Ground Delay Programs: Optimizing over the Included Flight Set Based on Distance

Michael O. Ball[†] and Guglielmo Lulli[‡]

[†] Robert H. Smith School of Businness - University of Maryland
College Park, 20742 Maryland, USA

[‡] Dpt. of Statistic, Probability and Applied Statistics - University of Rome "La Sapienza"
P.le A. Moro 5, I-00185 Rome, Italy

Abstract

The Ground Delay Program (GDP) is an air traffic flow management mechanism used to decrease the rate of in-coming flights into an airport when it is projected that arrival demand will exceed capacity. Under a GDP, a set of flights destined for a single airport is assigned ground delays. In this paper we investigate how the set of flights to which delays are applied is defined. Specifically, we define a "distance-based" GDP as one that only applies to flights whose origin airports are less than a prescribed distance, d, from the destination airport. We then investigate methods for setting the parameter d. This approach is different from the current approach which groups origin airports by air route traffic control center jurisdiction and restricts flight based on a center-based tier system.

1 Introduction

In recent years, air traffic has experienced a dramatic increase. This increase has not been supported by a corresponding development of airports and related systems. As a consequence, both the European Airspace System and the United States National Airspace System (NAS) are suffering from increased congestion. A short-term strategy for reducing or eliminating air traffic congestion is delay in the form of ground delay. The Ground Delay Program (GDP) is a mechanism used to decrease the rate of in-coming flights into an airport when it is projected that arrival demand will exceed capacity. Ground delay is the action of delaying take-off beyond a flight's schedule departure time. The motivation for doing so is that as long as an airborne delay is unavoidable, it is safer and cheaper for the flight to absorb this delay on the ground before take-off, rather than in the air.

Starting in 1998, GDPs have been planned and controlled using the Collaborative Decision Making (CDM) framework [1, 2, 4, 12]. CDM embodies a new philosophy for managing air traffic. It is based on the belief that air traffic flow management can be improved if there is a closer collaboration between the FAA and the airlines and other airspace users, with large benefits for both parties. This collaboration takes the form of mutual exchange of data and more flexible and efficient collaborative procedures. Recently, the concept of a "distance-based GDP" has emerged from the CDM working group (see [4]). Under a distance-based GDP, only flights whose origin airports are less than a prescribed distance, d, from the destination airport are included in the program (i.e. are eligible to receive a ground delay).

In this paper, we model the decision problem, related to the distance-based GDP, and analyze the GDP decision parameters. We highlight differences and similarities between the distance-based GDP and the tier-based GDP. We propose an approach for optimizing over the distance parameter. Our analysis employs several National Airspace System (NAS) test data sets.

The paper is organized as follows. In § 2, we provide a description of the GDP operations, methodologies and parameters. A description of our analyses is given in § 3. Here, we present both the analyses on the distance based and the departure-time based GDP deecision parameters. We also provide an high level optimization model to suggest a GDP initiative. The experimental study is given in § 4.

In § 5, we investigate more in details on the statistics currently used to evaluate the different GDP option, i.e. average and unrecoverable unrecoverable delay and average delay. Finally, the conclusions and future research are in § 6.

2 The Ground Delay Program

Whenever the number of flights projected to arrive over a 15-minute time interval exceeds the airport's predicted arrival capacity for that time interval, the Air Traffic Control System Command Center (ATCSCC) is required by regulation to take some form of corrective action. The response to long-term periods of capacity-demand imbalance is a Ground Delay Program.

GDP planning can be best conceptualized as a task of adjusting the *arrival times* of flights. A GDP is run when the number of flights scheduled to arrive at an airport

exceeds the number of arrival slots available over a certain period of time. To correct this imbalance, some flights are assigned to later arrival time slots. The ground delay assigned to a flight is simply the difference between the flight's assigned arrival time slot and the time slot it originally was scheduled to use.

The first algorithm used to plan GDPs was the Grover Jack algorithm. It formed a list of in-coming flights, sorted by the most recent estimated time of arrival (ETA). Controlled arrival times (CTAs) were assigned to flights in the list using a first-come first-served rule. In summary, the effect of the Grover Jack algorithm was to maintain the flight order associated with the most recent estimated arrival times while "stretching out" the list of in-coming flights over time. In the early 1990's, the Collaborative Decision Making began. Its first focus has been GDP enhancements and the improvement of the cancellation and substitution process used by the airlines. The driving philosophy behind CDM is that improved data exchange and communication between aviation transportation organizations will lead to better decision making in air traffic flow management. In this framework, air traffic management is moving in the direction of decentralized decision making, thus giving the scheduled carriers input to air traffic management and greater control over their operations. GDP operations, under CDM methodologies, are executed by means of a cycle of feedback between the service provider and users of the NAS. Once a capacity-demand inequity is forecasted, the ATCSCC implements a GDP and a delay is served to in-coming flights. Airlines react to the new situation with a cancellation/substitution process. At the end of this process, a new iteration is started. Note that the service provider (ATCSCC) can make an accurate situational assessment at each iteration

of a data exchange cycle only if the airlines are supplying updated data in the form of cancellations, revised ETA's, and so on.

The incentives for the airlines to provide timely data to the FAA during the formulation of a GDP have been housed in the two algorithms, ration-by-schedule (RBS) and compression. The RBS algorithm's purpose is to ration arrival slots according to original scheduled arrival times of flights and to serve as an initial assignment of CTA's for subsequent round of collaboration between the airlines and the FAA. The RBS algorithm is similar to Grover Jack algorithm with the exemption that it rations arrival slots by the original scheduled arrival time (computed as original gate time of arrival - OGTA - minus a standard ten-minute taxi time) instead of the estimated time of arrival (ETA). Once slots are assigned to flights, the corresponding airlines are said to "own" the slots. Then they can start a process of cancellation/substitution in order to adjust their schedules thus minimizing the damage of flights delays in a GDP. The compression algorithm's purpose is to move flights up in the schedule (earlier in time) to fill, whenever possible, all the slots that after the substitution/cancellation process are unused. In this filling process, flights from the controlling airline of a slot t are considered before all others when t is released. A time slot is released whenever the airline, which owns the slot, declines to substitute a flight into the slot or the slot is too early for any flight of the controlling airline. If the airline cannot use a slot, it is (eventually) compensated since it receives control of the slot vacated by the flight which moves into its slots. There is no way for an airline to involuntarily lose a slot reserved by RBS. For a more detailed treatment on the Ground Delay Program see [6, 11].

2.1 The Ground Delay Program Model and Parameters

In this section, we describe the parameters that define a GDP option. At the same time, we give some terminology used throughout the paper. The basic GDP parameters are illustrated in Figure 1.

