

# TECHNICAL RESEARCH REPORT

The Rate Control Index for Traffic Flow

*by Robert L. Hoffman, Michael O. Ball*

**NEXTOR TR 2001-1  
(ISR TR 2001-24)**

# NEXTOR

---

National Center of Excellence in Aviation Operations Research

---

*The National Center of Excellence in Aviation Operations Research (NEXTOR) is a joint university, industry and Federal Aviation Administration research organization. The center is supported by FAA research grant number 96-C-001.*

Web site <http://www.isr.umd.edu/NEXTOR/>

# The Rate Control Index for Traffic Flow

Robert L. Hoffman, Michael O. Ball \*

March 26, 2001

**Abstract:** The objective of Air Traffic Flow Management is to maintain safe and efficient use of airspace and airports by regulating the flow of traffic. In this paper, we introduce a single-valued metric for post-operatively rating the performance of achieved traffic flow against targeted traffic flow. We provide variations on the metric, one of which factors out stochastic conditions upon which a plan is formulated, and show how these improve on current traffic control analysis techniques. The core of the metric is intuitive and simple, yet leads to an interesting optimization problem that can be efficiently solved via dynamic programming. Numerical results of the metric are given as well as a sample of the type of analysis that should follow a low rating by the metric. Although this metric was originally developed to rate the performance of Ground Delay Programs, it is equally applicable to any setting in which the flow of discrete objects such as vehicles is controlled and later evaluated.

---

\*Robert Hoffman is a senior analyst with Metron Scientific Consulting, 11911 Freedom Drive, Reston VA 20190. Email: hoff@metsci.com. Michael Ball holds a joint appointment with the Robert H. Smith School of Business and the Institute for Systems Research at the University of Maryland, College Park, MD 20742. Email: mball@rhsmith.umd.edu.

# 1 Introduction

The flow of traffic through airspace sectors and through airports is regulated by air traffic flow managers to ensure that airspace components do not become overloaded and that throughput is maintained. Many of the short-term demand-capacity inequities in the NAS are smoothed out by FAA traffic flow managers by tactics such as vectoring and miles-in-trail restrictions. Longer-term demand-capacity inequities (on the order of several hours) are more problematic. These usually occur when an airport acceptance rate is drastically reduced because of inclement weather, airport construction, or special runway operations. In these instances, air traffic flow managers at the FAA employ ground holding strategies in which aircraft bound for an afflicted airport are held at their points of origin in lieu of costly and hazardous airborne holding that would occur if they were allowed to depart on schedule. The most prominent of these strategies is the ground delay program (GDP), which is an initiative taken by the FAA to lower the arrival rate to a level that can be safely handled by airport controllers.

The increase in air traffic in the United States over the last 20 years has necessitated more frequent use of traffic flow management initiatives. For instance, in 1998, there were 187 ground delay programs run at San Francisco airport alone [1]. As a result, there has been considerable interest both in the aviation and research communities regarding traffic flow management and its analysis. The issues of deepest concern have been the efficient use of airport landing resources and the equitable distribution of arrival slots among competing airlines [2]. Substantial efforts have been under way since the mid 1990's to revamp the manner in which traffic flow initiatives are

planned and executed. Most notable is the joint industry-FAA Collaborative Decision Making project (CDM), which has made major changes to GDP procedures. See [3] for details on ground delay program enhancements.

One can see the need to assess the quality of traffic control actions. In this paper, we provide practical solutions for three major aspects of post-operative, traffic flow analysis. The first aspect is the need for a simple way of contrasting aggregate traffic flow with desired traffic flow. We introduce the rate control index (RCI), which gives a single performance value to the flow of traffic into an airport or sector of space as compared to the planned flow of traffic. The rate control index bears an intuitive relation to the events that have lead to the deviation from the plan. In essence, the metric tracks the aggregate flight movements that have caused the realized traffic flow to differ from the planned traffic flow. Although the basic concept behind the metric is easily understood, the normalization of the metric leads to an interesting optimization problem.

The second aspect is the need to factor out from post analysis major stochastic factors upon which a control action is based. Every traffic flow initiative is based on forecasts of demand and resource capacities, which are often not realized because they depend upon highly stochastic conditions. Air traffic demand predictions, such as the number of arrivals to an airport, are vulnerable to airline operational deviations while capacity predictions for airspace sectors, runways, etc., are highly subject to weather conditions. Direct measurement of traffic control actions is not always the best way to judge the performance of a new program or initiative; naive or ill-chosen metrics are heavily influenced by the quality of the forecasts upon which a

plan of action was based. For instance, in a GDP, the objective is to deliver a specified number of aircraft to the airport during a fixed time horizon. The metric generally used by traffic flow managers for program evaluation is landings-per-hour [4]. If runway conditions turn out to be more severe than previously forecasted, then this metric will (correctly) reveal that the realized landing rate does not match the desired landing rate. However, to a large degree, the program performance was beyond the control of participating parties and is being judged in part by the forecast upon which it was based. For situations such as this, we show how to meter traffic flow into an airport independently of the ability of the airport to land aircraft.

The third aspect of traffic flow analysis we address is the need for a metric that rates the performance of a group of controlled flights on a more individual basis than aggregate metrics. Directives are given to individual flights (altering flight paths, arrival times, etc.) in order to affect the aggregate flow of traffic. Final success of these efforts is typically measured using aggregate metering of traffic. When the desired traffic flow is achieved, the question arises whether or not this was due to the adherence of aircraft to their controlled arrival times or due to the fortuitous cancellation of early arrivals with late arrivals. To address this issue, we present a nominal version of the RCI metric that acts as a complement to the aggregate version of the metric.

Section 2 of this paper provides motivation for the metric and an interpretation of its meaning. The rigorous formulation of the metric and its normalization process are deferred to Section 3. Section 4 demonstrates the application of both the nominal and aggregate versions of the metric to Ground Delay Program analysis. Conclusions are in Section 5.

