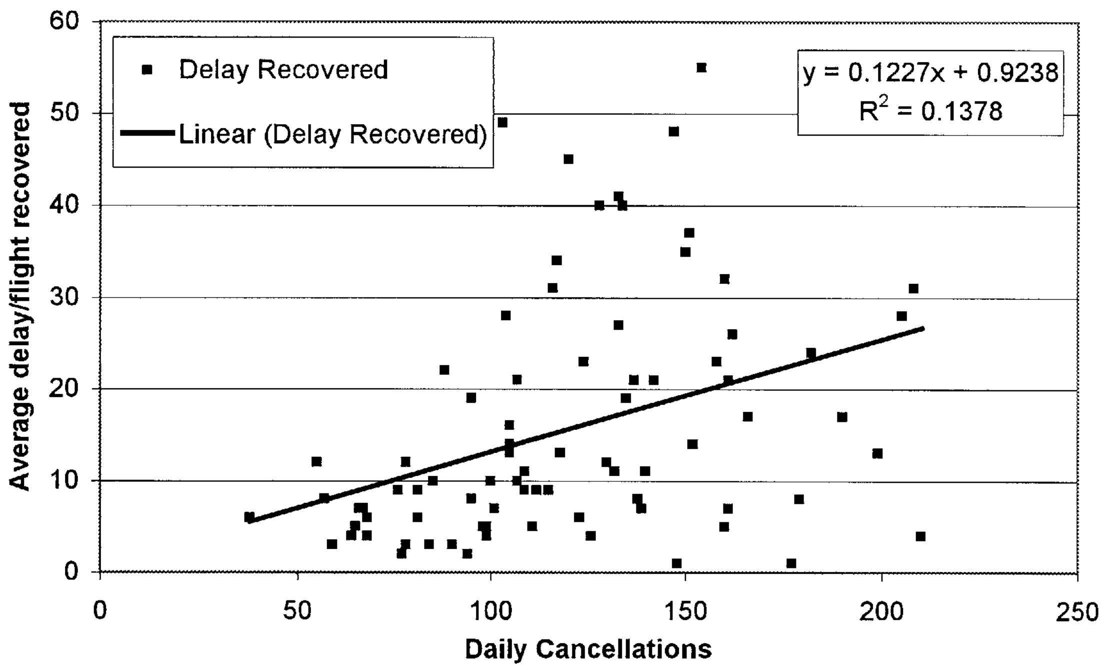

Figure 2-36: Correlating cancellations to average GDP savings from subs and compression

The more flights that are cancelled, the greater the opportunity to recover some of the delay imposed by the FAA arrival slot allocation process through substitutions and GDP compression on the remaining flights. Figure 2-36 shows the results of analysing this supposition, correlating the number of cancellations on GDP days to the average recovered delay per flight.

One possible interpretation of Figure 2-36 is to note that there appears to be a trade-off between delay savings and cancellations. Larger average delay savings comes at the expense of a higher overall cancellation total, and vice versa. The relationship, however, is again a very weak one.

The next chapter summarizes the principal conclusions from the analysis of the GDP-E prototype operation. It furthermore offers some suggestions for future research activities that can be undertaken to increase the overall understanding of the GDP programs, as well as to gauge the specific improvements that the introduction of CDM has brought about.

3 Conclusions and Suggestions for Further Analysis and Research

3.1 Conclusions

This section summarizes the principle results of this study. The results are organized using the same section headings as in the previous chapter for ease of cross-referencing.

3.1.1 Ground delay programs

SFO has a very high incidence of GDPs, with an average of 15.2 programs per month during the first nine months of 1998. This may have been caused in part by the severe weather conditions experienced in the early months of 1998 due to El Niño. The SFO GDPs exhibit a clear and consistent pattern, with a pronounced peak in the morning hours between 9:00 and 12:00. This pattern shows little seasonal variation, being repeated in each of the first three quarters of 1998.

The SFO airport activity level during a GDP remains almost constant throughout the day, which is due to the near-uniform use of an SFO AAR of 30 flights per hour. The average GDP activity level persists beyond 22:00, instead of tailing off significantly after 20:00, as is the case during non-GDP days. This can be attributed to accommodating flights that have been delayed by the GDP programs.

EWR has a much lower incidence of GDPs than SFO. It averages 4.6 programs per month. The data suggest a possible seasonal variation with GDP activity being much higher during the first two quarters. The hourly analysis indicates a peak in the incidence of GDPs between 16:00 and 18:00, but only for the first two quarters.

3.1.2 Substitution and Compression Analysis

Compression contributes positively to the overall reduction of delays imposed by the GDPs. The average savings achieved during the first nine months range from 3.4% at LGA, 5.2% at SFO, 6.4% at EWR and 7.85% at STL. The minutes saved per GDP flight are 2.2 at LGA, 6.1 at EWR, 6.5 at SFO and 10.0 at STL.

The savings due to compression seems to have increased over time at SFO, both in absolute terms and when measured relative to the contributions made by airline substitutions. During the third quarter of 1998, savings resulting from compression slightly exceeded those from airline substitutions, at 10.4 minutes versus 9.6 minutes per GDP flight. This stems entirely from improved compression results since airline substitutions have remained steady throughout.

The savings due to airline substitutions remained steady at SFO on a month-to-month basis, averaging 11.2 minutes per flight over the nine months studied. Since the average delay imposed by the FAA arrival slot allocation process seems to have decreased during the time period studied, savings due to substitutions actually increased in percentage terms. Due to data being available for just nine months, it has not been possible to determine if the observed decrease in the delay handed out by the FAA arrival slot allocation process is due to seasonal effects, or if it part of a more long-term improvement trend.

At SFO, the savings due to the combined effect of GDP compression and substitutions are substantially higher in the afternoon than in the morning. The two hours with the largest cumulative delays are 9:00-10:00 and 18:00-19:00 (see Figure 2-9). Compression plus substitutions reduces the delay in the former hour by 13.4% and in the latter by 27.5%, with compression making up 64.6% and 32.8% respectively of the total savings. The reason for the former value being so much larger than the latter is that substitutions generally do not contribute as much to the savings as compressions until after 12:00.

At EWR, the savings due to the combined effect of compression and substitutions follow the pattern of delays. During the hour associated with the largest delay, 18:00 to 19:00, GDP compression plus substitutions reduce the delay by 19.3%. Of this 19.3%, 7.8% can be attributed to compression savings, or about 40.5% of the total savings for that hour.

3.1.3 Arrival slot utilization

There is no obvious systematic under-utilization of capacity at SFO, based on the average difference between AAR and the actual number of arrivals at the airport. Although the average utilization may seem to indicate the possibility of one extra slot throughout most of the day, there are large stochastic fluctuations in the actual arrivals from day to day. Increasing the stated capacity (AAR) by one slot would increase the number of hours that experience unexpected airborne holding, as well as the length of time of the unexpected airborne holding itself.

The early morning situation at SFO should be carefully monitored, though. The three incidences of GDPs during the hour between 8:00 and 9:00 have shown an average under-utilization of 6.3 slots, but the limited number of events precludes any more definitive statements about this observation. The hour between 9:00 and 10:00 saw a total of 76 active-GDP days and showed an average under-utilization of 2.8 slots. On 51 of the 76 days, at least one arrival slot went unused. The under-utilization does not appear to be caused by a lack of demand, since there are an average of 36 scheduled arrivals during 8:00-9:00 and 46 during 9:00-10:00, and the AAR at SFO is almost always set at 30.

There are preliminary indications that slots are recovered rapidly at SFO when GDPs are cancelled. Due to the nature of the analysis performed so far, this is only a tentative observation. Increases of utilization have been measured to average 6 additional flights above the (cancelled) stated AAR capacity between 10:00 and 13:00, and 5 additional flights between 18:00 and 21:00.

3.1.4 Unexpected airborne holding

Unexpected airborne holding increases by an average of 2 minutes during GDPs at both SFO and EWR. This is a statistically significant increase given the large number of readings available. However, these additional minutes may serve to protect against under-utilizing scarce arrival capacity during GDPs.

STL showed particularly high unexpected airborne holding during active-GDP days in April (16.3 minutes) and May (8.3 minutes). These high UABH values appear to have been caused by GDPs not being instituted quickly enough or being cancelled prematurely, since the UABH during the no-GDP hours leading up to the start of the GDPs showed similar high UABH. In the two cases specifically analyzed, UABH was reduced dramatically within 1 and 2 hours after the start of the respective GDPs.

3.1.5 Airborne delay

The average airborne delay across all four airports and all nine months studied is 0.4 minutes per flight. The average airborne delay drops to -0.1 minutes when focusing exclusively on non-GDP flights. Similarly, active-GDP and cancelled-GDP flights have average airborne delays of 8.3 and 3.8 minutes respectively, representing an increase over the non-GDP stage by

8.4 and 3.9 minutes respectively. The per-GDP averages are all statistically significantly different from the overall average.

The highest average airborne delay is found at SFO with 2.5 minutes, followed by EWR at 0.6 minutes, LGA at 0.3 minutes and STL at -1.4 minutes. The high SFO average is in large part due to three bad months from January to March, when the average airborne delay measured 10.1 minutes, 9.7 minutes and 6.6 minutes respectively. Airborne delays at SFO were in line with other airports in the 2^{nd} and 3^{rd} quarters of 1998.