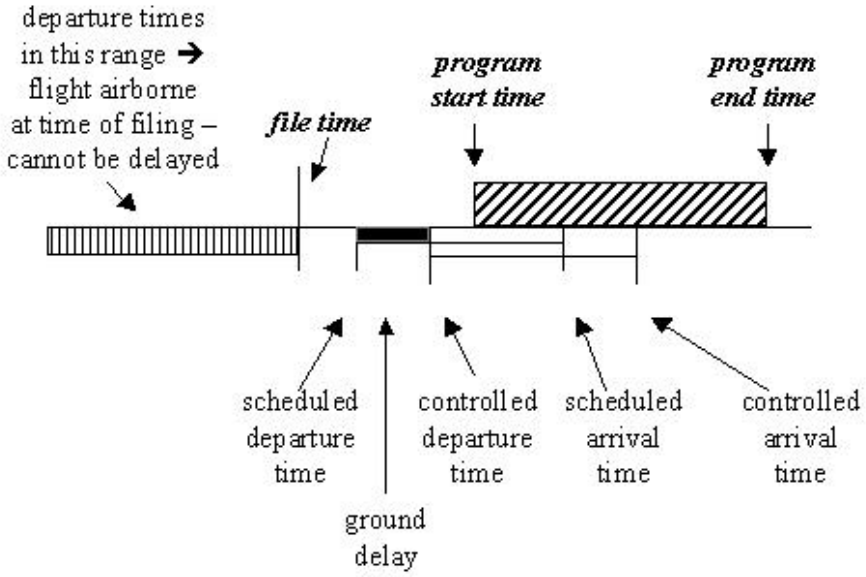

Figure 1: GDP parameters.

A GDP is motivated by a predicted arrival capacity-demand imbalance at an airport. The airport arrival capacity is referred to as the airport *acceptance rate (AAR)* and is usually measured in arriving flights per hours. A program is run when the predicted demand is substantially greater than the predicted AAR over a sustained period of time. The program *start time* and *end time* define the beginning and end of the predicted capacity-demand imbalance. The flights that are scheduled to arrive at the airport between the start and end time are said to be *in the GDP range*. For expositional simplicity here, we will usually assume that the predicted AAR is

constant throughout the GDP range, although this is certainly not always the case. Conceptually, all flights in the GDP range could potentially receive ground delays and adjusted arrival times. However, times are only adjusted for a restricted set of flights said to be *included the program*. Generally speaking the start time, end time and predicted AAR can be viewed as exogenous parameters that are determined based on weather forecasts and the airline flight schedules. On the other hand, choosing the set of flights to include in the program represents an important decision made by the specialists who plans the GDP and it is the problem we address in this paper. This set is restricted in two major ways. First, for policy reasons certain flights are designated as *exempt*. Exempt flights include, for example, international flights but others are included as well, according to their airport of departure. Second, flights that are already airborne clearly cannot be assigned ground delay. We note that the set of airborne flights is determined by a second key GDP decision parameter, namely, the program *file time*. This is the time at which the command center certain commits to running the GDP and after which time ground delay can begin.

We now define notation that represents the parameters just described.

$\mathcal{F}$ = the set of flights in the GDP range

$\hat{F}$ = the flights in $\mathcal{F}$ that are exempt flights

$\tilde{F}(t)$ = the flights in $\mathcal{F}$ that are airborne at time t.

$\overline{F}$ = flights included in the GDP

$\hat{t}$ = file time of program

The GDP planning problem can be defined:

choose a start time $\hat{t}$ and a set of flights, $\overline{F}$, to be included in the program such that $\overline{F} \subseteq \mathcal{F} - \hat{F} - \tilde{F}(\hat{t})$.

Of course, we have defined the problem variables but not the criteria by which to set these variables. In reality, this is a complex stochastic optimization problem that must take into account uncertainty associated with both the AAR (which depends on weather) and actual flight departure times (which depend on airline and FAA operations) and also a complex delay cost function.

We start be providing an intuitive discussion of the basic tradeoff involved. In a deterministic setting (i.e. assuming that the actual AAR exactly matches the planned AAR) the (minimum) total amount of delay necessary to resolve a capacity-demand imbalance is constant. Thus, assuming that no airborne delay is assigned, the total ground delay is approximately constant. However, it is important to consider that total ground delay and total ground delay cost may not be related in a simple, i.e. linear manner. For example, as the delay assigned to a flight increases, it becomes more likely that passengers will miss connections, that crews will timeout, that the delayed availability of aircraft will cause delays on subsequent flights, etc. Thus, the cost to an airline of 20 flights each incurring 15 minutes of delay, as a rule, is less than the cost of 5 flights each incurring 60 minutes of delay. Thus, we have our first effect:

as $\hat{t}$ increases and $|\overline{F}|$ decreases, the pool of flights that are assigned delays gets smaller, the average assigned delay gets larger and the cost of assigned ground delay gets larger.

We now consider the downside associated earlier file times (smaller values of $\hat{t}$) and larger sets $\overline{F}$. An earlier file time expands the pool of flights and enables ground delays to be assigned further in advance of the program start time. Generally speaking it will be possible to assign delay to longer haul flights. The disadvantage of assigning ground delays well in advance of the program start time is that if the weather forecast changes, then it is possible ground delay will be been unnecessarily incurred. By waiting as long as possible the program will be based on the best possible forecast.

as $\hat{t}$ decreases and $|\overline{F}|$ increases, the pool of flights that are assigned delays gets larger and includes long haul flights with earlier departure times; as a result delay assignments are based on less accurate weather forecasts and it becomes more likely that delays are assigned (and incurred) unnecessarily.

We can formalize these concepts by defining appropriate cost functions:

$f_1(\hat{t}, \overline{F})$ = cost of assigned ground delay

$f_2(\hat{t}, \overline{F})$ = expected cost of ground delay that was unnecessarily assigned.

The second cost function, f_2 models the stochastic aspects of the problem. Specifically, it captures the fact that due to the uncertainty associated with AARs, the AAR could increase to its normal level before the end time of the program (i.e. the

weather could clear earlier that expected). In such a case, more ground delay would have been assigned than was necessary to exactly balance demand and capacity.

Concluding, the decision of determining the included set for a GDP, $\overline{F}$ should be accommodated by considering a weighted combination of f_1 and f_2.

2.2 The Ground Delay Program Statistics

The problem of determining the included set for a GDP, $\overline{F}$ this should be done by considering a weighted combination of f_1 and f_2. However, these functions in general are difficult to evaluate. Under current operations, air traffic managers evaluate the different GDP initiatives using average delay and unrecoverable delay as surrogate measures for f_1 and f_2 respectively. The flights within the GDP range, $\mathcal{F}$, can be partitioned into two sets: those that are assigned a positive ground delay, $\overline{F}$, and the others, which receive no delay. Average delay is computed over the flights in $\overline{F}$. Since the total ground delay is approximately constant, an increase in the average delay corresponds to a concentration of the overall delay over a smaller set so that a larger set of flights receive no delay and those flight that do receive delay receive, on the average, more. As discussed earlier this corresponds to a higher overall delay cost. Unrecoverable delay is the amount of ground delay that is unnecessarily assigned in the event that the GDP is canceled at its start, i.e. all ground delay that is incurred before the start of the program. Thus, it clearly correlates with f_2. It fails to accurately represent f_2 because GDP may be canceled at any time, not just at the start.

For completeness of exposition, we define the other statistics used throughout

the paper and which have not been introduced yet:

Maximum delay: the longest delay assigned to any one flight in the GDP;

Delay variability: the standard deviation of the carrier's average delay. A small value of delay variability means average delays are quite similar for all carriers, while a large value shows a dissimilarity among the carriers' average delay;

Airborne delay: the en route delay that will be incurred if all flights depart at their planned departure times and if the actual AAR equals the projected AAR. Given an AAR vector and a set of flights with associated ETAs, each flight can be assigned an arrival time/slot in such a way that additional (airborne) delay is minimized. The airborne delay for an individual flight is the difference between its assigned arrival time and its ETA.

