RATION-BY-DISTANCE WITH EQUITY GUARANTEES: A NEW APPROACH TO GROUND DELAY PROGRAM PLANNING AND CONTROL

Robert Hoffman, Metron Aviation, Inc., Herndon, VA, hoffman@metronaviation.com,

Michael O. Ball, Robert H. Smith School of Business and Institute for Systems Research, University of Maryland, College Park, MD, mball@rhsmith.umd.edu

Avijit Mukherjee, University Affiliated Research Center, University of California (Santa Cruz), Moffett Field, CA, avijit@ucsc.edu

Abstract: In this paper we describe ration-by-distance (RBD), a new allocation method to be used in ground delay program (GDP) planning. We show that RBD minimizes total expected delay, under certain assumptions related to the manner in which GDP's are dynamically controlled that. On the other hand, RBD has poor characteristics with respect to the equity of the allocation it produces. To address this issue, we define a constrained version of RBD as a practical alternative to allocation procedures used in operations today, and we show that it has superior overall performance.

1. Introduction

One of the primary responsibilities of FAA air traffic managers is to formulate and apply strategic initiatives to alleviate anticipated demand-capacity imbalances at airports. When such an imbalance is expected at an airport, traffic managers apply ground delays to flights bound for the troubled airport commensurate with the delays they would (theoretically) receive in an airborne queue. This prevents air traffic controllers, who have limited delay options once aircraft are airborne, from being inundated with unmanageable numbers of airborne or diverting aircraft. In effect, anticipated airborne delay is transferred back to the ground, where it can be managed in a safe and orderly manner.

The processes for imposing these ground delays are embodied in a traffic flow management initiative known as a *ground delay program* (GDP). In 2005 [1], there were over ***1,350*** GDPs implemented in the US, which applied delays totaling more than ***16.8 million*** minutes, distributed over ***530,000*** flights. Efficient and equitable execution of GDPs is a paramount concern for air traffic management.

A principal concern in planning GDP's is the maximization of throughput into the airport [2]. However, preserving equity among the competing airspace users has also emerged as a second performance criterion (see [3], [4]). Under the Collaborative Decision Making (CDM) initiative, the ration-by-schedule (RBS) principle has been accepted as the standard for equitable allocation. However, RBS is not applied in its pure form. Select flights are exempted and are not assigned any delays. For instance, flights already in the air at the time the GDP is planned clearly cannot be assigned ground delay. Flights outside a certain radius from the airport are also exempted [5]. There are multiple motivations for this discretionary exemption policy, but the scientific basis most germane to our analysis, is mitigation of capacity uncertainty. Flights originating farther from the GDP airport must serve their ground delay well in advance of their arrival at the airport. The amount of ground delay is based on predicted capacity reductions, i.e. adverse weather conditions, several hours into the future. Overly pessimistic forecasts mean that (in hindsight) some of the ground delay is served unnecessarily. By assigning greater proportion of delays to shorter-haul flights, ground delay decisions can be reactively adjusted, and the overall delays can be reduced, based on near-term weather forecasts.

In this paper, we develop a stochastic model of the assignment of ground delays in the presence of weather uncertainty and show that a new allocation principle, *ration-by-distance (RBD)*, maximizes expected throughput into an airport, i.e. minimizes total delay, if the ground delay program cancels earlier than anticipated. RBD taken to the extreme generates inequities, which we will demonstrate. To offset this, we propose a constrained version of RBD that preserves a specified level of equity. We show that for any chosen equity level, constrained RBD produces a more efficient GDP policy than today's GDP policy, with respect to uncertainty of GDP end time. Furthermore, we show that in our algorithm we can vary the equity level more uniformly compared to the current practice of distance-based flight exemptions, and thus propose a more flexible approach to plan for a GDP.

2. GDP Planning and Control

The GDP planning and control process is currently supported by the Flight Schedule Monitor (FSM) decision support tool [1], [2]. FSM helps traffic managers determine and issue the appropriate amount of delay for flights involved in a GDP. The operational details of the tool and its use are too involved to cover in this paper. In this section, we provide just enough detail to convey the content and benefits of our approach to GDP planning and control.

2.1 GDP Planning

A *GDP plan* requires the assignment of ground delays to an *included* set of flights bound for a single airport with a predicted capacity-demand imbalance. The inclusion set is defined as those flights scheduled or estimated to arrive during the GDP planning horizon. Typically, the planning horizon coincides with a weather-induced period of reduced arrival capacity and is six hours in duration (the average for year 2006).

Let $f_1,..., f_n$ be the set of included flights. For each flight f_k, let d_k and a_k be the scheduled departure and arrival times, respectively, (For flights operating without a published schedule, estimated departure and arrival times are used as a surrogate.) When the GDP is planned, the FAA issues a *controlled time of departure time (CTD)* and a *controlled time of arrival* (CTA), denoted d'_k and a'_k, respectively. The *assigned ground delay* is given by $g_k = d'_k - d_k \geq 0$. We assume a known and deterministic en route travel time, $L_k = a_k - d_k$. Thus, $L_k = a'_k - d'_k$, and the CTA is shifted by the amount of the ground delay. That is, $a'_k = a_k + g_k$.