## 2 The Core of the RCI Metric

The perspective taken in this paper is one of post-analysis. We view a traffic flow initiative as a plan of action which is to be later analyzed for effectiveness. We assume that a traffic flow manager has set a goal for each time period  $t = 0, 1, \dots, T$  of a time horizon, meaning the number of vehicles that should be delivered to an airport (or more generally, pass through a region of airspace). A plan is enacted to achieve these goals. Once the time horizon has passed, the actual number of flights is recorded. This establishes two distributions: the planned distribution  $P = (p_0, p_1, \dots, p_T)$  and the realized distribution  $R = (r_0, r_1, \dots, r_T)$ , where  $p_t$  and  $r_t$  are the planned and realized number of flights during time period  $t$ , respectively. The question is how to weigh the performance of the realized distribution against the planned distribution.

We branch into two areas of concern, separated by a subtle but crucial distinction. One is the desire to measure program *performance*. Once a GDP is formulated, there are many factors that must fall into place in order for the GDP to be executed as planned. It is the duty of the controllers in the local traffic management units to hold flights at their departure airports until their controlled times of departures. The airlines must notify passengers of the imposed ground delay and yet have them boarded and ready to takeoff at their controlled departure times without further delay. Pilots must maintain an enroute airspeed commensurate with that was used in the forecast of their enroute times. One can see that there are many entities that participate in the execution of the plan other than the traffic flow specialists that have formulated the plan. A great deal of effort has been spent by the CDM com-

munity in improving GDP practices and performance. They have expressed a need to measure the execution of GDPs independent of the quality of its plan [5].

The other area of concern is the desire to measure program *impact*, meaning the effect that program performance has had on the system. Ideally, we would like a metric that produces a single, normalized value for the performance or impact of a program, so that programs can be tracked over long periods. This is especially valuable for discerning the effect of proposed program changes. Moreover, the metric should have a natural, intuitive interpretation to which cost benefits can be applied.

Given the planned and realized distributions,  $P$  and  $R$ , a straight forward approach to distribution comparison is to form a distribution of errors

$$E = P - R = [p_0 - r_0, p_1 - r_1, \dots, p_T - r_T] \quad (1)$$

by vector subtraction, then to analyze the variance of the errors by averaging over the time periods or summing the squares of the deviations. Consider the planned distribution  $P = (30, 30, 30, 30, 30)$  of airport arrivals, meaning that 30 flights are intended to arrive in each of five hours. Suppose that the actual number of arrivals per hour is given by  $R_1 = (30, 30, 30, 31, 29)$ . In this case, average error per hour is a poor measure; the excess of one flight in hour 4 cancels out with the shortage of one flight in hour 5 to produce a zero average. A more viable alternative is to compute the average absolute error, as below.

$$\begin{aligned} \text{Avg. Abs. Error} &= \frac{|0| + |0| + |0| + |1| + |-1|}{5} \\ &= \frac{2}{5} = 0.40 \text{ flights per hour} \end{aligned} \quad (2)$$

Although (2) correctly records an overdelivery in hour 4 and an underdelivery in hour 5, it fails to capture the events that have led to  $R_1$ . Specifically, since 31 flights have arrived in hour 4 and only 29 in hour 5, then there must have been (in the aggregate) a migration of one flight from hour 4 to hour 5.

The alternative measure of variance that we propose is to record the aggregate flight movement, or drift, that led to the deviation of  $R$  from  $P$ . This is the same as the minimum amount of flight movement that would be necessary to revert  $R$  to  $P$ . This can be viewed as an alternative way of measuring the difference between two distributions.

This type of approach makes distinctions between realized distributions that (1) does not. For instance, consider the flight movements necessary to convert  $R_2 = (31, 30, 30, 30, 29)$  into  $P$ . Since 31 flights arrived in the first hour when only 30 were intended, one of the flights intended to arrive in a later hour arrived in hour 1. Let's credit the participants with the least errors possible and say that this flight was intended to arrive in hour 2. Since 30 flights arrived in hour 2 instead of  $30 - 1 = 29$ , one of the flights intended to arrive in hour 3 must have arrived in hour 2. Similarly, one flight intended to arrive in hour 4 arrived in hour 3, and one flight intended to arrive in hour 5 arrived in hour 4. In all, we see that, as a minimum, there was a total migration of five flight-hours. Note that if we computed average absolute error, as in (2), then  $R_2$  would receive the same rating as  $R_1$ , 0.40 flights per hour. In contrast, the aggregate flight movement method recognizes that four times as much movement has taken place in  $R_2$  as in  $R_1$ .

The subtraction-based method (1) double-counts a single flight in hour 5. It can be argued that this correctly models the impact of the realized



distribution on the system, namely that the system is stressed in hour 4 and underutilized in hour 5. But this is based on the assumption that the airport capacity in each hour is as predicted, 30 aircraft per hour. Since 31 aircraft were accommodated in hour 4, the capacity of the airport clearly was not limited to 30 flights. Conversely, it is fairly common for an airport to accept fewer flights than were planned because airport acceptance rates are based not just on weather conditions and runway configurations, but also on departure demand and the abilities of air traffic service providers. For this reason, it can be difficult, if not impossible, to know what the actual airport capacity was, based solely on realized arrival rates. If these rates are known, metric (1) could be considered.

The approach described in this paper primarily compares the planned distribution against the realized distribution, and not against the airport acceptance rate distribution. The main idea is to assess how closely the plan was followed. For a given planned distribution,  $P$ , and a realized distribution,  $R$ , the method is based on shifting aircraft in the  $R$  distribution to transform it to the  $P$  distribution. The rate control index (RCI) metric is the outcome of this method. The following example shows how the RCI metric would be computed in practice and how cost parameters can be applied. In Section 4, we show how to remove the effects that the air traffic service providers at the airport may have had on the realized distribution. Section 3 covers the more complex aspects of the mathematics involved.