Traffic management specialists use projected airborne delay to determine the need for a GDP and (as we will show later) to determine certain parameter thresholds.

3 GDP Analysis

The decision of determining the set of flights to be included in the program, is accommodated by means of the GDP decision parameters described in § 2.1. Since the effect of those parameter is to establish if a flight is exempted or not, herein we also name them as criteria of exemption. In particular we refer to them as criteria

of exemption based on the flights' point of origin and on their departure time. The description of our analyses are given in § 3.1 and in § 3.2. In § 3.3, we are going to present the optimization model used to select the GDP option.

3.1 Distance-based GDP Analysis

Under the current tier-based procedures, the criterion of exemption, based on the flight point of origin, consists in including in the program particular airports or centers. In particular, first tier, second tier and entire NAS options are used. If the GDP airport is in a center of the western part of the NAS, the possible options are 6West, 10West and etc (see the Appendix).

The user community does not consider the current GDP system efficient and equitable. To overcome these feelings, the time-based GDP was defined. In the time-based GDP the different options are defined by including in the program all the flights whose en-route estimated time is shorter than an established value. In particular, it has been shown [9] that for any tier-based GDP action there is an equivalent time-based GDP action, i.e. an action, which has very similar statistics. This work called for a more detailed analysis. Although intuition behind of a time-based GDP met with approval, the following objection was raised. Using the time-based GDP, in a particular departure airport, there may be flights that are in the program and flights that are not. This situation is going to happen with high probability if in the departure airport there are both jets, which have shorter flight times, and props, which have longer flight times. The airline representatives considered this possibility confusing for airline people who manage the flight schedule and

unfair. Therefore, it has been proposed that the criteria be distance-based rather than time-based.

In this case, all the different distance-based GDP options are defined by drawing a circle around the GDP airport and including in the program all enclosed airports. All the flights whose departure airport is within the circle and whose arrival time is within GDP time period, might be served a ground delay. It is possible to get alternative programs by just increasing or reducing the radius of the circle. For each distance-based Ground Delay Program, there are infinitely many distances that can be selected. However, the finite set of airports to be included or excluded from the program naturally reduces these possibilities into a discrete set of options. There is no point in considering an additional distance if it does not encompass a new set of airports. Among this set of possible options, we instituted distance-based GDPs only for distances, such that we enlarge the number of flights coming from the airports included by about ten. In this way, we avoid possible scale effects in the statistics' graphs.

As one would assume, the distance-based GDP has a wider range of possibilities than the tier-based GDP. The set of tier-based GDP option is included in the set of distance-based GDP options. Hence, for each tier-based initiative, it is possible to find a corresponding distance-based.

3.2 Departure-time Analysis

The departure-time analysis considers the effects of file time and of the criteria of exemption based on the flight departure time, i.e. exemption by *time* and by *status*.

Exemption by time means that all flights whose departure time is before the file time plus a specified extension period, are exempted from the program. The extension period is usually fixed at 45 minutes, in order to avoid assigning a ground delay to a flight just about to take-off, when all the passengers are already boarded. With exemption by status we mean that all the flights still on the ground are not exempted. A statistical analysis [10] on 1953 GDPs shows that exemption by time with 45-minutes extension period has been used the 68% of the time, while exemption by status has been used 24.5% of time. The exemption by time using a 45-minute extension appears to be most frequently used as the standard operational convention, while the exemption by status is applied in situations of sudden or unexpected volatility in AAR.

Exemption by status is equivalent to exemption by time, whenever the extension period is set to 0. Moreover, modifying the file time appropriately the same GDP program, i.e. the same set of included flight, $\overline{F}$, is obtained independently of the criteria of exemption used.

Henceforth, in this analysis, we consider only the file time. We carry out a distance-based GDP analysis for each file time ranging in the five-hour time interval before the program start time.

3.3 Optimizing over the Included Set of Flights

In this section, we provide a high level mathematical description of the optimization model for the selection of the GDP included set $\overline{F}$. In case of a distance-based GDP the radius of the circle for included airports, is denoted by d. All airports that are

inside the circle are included in the program. Thus, each

GDP initiative may be identified by the furthest airport included in the program. To model distance-based GDPs, we use the following notation:

$\overline{F}(d) = \{f \in \mathcal{F}$: the distance of the origin airport of f to the GDP airport $\leq d$ $\}$

$AVG(\hat{t}, \overline{F})$ = {the average assigned ground delay with file time $\hat{t}$ and included flight set $\overline{F}$}

$UNR(\hat{t}, \overline{F})$ = {the unrecoverable delay incurred with file time $\hat{t}$ and included flight set $\overline{F}$}

$AHD(\hat{t}, \overline{F})$ = {the airborne delay incurred with file time $\hat{t}$ and included flight set $\overline{F}$}

τ = the maximum amount of planned airborne delay

α = weight associated with average delay.

β = weight associated with unrecoverable delay.

$tier1$ = the set of tier 1 airports.

$tier2$ = the set of tier 1 and tier 2 airports.

all = the set of all GDP origin airports.

$\overline{F}$ = the set of airports within distance d of the GDP airport.

We can now define an optimization model to choose the distance parameter for the GDP. This model minimizes a weighted sum of the unrecoverable delay and the

average delay:

$$Min_{d:AHB(\hat{t},\overline{F}(d))\leq\tau}\{\alpha\ AVG(\hat{t},\overline{F}(d)) + \beta\ UNR(\hat{t},\overline{F}(d))\} \quad (1)$$

Currently, under the tier-based system the ATCSCC specialists must restrict choice of $\overline{F}$ to a very small number of possibilities. Although an optimization problem is not formally considered, we can view the problem they implicitly solve as:

$$Min_{\overline{F}\in\{tier1,tier2,all\}:AHB(\hat{t},\overline{F})\leq\tau}\{\alpha\ AVG(\hat{t},\overline{F}) + \beta\ UNR(\hat{t},\overline{F})\} \quad (2)$$

As we can see, in the optimization model (1) the possible distance options are given by a continuous set as opposed to the optimization model (2), which represents the current management operations where only three options are used.

4 Experimental Study

In this experimental study we compare the distance-based GDP with the tier-based one and we analyze the effects of file time. We executed a GDP program for those airports, which are considered more interesting for GDP operations, according to our experience and to practitioners. We selected at least one airport for each center, with the exception of those centers in which GDPs are not typically based. For each selected airport, we ran a GDP program, that is representative of those programs executed in the period between September 8, 1998 and April 6, 2000. Table 1 reports the airport, date, file time, start time, end time and in the last column, the AAR for the program executed. During the GDP period, the AAR may assume different

Airport	Date	File time	Start time	End time	AAR
Atlanta	2 Nov 99	16.54	19.00	23.00	76
Boston	2 Feb 00	17.02	19.00	00.59	40
Dallas	10 Oct 98	12.48	14.00	19.59	64
Detroit	6 Apr 00	15.14	16.00	21.59	36
Newark	27 Feb 00	15.19	18.00	23.59	36 (3)/40 (3)
Houston	24 Feb 00	17.02	19.00	00.59	40
Chicago	3 Feb 00	16.51	18.00	01.59	76 (2)/80 (6)
Philadelphia	13 Aug 99	18.04	20.00	01.59	40
S. Francisco	18 Aug 99	13.32	16.00	21.59	30 (3)/40 (3)
St. Louis	19 Mar 99	12.58	16.00	21.59	32

Table 1: Ground Delay Programs executed

values, therefore, in brackets, we reported the number of hours that AAR assumes the specified value. When not specified, the AAR assumes the reported value for all of the ground delay period.