We model the reduced airport acceptance rate (AAR) by creating a set of slots, with corresponding slot times $s_1,...,s_n$. These are commensurate with the capacity reduction, e.g. for an AAR of 30 aircraft per hour, 30 slots would be defined in each hour. Of course, the AAR may vary over time. When stochastic planning models are used, the AAR is not known in advance, so a *planned AAR (PAAR)* is defined, based on stochastic information and strategy. Thus, the PAAR is quite likely to differ from the realized AAR.

Because each a'_k is a rigid translation of the corresponding d'_k, the GDP plan is completely determined once the a'_k variables have been chosen. Once the GDP planning horizon is set (not a topic of this paper), the planning problem reduces to choosing each a'_k from the slot times. This process can be formalized as an assignment model by setting the following binary variables,

$$x_{kj} = 1 \text{ if } f_k \text{ is assigned to } s_j\text{; 0 else}$$

under the following constraints:

$$\sum_{k:s_j \geq a_k} x_{kj} = 1 \quad \text{for } k=1,\ldots,n,$$

$$\sum_{j:s_j \geq a_k}^{n} x_{kj} \leq 1 \quad \text{for } j=1,\ldots,n,$$

These constraints ensure that each flight receives a slot, and that no slot is used more than once. From an efficiency standpoint, planners would like to formulate a GDP plan that minimizes a weighted combination of total expected ground delay and total expected airborne delay subject to the above constraints.

Thus, on a basic level, the GDP planning problem is a very simple one. However, when one considers explicitly dynamic and stochastic aspects, the problem and its analysis becomes considerably complex, as we will see in the next subsection.

2.2 GDP Control Dynamics

After the initial GDP plan is developed and the CTD's are issued, various stochastic elements invoke changes in the PAAR and, therefore, positive and negative variations in the CTD's. Several models from the literature, e.g. [6], [7], [8], address problems of this type. While all these models can be classified as stochastic, they differ in the degree and manner in which they model GDP dynamics. In this paper, we consider a specific GDP dynamic and stochastic planning problem, namely, the strategy for terminating and exiting a GDP.

The stochastic problem of concern for us is that the AAR is generally unknown. In particular, the time at which the AAR is restored (full arrival capacity) is not known precisely until it happens. There are also minor variations in the AAR during periods of reduced capacity before full restoration but these are usually very minor and are swamped by the uncertainty in demand. From a stochastic planning perspective, the PAAR should be changed to match the AAR.

The dynamic aspect of concern is that the ability to achieve a desired PAAR (i.e. deliver aircraft at a targeted rate) is affected by prior choices of CTDs and CTAs. That's because the airborne and ground-based 'inventory' of flights at any given time is determined by the CTDs executed in the past. For instance, suppose that at a time of significant capacity increase, all the short-haul flights have already departed (under the CTDs issued to them). Now if a sudden surge of flights is called for, this can come only from the ground-based inventory. But with only long-haul flights on the ground (far away), it will take considerable time for the surge to arrive. In turn, this leads to a period of airport underutilization realized as excessive ground delay. Of course, flight stage lengths do not fall neatly into two categories; we are merely highlighting a basic principle of inventory control.

Conversely, the airport might experience a longer duration of reduced capacity, or even a sudden drop in capacity; in either case, the GDP needs to be extended However, in this paper, we acknowledge only capacity increases once the GDP is in progress. The validity of this assumption is surprisingly well supported by analysis of GDPs over the last six years. For sake of brevity, we reserve that analysis and a relaxation of the capacity increase assumption for a future presentation. (Although later in Section 4, we discuss the consequences of various slot allocation strategies when the program is extended.) Here, the reader need only appreciate that the GDP exit strategy is a major concern in GDP control.

Before we present our proposed solution, we need to state more precisely what we mean by an exit strategy and impose one more assumption. By *exit strategy*, we mean an adjustment of CTDs (and therefore CTAs) for flights still on the ground, in anticipation of GDP termination. The added assumption is that traffic managers follow a GDP *cancellation policy* (CP), meaning that once capacity has been restored, all flights currently being held by the GDP (past their scheduled departure time) are free to depart. CP is in marked contrast to a policy in which flights are released gradually. Also, CP clearly assumes that the rise in capacity is sufficient to accommodate the pent-up demand in future time periods. Again, we emphasize that the validity of our results do not require the capacity rise or CP assumptions, but their acceptance greatly facilitates presentation.

Although the time of cancellation, T_c, is at the discretion of the FAA traffic managers, it directly depends on changes in weather conditions. Therefore, it can be modeled as a random variable. We associate a discrete probability distribution with the cancellation time $p_t = \Pr[T_c = t]$. In general, this distribution depends on weather characteristics.