**Example:** Suppose that a 4-hour GDP is planned and that the planned arrival acceptance rate (PAAR) for each of those 4 hours is 30 flights. Then the planned distribution is  $P = (30, 30, 30, 30)$ . Further suppose that the

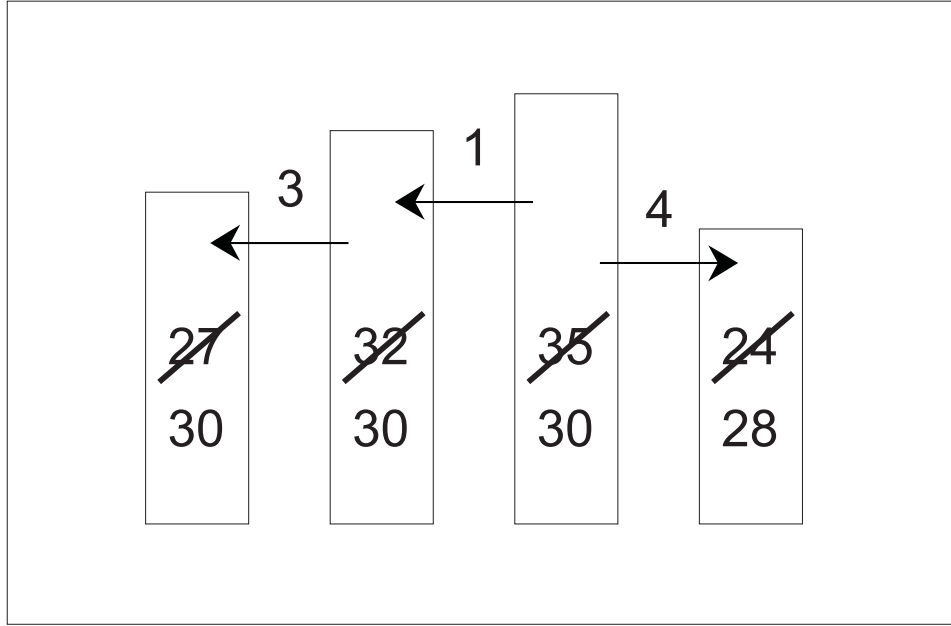


Figure 1: Flight movements necessary to transform  $R$  into  $P$

actual number of flights that arrived at the airport (or the airport terminal space) is given by  $R = (27, 32, 35, 24)$ . Then the rate control index for  $R$  (relative to  $P$ ), denoted  $RCI(P, R)$ , is computed via the following two-part calculation.

**Part 1: Compute the difference  $P - R$ .** This is the minimum amount of flight movement that is necessary to turn  $R$  into  $P$ . We sweep left to right through  $R$  (increasing  $t$ ), moving as many flights (say  $f_{t+1}$ ) as necessary from  $r_t$  to  $r_{t+1}$  to achieve  $r'_t = r_t - f_{t+1} = p_t$ . Counting left-hand movements as negative and right-hand movements as positive, we must move 3 flights from hour 1 to hour 0 ( $f_1 = -3$ ), move 1 flight from hour 2 to hour 1 ( $f_2 = -1$ ), and move 4 flights from hour 2 to hour 3 ( $f_3 = 4$ ). See Figure 1.

Movement	Notation	Flight-hours
Hr 0 $\longrightarrow$ Hr 1	$f_1$	-3
Hr 1 $\longrightarrow$ Hr 2	$f_2$	-1
Hr 2 $\longrightarrow$ Hr 3	$f_3$	4
Hr 3 $\longrightarrow$ Hr 4	$f_4$	-2

Table 1: Summary of flight movements through time

So far,  $R$  has been transformed into the distribution,  $R' = (30, 30, 30, 28)$ . Note that this is 2 flights short of the desired distribution, because  $\sum p_t - \sum r_t = 120 - 118 = 2$ . We create a “slush fund” of 2 flights at the end of  $R$  to compensate for the lack of conservation of flights. Equivalently, we could have started with an augmented distribution  $R' = (27, 32, 35, 24, 2)$ . Similarly, we extend  $P$  to a five-hour distribution,  $P' = (30, 30, 30, 30, 0)$ . To complete the example, we move 2 flights from hour 4 to hour 3 ( $f_4 = -2$ ). The summary of flight-movements is given in Table 1. We defer to Section 3 the formal proof that this correctly computes the minimum cost transformation.

Now we can compute the cost of transforming  $R$  into  $P$ . Let  $c^-$  be the average cost (say, in dollars per hour) of delaying a flight for one unit of time and let  $c^+$  be the average cost (say, in dollars per hour) of a flight arriving early by one unit of time. Then we break the flight-movements into left-hand movements  $M^- = |-3| + |-1| + |-2| = 6$  and right-hand movements  $M^+ = 4$ , corresponding to tardiness and earliness, respectively. The final cost is given by

$$c^- M^- + c^+ M^+ = 6c^- + 4c^+ . \quad (3)$$

In this example, we opt to set  $c^- = c^+ = 1.0$  to obtain pure units of 4 +

6 = 10 flight-hours. In other words,  $R$  was off from  $P$  by 10 flight-hours. Intuitively, this means that in order to turn  $R$  into  $P$ , one would have to do the work equivalent to moving 1 flight 10 hours, or 2 flights 5 hours, etc. One variation on the metric is to retain the positive and negative sums to show the breakdown of this total.

**Part 2: Normalize the distribution error  $P' - R'$ .** To compare GDPs of differing lengths and number of flights, we normalize by dividing the distribution error by the cost of the worst-case scenario. This means we must find the five-hour redistribution  $W$  of the 120 flights in  $P'$  with the highest cost difference,  $P' - W$ . In general, this involves solving a max-min problem, which is covered in Section 3. For now, we take as a given that  $W = (0, 0, 0, 0, 120)$  with a cost of

$$\begin{aligned} P' - W &= (|-4| + |-3| + |-2| + |-1|) \times 30 \\ &= (10) \times 30 = 300 \text{ flight-hours.} \end{aligned} \tag{4}$$

( $W$  corresponds to the scenario in which all the flights land in the final hour.) Thus, the rate control index we assign to the distribution  $R$  is

$$\frac{10 \text{ flight-hours}}{300 \text{ flight-hours}} = 0.03\bar{3}. \tag{5}$$

Since we have kept pure cost parameters of 1.0, this is a pure ratio that indicates the error of the realized distribution. In order to make the index more palatable, we phrase the performance of the realized distribution in terms of what was achieved rather than what was not achieved. Subtracting from 1.0, we obtain a final rate control index (RCI) of

Since we have kept pure cost parameters of 1.0, this is a pure ratio that indicates the error of the realized distribution. In order to make the index

more palatable, we phrase the performance of the realized distribution in terms of what was achieved rather than what was not achieved. Subtracting from 1.0, we obtain a final rate control index (RCI) of

$$RCI(P', R) = 1.0 - .03\overline{33} = 0.9\overline{66} \quad (6)$$

If preferred, this final index can be transformed into a percentage, 96.67%. The interpretation of the index is that the realized distribution achieved 96.67% of the intended (planned) distribution. This number can then be used for comparing GDP performance on different days and lends itself nicely to trend analysis (see Section 4).