Since the results of our analysis did not vary substantially by airport, in this paper, we report only the results obtained for the program executed at Boston's Logan airport on February 2, 2000. Those interested in more details are referred to the presentation given at the CDM Meeting on June 28, 2000 [8]. Included in the presentation are the results of the optimization model, described in § 3.3.

In all the graphs presented in this paper, we report the value of the statistics on the ordinate, while either the distances, expressed in nautical miles, or the file times are reported on the abscissa.

4.1 Varying the Distance Parameter

When an imbalance between demand and capacity takes place, the total amount of delay required to balance demand and capacity, i.e. the sum of airborne delay and

ground delay, is constant. The total delay depends only on the AAR values and the flight demand at the airport. Therefore, the ground delay and the airborne delay are complementary with respect to total delay (see Figure 2). The purpose of a GDP program is to shift airborne delay to cheaper and safer ground delay. To accomplish this task a certain number of flights should be included in the program. This can be viewed in Figure 2, where, for the shortest distance program, airborne delay has a spike, because the number of included flights is too small to mitigate the imbalance between demand and capacity. Increasing the radius of the circle, i.e. including more airports in the program, enlarges the pool of flights that receive ground delay and leads to a decrease in airborne delay. Beyond a certain distance, the airborne delay is almost constant. This level of airborne delay provides a threshold on the set of feasible distances, since a program distance shorter than the point where airborne delay levels off, is not allowed, because the amount of airborne delay is viewed as being unnecessarily high.

Once the necessity of implementing a GDP program is recognized, alternate initiatives are evaluated based on average delay and unrecoverable delay. These statistics are the surrogate measures for the cost of assigned ground delay and the expected cost of ground delay that is unnecessarily assigned. Average delay, as well as maximum delay, decreases with distance (see Figure 3), while unrecoverable delay has an increasing trend. It is important to note that unrecoverable delay has some singularities. For some particular distances, increasing the radius of the circle, i.e. adding airports to the program, leads to a decrease in unrecoverable delay rather than a increase, i.e. it is not a monotonically increasing function as

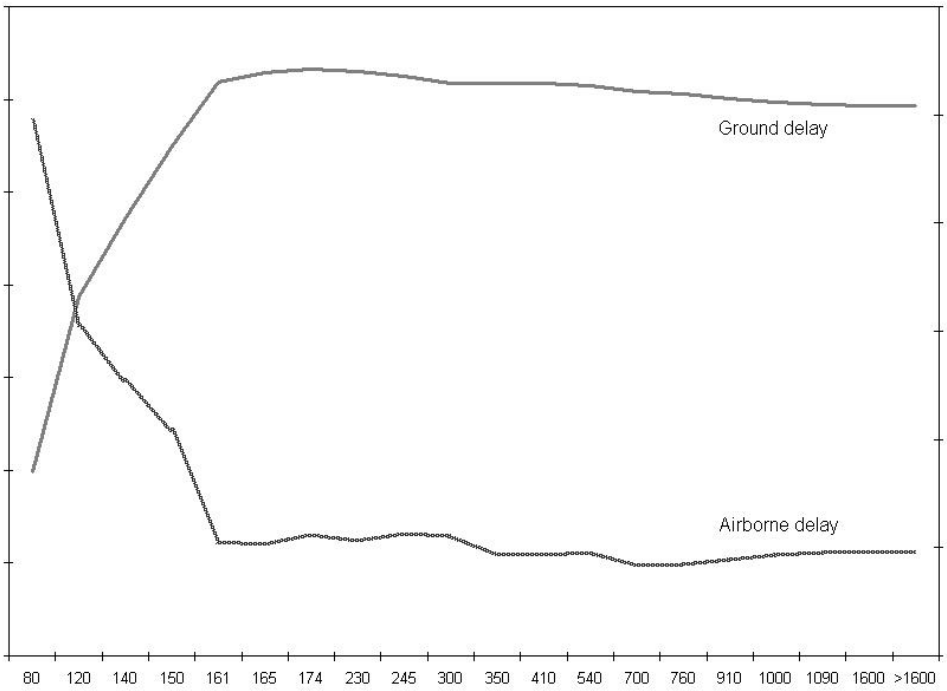

Figure 2: Airborne holding delay and total ground delay (dashed line).

one might expect. Unrecoverable delay is the part of delay that will be incurred even if the program is cancelled at its start time. That is, it consists of the portion of the ground that is taken prior to the start of the program. To understand this phenomena recall that as program distance is increased, more flights are included in the program and consequently each flight receives less ground delay and also the unrecoverable part is reduced. The impact on total unrecoverable delay is negative. At the same time, the added flights contribute to unrecoverable delay in a positive way. Generally, since the added flights are further out, the departure times tend to be earlier, which lead to greater unrecoverable delay. However, in some cases, the net variation is negative and the total amount of unrecoverable delay is reduced. These "anomalies" allow us to detect inefficient programs, i.e. programs that are not Pareto optimal. There is no reason to choose a lower program distance over a

larger one, if the larger distance program has a smaller average delay and a smaller unrecoverable delay. For example, in Figure 3, programs corresponding to 174, 410 and 1090 nautical-mile distances are efficient respect to their neighborhood. The statistic of delay variability is the only one of those listed in Section 3.1 that does not show a well-defined trend since it depends on many other factors that are not

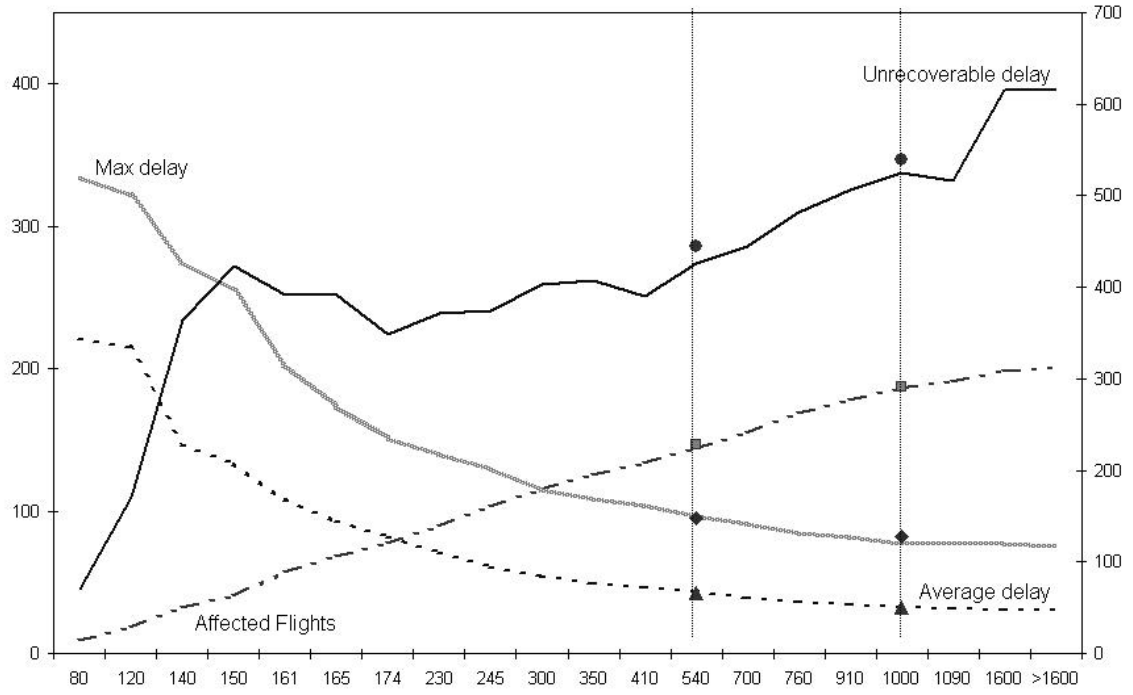

Figure 3: Tier vs. distance-based GDPs.