SFO shows great variation in airborne delays between no-GDP flights and active-GDP flights. The airborne delay is 2.5 minutes on average, constituted of 1.4 minutes for no-GDP flights, 4.1 minutes for cancelled-GDP flights and 9.0 minutes for active-GDP flights. The increase from non-GDP flights to cancelled-GDP flights and active-GDP flights are 2.7 minutes and 7.6 minutes respectively. These increases are statistically significant at all test levels.

EWR displays a very similar picture to that found at SFO, albeit from a lower base level. The airborne delay is 0.6 minutes on average, constituted of 0.2 minutes for no-GDP flights, 4.0 minutes for cancelled-GDP flights and 7.8 minutes for active-GDP flights. The increase from non-GDP flights to cancelled-GDP flights and active-GDP flights are 2.8 minutes and 7.6 minutes respectively, almost identical to the increases measured at SFO. The increases are statistically significant at all test levels.

The airborne delay was also measured relative to the scheduled flight time. SFO again shows the biggest average delay at 0.4% of scheduled time, followed by LGA at 0.0%, EWR at -0.6% and STL at -2.8%. Distributing the data across the three GDP states (no-GDP, cancelled-GDP and active-GDP), one finds averages for SFO of -0.4%, 2.0% and 5.4%, and for EWR of -0.8%, 2.1% and 4.7% respectively.

A detailed investigation focused on flights arriving at SFO from six different origin airports (LAX, SEA, DEN, DFW, ORD and EWR), chosen to provide a sample of different travelling distances. At one extreme, flights from DEN arrived with an average airborne delay of -3.7 minutes or -3.0% of scheduled flight time. At the other extreme, flights from EWR arrived with an average airborne delay of 9.9 minutes and 2.8% of scheduled flight time. The other four airports were bracketed within these values. SEA-SFO flights were found to arrive within the scheduled airborne time 62.8% of the time for no-GDP flights, dropping to 53.7% for cancelled-GDP flights and 34.2% for active-GDP flights. Similarly, ORD-SFO flights arrived within the scheduled airborne time 53.9% of the time for no-GDP flights, dropping to 49.6% for cancelled-GDP flights and 33.6% for active-GDP flights.

It is not possible to make any inferences about a possible relationship between the significant active-GDP airborne delay found and the delay savings achieved from GDP compressions. Doing so would require additional data to determine if the airborne delay measured in this study is a recent occurrence, or if it predates the introduction of GDP-E.

3.1.6 Cancellations

Flight cancellations increase on GDP days. The increase ranges in percentage terms from 24% at SFO to 85% at LGA. The increases are all statistically significant at the 5% level. The relationship between the number of cancellations and a host of factors related to the GDPs (such as number of FAA arrival slot allocation changes, average GDP delay imposed and net GDP delay) were investigated. All showed the expected positive correlation, but these correlations are weak.

3.2 Suggestions for further analysis and research

This study has investigated several different aspects of the GDPs, but many more could have been pursued. Several of these possibilities have already been mentioned briefly in the previous chapter. These possibilities, along with several additional ones, are discussed in this section.

3.2.1 General

The topic of seasonal variations can be more fully explored when more data become available. The current analysis has been performed with nine months' worth of data. A full year's data would give a better picture of the seasonal variations. Year-to-year trend analysis will require data for multiple years. This should be pursued in the future. It will allow us to investigate whether some observed trends (such as the drop in the number of incidences of GDP at SFO in the 2^{nd} and 3^{rd} quarters of 1998, or the low number of GDP incidences at EWR in the 3^{rd} quarter) are seasonal or represent a more long-term change.

The current detailed analysis has been focused primarily on SFO, due to the availability of data and the high number of GDP events at this airport. Similar detailed analyses can be applied to other airports as well, when the data become available. This will enable a better understanding of the impacts of the GDP programs on a national basis.

3.2.2 *Ground delay programs*

A day-of-week analysis may further enhance the analysis of the GDPs. Since the scheduled number of arrivals vary by day of week at many of the airports studied, capacity restrictions will have different effects on different days. One might expect days with heavy traffic to be both more frequently and more severely impacted than days with low traffic, since the system will have less spare capacity with which to accommodate disturbances. This may have implications for the choice of the AARs. A slightly higher AAR could perhaps be used on low-volume days, since it would be easier to recover from any mistakes due to the smaller amount of traffic.

3.2.3 *Substitution and Compression Analysis*

The substitution and compression analysis conducted so far has made few attempts to distinguish between different classes of GDP flights. The only two characteristics that have been used are the arrival airport and the arrival time, with the latter being further sub-divided into months or arrival time of day. However, every flight is characterized by many other attributes that may influence the choice of whether or not that flight is moved up or down in the sequence of arrivals through the substitution process. Some attributes that may have a direct bearing on this question are (1) the origin of the flight, (2) the frequency of the schedule between the origin and destination, (3) the length of the flight, and (4) the type or class of equipment used.. These can be explored further by performing similar types of analysis as the ones already conducted, but using one or more of the criteria given as a basis of stratification.

Once a GDP is cancelled, flights previously subjected to the GDP are cleared to depart. How quickly they can do so has a direct impact on their ability to recover some of the delay allocated. This recovery of a portion of the previously allocated delay has not been measured in this study, yet it may be important to the airlines, since it helps them reduce the net amount of delay for a given flight. With the data available, it would be possible to determine how large the actual recovery is, both in absolute and relative terms.

Most importantly, the full impact of CDM on delays has not been fully measured by this study. An analysis of the situation prior to the introduction of CDM has not been possible, due to the lack of data on earlier (pre-1998) periods. This makes it impossible to make any before-and-after comparisons. The situation may be remedied, if reliable data covering the period prior to the introduction of CDM can be made available.

3.2.4 Arrival slot utilization

As mentioned in section 2.4, it is possible to obtain a better picture of the rate of capacity recovery after a GDP cancellation by basing the analysis on the time elapsing from GDP cancellation to flight arrival. The analysis can then be extended beyond the end of the original GDP program. Since the hours beyond the end of the original GDP may also experience the after-effects of the GDPs, it is relevant to look at a longer span of time. Our analysis did not do this, since it used data only from the cancelled-GDP hours. In order to conduct this new type of analysis, it would be necessary to measure the flight arrival times relative to the time of the cancellation, and adjust for the availability of demand to fill the available capacity, in much the same way that the adjusted AAR (AAR') was used in section 2.4. However, a specific data-related issue needs to be addressed before this can be pursued further: Currently there is no information available about airport capacity or runway configuration in use for non-GDP hours. Without this information it is not possible to determine how well the capacity is being utilized, only that an increase in the number of arrivals above and beyond the AARs imposed by the GDPs can be seen.

The cancellation of a GDP and the resulting removal of restrictions on an airport's capacity is only the most extreme example of a change in the AAR. Other changes take place as part of GDP revisions and extensions. One interesting topic would be to determine whether or not a decrease in an AAR leads to additional unexpected airborne holding, and similarly, if an increase in an AAR leads to greater throughput of the system. These topics clearly are a function of the lead-time of the change, since any change in the AARs will take some time to be negated or utilized. It would be useful to gain an understanding of how late a change can be made and still have the desired effect.

The AAR set for a given hour and airport determines how many flights will be allocated to arrive at that airport for that hour. The AAR is set by the FAA according to the local conditions prevailing at the airport, such as inclement weather or airport construction work. It is not, however, influenced by the scheduled number of arrivals and departures for that hour, implicitly making the assumption that the ratio of arrivals and departures for that hour and its neighboring hours are reasonably similar. This is definitely not so at major U.S. hub airports such as DFW, ATL and PIT. At these airports, there are distinct banks of flights arriving during one hour, followed by a bank of departing flights an hour or so later. In such cases, the use of a constant AAR for successive hours may not be appropriate, since it has the effect of spreading flight arrivals out over a longer period of time. This may force the airlines to have flights depart without waiting for connecting passengers, or alternatively, to keep departing flights on the ground for a much longer period of time to accommodate the connecting passengers. This suggests two topics of investigation. First, a detailed analysis should be performed on one or more major hub airports to determine, if the available total capacity for arrivals and for departures is indeed being fully utilized during hours of restricted operations. The type of detailed capacity utilization analysis performed on SFO in this study may be appropriate here. Second, a detailed analysis should be performed to evaluate the effects of changing from a capacity control method based on limiting the number of *arrivals* to one based on limiting the number of *operations*. This latter scheme would allow the airlines to better tailor their allocations to suit the structure of their arrivals and departures than the existing system. It would enable them to use all or most of their allocated slots within a given hour as arrival slots, and then as departure slots during the following hour. Furthermore, this scheme would be very much in the spirit of CDM, since it would vest the operational choice with the airlines, rather than with the FAA.

The method we have used to determine if capacity is being fully utilized is based on the difference between the number of desired (under the GDP) and actual arrivals for a given hour during a GDP. This may sometimes be too sensitive a measure. A shift of one minute in actual arrival time may move a flight from one hour to an adjoining hour, possibly resulting in reporting one hour as being over-utilized and the following hour as under-utilized. This problem may be avoided by applying data-smoothing techniques to the method of analysis. The net result would be a more accurate determination of the variability of the hourly capacity utilization.