For a flight *k*, we define the random variable $D_k(i,t)$ as the departure time of flight *k* with $T_c = t$ and $a'_k = i$, and we define $G_k(i,t)$ as the random variable corresponding to the ground delay faced by that flight. Now, it is easy to see that:

$$D_k(i,t) = Min\{i - L_k, Max\{t, d_k\}\}$$

Further, we have the following equalities:

$$\begin{aligned} G_k(i,t) &= D_k(i,t) - d_k \\ &= Min\{i - L_k, Max\{t, d_k\}\} - (a_k - L_k) \\ &= Min\{i - L_k + L_k, Max\{t, d_k\} + L_k\} - a_k \\ &= Min\{i, Max\{t + L_k, a_k\}\} - a_k \qquad (1) \end{aligned}$$

Under the CP assumption, there is no airborne delay so the efficiency metric of interest is total ground delay. We define $GT(t)$ as the total ground delay incurred with $T_c = t$. Of course, $GT(t)$ depends on the GDP plan so that $GT(t) = \sum_k G_k(a'_k, t)$. A GDP plan that maximizes expected efficiency would be a set of valid a'_k variables that minimizes total expected ground delay, GT, which is computed via $GT = \sum_t p_t GT(t)$.

3. Ration-by-Distance

We derive a ration-by-distance (RBD) rationing algorithm that yields minimal expected delay. This policy is actually quite close to, but improves upon, the approach used in practice. Later in the paper, we discuss the degree to which practice deviates from this policy and the potential implication. We also discuss and model the very important issue of equity and develop a practical approach that considers equity. First, we precisely define the distance-based RBS allocation algorithm (DB-RBS), which is in use today.

<u>DB-RBS Algorithm</u>

Step 0. Choose a radius *r* about the GDP airport. For convenience, assume *r* is in minutes of flying time (rather than miles). Mark as exempt all flights f_k with estimate en route travel time greater travel than *r*, that is, $L_k > r$.

Step 1. Assign each airborne and exempt flight, f_k, to the slot closest to a_k. Let *F* be the list of remaining flights and let *S* be the list of remaining slots, sorted by increasing slot time and re-indexed by $j = 1,...,m$. Mark each $f \in F$ as unassigned.

Step 2. Process the slots in *S* as follows. For $j = 1,...,m$, find the unassigned flight $f_k \in F$ with the least a_k such that $a_k \le s_j$. Assign f_k to slot

s_j. That is, set $a'_k = s_j$. (If no such flight exists, leave s_j empty.)

End algorithm.

The two important features of DB-RBS to observe are that flights outside the radius r are exempt from delay (Step 1) and that flights inside the radius r receive slots according to earliest schedule arrival time (Step 2).

We now define our RBD allocation algorithm.

RBD Algorithm

Step 1. Assign airborne flights to slots as in Step 1 of the DB-RBS algorithm.

Step 2 *:* For remaining slots (ordered by increasing slot time) $j = 1,...,m$, find the unassigned flight, f_k, with the largest flying time L_k such that $a_k \le s_j$. Assign f_k to slot s_j. That is, set $a'_k = s_j$. (If no such flight exists, leave s_j empty.)

End algorithm.

The important differences between DB-RBS and RBD are that under RBD no flights are exempted and the priority ruled used in Step 2 changes from smallest a_k to largest L_k.

We will now show that RBD has a very significant property: it produces a GDP plan that minimizes total expected delay. To prove this, we first we show that an elementary slot exchange, called *a long-short (LS) swap*, always improves or preserves total ground delay. Given any allocation of flights to slots, a LS swap is an exchange of the assigned slots between two flights, f_1 and f_2, such that $L_1 \ge L_2$, $a'_1 > a'_2$ and $a_1, a_2 \le a'_2$. In other words, in the initial assignment, a longer-haul flight (f_1) has been assigned to a later slot than the shorter-haul flight (f_2), and the LS swap reverses the assignment. To prove our main results we need the following elementary inequality.

Lemma: *Let* u_1, u_2, v_1, v_2 *be real numbers with* $u_1 \le u_2$ *and* $v_1 \le v_2$. *Then we have the following inequality*:

$$\min(u_1, v_2) + \min(u_2, v_1) \le \min(u_1, v_1) + \min(u_2, v_2).$$

Proof: Suppose that $u_1 \le v_1$. By assumption, $v_1 \le v_2$, so we have $u_1 \le v_1 \le v_2$. This means that $\min(u_1, v_2) \le \min(u_1, v_1)$. Since $v_1 \le v_2$, we also have that $\min(u_2, v_1) \le \min(u_2, v_2)$. Summing the last two inequalities, we obtain the desired result. Now we must consider the opposite case, in which $v_1 \le u_1$. But this case is symmetric, meaning that the same logic can be applied by reversing the roles of the u's and the v's. ■

Proposition:

If $L_1 \ge L_2, i < j$ *and* $a_1, a_2 \le i$, *then*

$$G_1(i,t) + G_2(j,t) \le G_1(j,t) + G_2(i,t).$$

Proof: From equation (1) we have:

$$G_1(i,t) + G_2(j,t) = \min\{i, \max\{t + L_1, a_1\}\} - a_1 + \min\{j, \max\{t + L_2, a_2\}\} - a_2 \quad (2)$$

Case 1: $t + L_2 \ge a_2$. In this case equation (2) can be rewritten as:

$$G_1(i,t) + G_2(j,t) = \min\{i, \max\{t + L_1, a_1\}\} - a_1 + \min\{j, t + L_2\} - a_2 \quad (3)$$

Since $L_1 \ge L_2$, $t + L_1 \ge t + L_2$ and it follows that $\max\{t + L_1, a_1\} \ge t + L_2$. Thus, equation (3) has the following form:

$$\min\{u_1, v_2\} + \min\{u_2, v_1\} - a_1 - a_2$$

where $u_1 \le u_2$ and $v_1 \le v_2$. If follows from the Lemma that

$$\min\{i, \max\{t + L_1, a_1\}\} + \min\{j, t + L_2\} \le \min\{j, \max\{t + L_1, a_1\}\} + \min\{i, t + L_2\}$$

which completes the proof for case 1.