In Figure 3, we also report the statistics of the first-tier and second-tier initiatives, for the same GDP program. They are very similar to those of the 540 and 1000 nm distance programs, as evidenced by the vertical lines on the same figure.

Let us now consider applying the optimization model proposed in § 3.3. The weighted sum of average delay and unrecoverable delay with weights $(\alpha, \beta) = (50, 1)$ is plotted in Figure 4. The 1090 nm GDP is the optimal initiative. We note that

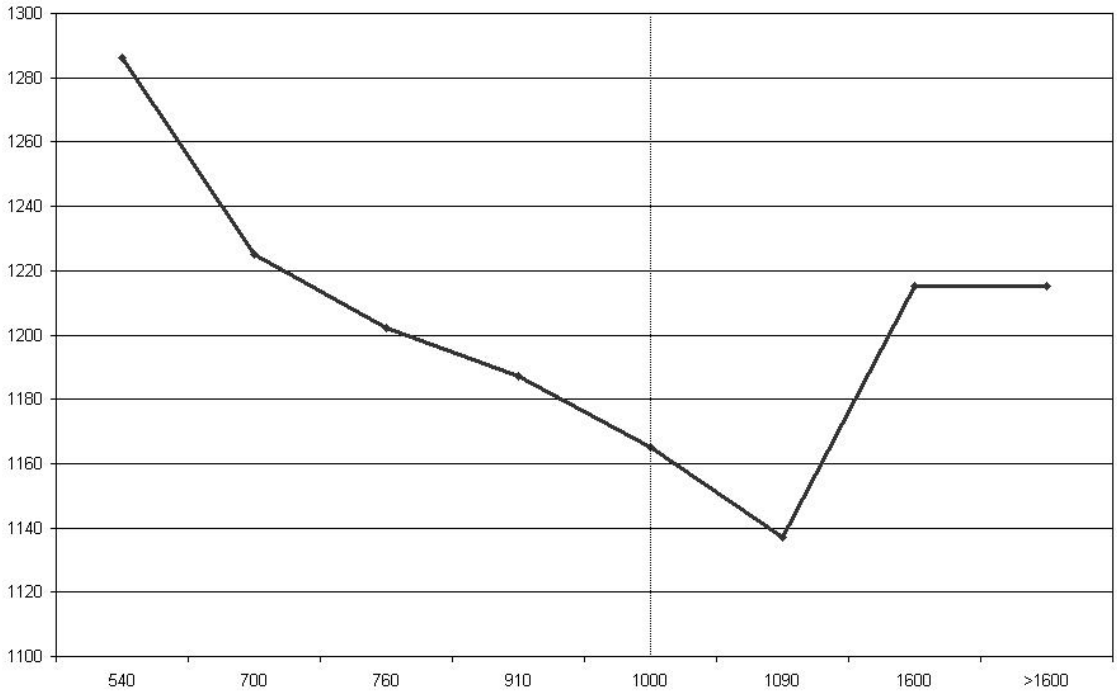

Figure 4: Weight trade-off between average delay and unrecoverable delay.

this solution is quite robust with respect to the weights used. In fact, the 1090 nm GDP is optimal for values of the ratio $\alpha/\beta \in [0, 98]$ (with $\beta \neq 0$) as reported in Table 2.

α/β	Distances		
	1000	1090	1600
0	525	517	615
98	3661	3555	3555

Table 2: Weighted trade-off for weights 0 and 98.

The GDP program suggested by the optimization model has a greater distance than the second tier one. Hence, not only can we find a distance-based initiative that is almost equivalent to each tier-based initiative, but, in addition, the distance-

based GDP allows air traffic managers more flexibility which can lead to superior programs.

4.2 Varying the Distance Parameter and File Time

File time also impacts the pool of flights included in a GDP program. In this section, we analyze varying both program and file time.

We now consider how GDP statistics vary as a function of file time. A late file time means a reduced number of included flights, and consequently, a larger value of airborne delay. As in the distance-based analysis, the airborne delay curve provides a threshold value that, in this case, indicates the latest admissible file time. Filing a program after the admissible time, leads to unacceptably high airborne delay. As depicted in Figure 5, airborne delay has a sharp increase for later file times.

In Figure 5, in addition to airborne delay, ground delay, average delay and unrecoverable delay are plotted for the 1090 nm distance GDP option. The scale of each statistic is different so we report the value of each statistic for the first and last File times. By definition, unrecoverable delay is 0 for a File time equal to the program's start time. For each distance option, the statistics show a similar trend, with the exception of average delay, which is the only statistic which has irregularities. In fact, the trend of average delay depends on which of total ground delay and number of affected flights decreases more quickly, since both decrease as the file time increases.

As in the distance-based GDP analysis, we may evaluate the best file time by applying the optimization model described in § 3.3. In Figure 6, the weighted

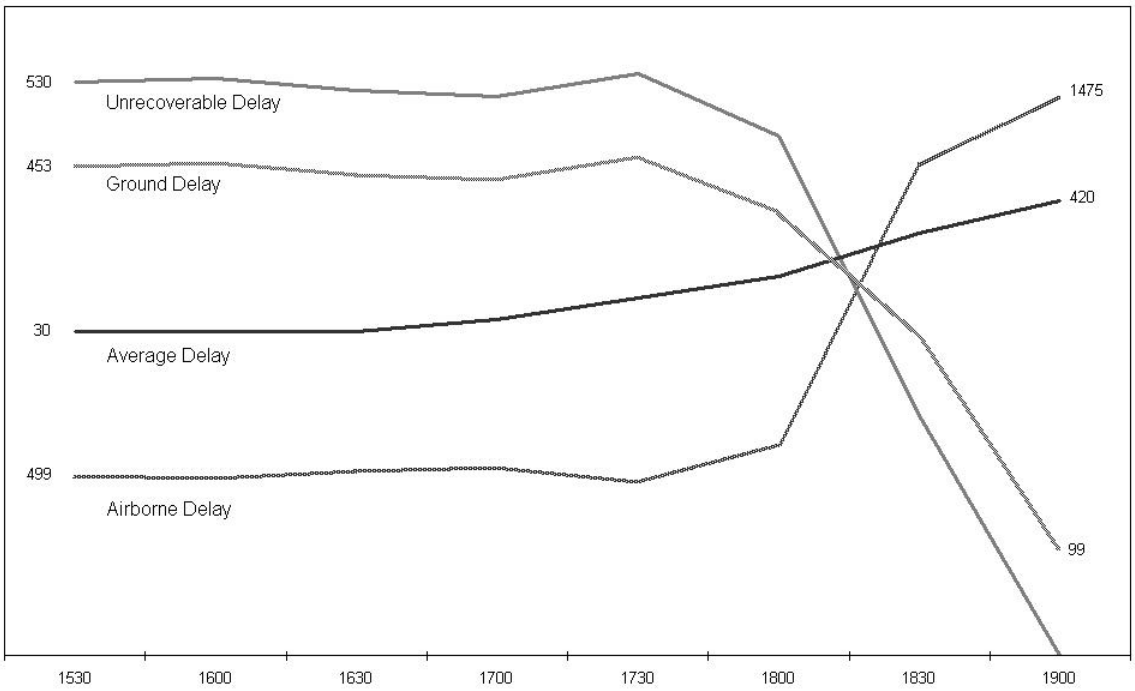

Figure 5: Statistics' trend in the departure-time analysis.