Finally, any difference found between desired and actual arrivals (i.e., how well the capacity is utilized) may be due to two very different reasons. First, it may reflect the ability of adequate demand by the airlines to make use of the available capacity. Second, it may give an indication of how well the FAA is able to set an accurate value for the AAR. An indication of which of these two factors is the governing one may possibly be acquired by utilizing the unexpected airborne holding in conjunction with the arrival slot utilization measure. If the system is being driven close to its actual capacity, one would expect to see UABH increase. This would indicate that demand is indeed available to fully utilize the AAR. Performing this type of study would help resolve if any observed under-utilization is caused by lack of demand, or if the AAR is being set unattainably high for the prevailing weather conditions.

3.2.5 Unexpected airborne holding

An issue that has not been addressed yet is the variability of unexpected airborne holding from hour to hour. One might hope that it remains quite steady, meaning that a given hour with extensive UABH is followed by another hour exhibiting similar characteristics. If this is the case, then it may be possible to lower the UABH. This may be done by either lowering the initial AAR set, or by lowering the AAR some number of hours into the program. However, this is related with the need to maintain an adequate MAR available to ensure full utilization of the scarce capacity.

Figure 2-20 shows that SFO UABH increases substantially during the late evening hours. This observation has been made for active-GDP, cancelled-GDP and non-GDP hours, and is rather unexpected as far as the non-GDP hours are concerned. Since the scheduled number of arrivals (see Figure 2-3) during that period of the day is less than the stated AAR capacity normally used (30 flights per hour), the airport should have adequate capacity to accommodate all arrivals expeditiously. (Using the normal stated AAR as a proxy for the capacity is reasonable in this case, since all that is needed is a *lower* bound for the capacity.) However, the observed situation indicates that substantial and persistent delays are incurred during the late hours. A detailed analysis could be performed to determine if this is a local SFO phenomena, or if it is part of a more general case valid for other airports as well.

3.2.6 Unexpected airborne holding and airborne delay

The current analysis methodology used determines scheduled flight airborne time by using the originally estimated times of departure and arrival. This information is typically sent to the FAA several days in advance of the actual departure of the flights. As such, the current scheduled airborne time cannot possibly take into account any day-to-day issues. This may be remedied by using information that is set closer in time to the actual departure date and time, such as by using the estimated time of arrival available at the day of departure, or possibly at the actual time of departure. Doing so would allow specific, day-to-day conditions to be factored out of the scheduled flight time. The analysis could then be extended to use both methods of calculating the scheduled flight time, enabling two additional types of comparisons. First, how much do the estimated flight times change as the time of departure approaches. Second, how accurate are either of the estimated flight times when compared with the actual flight time.

Very little use has been made in this study of some of the information stored in the database. Two such items are the airborne time and the type of equipment operating a given flight. It would be interesting to determine if all types of flights are equally prone to delays, or if some are delayed more often than others. This may vary from airport to airport, since the local configurations may favor one type of equipment over another, for instance due to the inability of jet aircraft to land on certain runways because of environmental or community considerations.

3.2.7 Cancellations

GDP-E has introduced an additional method (the FX message) for the airlines to notify the FAA of flight cancellations, and has also provided a simpler way of doing so (See [2]). As discussed in [1], the result was that *"flight cancellation notices were received under CDM, on average, at least 76 minutes earlier at EWR and at least 63 minutes earlier at SFO, than they would have been without CDM"*. This analysis can be extended with the current data to cover LGA and STL and, with the future inclusion of additional data, other airports as well.

3.2.8 Computational model

The GDP control information is currently made available for use in the database through a manual process. The information is read from a web page and transcribed into a file with a standardized, fixed-format record layout. This process is tedious, error-prone and time-consuming, and often is the controlling factor in how quickly data can be made available for use in the computational model. As of 7/22, the GDP control information became part of the ΔADL file content, and so is now directly available. It would be a straightforward matter to create an addition to the current "data capture" computer program that would also pick up the GDP control information, and would thus obviate the need for the manual transcription process.

Currently the slot allocation information found in the slot files is only used to determine how to distribute the delays and delay savings across the FAA arrival slot allocations, the airline substitutions and the compressions. It is possible to envision a more extensive role for this information. One example is its use in this study to determine if the times when the FAA arrival slot allocations were set or revised could be found automatically. This was answered in the affirmative.

Many of the reports generated by this study could quite easily be retrofitted to run on a web-site. Technically, the biggest issue to be addressed is that of performance, since many of the reports today take several minutes to produce, if all nine months worth of data are being analyzed. To a large extent, this can be addressed through the acquisition and use of a faster PC than the one currently being used. Also, using the capability of the database to distribute data across multiple disk-drives would result in access time improvements that scale nearly linearly with the number of independent disks in use. However, since the information displayed would be of a sensitive nature, this would involve addressing a host of institutional issues as well, to determine who should have access to the information, and if so, to what parts.

References

[1] Ball, M., R. Hoffman et al. 1998, "Collaborative Decision Making in Air Traffic Management: A preliminary Assessment", NEXTOR Project Report to FAA, August, 1998.

[2] Hoffman, R., W. Hall, M. Ball, A. Odoni, M. Wambsganss. 1999, *"Collaborative Decision Making in Air Traffic Flow Management"*, forthcoming.

[3] Howard, K. 1998, "ADL format changes", memo from Ken Howard, DOT Volpe NTSC, dated 6/30/98.

[4] Williams, B. 1998, *"The 51 fields in a CDM-string analysis file flight record"*, memo from Brian Williams, Metron, Inc, dated February 3, 1998.

APPENDICES

A Acronyms

AAR - Arrival Acceptance Rate
AAR' - Adjusted Arrival Acceptance Rate
ABH - Airborne Holding
ADL - Aggregate Demand List (file)
ΔADL - Delta Aggregate Demand List (file)
CDM - Cooperative Decision Making
CTA - Controlled Time of Arrival
CTD - Controlled Time of Departure
DEN - Denver International Airport
DFW - Dallas/Fort Worth International Airport
ETA - Estimated Time of Arrival
ETD - Estimated Time of Departure
EWR - Newark International Airport
FAA - Federal Aviation Administration
GA factor - General Aviation Factor
GDP - Ground Delay Program
GDP-E - Enhanced Group Delay Program
LAX - Los Angeles International Airport
LGA - LaGuardia Airport
MAR - Managed Arrival Reservoir
NAS - National Airspace System
OABD - Overall Airborne Delay
OETA - Originally Estimated Time of Arrival
ORD - O'Hare International Airport (Chicago)
RBS - Ration by Schedule
SEA - Seattle-Tacoma Airport
SFO - San Francisco International Airport
STL - St. Louis International Airport
UABH - Unexpected Airborne Holding

B Computational Analysis Model

The Air Traffic Control system collects and processes a large amount of data on a daily basis. This information provides the underpinnings for the effective operational management of the U.S. airspace on a day-to-day basis. It is, however, structured conveniently to serve as a vehicle for longitudinal and cross-sectional analyses. There are two primary reasons for this:

1. The sheer volume of data is huge. Being able to store and process the complete set of data would require computer resources that are beyond the reach of this study.

2. Information of interest is scattered throughout. The core elements of the information needed to answer the questions posed in this study are found at the composite flight record level. However, the flight information must be accompanied by information about the GDPs to enable analysis of the effects of CDM.

The remainder of this chapter deals with two specific questions: first, how the data is transformed to form a suitable basis for the analysis of the information; and second, once the data have been converted to such a form, how is the analysis conducted.

B.1 Overview and Environment

B.1.1 Data flow

This section provides a high-level overview of the flow of data at it is processed in this study. This flow of data is illustrated in Figure B-1. There are three paths in the figure, all converging on the central database. The upper-left-hand corner deals with the capture of the flight information, and the lower-left-hand corner similarly deals with the capture of GDP control information. The right hand side represents the analysis component, where extracted data is fed into standard spreadsheets for final analysis.