Case 2: $t + L_2 < a_2$. In this case, equation (2) can be rewritten as:

$$G_1(i,t) + G_2(j,t) = \min\{i, \max\{t + L_1, a_1\}\} - a_1 + \min\{j, a_2\} - a_2 \quad (4)$$

By hypothesis, $a_2 < j$ so the last two terms of equation (4) sum to zero. Also by hypothesis, $i \le j$, so it easily follows that:

$$\min\{i, \max\{t + L_1, a_1\}\} - a_1 \le \min\{j, \max\{t + L_1, a_1\}\} - a_1$$

which completes the proof for case 2. ■

We can now give the basic result concerning RBD.

Theorem: *Let* $\{a'_k{}^*\}$, *the set of CTA's output by RBD, be used to compute* $GT(t)$ *then,*

(i) for any given cancellation time t, $GT(t)$ *achieves its minimum value;*

(ii) for any given cancellation time distribution $\{p_t\}, GT$ *achieves its minimum value.*

Proof: We prove (i) by contradiction and, thus initially assume it is not true. If this is the case there is an optimal CTA assignment, *AS**, different from the RBD assignment that has a smaller value of $GT(t)$. Proceeding from the earliest to the latest slot, consider the associated ordered list of flights in *AS**. Consider a similar list for the RBD solution. Since the solutions are different, there will be an earliest slot where they differ. Let f_2 be the flight that occupies that slot in *AS** and let f_1 be the flight that occupies it in the RBD solution. By the basic structure of RBD it must be the case that $L_1 \geq L_2$. Further, in *AS**, f_1 must be assigned to a later slot (otherwise, the first slot where the solutions differed would have occurred earlier). It follows that f_1 and f_2 can be interchanged in the optimal solution. Moreover, this is an LS swap so that by the Proposition, it will leave the value $GT(t)$ the same or reduced. We can continue this process until the two solutions are the same. If at any step $GT(t)$ is reduced, we have a contradiction on the optimality *AS**. Otherwise, the process ends with a proof that the RBD solution has the same value of $GT(t)$ as *AS**, also a contradiction, proving part (i).

It can easily be seen that part (ii) follows directly from part (i) since the $GT(t)$ are the variable coefficients in the expression for GT. ■

4. Equity Considerations and a Practical Approach

There are two general areas of concern that should be considered when evaluating the results of the previous section and the potential applicability of RBD. First, the Theorem relies on a very specific model of GDP dynamics, which is certainly not always followed under real conditions. Second, the Theorem indicates that RBD optimizes a system efficiency metric but says nothing about equity, the second important performance criterion. We will now argue that the first concern is actually somewhat minor and should not blunt in any significant way the impact of our Section 3 results. The second concern, on the other hand, is of great importance and motivated us to create a constrained version of RBD, which we feel has practical applicability.

As stated in Section 2, there can be a wide range of AAR adjustments over the course of a GDP. However, from a practical perspective the most common are program cancellations and program extensions. A program extension occurs when the planned end time is pushed further into the future relative to the originally planned end time. Generally speaking, the PAAR remains the same as the value specified for the original program. An extension largely involves the assignment of CTA's to flights that originally were scheduled to arrive after the planned end time. In fact, since these flights had previously received no assigned delays, the ability to assign delays to them is totally unaffected by the choice between DB-RBS and RBD. Additionally, flights that had received CTA's are sometimes assigned additional delay. The choice between DB-RBS and RBD does impact which, and how many, of these flights are on the ground when a decision to extend is made. Generally speaking, under RBD, there are fewer flights on the ground in the early stages of a GDP while there are more on the ground in the later stages of a GDP. We maintain that the decision to extend programs is usually made in the later stages of a GDP and thus RBD in fact improves the flexibility and the options available in planning an extension.

Thus, the RBD slot allocation policy has been proven to maximize efficiency (expected total delay) in the event of program cancellations and it also offers advantages under the most typical extension scenarios. We should note one case in which it can be detrimental. This involves the situation where the AAR drops to a value below the value for which the program was planned. In fact, RBD is counter productive in this case, as one would like to have the flexibility to dynamically assign delay or additional delay to short haul flights on top of the delay already assigned to long haul flights. Our analysis of GDP history has shown this case to be quite rare, especially relative to the number of instances of program cancellations and extensions.