sum of average delay and unrecoverable delay for each distance option is plotted. For each File time, we are interested in the optimal distance parameter, which provides the minimum value for the weighted sum. The lower envelope of the trade-off curves defines the pairs of parameters, distance and file time, which produce the optimal GDP initiatives. As discussed earlier, we should only consider those pairs of parameters which are within the airborne delay thresholds.

In analyzing the trade-off curves, we may make several interesting observations. As evidenced by the optimal trade-off curves, when we consider a file time close to the start time of the program the optimization model always yields a program that includes the entire NAS, i.e. the largest possible distance value. This result because unrecoverable delay is not able to capture the true effects of a program cancellation. In fact, even though we file the program at its start time, there still is a possibility

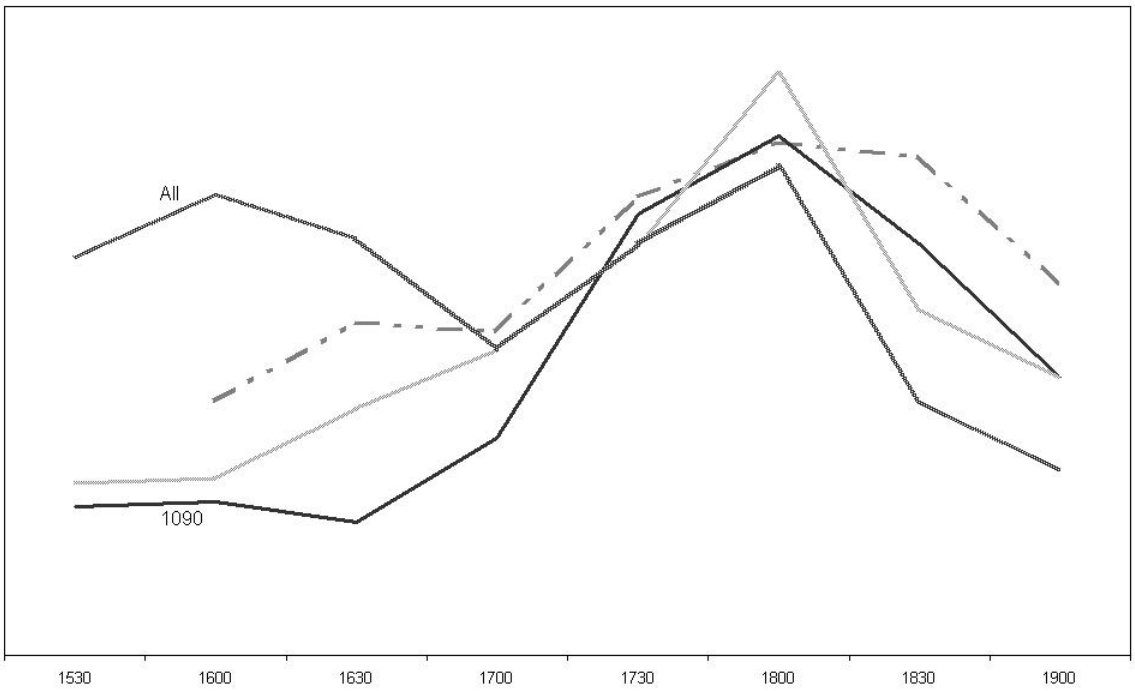

Figure 6: Optimal distance-based GDP options.

of canceling the program which results in an amount of unnecessary delay, while unrecoverable delay is 0 by definition. Hence, in this case, the optimization model minimizes only average delay.

5 Better Surrogate Measures.

Under current operations, average delay and unrecoverable delay are used to evaluate alternate GDP initiatives, i.e. they are the surrogate measures for the cost of assigned ground delay and the expected cost of ground delay that was unnecessarily assigned. Even though they are easy to compute, they fail to accurately represent GDP costs. In this section, we investigate possible drawbacks of using unrecoverable delay and present a new statistic named *unnecessary delay*.

5.1 Unnecessary Delay

In § 4.2, we have shown a possible drawback of using unrecoverable delay. To overcome this pitfall, we propose a more general statistic called *unnecessary delay*. The unnecessary delay for a flight is defined as the delay that has already been incurred, in vain, when a program is cancelled. We define the unnecessary delay for a flight as

$$unnec(CT, f) = \begin{cases} min\{CT - OTD_f, ETD_f - OTD_f, ETA_f - CT\} \\ \qquad\qquad\qquad\qquad \text{if} \quad OTD_f < CT < CTA_f; \\ 0 \qquad\qquad\qquad\qquad \text{otherwise.} \end{cases}$$

where CT is the cancellation time, OTD and ETD are the original and the estimated time of departure and ETA is the estimated time of arrival. Unnecessary delay for a GDP is defined as

$$unnec(CT) = \sum_{f \in \overline{F}} unnec(CT, f) \tag{3}$$

By definition, *unnec(CT)* is a function of cancellation time, and, since CT is a random variable, we are led to consideration of stochastic models. We note that unrecoverable delay is equal to *unnec(ST)* where ST is the start time of the program.

Figure 7 reports the unnecessary delay curves for different cancellation times for the GDP program executed at San Francisco on August 18, 1999. Note that in most the cases, *unnec(CT)* increases with program distance (this will not always hold).

With GDP distance fixed, we can observe that unnecessary delay is a concave function of the cancellation time (this is a general property that can be easily deduced from the definition). Unnecessary delay initially increases in value as a func-