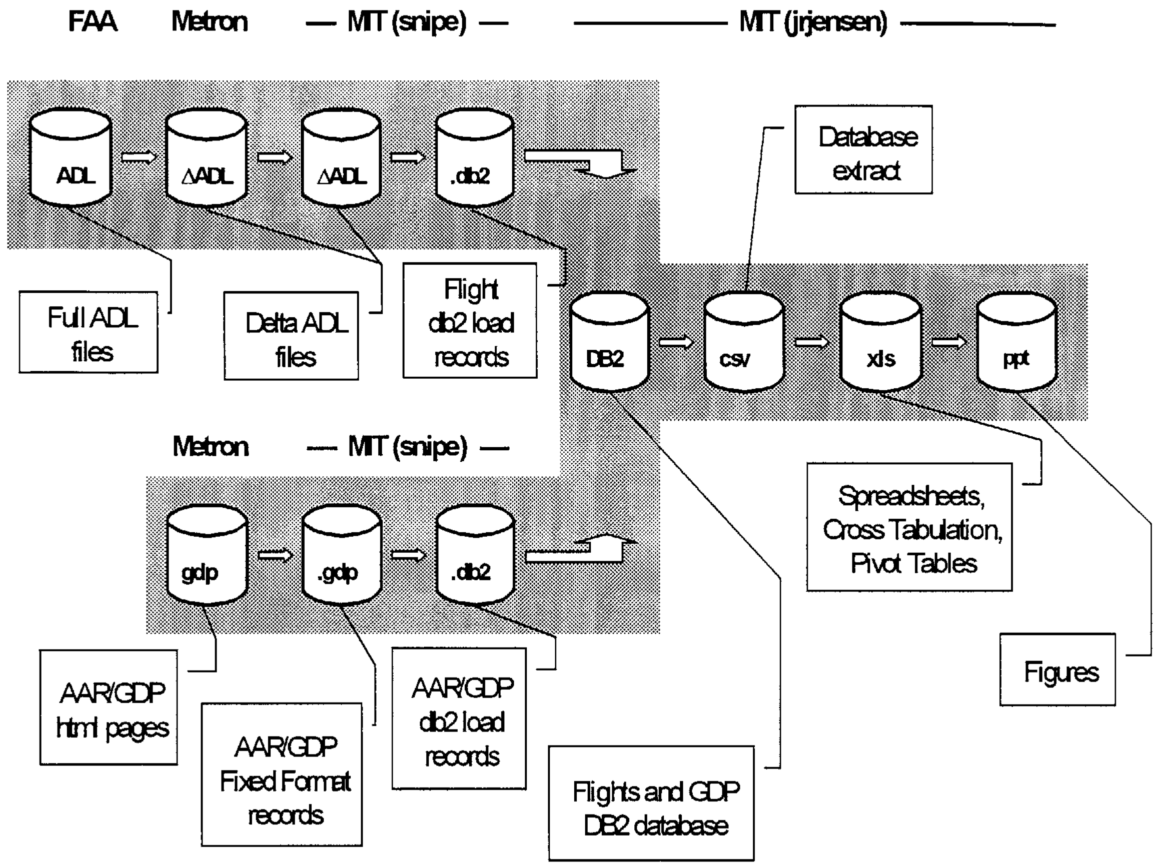

Figure B-1: Computational Analysis model

B.1.2 Collecting flight information

Information about every flight in the system is captured by ETMS every 5 minutes and saved in the *ADL files*. A record is generated for each flight, and each of those records contains 55 distinct types of information. In addition to this, information about the Arrival Acceptance Rate (AAR) and other Ground Delay Program (GDP) parameters are also captured within this file.

The ADL files contains a tremendous amount of redundant information, since the flight records are generated every 5 minutes whether or not they have changed or not. The *ΔADL files* reduce this redundancy by only retaining those records where at least one field changes from 5-minute time period to 5-minute time period.

Based on the information in the ΔADL files, two separate data-files are created. The first of these is the *slot file*, containing information about which flight is assigned to which slot, or equivalently, the series of slot allocations to a given flight during its presence in the system. Each change is tagged with information about the reasons for the change, enabling later summary information to be amassed. The second data-file created at this stage is the *flights file*, which contains one summary record of information for each flight in the system.

The flights file contains all the information needed for the database. To make it useable for the database it is reformatted into a *flights.db2* file, which can be bulk-loaded into the database, assuring the fastest processing possible of the database load.

B.1.3 Collecting GDP and AAR information

The lower-left side of the data-collection leg proceeds through an analogous series of steps, although the amount of data is much smaller. The initial source representation of the GDP information is found in the Metron html web-pages that are produced on a weekly basis detailing the GDP events of the past week.

The information from these web-pages is then transcribed manually into a daily *gdpevent* file. These files contain the same information content, but their internal syntax is better suited to the required follow-on automated processing.

Three separate db2 load files are created from the gdp-event file to feed into the database tables. First, the *aar.db2* file contains all the AAR information, so it is possible to reconstruct the AAR settings for any hour of the day. Secondly, the *gdp.db2* file contains summary information for each GDP, and *gdpevent.db2* contains a reformatted version of the gdp-event file above suitable for loading into the database.

B.1.4 Database

The database currently consists of five tables: *fights* holds the flights information; *gdp* holds the GDP summary information; *gdpevent* holds the GDP detail information; *aar* holds the Arrival Acceptance Rate information; and *airport* holds the airport time zone information. Of these five tables, the flights table is by far the biggest. Its organization is critical to achieving good performance from the database during the queries to it.

B.1.5 Analyzing the data

All analysis is done based on the content of the database. This typically involves three different steps. *Consolidated data* is retrieved from the database. The consolidation performed and the set of fields returned in the result varies on the type of analysis being performed. It always involves the use of the flights table, and may in addition involve several of the auxiliary tables as well.

Next, the consolidated data is imported into a Microsoft Excel *spreadsheet* to perform the actual analysis. The spreadsheet environment provides many useful functions that are not available with the database, such as cross-tabulating the data using the Excel pivot function.

Often the true nature of the underlying data is only revealed when the data is charted in a graph. This process can either be performed within the spreadsheet environment, or it can be done as part of a presentation environment. A combination of these two approaches has been used in this study.

B.1.6 System Environment

Four different computers are indicated on Figure B-1. The full ADL files are originally collected on an FAA system. Metron collects these files on a daily basis and produces the ΔADL files. This data is then shipped to MIT on an ad-hoc basis, but never more frequently that once a month. The initial processing takes place on snipe.mit.edu, a Linux-based system, and the data is prepared for the database load. The database itself is installed on a separate, Windows-based system, jjensen.mit.edu, since the database used, IBM's DB2, does not at the time of writing have a version available for Linux. This latter system also hosts the needed spreadsheet and presentation tools.

B.2 Flight Data Extraction

This section elaborates on the creation of the flight data in the database. Data is initially received in the form of a series of ΔADL files from Metron[6]. Each file contains data for one destination airport for one day. The data is transferred using a standard Unix ftp file transfer

[6] The data formats for the ΔADL files are given in appendix C.2.1, *ADL file format.*

from the Metron ftp site vivaldi.metsci.com. Upon receipt the files are placed in a directory structure where there is one directory per date.

An example of this directory structure is given below. Data-files for the four airports being analyzed are present in compressed form:

```
/data/adl
|
└───19980131
        ewr.adl.gz
        lga.adl.gz
        sfo.adl.gz
        stl.adl.gz
```

This directory structure has proven to be very flexible, in that it is easy to add additional control and data files as the need arises.

The data extraction process is conducted by three primary programs, *extract, savings* and *prepdb2.* These and all other programs used in this body of work are all written in Perl, a scripting language commonly used in Unix but also available for Windows.

As its name implies, *extract* extracts the salient pieces of information from the ΔADL files. For each ΔADL file, two files are generated as its output, the *flights* file[7] and the *slot* file[8].

1. The *flights* file contains a record of information about each flight found in the ΔADL file. The information collected can be grouped into three categories. The first category contains basic flight information such as date, flight identifier, source and destination, scheduled and actual departure and arrival information, arrival slot (if any) and cancellation status. The second category of data is the data related to cancellations, to enable the analysis of any advantages accrued from the introduction of the FX message. The third and final category of data is the data related to unexpected airborne holding.

2. The *slots* file contains information about the allocation of arrival slots to specific flights during the running of GDP programs. Each record in this file represents a slot allocation

[7] The data formats for the flights files are given in appendix C.2.3, *Flight and Flight.db2 file formats.*

[8] The data formats for the slot files are given in appendix C.2.2, *Slot file format.*

to a particular flight, and includes information about the flight identifier, the destination airport, the assigned slot and when it was assigned. A given flight may have multiple records associated with it, representing revisions of its arrival slot allocation over the course of the GDP. Likewise, a given slot may be assigned to several different flights over the course of the day, or at times to no flights at all.

These two files are used as input to the next step in the process, handled by the *savings* program. This program first determines the underlying cause for each arrival slot allocation change (FAA, airline substitution, compression, or other). It then tallies the result for each of the four categories on a per-flight basis, and finally updates the *flights* file with this information. The 'other' category was introduced to handle those rare instances where a slot allocation change cannot be attributed to any of the first three causes; it typically is used by less than three changes per day.

The final step in the process is to convert the information in the *flight* file, so it is syntactically amenable to the database load application. The *prepdb2* program does this, producing a database-ready load file named *flight.db2*. No new information is added at this step, but all timestamps are changed from the Metron format used in both the *flight* and *slot* files (day-hour-minute) to a pure minute offset from the starting date. Keeping the timestamps in the latter format makes any difference calculations much easier to perform later on. It also has the beneficial side-effect of taking up less space, since the offsets can be stored in a two-byte integer as opposed to a four-byte integer needed for the Metron format.

B.3 GDP Program Parameters

The GDP control information is used in several places within this body of work. It is used to generate the flights files above, and it is also stored into the database and used as part of some of the database reports, e.g., slot utilization.

The GDP data is retrieved from weekly listings on the Metron web site, http://www.metsci.com/cdm/members/gdp.html. These files contain information about all the GDP programs during the week. To make this information more useful it is transcribed manually into files named *events.gdp* and stored on a per-day basis as appropriate in the normal, daily directory.[9] The transcription produces a standardized record layout that is more suited to automated processing. A file contains information for all GDP programs run on a particular day; and so will often contain information from several different airports. It may also hold information from several different programs run on the same day for the same airport.

The events.gdp files are used as part of the savings calculation described in the previous section. They are also used by the *prepgdp* program. This program creates three separate database load-files based on the information found in the events.gdp file:

- *gdp.db2* contains a single record for each gdp program run on that day, with statistics on when it was initially created and when it was active, the number of compressions, revisions and extensions, and whether it ran to completion or was cancelled early.[10]

- *gdpevent.db2* contains the version of the event.gdp file that is suitable for loading into the database.[11] It holds the same information except that it omits the AAR and Gafactor fields.