### 3 Formulation and Computation of the RCI Metric

In Section 2, we provided motivation and intuition for the RCI metric. In this section, we provide a more rigorous formulation of the RCI metric and show how to correctly compute the cost of the worst-case distribution in the normalization process. Also, we discuss the nuances behind normalization of the metric with respect to distribution length and mass.

#### 3.1 The Raw RCI Score: Comparing Two Distributions

We define a *distribution* to be a  $(T + 1)$ -tuple of nonnegative numbers and we define its *mass* to be the sum of its entries. Let  $D$  and  $D'$  be two distributions

that we wish to compare using the RCI metric. Based on the motivation provided in Section 2, we have chosen to represent the transformation of  $D$  into  $D'$  by a  $T$ -tuple,  $F = (f_1, f_2, \dots, f_T)$ , where  $f_t$  is the amount of mass necessary to move from  $d_{t-1}$  to  $d_t$ . Specifically,  $d_t$  receives an increment of  $f_t$  units from  $d_{t-1}$  and sends  $f_{t+1}$  units to  $d_{t+1}$ . The net increment for period  $t$  is  $\Delta_t = d_t - d'_t$ , as illustrated in Figure 2. Thus, in order to transform  $d_t$  into  $d'_t$  the following condition must hold:

$$d_t + f_t - f_{t+1} = d'_t \quad (7)$$

Note that  $f_t$  can be positive or negative. Given  $D$  and  $D'$ ,  $F$  can be computed by initializing  $f_1$  using

$$f_1 = d_0 - d'_0 \quad (8)$$

and then recursively applying,

$$f_t = d_{t-1} - d'_{t-1} + f_{t-1} \quad (9)$$

for  $t = 2$  to  $T$ . Alternatively, we can compute  $F$  directly from

$$f_t = \sum_{s=1}^t (d_{s-1} - d'_{s-1}) \text{ for } t = 1, 2, \dots, T. \quad (10)$$

If  $D$  and  $D'$  do not have the same mass then, prior to computing  $F$ , their masses should be equalized by adding a  $(T + 2)^{nd}$  component to each distribution. This extra component should have zero mass in the larger of the two masses and should have a mass equal to the nonnegative difference of their masses in the smaller of the two distributions.  $T$  is then incremented by one.

Intuitively,  $\sum_{t=1}^T |f_t|$  is the amount of work necessary to transform  $D$  to  $D'$ . If we multiply the positive and negative values of  $f_t$  by cost parameters

$c^+$  and  $c^-$ , respectively, then the cost of  $f_t$  is given by

$$cst(f_t) = c^+ \max\{f_t, 0\} + c^- \max\{-f_t, 0\} \quad (11)$$

and the total cost of the transformation is computed via

$$C(D, D') = \sum_{t=1}^T cst(f_t) \quad (12)$$

which serves as the unnormalized, or raw, RCI score. Note that this expression is well-defined since for a given pair of distributions,  $D$  and  $D'$ ,  $F$  is unique. One can see from (12) that if  $C(D, D')$  is given by  $ac^+ + bc^-$ , then  $C(D', D)$  is given by  $bc^+ + ac^-$ . This means that whenever  $c^+ \neq c^-$ , the cost of transforming  $D$  into  $D'$  may not be the same as the cost of transforming  $D'$  into  $D$ . For this reason, care must be taken when comparing multiple distributions against a baseline distribution: the direction of transformation should be made consistent.

It is interesting to note that the computation of  $F$  could be recast as a network flow problem defined on a simple linear network with non-negative flow variables,  $x_t$  and  $y_t$ , where  $f_t = x_t + y_t$ . In Figure 2, one would replace each arrow  $f_t$  with a right arrow ( $x_t$ ) and a left arrow ( $y_t$ ). The vector  $F$  corresponds to an acyclic flow, i.e., a flow in which for each  $t$ , at most one of  $x_t$  and  $y_t$  is nonzero. An acyclic flow is optimal as long as the flow costs are non-negative (see [6] for an analysis of flow problems on linear networks and see [7] for a treatment of general network flow problems).

## 3.2 Normalization of RCI

### 3.2.1 The Worst-case Scenario

In the motivating example of Section 2, we saw that the normalization step of the RCI computation requires that we find the redistribution  $W$  of a fixed distribution  $P$  with the worst-case (maximum) value,  $C(W, P)$ . Consequently, we must consider any restrictions that may exist on the worst-case  $W$ . In the air traffic management case, we can assume that there is an earliest possible arrival time for each aircraft. This time is usually the scheduled arrival time of the flight. Let  $B = (b_0, b_1, \dots, b_T)$  be the distribution of earliest possible arrival times of the flights. That is,  $b_t$  is the number of flights with an earliest arrival time in the  $t^{\text{th}}$  time period. Note that for a fixed time interval  $t$ , the cumulative number of aircraft that can arrive in time periods  $s = 0, 1, \dots, t$  is limited by  $\sum_{s=0}^t b_s$ . This means that vector  $W$  is restricted by

$$W_s \leq B_t \text{ for all } t, \quad (13)$$

where for any distribution  $E = (e_0, e_1, \dots, e_T)$ , we have defined the following partial sum notation:

$$E_t = \sum_{s=0}^t e_s \text{ for } t = 0, 1, \dots, T. \quad (14)$$

$B$  is said to be a (leftward) *bounding distribution* for  $P$ . Since there is no limit to how late a aircraft could arrive, there is no need to analogously define a rightward bounding distribution. Let  $\Omega^B$  be the set of redistributions of  $P$  that are restricted by (13). Then  $E \in \Omega^B$ , if and only if  $E$  can be obtained from  $B$  by making strictly rightward shifts in  $B$ . For example, if  $P = (0, 3, 3)$  and  $B = (1, 2, 3)$ , then  $(0, 0, 6) \in \Omega^B$  while  $(3, 0, 3) \notin \Omega^B$ .