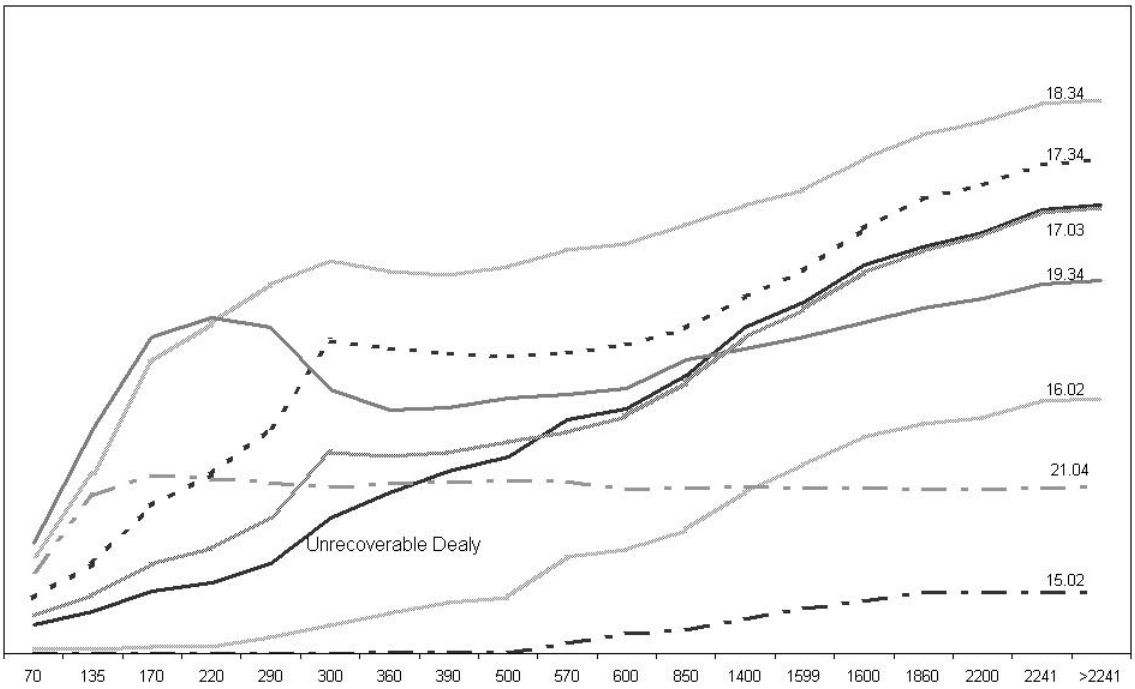

Figure 7: Unnecessary delay.

tion of cancellation time. But, when the cancellation time gets close to the end of the program, unnecessary delay decreases, because the majority of flights either have already arrived or are close to arriving, thus reducing their contribution to unnecessary delay.

Finally, note that, unnecessary delay is zero if the program is not cancelled or file time equals cancellation time, i.e. the cancellation time is equal to the end time of the program. This is true for any distance option.

5.2 Optimization Model Results Using Unnecessary Delay

In this section, we present the results obtained using unnecessary delay as surrogate measure for the expected cost of ground delay that is unnecessarily assigned in the optimization model.

As seen before, unnecessary delay is defined as a function of GDP's cancellation time. Cancellation time is represented by a random variable with a specific probability distribution for each GDP program. This probability distribution can be computed using historical data. Given a distribution for CT, we can easily compute the corresponding distribution for *unnec(CT)*. Using a probability distribution for CT for San Francisco airport(see Innis [7]), we were able to compute the expected value of unnecessary delay.

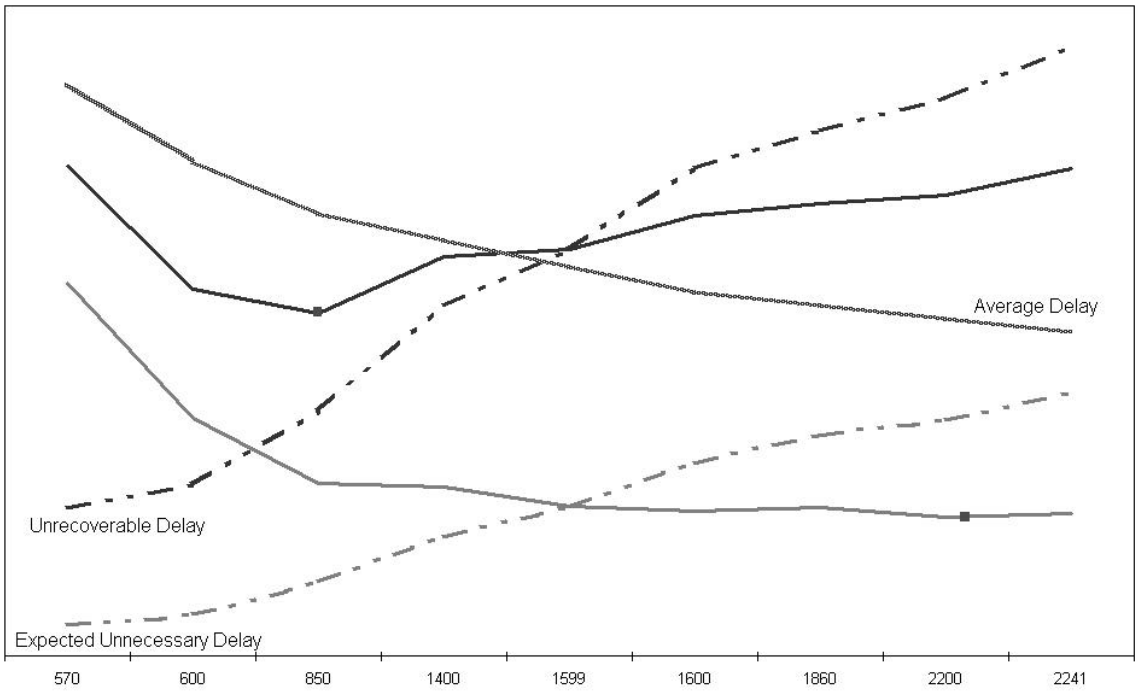

Figure 8: Trade-off between average delay and unnecessary delay.

Figure 8 gives the weighted sum of unnecessary delay and average delay to which our optimization model was applied. The black line represents the weighted sum of average delay and unrecoverable delay, while the grey one represents the weighted sum of average delay and expected unnecessary delay. The black and grey lines represent the weighted trade-off between the average delay and unrecoverable delay

and average delay and expected unnecessary delay respectively.

The optimal distance-based GDP option is designated in the figure with a square for both of the weighted sum curves. As depicted in the figure, the optimal distance based on unrecoverable delay is less than the optimal distance based on expected unnecessary delay. This situation is specific to the program at San Francisco airport and there is no reason to expect this result in general. We are not able to experiment with other airports, since the required cancellation time probability distribution are not available to our knowledge.

6 Conclusions and Future Research

In this paper we described a new approach for defining the included set of a Ground Delay Program. Our analysis has shown that the use of distance-based GDP can substantially improve GDP quality. This approach has been implemented and is being integrated into a future release of FSM, the CDM decision support tool.

As part of our analysis, we have shown the criteria used to evaluate GDPs, unrecoverable delay and average delay, are surrogates for more complex measures. We have provided some insights into possible better approximations and feel that further research in this direction is merited. An other worthwhile direction is to apply optimization to a larger GDP parameter set, e.g. as we here suggested for file time as well as distance.

Acknowledgements

This work was supported by the National Center of Excellence for Aviation Operations Research, under FAA Research Grant Number 96-C-001 and contract numbers DFTA03-97-D00004 and DTFA0101C00031. Any opinions expressed herein do not necessarily reflect those of the FAA or the U.S. Department of Transportation.

References

[1] Ball, M.O., C.Y.Chen, R.L.Hoffman, T.Vossen: *Collaborative Decision Making Air Traffic Management: Current and Future Research Directions*, to appear in New Concepts and Methods in Air Traffic Management, Bianco-Dell'Olmo-Odoni (eds.), Springer Verlag, 2001.

[2] Ball, M.O., R.L.Hoffman, D.Knorr, J.Wetherly, M.C.Wambsganss: *Assessing the Benefits of Collaborative Decision Making in Air Traffic Management*, 3^{rd} USA/Europe Air Traffic Management R&D Seminar, Naples , June 2000.