- *aar.db2* contains the aar and gafactor information.[12] A given GDP hour may have several aar's associated with it, since these may change when the program is revised.

B.4 Database Processing

The database is at the nexus of the computational model. This section will deal with several important topics related to this database.

B.4.1 Database table setup

The database itself is defined using the *mkcdmdb* program. It is possible to create several different and independent versions of the database by specifying an instance name, but by default there is only one instance named cdm defined. This feature has not been used with the current work, but could be used to define a test version to go alongside a production version.

---

9 The data format for the flights files are given in appendix C.2.3, *Flight and Flight.db2 file formats.*

10 The data format for the gdp.db2 files are given in appendix C.2.5,.*Gdp.db2 file format*

11 The data format for the gdpevent.db2 files are given in appendix C.2.6, *Gdpevent.db2 file format.*

12 The data format for the aar.db2 files are given in appendix C.2.7, *Aar.db2 file format*

With the database defined it is then possible to create the individual tables within the database. Programs have been set up to do this for each of the five tables defined so far:

1. *Flights:*[13] This table contains the bulk of the actual data stored in the database. It consists of one record per known, registered flight over the nine months analysis period for the airports being studied. Each record contains twenty separate items, which can be divided into four separate categories:

   Common: Flight characteristics such as flight identifier, date, origin and destination airport, equipment and slot information.

   Ground Delay: Summary information related to the ground delays. This information will be used to judge the effects of compression.

   Unexpected airborne holding: Summary information related to unexpected airborne holding.

   Cancellations: Information about cancellations and the advantages that the FX message has introduced.

2. *AAR:*[14] The AAR table holds the AAR settings for the analyzed airports. When used in conjunction with the flight information, this can help answer the question of whether or not the available capacity is being fully utilized.

3. *GDP:*[15] The GDP table is used to assign flights to one of three states: (1) those that land under active slot control; (2) those that at some point were influenced by slot control; and (3) those that never were impacted by slot control.

4. *GDPevent:*[16] The GDPevent table is used as part of the compression benefits calculation.

5. *Airport:*[17] This table currently only provides the time zone offset, but other airport-specific characteristics could conceivably be added to this table later on.

---

13 The table definition for the flights table is given in appendix C.1.1 *Flights.*

14 The table definition for the aar table is given in appendix C.1.2 *AAR*

15The table definition for the gdp table is given in appendix C.1.3 *GDP*

In addition to these fixed tables, several of the extraction programs will create views (i.e., dynamic tables) as part of their processing.

The table name is reused as the name of the program creating the table and loading the data into the table, as well as for the database loadfiles containing the appropriate data, thus providing consistency across the application domain

B.4.2 Loading the database tables

The five database tables are each loaded with data stored in data loadfiles. With the large disparity in the size of the data to be loaded, two different strategies are used.

The first and simplest strategy is to load an unlimited number of days and/or airport data. This strategy works well for the aar, gdp, gdpevent and airport data, since there are relatively few rows of information to be loaded. The largest of these files is the gdpevent file, which contains information about 1000 or so separate events. This does not present a problem for the database, even on a relatively slow PC constrained by the amount of disk-space available.

The second strategy is used for loading the flight data. Flight data are much more voluminous than any of the other four types of data loaded, even after applying the extraction scheme described earlier. This can be illustrated by the number of rows of data to be loaded for 9/30, the last day of the sampling period. Table B-1 lists the number of flights for each airport that contributed data on that particular date:

Airport	Flights	Airport	Flights	Airport	Flights	Airport	Flights
ATL	1347	DTW	787	LGA	555	PIT	671
BOS	838	EWR	737	ORD	1403	SFO	664
DFW	1306	LAX	1178	PHL	723	STL	804

Table B-1: Scheduled number of flights at 12 airports for one day (9/30/98)

Two conflicting factors influence the number of rows that can be loaded in a single operation. First, the number of rows is limited by the temporary space available. Temporary space is used by the database for many different purposes, but one of the main uses is to be able to do a rollback, i.e., restore the database table to the state before the load operation commenced. The

[16]The table definition for the gdpevent table is given in appendix C.1.4 *GDPevent*

[17]The table definition for the airport table is given in appendix C.1.5 *Airport*

larger the number of rows to be loaded, the larger the temporary space needs to be. This argues for loading a limited number of rows at a time, e.g., one airport and one date.

Second, loading data into a database table requires rebuilding of the table indices. Every table has at least one set of indices associated with it, but some may have more. These indices speed up the retrieval time considerably, so judicious use of these is important. In the case of the flight table, a joint index has been defined on the date, destination and flight columns. This combination serves as a unique index. Since most reports are created over particular sets of airports, dates or both, this index is also useful during the query process. However, these indices must be rebuilt whenever data are added to a database, whether a single row or many thousands. Furthermore, the time to rebuild them is roughly proportional to the number of records in the table. All told, this argues for loading the data into the database table in a few, large chunks.

Taking these two factors into consideration, the flight data are loaded into the database in chunks that combine data from all airports for three days into a single load. If more dates are specified on a load request, then the data will automatically be divided into several independent loads. Using this chunking scheme has proven to be efficient and to run within the confines of the system resources at hand. The overall speed of loading the database was improved from in excess of one day to load about 5 months of data to loading the full nine months in less than two hours, using a 166MHz PC with 200 megabytes of disk-space allocated to the database.

The size reduction achieved can be illustrated by listing the size of the files and database tables used to hold the data from SFO on 9/30, a date where two GDPs were in effect at SFO. Table B-2 shows this comparison. *File* refers to the different types of files in use, *compressed* and *uncompressed* gives the size of the files in kilobytes (Kb) in compressed and uncompressed format respectively, and *% of ΔADL* compares the sizes to the ΔADL file. The size given for the database table also includes the space used by the auxiliary tables and database indices.

File	Compressed (Kb)	Uncompressed (Kb)	% of ΔADL
ΔADL	612	4355	100.0%
flights	20	118	2.7%
slots	13	101	2.3%
flights.db2	23	69	1.6%
Database table	n/a	106	2.4%

Table B-2: File size comparison

Table B-2 clearly indicates the results of the data reduction process. Overall, the entire database consumes 140 Mb of disk-space.

B.5 Querying the database

The analysis starts once the data is stored into the database. It consists of two or three separate steps, each performed using a separate set of tools.

The first step extracts data from the database tables, using DB2 SQL statements to define the data to be extracted. The result from this step can be thought of as a table itself, usually consisting of one or more category/group columns and several value columns. Examples of category fields are destination, month and hour. Examples of value fields are a count of arriving flights within a given time period, number of compressions run, or the average unexpected airborne holding.

The second step uses the results from the first step as its input. It is typically performed using a spreadsheet environment such as Microsoft Excel. It is generally used to take the tabular data generated in the first sub-phase and reformat it into a cross-tabulation report, i.e., a report with a category fields defining the column values, another defining the row values, and one of more cell values being calculated at the intersection of each column and row. Microsoft Excel's pivot table feature has been used extensively here.

Finally, the third step creates a visual representation of the results produced in the second sub-phase, in the belief that this helps detect or highlight trends that might otherwise go undetected. This sub-phase could be integrated with the previous one -- and sometimes is -- but it is generally kept distinct from it by using Microsoft Powerpoint as the presentation tool of choice.

The remainder of this section will focus on four specific computer programs that falls in the domain of the first sub-phase. The second and third sub-phase will be covered in the section B.6, *Finalizing the results.*

B.5.1 Compression Benefits

Compression benefits are calculated by the *comprep* program. Table B-3 shows and example of a call to this program and a partial listing of the result it produces. The program groups the result by one grouping fields, here airports (*dest*) and the originally scheduled arrival hour (*oetatz*). Basic scheduling information is included, i.e., scheduled number of flights (*flts*) and number of flights cancelled (*fcnx*). The program tallies the effects of running a GDP program, returning how many flight arrival slots were changed and what where the cumulative results of these changes, broken down by actions attributable to the FAA arrival slot allocations (*fchg* and *fdel*), airline substitutions (*achg* and *adel*) and compressions (*cchg* and *cdel*). Information about the GDP control information is also included, i.e., how many programs (*pgms*), revisions (*rev*), compressions (*com*) and cancellations (*pcx*) were executed. Additional examples are given in appendix C.3.7, comprep.

```
perl comprep sfo 19980930
```

Dest	Oetatz	Flts	Fcnx	Fchg	Fdel	Achg	Adel	Cchg	Cdel	Pgm	Pcx	Rev	Com
sfo	12	31	7	39	2003	16	-370	8	-104	1	1	1	0
sfo	13	34	10	42	2256	23	-544	25	-42	0	0	0	1
sfo	14	37	4	59	1790	28	-662	18	-184	0	0	0	1
sfo	15	35	5	53	2368	19	-660	9	-187	0	0	0	1

Table B-3: comprep example output (partial results from 'comprep sfo 19980930')

For the noon hour, there were 31 flights flown (*flts*) and 7 flights cancelled (*fcnx*). There were 39 FAA arrival slot allocation changes (*fchg*) resulting in a cumulative delay of 2003 minutes (*fdel*), but this was partially offset by 16 airline substitutions (*achg*) resulting in -370 minutes of delay (*adel*) and 8 compression changes (*cchg*) for a further -104 minutes of delay (*cdel*). 1 program was started this hour (*pgm*) and another (a prior one) cancelled. In total there was one revision (counting the initial program as a revision) (*rev*) and no compressions (*com*).