Recall that our approach to the normalization of the RCI metric is to divide the cost of transforming a distribution  $E$  into a target distribution  $P$ ,  $C(E, P)$ , by  $C(W, P)$ , where  $W$  is the worst case (highest cost) scenario. We assume that all distributions are rightward shifts of the distribution  $B$  so we seek a solution to

$$\max \{C(W, P) \mid W \in \Omega^B\}. \quad (15)$$

We note that (15) is a max-min problem, since  $C(W, P)$  is the optimal objective value of a minimization problem. In the next section, we show how to solve this max-min problem by dynamic programming.

### 3.2.2 A Dynamic Programming Solution

Let distribution  $P = (p_0, p_1, p_2, \dots, p_T)$  be given, along with bounding distribution  $B = (b_0, b_1, b_2, \dots, b_T)$ . We wish to write a recursive relation for finding a solution to (15) in the case where all distributions are integer. Fix an index  $t \leq T$  and an integer amount of mass  $\beta \geq 0$ . Consider the truncated vector  $P^t = (p_0, p_1, p_2, \dots, p_t)$ . The subproblem we define is to find a truncated vector  $W^t = (w_0, w_1, w_2, \dots, w_t)$  of mass  $\beta$  such that  $C(W^t, P^t)$  is as expensive as possible. We denote this maximum cost by  $Cmax(t, \beta)$ . To form the basis of a dynamic programming algorithm, we now derive a recursive relation to compute  $Cmax(t, \beta)$ . See Figure 3. Based on our notation,  $(W^t)_{t-1}$  is the  $(t-1)^{st}$  partial sum of the vector  $W$  truncated at  $t$ . For a fixed  $\beta$ , the mass of  $(W^t)_{t-1}$  can range over the values  $m = 0, 1, \dots, B_{t-1} \leq \beta$ . For each value of  $m$ , (10) dictates that the value of  $f_t$  is uniquely determined by  $f_t = (W^t)_t - (P^t)_{t-1} = m - P_{t-1}$ . This means that the value of  $Cmax(t, \beta)$

is related to  $Cmax(t-1, m)$  via the following recursion:

$$Cmax(t, \beta) = \begin{cases} -\infty, & \text{if } \beta > B_t \text{ or } t > T \\ 0, & \text{if } t = 0 \text{ and } \beta \leq B_t \\ \max_{m=0,1,\dots,\beta} \chi, & \text{otherwise} \end{cases} \quad (16)$$

where  $\chi = \{C(t-1, m) + cst(f_t) \mid f_t = m - P_{t-1}\}$ . The first case in (16) prevents  $\beta$  and  $t$  from taking on ill-defined values and the second case initializes the recursion. The third case expresses the fundamental recursion. The solution to our problem, (15), is given by  $Cmax(T, P_T)$ .

This recursion might seem simple, even trivial, since  $f_t$  does not vary in the max operation. However, the complexity of this problem lies within the restrictions imposed by the bounding distribution  $B$ , which limits the values to which the max operation is applied. In fact, without, this restriction, extreme or trivial solutions would always result.

Note that for a given  $\beta$  and  $t$ , the computation of  $Cmax(t, \beta)$  is dependent upon the values

$$Cmax(t-1, 0), Cmax(t-1, 1), \dots, Cmax(t-1, \beta) \quad (17)$$

This structure leads naturally to a forward recursion, dynamic programming algorithm. If we let  $S$  be the total mass in  $P$ , then the running time of such an algorithm is at most  $O(TS^2)$  since there are at most  $O(TS)$   $Cmax(\cdot)$  values to be computed and the computation of each one requires at most  $O(S)$  comparisons.

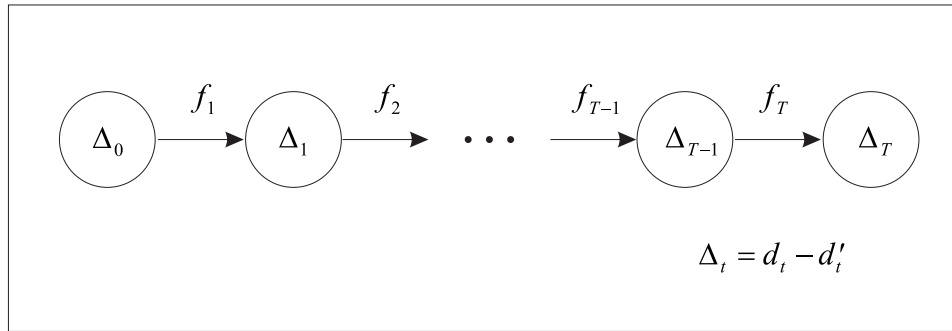


Figure 2: Transformation of distribution  $D$  into distribution  $D'$  by vector  $F$

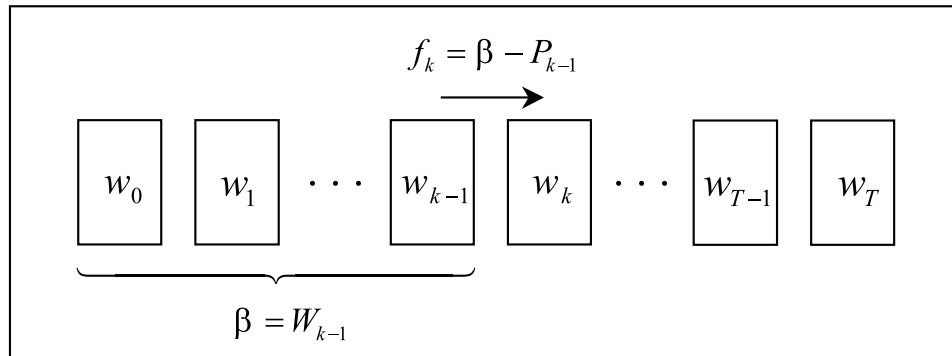


Figure 3: Finding the highest cost redistribution  $W$  when the partial sum  $W_{t-1}$  is set to  $\beta$

## 4 Ground Delay Program Analysis

In this section, we demonstrate the application of the RCI metric in the context of ground delay programs.