A simple example will illustrate the potential shortfalls of RBD from an equity perspective. Consider a situation where a flight with one of the smallest value of L_k has a scheduled arrival time early within a 4-hour program. Such a flight will have the lowest priority throughout the allocation process and will very likely receive a very long delay, e.g. close to four hours. Such a situation would certainly be deemed unacceptable in practice. To effectively address equity considerations, we define an equity metric and associated constraint to maintain equity within the RBD process. We follow the approach of [3], which defines the inequity of a given allocation as its deviation from an ideal allocation (which also must be defined). In this case, we define the ideal as the (pure) RBS allocation. For each flight f_k, let a''_k denote the RBS slot assignment. We define the inequity imposed on f_k as the maximum positive deviation between a flight's assigned slot in its RBS slot:

$$Max_k(a'_k - a''_k)$$

We now define a constrained version of RBD that enforces an upper bound on this inequity metric. We call this allocation method *equity-based RBD (E-RBD)* . This upper bound leads to some very significant changes to RBD. Rather than directly creating an allocation, E-RBD first gives each flight a temporary slot assignment based on the application of RBS. A set of assignment exchanges is executed where each assignment-exchange gives a permanent slot assignment to one flight and adjusts the temporary assignments of one or more others. The assignment-exchanges are identified by choosing flights in order of decreasing value of stage length (L_k) and then executing an operation that assigns the chosen flight the earliest feasible slot. Before describing the algorithm it is useful to illustrate the assignment/exchange operation.

The most elementary form of the assignment-exchange moves the identified flight to an earlier slot and then "bumps forward" each of the intermediate flights to "make room" for the move. Such an operation is illustrated in Figure 1A. In this example, the delay on the targeted flight f_4, is reduced by the width of 3 slots and the delay on each of the 3 intermediate flights (f_1, f_2, f_3) is increased by the width of 1 slot. This exchange would give f_4 a permanent slot assignment and adjust the temporary slot assignments of (f_1, f_2, f_3).

Figure 1B illustrates a more complex assignment-exchange. In this case, the existing assignment of one of the intermediate flights, f_2 , is permanent. Therefore, adjustments to the temporary assignment of f_1 must take this into account. In this case, the delay of f_1 increases by the width of 2 slots. If s_j is the slot to which f_1 is reassigned we call this operation an f_1-to-s_j assignment/exchange. This operation is δ-feasible provided that:

(a) the current flight assignment for s_j is temporary, i.e. $a_k \leq s_j$;

(b) none of the delay increases to the intermediate flights induces a violation to the equity constraint with a right hand side of δ.

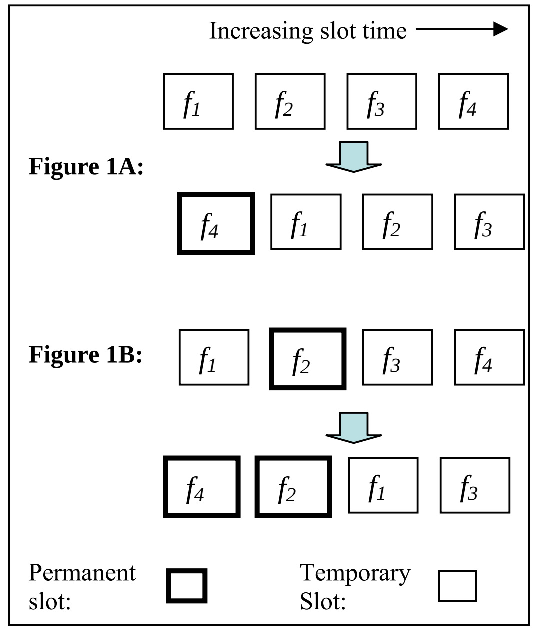

Figure 1: Examples of Assignment Exchange

We now define the equity version of the RBD algorithm.

<u>E-RBD Algorithm</u>

Step 0. Choose an equity deviation limit δ.

Step 1. Assign each airborne flight, f_k, to the slot closest to a_k and remove these flights and slots from the respective lists. Give each remaining included flight f_k a temporary slot assignment by setting a'_k to its (unconstrained) RBS slot. Order the remaining m flights by decreasing value of L_k .

Step 2. For $k = 1,...,m$:

find the earliest slot s_j such that the f_k-to-s_j assignment/exchange is δ-feasible;

execute this exchange and permanently assign the f_k to s_j .

End Algorithm.

It should be noted that in the later executions of Step 2, it will typically be the case that the earliest slot identified will be the one to which the flight is temporarily assigned. Thus, in such cases, the net effect of the Step 2 iteration will be to make the existing temporary assignment permanent. Of course, the early executions of Step 2 will implement the types of operations illustrated in Figure 1.

5. Experimental Results

We conducted a set of experiments to gain insight in the differences between the three rationing policies: DB-RBS, RBD and E-RBD.

5.1 Test Data and Scenarios

We constructed a test data set based on demand data from the ASPM database for San Francisco International airport (SFO) on August 11, 2005. Our scenario mimics a typical SFO morning GDP induced by the late burn off of marine stratus. The marine stratus conditions effectively eliminate side-by-side landings, thereby reducing the AAR from approximately 60 flights per hour to 30 flights per hour. Our demand data exceeded 30 flights per hour for three consecutive hours, from 9:00 to 12:00 local time. To accommodate this imbalance, we planned a 4-hour GDP from 9:00 to 13:00. The control times in the 12:00 hour were necessary to accommodate the pent-up demand from earlier hours. We evaluated five GDP cancellation times, one for the top of each hour during the program (9:00, 10:00,…, 13:00). Each of the three algorithms – DB-RBS, RBD, and E-RBD -- was evaluated under the five cancellation scenarios.