[3] Beatty, R., R.Hsu, L.Berry, J.Rome: *Preliminary Evaluatin of Flight Delay Propagation trough an Airline Schedule*, April 1998, http://www.metsci.com .

[4] *Collaborative Decision Making in Aviation Transportation* Web Pages, October 2000, http://www.metsci.com/ .

[5] Free Flight Phase I Office, Federal Aviation Administration: *An Operational Assessment of Collaborative Decision Making in Air Traffic Management*, DCN: R90145-01, 1999.

[6] Hoffman, R.L.: *Integer Programming Models for Ground-Holding in Air Traffic Flow Management*, PhD dissertation, University of Maryland, 1997.

[7] Inniss, T.R.: *Stochastic Models for the Estimation of Airport Arrival Capacity Distributions*, PhD dissertation, University of Maryland, 2000.

[8] Lulli, G., M.O.Ball, R.L.Hoffman: *GDPs - Based on Distance vs. Tiers*, CDM Meeting, June 2000.

[9] Shisler, L.: *Mileage vs. Tier Initiatives*, CDM Meeting, March 2000.

[10] Shisler, L.: *Historical Use of "Exempt by Time/Status"* , CDM Meeting, June 2000.

[11] Wambsganss, M.C.: *Collaborative Decision Making Through Dynamic Information Transfer*, Air Traffic Control Quarterly **4**, 1996, pp. 107-123.

[12] Wambsganss, M.C.: *Collaborative Decision Making Air Traffic Management*, to appear in New Concepts and Methods in Air Traffic Management, Bianco-Dell'Olmo-Odoni (eds.), Springer Verlag, 2001.

Appendix

The NAS is divided into 20 centers as shown in Figure 9. For each center a first tier and second tier are defined. The first tier is the set of all centers immediately adjacent to

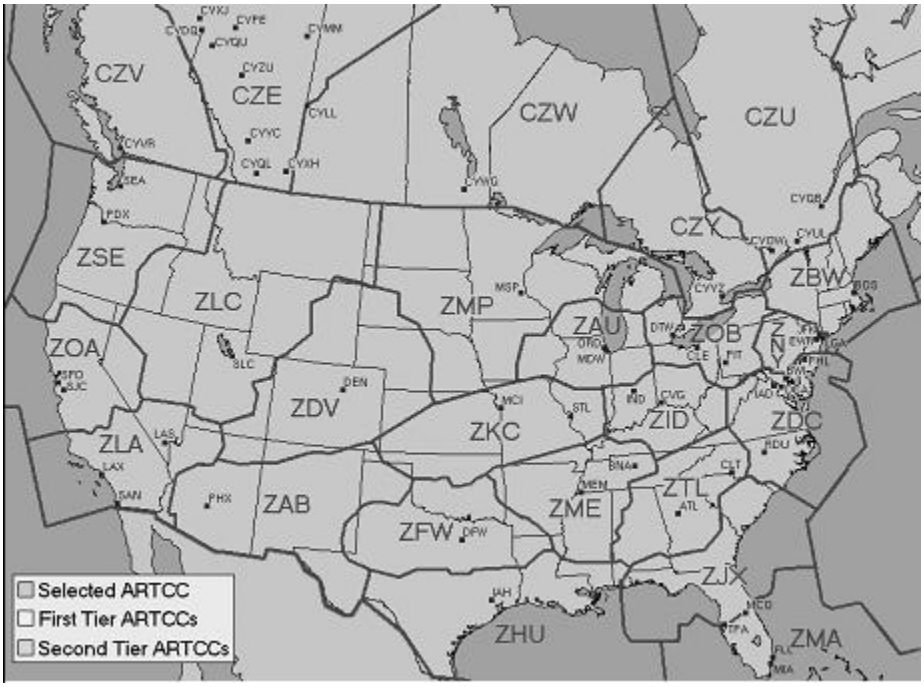

Figure 9: US National Airspace.

the center in consideration, while the second tier is the first tier plus all centers immediately adjacent to the first tier centers, and so on. For the western part of the NAS centers are also grouped in groups called 6West, 10West and 12West. In the group 6West, the centers that are grouped are the six most western sectors; the other groups are defined analogously. In Table 3, the centers that belong to the first tier and to the second tier are listed for each center.

Center	First Tier	Second Tier
ZAB	ZDV ZFW ZHU ZKC ZLA	ZAU ZID ZJX ZLC ZME ZMP ZOA ZTL
ZAU	ZID ZKC ZMP ZOB	ZAB ZBW ZDC ZDV ZFW ZLC ZME ZNY ZTL
ZBW	ZDC ZNY ZOB CYYZ CYOW CYUL	ZAU ZID ZJX ZMP ZTL
ZDC	ZBW ZID ZJX ZNY ZOB ZTL	ZAU ZHU ZKC ZMA ZME ZMP
ZDV	ZAB ZKC ZLA ZLC ZMP	ZAU ZFW ZHU ZID ZME ZOA ZOB ZSE
ZFW	ZAB ZHU ZKC ZME	ZAU ZDV ZID ZJX ZLA ZMA ZMP ZTL
ZHU	ZAB ZFW ZJX ZME ZTL	ZDC ZDV ZID ZKC ZLA ZMA*
ZID	ZAU ZDC ZKC ZME ZOB ZTL	ZAB ZBW ZDV ZFW ZHU ZJX ZMP ZNY
ZKC	ZAB ZAU ZDV ZFW ZID ZME ZMP	ZDC ZHU ZLA ZLC ZOB ZTL
ZJX	ZDC ZHU ZMA ZTL	ZAB ZBW ZFW ZID ZME ZNY ZOB
ZLA	ZAB ZDV ZLC ZOA	ZFW ZHU ZKC ZMP ZSE
ZLC	ZDV ZLA ZMP ZOA ZSE CYYC CYEG	ZAB ZAU ZKC ZOB
ZMA	ZJX ZHU* ZNY*	ZDC ZFW ZHU ZME ZNY ZTL
ZME	ZFW ZHU ZID ZKC ZTL	ZAB ZAU ZDC ZDV ZJX ZMA ZMP ZOB
ZMP	ZAU ZDV ZKC ZLC ZOB CYYC	ZAB ZBW ZDC ZFW ZHD ZLA ZME ZNY ZOA ZSE
ZNY	ZBW ZDC ZOB ZMA*	ZAU ZID ZJX ZMA ZMP ZTL
ZOA	ZLA ZLC ZSE	ZAB ZDV ZMP
ZOB	ZAU ZBW ZDC ZID ZMP ZNY CYYC	ZDV ZJX ZKC ZLC ZME ZTL
ZSE	ZLC ZOA CYVR CYYC CYEG	ZDV ZLA ZMP
ZTL	ZDC ZHU ZID ZJX ZME	ZAB ZAU ZBW ZFW ZKC ZMA ZNY ZOB
6 WEST	ZAB ZDV ZLA ZLC ZOA ZSE	
10 WEST	ZAB ZDV ZFW ZHU ZKC ZLA ZLC ZMP ZOA ZSE	
12 WEST	ZAB ZAU ZDV ZFW ZHU ZKC ZLA ZLC ZME ZMP ZOA ZSE	

*Oceanic area only

Table 3: First tier and second tier centers