### 4.1 Modelling Air Traffic Flow into an Airport

A Ground Delay Program (GDP) is an FAA traffic flow action to reduce the flow of aircraft into an airport. A GDP is implemented whenever it is predicted that the arrival demand at an airport will exceed the arrival capacity for a significant period of time. In essence, flights bound for a single airport are held at their origin airports so that they can be accepted to enroute and terminal area traffic flows without delay. A light amount of airborne holding during times of reduced capacity is considered desirable because it ensures that airport arrival resources are being fully utilized. However, excessive airborne holding queues are considered undesirable because of the added workload on air traffic controllers. Aircraft must be physically separated, placed in and out of layered holding patterns, and monitored for sufficient fuel reserves.

Most GDPs are prompted by adverse weather conditions that can dramatically reduce the airport acceptance rate (AAR). Other causes are runway construction and special airport operations. GDPs are planned several hours in advance and can run for periods as long as 12 hours. For more background on ground delay programs and the ground holding problem, see [8], [9], [10], [11], [12].

Since the primary purpose of a GDP is to control the rate of aircraft flow

into an airport, the typical metric for evaluating the performance of a GDP is to measure the actual landings-per-hour ( $LAND_t$ , where  $t$  varies over the discretized time intervals over which the GDP operated) against the planned landings per hour ( $PAAR_t$ ). Although this is often taken to be the “bottom line” in a GDP, it is in fact a hybrid analysis that blurs the appropriateness of the plan with the execution of the plan.

The ability of air traffic service providers to land aircraft is directly related to airport conditions and runway configuration. These are, in turn, dependent upon weather conditions, which are highly stochastic. If weather conditions turn out to be worse than expected, then airport capacity is lower than forecasted and the desired arrival rate will not be achieved. If we measure program performance strictly by landings-per-hour, then we are effectively holding program participants responsible for the quality of the forecasts upon which the program was based. There needs to be a mechanism to analyze the success with which the plan was executed, independently of the appropriateness of the plan and the forecasts upon which it was based. (See [13] for treatment of stochastic airport capacity.)

The solution to removing stochastic airport conditions from measurement is to meter the traffic flow at a point close to the airport but before it could be directly affected by airport capacity. Ideally, we would like to know for each flight how much airborne holding it incurred as a result of restricted airport capacity. From this we can compute the time it would have arrived at the airport in the absence of airport capacity restrictions. Holding data is recorded by air traffic service providers, but only when work conditions allow them sufficient time to do so. Hence, FAA flight holding databases are

often incomplete.

As a surrogate for this holding data, several approaches can be used. A general idea is to define a radius of geographical distance (or time) about the airport and declare that once a flight has passed over this boundary, it has been *delivered* to the airport. This should be the smallest possible radius that captures all, or most of, the airborne holding. For many airports, the holding location for a flight varies with the airport arrival fix over which the flight will pass. The radius may have to be flight-path specific. Such a model can use as input readily available flight data, such as that provided by the Enhanced Traffic Management System (ETMS). For each flight, ETMS provides a runway departure time, a runway arrival time, and a sequence of estimated arrival times (ETAs), as the flight moves along its flight path [14].

We assume the existence of a model that will estimate, post facto, for each flight  $f$  the amount of airborne holding that it incurred, which we denote  $ABH_f$  (see [15] for one such model). From this, we can deduce that  $f$  was in a state of airborne holding at a given time  $t < ARTA_f$  if and only if  $ABH_f \geq ARTA_f - t$ , where ARTA is the actual runway time of arrival as recorded by ETMS.

For a set of contiguous time intervals  $t = 0, 1, \dots, T$ , let  $DEL_t$  be the number of flights delivered to the airport during time interval  $t$ , meaning, the number of flights that would arrive given unlimited capacity. Let  $ABH_t$  be the number of flights that are in a state of airborne holding at the end of interval  $t$ . Let  $LAND_t$  be the number of flights that land during time period  $t$ . If we view the airport as a closed network, then we have the following

elementary relationship for each  $t$  (see Figure 4).

$$DEL_t = (ABH_t - ABH_{t-1}) + LAND_t \quad (18)$$

After a GDP, three distributions can be assembled:  $PAAR$ ,  $DEL$ , and  $LAND$ . In the case study that follows, we use  $RCI(PAAR, LAND)$  to measure general program performance and we use  $RCI(PAAR, DEL)$  to measure program performance independent of stochastic airport conditions.

## 4.2 The Aggregate Version of RCI

Figure 5 shows a graphical analysis of the performance of a ground delay program (GDP) conducted at San Francisco airport (SFO) on March 5, 1998. The planned acceptance rate of flights (the  $PAAR$  distribution) was set at 32 flights per hour for each of the six hours of the GDP. The RCI value assigned to this GDP was 88.24%. Since this is below the 1998 RCI average for SFO, approximately 92.0%, it would be a likely candidate for further analysis. The RCI value of 88.24% was computed based on  $PAAR$  versus  $DEL$ , the distribution of “flights delivered” to the terminal airspace of the airport (but not necessarily landed). The cost parameters were set to  $c^+ = c^- = 1.0$ , to obtain pure flight-movement units. Also shown are the distributions  $LAND$ , flights landed at the airport, and  $ABH$ , the size of the airborne holding queue at the end of each hour.

Figure 5 shows why the  $RCI(PAAR, DEL)$  was not closer to the optimal value, 100%. There were too many flights delivered to the airport early in the program: in the first hour, 41 flights were delivered when only 32 were intended. One possible explanation for this is that the GDP was implemented

too late and some of these 41 flights were already airborne when the program was planned, hence, they could not have been held on the ground. Some of these flights may have been assigned ground holds but departed too early, but this is a less likely explanation. Moreover, there were drastically too few flights delivered in the 2100 (zulu) hour: 16 flights compared to the desired 32. Some of the flights intended to arrive in hour 2100 may have arrived in hour 2000. Note the increase in the airborne holding queue size in hour 2100 as a result of high arrival volume in hour 2000.

One can see that the *LAND* distribution more closely follows *PAAR* than does *DEL*. Thus, the value of  $RCI(PAAR, LAND)$  (not computed) would be slightly better than  $RCI(PAAR, DEL)$ . This is not uncommon: some of the arrival flow is smoothed out by airborne holding, resulting in a smoother distribution.