5.2 Experimental Results

Figure 2 shows the equity and efficiency evaluation of DB-RBS. The horizontal axis specifies the exemption distances in nautical miles that we tested, with higher values (i.e. smaller number of exempt flights) to the left. The vertical axis has two scales, the left one being for efficiency, measured as total minutes of delay, and the right vertical axis being for equity, measured as maximum deviation from the RBS allocation over all flights. The right axis is for the bars, our measure of equity. Note that the equity deviation ranges from nearly zero minutes (leftmost bar) to just over 160 minutes (rightmost bar). The dramatic increase in deviation (inequity) is to be expected, since the total amount of delay is absorbed by a shrinking pool of (non-exempt) flights. The rate of decline is essentially quadratic (for reasons we don't entirely understand). The plateaus correspond to ranges of distance in which no new flights bound for the GDP airport are encountered.

The efficiency of DB-RBS is measured by the five line plots – one for each cancellation time. Note that these are vertically stacked with the latest cancellation time being on top. This means that for any fixed exemption distance, the delay minutes drops each time the program is cancelled earlier. This makes intuitive sense, since more flights can be released earlier than their controlled times, hence reducing the total amount of delay.

Scanning any of the line plots left to right, we see that as the distance parameter decreases (thereby exempting longer-haul flights), the total delay decreases also, or levels off, in a nearly linear manner. This phenomenon confirms the fundamental principle of ration-by-distance: delay can be saved under early GDP cancellation by assigning a greater proportion of delays to the short haul flights as compared to long haul flights; the earlier the cancellation time, the greater the savings. Note also the equity-efficiency tradeoff in Figure 2: as equity goes up, efficiency goes down. This is the tradeoff associated with the distance parameter. It becomes more pronounced for earlier cancellation times.

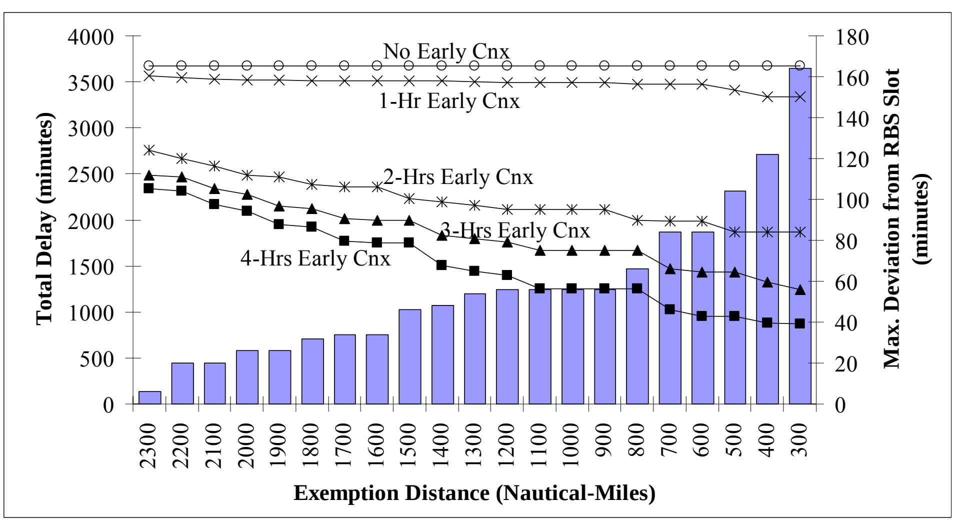

Figure 2: Performance Statistics for Distance-based Ration-by-Schedule (DB-RBS) Algorithm

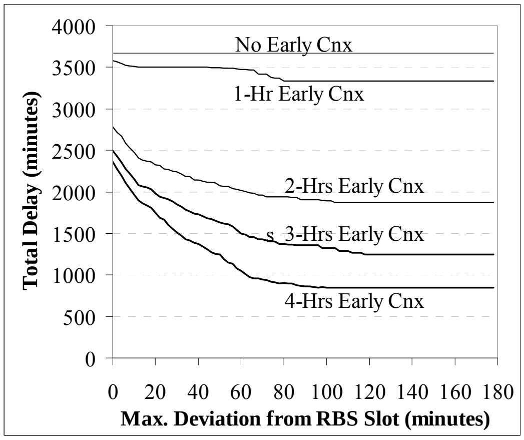

Figure 3: Performance of E-RBD

Figure 3 shows the performance of E-RBD: the horizontal axis is the maximum deviation parameter, δ, while the vertical axis is the total delay, in minutes. We did not 'plot' the equity of E-RBD (as we did in Figure 2) because, by our metric, it is directly controlled by parameter δ. The five line plots in Figure 3 correspond to the five possible GDP cancellation times. Note that the total delay decreases as δ increases. This is because higher values of δ allows more long-short swaps to take place. Also, as the program cancellation time gets earlier, the total delay decreases. The line plots in Figure 3 are very much like those of Figure 2, with two important exceptions: in Figure 3, they are nearly quadratic, while in Figure 2, they were nearly linear. (That's because the horizontal axis of Figure 3 is a quadratic translation of the horizontal axis of Figure 2.) Also, the total delay resulting from E-RBD is generally lower than that of DB-RBS (more on this later). Figure 3 shows the same equity-efficiency tradeoff that we saw in Figure 2.