B.5.2 AAR utilization

The *aarrep* program calculates the basic AAR utilization characteristics. The program first calculates a number of salient characteristics on a per-hour basis for a user-specified set of airports and dates, such as flights scheduled, flown, cancelled and landed, as well as actual and adjusted AAR capacity. This is followed by the second and final step, where the information is grouped according to one or more user-specified criteria, such as by hour, date, month, quarter, destination or GDP state (active or cancelled).

Using the actual AAR during time periods of low demand result in reporting under-utilized capacity, although this under-utilization is in actuality caused by the low demand. The adjusted AAR' corrects for this by adjusting the AAR downward, when demand is too low to fill the actual AAR. It is defined as (1) AAR' = AAR, when scheduled arrivals exceeds or equals the available AAR capacity, and (2) AAR' = flights landed, when scheduled arrivals is less than the available AAR capacity.

Table B-4 shows an example of a call to this program and a partial listing of the result it produces. Additional examples of the use of the program can be found in appendix C.3.5, *aarrep*.

```
perl aarrep sfo 19980930
```

Dest	Hour	Count	Sched	Flown	Cnx	Land	AAR	AAR'	Pct
SFO	20	1	48	36	12	31	30	30	-3%
SFO	21	1	36	25	11	32	30	30	-6%
SFO	22	1	24	18	6	28	30	28	0%

Table B-4: aarrep example output (partial results from 'aarrep sfo 19980930')

The program calculates the results per destination (*dest*) and hour in the local time zone (*hour*). The hour starting at 20:00 covers 1 hour of GDP (*count*). The hour originally had 48 flights scheduled to arrive at that hour (*sched*), out of which 36 flights were flown (*flown*) and 12 were cancelled (*cnx*). 31 flights actually landed during this hour (*land*). The results provide no information about when these particular 31 flights originally were scheduled to land. The stated AAR capacity was 30 flights per hour (*AAR*), which is the last set value for the AAR capacity for this hour. The AAR for a particular hour may be changed several times during the day as part of a GDP program initialization, revision or extension, but the *aarrep* program in its current version only uses the last value set. AAR', the adjusted AAR capacity for this hour is also 30 (*AAR'*). Finally, the slot capacity under-utilization is -3% (*pct*). This is calculated as *(AAR - Land)/AAR*, and so represents the percentage of slots that go unused during this hour.

Note that only hours with GDPs (and hence AARs) are included in this report.

B.5.3 Unexpected airborne holding

The unexpected airborne holding data analysis is generated using the *abhrep* program. Table 2-8 shows an example of a call to this program and a partial listing of the result it produces. The report includes the airport (*dest*), the actual arrival time given in local time (*etatz*), the number of flights (*count*), average UABH (*avg*) and UABH standard deviation (*stddev*), minimum (*min*) and maximum (*max*) UABH values, and the UABH sum (*sum*) and sum of squares (*sum2*). Additional examples are given in appendix C.3.6, *abhrep*.

```
perl abhrep -g etatz sfo 19980930
```

Dest	Etatz	Cnt	Avg	Stdev	Min	Max	Sum	Sum2
SFO	8	24	1.5	2.9	0	11	38	272
SFO	9	24	3.0	6.9	0	32	74	1390
SFO	10	30	6.9	9.7	0	29	209	4297

Table B-5: abhrep example output (partial results from 'abhrep -g etatz sfo 19980930')

The report extract shown indicates that the average unexpected airborne holding increases from 1.5 to 6.9 minutes in the three hours measured. Other basic statistical measures are also calculated and listed in the report, such as an attendant increase in the standard deviation over the same three hours.

B.5.4 Airborne times

The *eterep* program calculates scheduled and actual flight en-route times. Table B-6 show an example of a call to this program and a partial listing of the result it produces. The scheduled airborne time is calculated as the difference between the original scheduled time of arrival (*oeta*) and departure (*oetd*) respectively. Both of these fields come straight out of the database. Similarly, the actual airborne time is calculated as the difference between the final estimated time of arrival (*last eta*) and departure (*last etd*) respectively. These fields are also directly available in the database.

```
perl eterep
```

Origin	GDP State	OETE	ETE
LAX	ActiveGDP	61	57
SEA	CnxGDP	95	98
DEN	NoGDP	141	118

Table B-6: eterep example output (partial results from 'perl eterep')

The *etarep* program is not nearly as fully developed as the three discussed earlier in this section, *abbrep, comprep* and *aarrep*. It is currently hard-coded to produce output for one specific destination airport (sfo) and six specific origin airports (lax,sea,den,dfw,ord,ewr), either for two specific days (9/1 and 9/30) or for the month of September.

B.6 Finalizing the results

The database has proven itself as a very useful repository of information. It has greatly facilitated an iterative analysis methodology, since the acquisition and extraction of relevant data from a database is much simpler than when working directly on the source data-files, as represented by the ΔADL files. It does, however, have shortcomings in some areas, especially in the ability to create cross-tabulations and to create charts and graphs. This latter functionality has been needed extensively in this body of work. It has been achieved by importing the data extracted from the database into Microsoft Excel spreadsheets.

Excel pivot-tables have been used to generate many of the analysis tables. The pivot-tables allow for the dynamic creation of cross-tabulation tables, where the column and row values range over all values found in the imported data. One style of table that have been used repeatedly is a table that employs a date or time based set of row values (e.g., quarters, months, dates, hours) along with the set of airports as the column values. Examples of such tables can be found in *Table 2-1: Ground delay programs, Table 2-5: SFO average delay and delay savings per GDP flight, by month* and *Table 2-24: Average daily cancellations for GDP and non-GDP days,* among others. Furthermore, the spreadsheet environment allows for easy addition of additional calculated fields, and so have proven to be invaluable.

This latter observation is especially true in the creation of charts and graphs. These are easily created directly from spreadsheet data. The sophisticated formatting features found in both Microsoft Excel and Microsoft PowerPoint have been more than adequate to render the information in an illuminating fashion, as the figures found in this report hopefully will substantiate.

C Database Tables and File Formats

C.1 Database tables

C.1.1 Flights

Column	Type	Size	Description
Acid	Character	7	Flight identifier
Date	Date	4	Flight date
Dest	Character	4	Flight destination (for this flight leg)
Src	Character	4	Flight source (for this flight leg)
Equip	Character	4	Equipment used (unfortunately not standardized values)
Time	Smallint	2	Time of last GDP update of 'slot' field below
Slot	smallint	2	Assigned arrival slot, if any
Oetd	smallint	2	Original estimated time of departure
Oeta	smallint	2	Original estimated time of arrival
Etd	smallint	2	Actual time of departure
Eta	smallint	2	Actual time of arrival (last known eta)
cnx1	character	2	First type of cancellation received (e.g., FX)
cnx2	character	2	Second type of cancellation received (e.g., SI)
tmcx2	smallint	2	Time of second cancellation
Tmfx	smallint	2	Time of FX cancellation
Tmcxdl	smallint	2	Time 'lost' if FX messages not present
Etad	smallint	2	Estimated time of arrival at time of departure + 15 minutes
Etaa	smallint	2	Estimated time of arrival at arrival - 30 minutes
Etal	smallint	2	Last estimated time of arrival
Abh	smallint	2	Unexpected airborne holding (= etal - etaa)
Tdel	smallint	2	Total (Net) amount of GDP delay
Tchg	smallint	2	Total count of changes made
Fdel	smallint	2	Cumulative FAA arrival slot allocation delay
Fchg	smallint	2	Count of FAA arrival slot allocation changes made
Adel	smallint	2	Cumulative airline substitution savings
Achg	smallint	2	Count of airline substitutions made
Cdel	smallint	2	Cumulative compression savings
Cchg	smallint	2	Count of compressions made
Odel	smallint	2	Cumulative other delays/savings
Ochg	smallint	2	Count of other delays/savings changes made

Table C-1: flights table definition

C.1.2 AAR

Column	Type	Size	Description
Date	date	4	Date of AAR change
Time	smallint	2	Time of AAR change
Dest	character	4	Airport affected
Tbeg	smallint	2	Time of GDP program start
Hour	smallint	2	Hour of setting
AAR	Smaillint	2	AAR setting for this hour

Column	Type	Size	Description
Gafactor	Smallint	2	GA factor setting for this hour

Table C-2: AAR table definition

C.1.3 GDP

Column	Type	Size	Description
Date	date	4	Date of GDP change
Dest	character	4	Airport affected
Tinit	smallint	2	Initial time of GDP start
Tcnx	smallint	2	Time of GDP cancellation/expiration
Tbeg	smallint	2	Final time of GDP start
Tend	smallint	2	Final time of GDP end
Revs	smallint	2	Count of number of revisions run
Comps	smallint	2	Count of number of compressions run
Cnx	smallint	2	Cancelled (1) or Expired (0)

Table C-3: GDP table definition

C.1.4 GDPevent

Column	Type	Size	Description
date	date	4	Date of change
time	smallint	2	Time of change
dest	character	4	Airport affected
tbeg	smallint	2	GDP initial start time
Event	character	4	Type of change (e.g., INIT, REV, COMP, CNX)

Table C-4: Gdpevent table definition

C.1.5 Airport

Column	Type	Size	Description
Airport	character	4	Airport
Tz	smallint	2	Time zone offset

Table C-5: Airport table definition

C.1.6 Notes

- All time fields in the database tables are specified in minutes GMT relative to the start of the date. This allows for easy comparisons and difference calculations.