An interesting feature of this GDP was that, during hour 2100, even with the pressure of a substantial airborne holding queue, the airport was able to land only 23 flights, instead of the forecasted number, 32. Airport tower traffic counts confirm that the controllers at the airport favored departures during the 2100 hour.

### 4.3 The Nominal Version of RCI

In this section, we demonstrate the use of a nominal version of the rate control index. For each flight  $f$ , we compute the amount of arrival delay,  $M_f$  via

$$M_f = |\text{actual arrival time} - \text{planned arrival time}|. \quad (19)$$



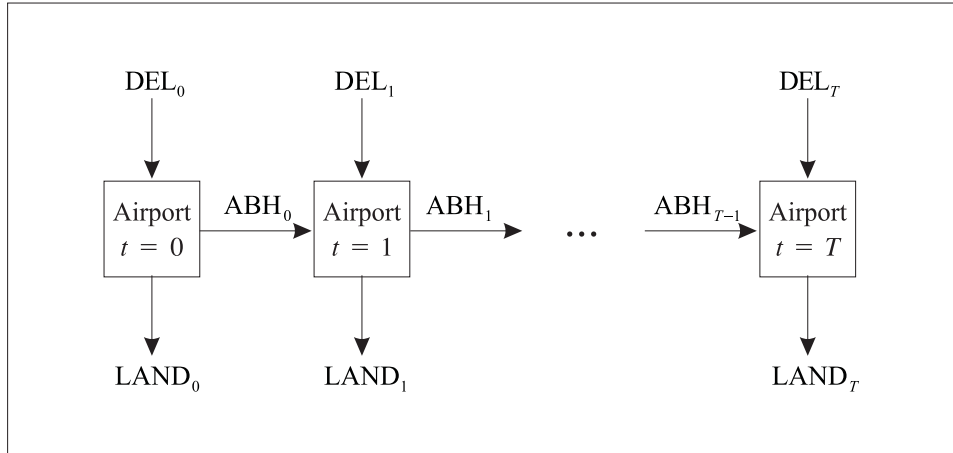


Figure 4: Model of an airport as a closed system

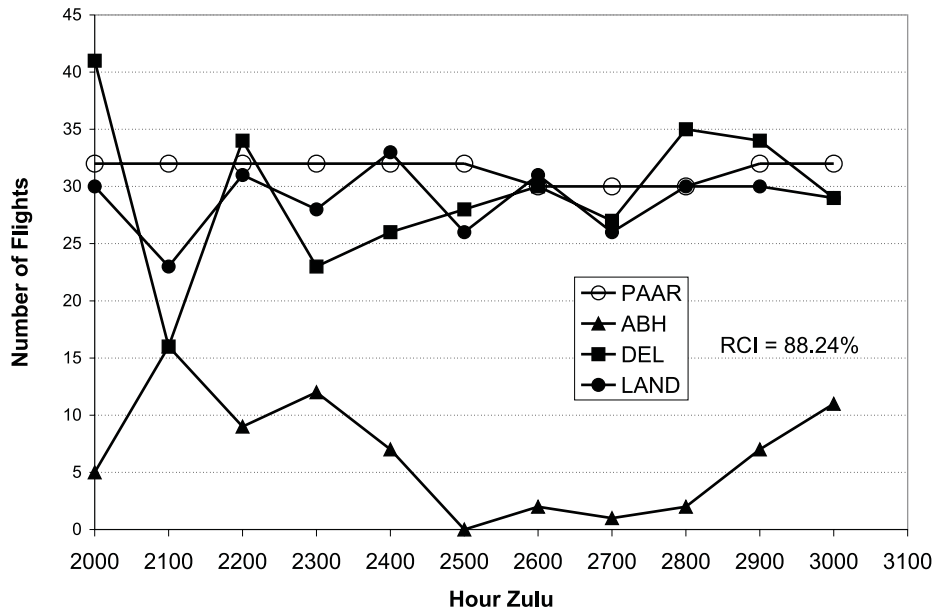


Figure 5: Analysis of ground delay performance at San Francisco Airport, 3/5/98

The unnormalized RCI score is  $\sum_f M(f)$ , which represents the amount of flight movement through time that would be necessary to restore all flights to their planned arrival time. This number is normalized by dividing by the cost of the worst-case scenario,  $\sum_f W(f)$ , where  $W(f)$  is the largest deviation that could have occurred for flight  $f$  (i.e., the farthest time period in either direction from its planned period of arrival). The final value is subtracted from 1.0 and multiplied by 100%. If desired, the positive delays ( $M_f > 0$ ) and negative delays ( $M_f < 0$ ) can be weighted with cost parameters  $c^+$  and  $c^-$  before summing, as was done with the aggregate version of the metric. In the examples shown in this section, we have elected to set  $c^+ = c^- = 1.0$  so that  $\sum_f M(f)$  represents the (absolute) variance of flight arrivals from their planned arrival times.

Each point in Figure 6 is an ordered pair,  $(RCI^{Nom}, RCI^{Agg})$ , for a GDP run at SFO during the period January 1 to October 28, 1998. Consider the point (64, 92) for July 10. The value  $RCI^{Agg} = 92\%$  indicates that, in the aggregate, the distribution of landed flights closely matched the desired distribution. However, the low value of  $RCI^{Nom}$  reveals that, too often, the flights that landed in a given time interval were not the flights that were intended to land in that time interval. Some of the flights arrived earlier than planned while others arrived too late. On the whole, for any given time interval, the number of flights that migrated out of a time interval was almost equal to the number of flights that migrated into that time interval, hence, the aggregate numbers of flights were preserved. The stochastic processes cancelled each other out and there was a great deal of ‘luck’ involved in achieving the high value of  $RCI^{Agg} = 92\%$ .