We do not have a separate figure to evaluate the performance of RBD, because its performance is imbedded in Figure 3 as an extrema of E-RBD. The total delay resulting from RBD can be found at the far right of each cancellation-time line plot.

It is instructive to further compare the performance of RBS, RBD and E-RBD from another equity perspective. In Figure 4, we show the sum over all flights, the squared deviation from RBS under four different allocation strategies. Thus, the units are minutes squared. For E-RBD, we chose two representative values of the δ parameter: 20 minutes and 80 minutes. (These are values where the marginal decrease in delay became small as δ increased further.)

Figure 4 yields several interesting results. First, when the program is not cancelled, RBS has perfect equity (zero deviation). But when the program is cancelled early, RBS deviates from perfect equity because the RBS allocation is based on the planned 4-hour AAR reduction, and not the AAR that results from early cancellation of the GDP.

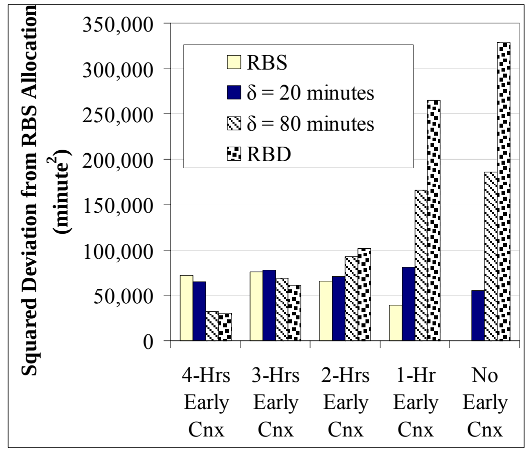

Figure 4: Equity Measures for RBS, RBD, and E-RBD

Second, as expected, RBD has very poor equity when the program is not cancelled; in hindsight, short-haul flights have been penalized to no end. On the other hand, since RBD saves significant delay under early cancellation times, it also registers less deviation from the ideal allocation. This is because more flights are allowed to depart closer to their schedule departure times. Third, as δ increases from 20 to 80 min under E-RBD, equity worsens, if the GDP does not cancel earlier.

Figures 3 and 4 provide an assessment of the efficiency and equity tradeoffs under E-RBD. If the GDP cancels two hours earlier than planned, E-RBD with δ=0 (i.e. RBS), δ=20 minutes, and δ=80 minutes, yields 49%, 25%, and 4% additional total delay compared to the RBD allocation. The equity gained from the three allocation policies, in lieu of the loss in efficiency, are 35%, 30%, and 9% respectively, compared to the RBD allocation. Clearly, under RBS, the 35% gain in equity, compared to RBD, is outweighed by the 49% loss in efficiency, if the GDP cancels two hours earlier than anticipated. Whereas, setting the parameter δ to a value of 80 minutes produces less percentage loss in efficiency than the gain in equity metric. In case of even earlier cancellation of the ground delay program, the RBD algorithm produces the most efficient and equitable allocation of slots to flights.

Our most prominent results are shown in Figure 5, which directly compares E-RBD to today's policy, DB-RBS, with respect to both equity and efficiency, for a 2-hour early cancellation time (chosen because it is typical in

GDPs). Each point shows the total delay (vertical axis) for a given of level of max deviation from RBS (horizontal axis). More efficient points sit lower on the graph, while more equitable points lie farther to the left.

Note that the E-RBD curve dominates the DB-RBS curve. Delay savings are in the order of 10%. Also note that unlike the E-RBD algorithm in which the parameter δ can be varied over a wide range of values, the DB-RBS produces only a small set of efficient points; this is because airports are not evenly distributed over distance. Often times, a slight increase or decrease in the exemption distance results in inclusion or exclusion, respectively, of flights from a set of airports. This can cause significant changes in the efficiency (i.e. total delay) and equity (i.e.δ) metrics; whereas, in the E-RBD algorithm, the efficient frontier varies more uniformly.

We have computed comparable results for the other four cancellation times as well (charts omitted for sake of brevity). The delay savings generated by E-RBD range from 0 to 19%, with the greatest savings occurring at the 4-hour cancellation time. This demonstrates the basic advantage of E-RBD over today's rationing policy.

6. Conclusions

We have described a new GDP slot rationing scheme, RBD, and we have shown that it minimizes total expected delay under the most typical GDP dynamic scenario. We have further described a second rationing scheme, E-RBD, that is practical in the sense that, unlike (pure) RBD, it takes into account both equity and efficiency factors. Our computational experiments show that not only is E-RBD comparable to the DB-RBS algorithm used in practice today, but in fact, it provides an efficiency advantage.

E-RBD has a second important advantage over DB-RBS. DB-RBS is driven by its distance parameter. As this parameter is increased, additional airports fall into the scope of the GDP, meaning that flights departing those airports must share in the total assigned ground delay. Often times, a slight change in the distance parameter can affect a large number of flights of one airline an entire (e.g. when the airport is a hub for one airline). This sensitivity of the distance parameter incites airlines to argue for or against specific distance parameters on a daily basis. In contrast, our E-RBD policy is driven by a very natural parameter (δ = maximum deviation from RBS), with a clear performance interpretation: a measure of equity. As such, it can be set based on more objective principles or on a value specified by FAA policy.