C.2 File Formats

C.2.1 ADL file format

The two sources for the ADL file format description are references [3] and [4]. Table C-6 describes the current semantics of the *Arrivals* section of the ADL files:

# Field	Description
0 ACID	Flight Identifier or Call sign
1 DEST	Current arrival airport
2 ACENTR	Current arrival center
3 SRC	Current source (departure) airport
4 DCENTR	Current departure center
5 AFIX	Arrival fix
6 EFTA	???
7 TYPE	Aircraft Type
8 AC_CAT	Aircraft Category
9 CLAS	Aircraft weight class (J/T/P)
10 ETD	Estimated time of departure, based on - S: OAG data - P: Flight plan data - L: Airline-generated CDM msg - C: Controlled departure time - A: Actual (NAS activation message) - E: Extrapolated; flight is airborne
11 ETA	Estimated time of arrival, based on - L: Airline-provided runway arrival time - E: ETMS modeling - A: Actual (from AZ message)
12 ARTD	Actual runway time of departure
13 ARTA	Actual runway time of arrival
14 SGTD	Scheduled gate time of departure
15 SGTA	Scheduled gate time of arrival
16 OETD	Original estimated departure time
17 OETA	Original estimated arrival time
18 OGTD	Original gate time of departure
19 OGTA	Original gate time of arrival
20 PGTD	Proposed gate time of departure
21 PGTA	Proposed gate time of arrival
22 OCTD	Original controlled runway time of departure
23 OCTA	Original controlled runway time of arrival
24 CTD	Controlled runway time of departure
25 CTA	Controlled runway time of arrival
26 LRTD	Airline runway time of departure
27 LRTA	Airline runway time of arrival
28 LGTD	Airline gate time of departure
29 LGTA	Airline gate time of arrival

# Field	Description
20 ERTD	Earliest runway time of departure
31 ERTA	Earliest runway time of arrival
32 PETE	Filed ETE (Estimated time en-route)
33 CNX_SI	Flight cancelled by an SI msg
34 CNX_FX	Flight cancelled by an FX msg (CDM)
35 CNX_RZ	Flight cancelled by an RZ msg (NAS)
36 CNX_RS	Flight cancelled by an RS msg (OAG)
37 CNX_TO	Flight cancelled by an TO msg (ATMS time-out)
38 CNX_DV	Flight cancelled by an DV msg (dest chg)
39 CNX_ID	Flight cancelled by an ID msg (call sign chg)
40 DLY_ALD	Delay reported by airline CDM msg
41 DLY_GDP	Delay by GDP program
42 DLY_FA	Delay by an FA delay ?
43 DLY_GSD	Delay by ground stop
44 DLY_TOD	Delay by ATMS logic
45 SLOT_MAT	Slot maturity
46 SLOT_CLS	Slot class (1,2,-)
47 SLOT_REL	Slot released for compression
48 ASLOT	Assigned slot
50 USR	?????
51 ALM	?????
52 CDM_MBR	CDM member
53 SUB	?????
54 EXMPT	?????

Table C-6: ADL file format (arrivals section)

C.2.2 Slot file format

#	Field	Syntax	Description
0	Tag	'SLOT'	Record tag-field. Always 'SLOT'.
1	Dest	ccc	Destination airport
2	Date	yyyy-mm-dd	Date
3	Time	ddhhmm	Time of slot change, in ADL time format, or '-' if null
4	Acid	Char7	Flight identifier, or '-' if null
5	Slot	ddhhmm	Slot time, or '-' if null
6	Cnx	cc	Flight cancellation code, e.g., FX, or '-' if not cancelled

Table C-7: Slot assignment file format

C.2.3 Flight and Flight.db2 file formats

The *flight* and *flight.db2* file formats are identical as far as their fields are concerned, but the individual field syntax varies somewhat.

#	Field	Syntax (flight)	Syntax (flight.db2)	Description
0	Tag	'flight'	N/a	Record tag field - not in flight.db2 record
1	Date	yyyy-mm-dd	"yyyy-mm-dd"	Date of flight
2	Acid	char7	"Char7"	Flight #
3	Dest	cccc	"cccc"	Destination
4	Src	cccc	"cccc"	Origin
5	Equip	char5	"char5"	Equipment
6	Time	ddhhmm	min	Time of last update
7	Slot	ddhhmm	min	Last assigned slot
8	Oetd	ddhhmm	min	Original Estimated Time of Departure
9	Oeta	ddhhmm	min	Original Estimated Time of Arrival
10	Agtd	ddhhmm	min	Actual Time of Departure
11	Agta	ddhhmm	min	Actual Time of Arrival
12	Cnx	cc	"cc"	Earliest cancellation type
13	Cnx2	cc	"cc"	Earliest non-FX cancellation type
14	Tmcx2	ddhhmm	min	CNX other than FX first time
15	Tmfx	ddhhmm	min	CNX FX first time
16	Tmcxdl	min	min	FX cancellation savings (neg => better)
17	Etad	ddhhmm	min	Estimated time of arrival after 15 mins of flight
18	Etaa	ddhhmm	min	Estimated time of arrival at ETAD - 30 min
19	Etal	ddhhmm	min	Estimated time of arrival last recorded
20	Abh	min	min	Unexpected airborne holding (neg => better)
21	Tdel	min	min	Total cumulative delay
22	Tchg	n	n	Total number of changes
23	Fdel	min	min	FAA-assigned delay
24	Fchg	n	n	FAA-assigned number of changes
25	Adel	min	min	Airline substitution delay
26	Achg	n	n	Airline substitution number of changes
27	Cdel	min	min	Compression delay (savings hopefully)
28	Cchg	n	N	Compression number of changes
29	Odel	min	min	Other delay
30	Ochg	n	N	Other number of changes

Table C-8: flight and flight.db2 file format

Note the following differences:

- Flight.db2 does not have a tag field as the first field
- Flight.db2 transposes the date and acid fields, i.e., the first two fields in the record are acid and date.
- Character fields are enclosed in double-quotes

- Time fields are specified in minutes offset from the date field.
- Fields are separated by one of more blanks in flights (to line each field up visually), and by commas (',') in flights.db2.
- Null fields are specified by a '-' in flights, by an empty field in flights.db2.

C.2.4 Gdp.event file format

#	Field	Syntax	Description
0	Tag	'GDP'	Record tag field; must be 'GDP'
1	Event	keyword	Event type descriptor
2	Airport	cccc	3- or 4-char airport code
3	Date	yyyy-mm-dd	Change date
4	Time	ddmmhh	Change hour, in Metron time
5	Desc	char	Program extent, e.g., 12West
6	Tbeg	ddmmhh	New AAR valid from Tbeg
7	Tend	ddmmhh	New AAR valid to Tend
8	AAR	n or n/n/...	Hourly AAR setting (one or more; last replicated as needed)
9	Gafactor	n	Gafactor setting

Table C-9: gdp.event file format

Keyword	Description
INIT	New GDP initiated
REV	GDP revised
RBS	GDP revised using RBS
RBS+	GDP revised using RBS++
EXT	GDP extended
R+E	GDP revised and extended
COMP	Compression run
CNX	GDP cancelled
EXP	GDP expired

Table C-10: GDP event types

Example:

```
GDP RBS  SFO 1998-09-30 302057 12West 310000 310759 30 2
GDP COMP SFO 1998-09-30 302226
GDP COMP SFO 1998-09-30 302343
GDP COMP SFO 1998-09-30 310037
GDP RBS  SFO 1998-09-30 310113 12West 310100 310759 0/35/30/30/30/45/60 0
GDP COMP SFO 1998-09-30 310141
GDP CNX  SFO 1998-09-30 310330
```

C.2.5 Gdp.db2 file format

#	Field	Syntax	Description
0	Date	"yyyy-mm-dd"	Date
1	Dest	"cccc"	Airport
2	Tinit	Min	Program Creation time
3	Tcnx	Min	Program Cnx/Exp time
4	Tbeg	Min	Program begin time
5	Tend	Min	Program end time
6	Revs	Int	Count of revisions done
7	Comps	Int	Count of compressions done
8	Cnx	Int	Program cancelled (1) or expired (0)

Table C-11: GDP.db2 file format

C.2.6 Gdpevent.db2 file format

#	Field	Syntax	Description
0	Date	"yyyy-mm-dd"	Date
1	Time	Min	Time of change
1	Dest	"cccc"	Airport
4	Tbeg	Min	Program begin/revise time
5	Event	Char 4	Event type descriptor (see *Table C-10: GDP event types*)

Table C-12: gdpevent.db2 file format

C.2.7 Aar.db2 file format

#	Field	Syntax	Description
0	Date	"yyyy-mm-dd"	Date
1	Time	Min	Time of change
2	Dest	"cccc"	Airport
4	Tbeg	Min	Program begin/revise time
5	Hour	Int	Hour of AAR (from date start)
6	AAR	Int	Stated AAR
7	GAFactor	Int	Stated GA factor

Table C-13: aar.db2 file format

C.3 Programs

C.3.1 Extract

The *extract* program extracts the flight and arrival slot allocation information from the ΔADL files, in the process creating the *flight* and *slot* files. This is the first step in the process of getting the data ready for the database. An example of a call to this program is given below:

```
perl extract sfo 19980930
```

This example will process one airport (SFO) and one date (9/30/1998), creating two files, /data/adl/19980930/sfo.flights and /data/adl/19980930/sfo.slot. It assumes the presence of the input ΔADL file in /data/adl/19980930/sfo.adl.