In addition, as it is changed, the impact on delays allocated to flights should be less abrupt than the impact of changing the DB-RBS distance parameter. Thus, the use of E-RBS admits a more scientific, and less political, basis for GDP planning. While a wider range of experiments and scenarios are certainly needed, our proven principles and demonstrated results provide a strong case for the adoption of E-RBD.

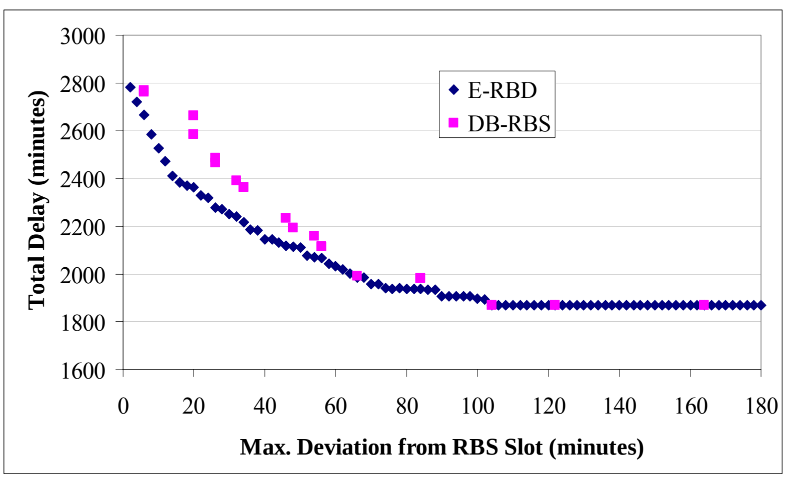

Figure 5: Efficient Frontiers for E-RBD and DB-RBS

Acknowledgements

This work was supported by the National Center of Excellence for Aviation Operations Research (NEXTOR), under FAA cooperative agreement number 01CUMD1. Opinions expressed herein do not necessarily reflect those of the FAA or the U.S. Department of Transportation.

References

[1] Federal Aviation Administration, CDM Web Site: http://cdm.fly.faa.gov/index.html.

[2] Wambsganss, M., 1996, "Collaborative Decision Making through Dynamic Information Transfer", *Air Traffic Control Quarterly*, **4**, pp. 107-123.

[3] Vossen, T., M. O. Ball, R. Hoffman, and M. Wambsganss, 2003, "A General Approach to Equity in Traffic Flow Management and its Application to Mitigating Exemption Bias in Ground Delay Programs", *Air Traffic Control Quarterly*, **11**, pp. 277-292. (a preprint was published in *Proceedings of 5th USA/Europe Air Traffic Management R&D Seminar*, 2003).

[4] Vossen, T. and M. O. Ball, 2006, "Optimization and Mediated Bartering Models for Ground Delay Programs", *Naval Research Logistics*, **53**, pp. 75 – 90.

[5] Ball, M.O. and G. Lulli, 2004, "Ground Delay Programs: Optimizing over the Included Flight Set Based on Distance", *Air Traffic Control Quarterly,***12**, pp. 1 – 25.

[6] Richetta, O. and A. Odoni, 1993, "Solving Optimally the Static Ground-Holding Problem in Air Traffic Control", *Transportaion. Science,* **27**, pp. 228-238.

[7] Ball, M.O., R. Hoffman, A. Odoni, and R. Rifkin, 2003, "A Stochastic Integer Program with Dual Network Structure and its Application to the Ground-Holding Problem", *Operations Research*, **51**, pp. 167 – 171.

[8] Mukherjee, A. and M. Hansen, 2004, "Dynamic Stochastic Model for A Single Airport Ground-Holding Problem", submitted for publication.

Key Words

Traffic flow management, ground delay program, stochastic model, dynamic model, collaborative decision making.

Biographies

Robert Hoffman is a principal analyst at Metron Aviation, Inc. He holds a Ph.D. in applied mathematics with a concentration in operations research from the University of Maryland. Currently, he is involved in modeling and analysis of strategic ATM initiatives - specializing in resource allocation and optimization issues. Dr. Hoffman is also a Visiting Research Scientist at the University of Maryland's Institute for Systems Research.

Michael O. Ball is the Orkand Corporation Professor of Management Science in the Robert H. Smith School of Business at the University of Maryland. He also holds a joint appointment within the Institute for Systems Research (ISR) in the Clark School of Engineering. Dr. Ball received his PhD in Operations Research in 1977 from Cornell University. He is co-Director of NEXTOR, the National Center of Excellence for Aviation Operations Research, and he leads the NEXTOR Collaborative Decision Making project.

Avijit Mukherjee is an Associate Project Scientist at the University Affiliated Research Center, University of California (Santa Cruz), Moffett Field, CA. He is involved in the research on Air Traffic Management at the Aviation Systems Division of the NASA Ames Research Center. Dr. Mukherjee received his Ph.D. in Civil and Environmental Engineering from University of California, Berkeley, in 2004.