C.3.2 Savings

The *savings* program attributes the delays and delay savings to their underlying causes, i.e., FAA arrival slot allocation delays and savings, airline substitution delays and savings, compression savings, and other delays and savings. It uses the *flight* and *slot* files as input, and updates the flights file with additional information related to the ground delay programs. An example of a call to this program is given below.

```
perl savings -n -z -hdr sfo 19980901..19980930
```

This example will process one airport (SFO) and a range of dates (9/1/1998 through 9/30/1998), updating the *sfo.flight* files in each of the appropriate directories, i.e.., /data/adl/19980901 through /data/adl/19980930. It also produces output summarizing the effect of the GDP program for each of the days along with a column header (-hdr), including non-GDP days as well (-n). The output files are compressed (-z) to save space.

C.3.3 prepdb2

The *prepdb2* program converts the flights information to a format suitable for loading into the database. It uses the *flights* file as input. An example of a call to this program is given below:

```
perl prepdb2 sfo 19980901..19980930
```

This example will process one airport (SFO) and a range of dates (9/1/1998 through 9/30/1998), creating *sfo.flights.db2* based on the input from *sfo.flights*. This *flights.db2* file can be used directly by the database loader.

C.3.4 Library routines.

A number of subsidiary library programs have been created to handle common tasks or tasks related to a particular file type. *flight.pl* handles the characteristics of the *flights* and *flights.db2* files. *adl.pl* handles the characteristics of the ΔADL files. *slot.pl* handles the characteristics of the *slot* files. *gdp.pl* manages all the GDP-related files, i.e., *events.gdp, gdp.db2, gdpevents.db2*, and *aar.db2*. *codeshr.pl* maps airline code-shares of partner airlines, since the slot allocation process is managed as a single entity for each pool of partner airlines. *cdmutils.pl*

C.3.5 aarrep

The *aarrep* program extracts data from the database as the basis for analysis of the AAR utilization. The resulting output data is then imported into a spreadsheet for final analysis. The following examples show the capabilities of the *aarrep* program.

```
perl aarrep -g etatz,gdp sfo 19980930
```

Extract the data from one airport (SFO) for one date (9/30/1998). The report groups the data by two key-fields, actual arrival time (etatz) and GDP state (gdp). The GDP state will either be 'ActiveGDP", if the GDP is active at any point within that hour, or 'CnxGDP' if the GDP has been cancelled for that hour.

```
perl aarrep -g date sfo 19980901..19980930
```

Produce a report for SFO for the month of September, with one line per date for each date that had a GDP program defined.

C.3.6 abhrep

The *abhrep* program extracts data from the database as the basis for analysis of the unexpected airborne holding. The resulting output data is then imported into a spreadsheet for final analysis. The following examples show the capabilities of the *abhrep* program.

```
perl abhrep sfo,ewr 19980901..19980930
```

Run the default report for two airports (SFO and EWR) for the month of September, 1998. Multiple airports and dates can be specified using comma-separated values. Dates are given as yyyymmdd format throughout. Date-ranges can be given by using a from-date and a to-date separated by two periods ('..').

```
perl abhrep * *
```

Produce the default report for all known airports and dates. The asterisk ('*') is used as an short-cut indicator to specify all know instances for both fields.

```
perl abhrep -csv jj.csv * *
```

As the previous example, but write the result to a comma-separated file named jj.csv instead of producing a tabular output result on the display. This is a useful option, if the results are to imported elsewhere for further analysis.

```
perl abhrep -g month,etatz -csv jj.csv * *
```

As the previous example, but group the result by month and local arrival hour (etatz). Many different grouping fields exist. Note that the destination is an implicit grouping field that cannot be unspecified.

C.3.7 comprep

The *comprep* program extracts data from the database, as the basis for analysis of the compression benefits and of the cancellations. The resulting output data is then imported into a spreadsheet for final analysis. The following examples show the capabilities of the *comprep* program.

```
perl comprep -g date * 19980930
```

Produce a report for all airports for a single date. One line per date and airport (-g date) is produced, since the destination is an implicit grouping field.

```
perl comprep -g month * *
```

Produce a monthly report for all airports, with one line per month and airport.

```
perl comprep -g month -csv jj.csv * *
```

As the previous example, but save the result in a comma-separated file named jj.csv.

C.3.8 eterep

The *eterep* program currently generates a file containing the scheduled and actual airborne times for flights arriving at SFO from six specific origins (LAX, SEA, DEN, DFW, ORD and EWR) for the month of September, 1998. This file is subsequently imported into a spreadsheet for further analysis of the scheduled versus actual airborne times.

C.4 General computational issues

A substantial amount of time and effort have gone into establishing the environment necessary to answer the research questions of how well the Ground Delay Programs are operating. Several features have turned out to be useful to the smooth and successful operations of this analysis environment, but four stand out in particular: These four are described in this section.

C.4.1 Storing the data in the database

Storing the data in a true relational database was probably the single most important design decision. It had two especially important consequences. First, creating the specific data extractions needed to perform a particular analysis became relatively straightforward, instead of requiring extensive programming effort. Secondly, the speed of the data extraction can be measured in minutes as opposed to hours, enabling repetitive runs and drill-down analysis.

C.4.2 Db2show for databases

All the programs that access the database have a special option named '-db2show' added. This option makes the programs display the database access logic, as opposed to running it, and so helps immeasurably in tracking down errors in the access logic. A simple example of this is given below:

```
E:\cdm\bin\perl> perl eterep -db2show

connect to cdm;

export to 'e:\cdm\results\eteoete2.csv' of del
    select   src
            ,case when tchg is null then 'NoGDP'
                        when slot is null then 'CnxGDP'
                        else                   'ActiveGDP' end as sslot
            ,oeta-oetd as oete
            ,eta-etd   as ete
    from     cdm.flights
    where    dest='SFO'
      and    src in ('EWR','ORD','DFW','DEN','SEA','LAX')
```

```
     and     (date = '1998-09-01' or date = '1998-09-30')
     and     eta-etd   is not null and eta-etd > 0   and eta-etd < 1000
     and     oeta-oetd is not null and oeta-oetd>0 and oeta-oetd<1000
     and     abs(eta-etd - (oeta-oetd)) < 200
   order by 1,2,4;

connect reset;
```

C.4.3 SQL

Accessing the data in the database required extensive use of the SQL language. SQL is the common access language used by all relational databases. Having an ability to quickly try out new SQL constructs before embedding these into actual programs solved many problems quickly and effortlessly. A small program call *sql* was written to facilitate this type of development effort. A simple example of the use of this program is given below:

```
E:\cdm\bin> perl sql cdm "describe table cdm.aar"

connect to cdm

   Database Connection Information

 Database product        = DB2/NT 5.0.0
 SQL authorization ID    = JOHN
 Local database alias    = CDM


describe table cdm.aar

Column        Type         Type
name          schema       name        Length   Scale    Nulls
------------- ------------ ----------- -------- -------- --------
DATE          SYSIBM       DATE               4        0 No
TIME          SYSIBM       SMALLINT           2        0 No
DEST          SYSIBM       CHARACTER          4        0 No
TBEG          SYSIBM       SMALLINT           2        0 No
HOUR          SYSIBM       SMALLINT           2        0 No
AAR           SYSIBM       SMALLINT           2        0 No
GAFACTOR      SYSIBM       SMALLINT           2        0 No

  7 record(s) selected.

connect reset
DB20000I  The SQL command completed successfully.
```

C.4.4 Inline help

All programs written have a standard help feature. This feature provides a brief description of the purpose of the programs, what options are available, and a few succinct examples. A simple example of its use is given below. The example uses the *comprep* program, displaying its built-in help:

```
E:\cdm\bin> perl comprep -h

Purpose:  Compression benefits

Usage:    comprep  [options+]  {airport daterange}+

options:  None or more of:
  -csv file      : Export data to csv <file> (dflt: to stdout)
  -g grp-flds    : Comma-separated list of fields to group on.  One+ of:
                     dest,date,month,etahr,gdpstate,gdpstat2  (dflt:
oetatz)
  -r             : Apply rollup (subtotals) to data
  -w <sel=val>   : Specify specific values only.  <sel> same as grp flds
  -db db         : Set database to be used (and implicitly the instance)
  -di inst       : Set db2 instance to be used
  -db2show       : Show db2 commands generated (but do not execute)
  -h|-H          : Print usage info with (-H) or without (-h) common opt

Arguments:
# db             : Database to hold tables
# schema         : Schema (Owner) of tables
  airport        : 3-letter airport code
  daterange      : yyyymmdd or yyyymmdd..yyyymmdd

Examples:
   # Analyze compression benefits one airport for one date
   comprep sfo 19980908

Notes:
  1. Dest will always be the first group field.  This cannot be changed
```