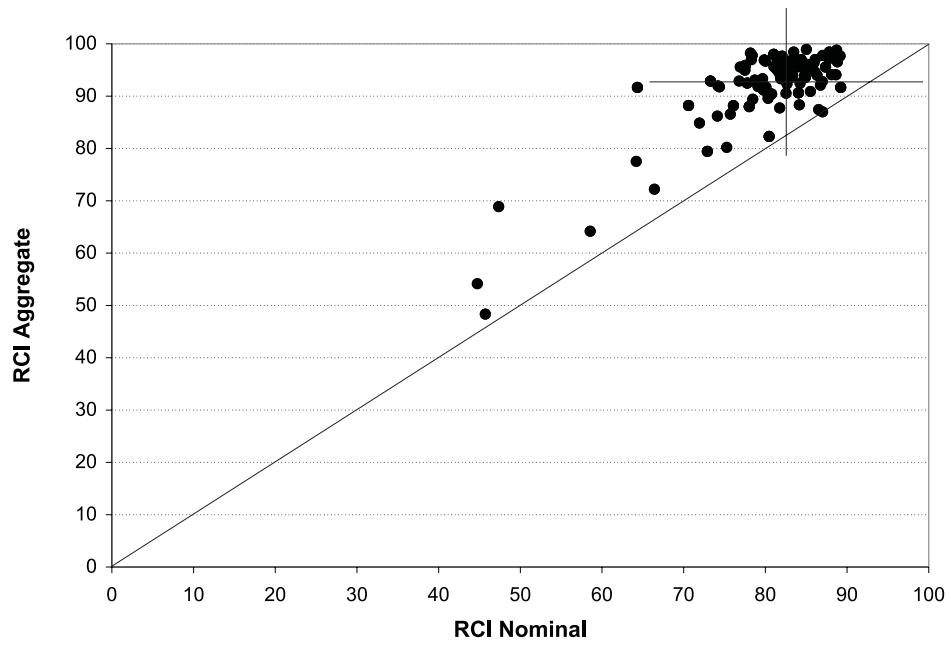


Figure 6: Nominal and aggregate RCI values for Ground Delay Programs at SFO, January - October, 1998

In general, points in the upper right quadrant of Figure 6 correspond to programs in which the aggregate distribution of flights was achieved and most of the flights arrived in their expected time periods. These are the best-run programs, the goal of each program being (100%, 100%). Given two points  $(x_1, y_1)$  and  $(x_2, y_2)$ , with  $x_1 < x_2$  and  $y_1 = y_2$ , we can say that they had the same level of aggregate success but that the first program involved more ‘luck’ while the second program involved more ‘skill’.

The center of mass of the squares lies at about (82, 92), marked by the cross. This indicates that the overall ground delay performance at SFO is quite good. In general,  $RCI^{Nom}$  is about 10% less than  $RCI^{Agg}$ . Note that all of the points lie above the 45-degree line. This is to be expected since as  $RCI^{Nom}$  increases, more flights are arriving in their planned arrival periods and so the aggregate distribution of flights is more likely to match the planned distribution, which also increases the value of  $RCI^{Agg}$ . Note that when  $RCI^{Nom} = 100\%$ , the only way for  $RCI^{Agg}$  to fall below 100% is for there to be arrivals that were not anticipated by the GDP.

## 5 Closing Remarks

We have introduced a new metric for the evaluation of planned versus realized traffic flow for a region of space or an airport. In its most general form, the metric generalizes to a comparison of two finite distributions, hence, has the potential for use in any area of traffic management in which vehicular movement through time is regulated. This represents a substantial improvement over the standard techniques for comparing two finite distributions. The met-

ric naturally lends itself to intuitive interpretation (when decomposed into left and right-handed movements) and to cost evaluation. The development of an aggregate and nominal version of the metric captures the two crucial aspects of traffic flow management: *how many* flights flowed through a region of space and *which* flights flowed through the region. Also, we have shown how to apply the metric to factor out the effects of inaccurate forecasts from performance analysis of a traffic flow initiative.

We showed how to compensate for controlled flights that arrive outside the planned time horizon by augmenting the realized distribution. The intuitive justification for this was that the cost added is assessed at a rate commensurate with the wasted capacity due to these flight cancellations. However, there doesn't seem to be an analogous compensation for the case in which unanticipated flights arrive within the planned time horizon. For consistency with flight shortages, the logical adjustment for these pop-up flights would be to augment the planned distribution. Unfortunately, the added cost would be assessed at a rate commensurate with the movement of flights backward in time, for which there is no intuitive obvious justification. In practice, pop-up flights comprise a small enough percentage of the total flights that their effect on the RCI metric is minimal when they are ignored. However, future development of the metric should incorporate pop-up flights.

## 6 Acknowledgements

This work was funded by the Federal Aviation Administration (FAA) through NEXTOR, the National Center of Excellence for Aviation Operations Re-

search. The opinions expressed herein are solely those of the authors and do not reflect the official position of the FAA.

## References

- [1] Ground delay program summaries, available at the CDM web site, <http://www.metsci.com/cdm>.
- [2] M. O. Ball, R. L. Hoffman, C. Y. Chen, and T. Vossen, "Collaborative decision making in air traffic management," in *Proceedings of Capri ATM-99 Workshop*, 1999.
- [3] R. L. Hoffman, M. O. Ball, A. R. Odoni, W. D. Hall, and M. Wambsganss, "Collaborative decision making in air traffic flow management," available at the NEXTOR web site, <http://www.nextor.umd.edu>.
- [4] Personal conversations and interviews with traffic flow managers at the FAA's Air Traffic Control System Command Center in Herndon, Virginia.
- [5] Minutes of the Collaborative Decision Making Analysis Subgroup, available at <http://www.metsci.com/cdm>.
- [6] M. O. Ball and M. J. Magazine, "Sequencing of insertions in printed circuit board assembly," *Operations Research*, 1998.
- [7] R. Ahuja, T. Magnanti, and J. Orlin, *Network Flows*, Prentice Hall, 1993.

- [8] R. L. Hoffman, *Integer Programming Models for Ground-Holding in Air Traffic Flow Management*, Ph.D. thesis, University of Maryland, 1997.
- [9] A. R. Odoni, L. Bianco, and G. Szego, Eds., *The Flow Management Problem in Air Traffic Control; Flow Control of Congested Networks*, Springer-Verlag, 1987.
- [10] O. Richetta and A. R. Odoni, “Solving optimally the static ground-holding policy problem in air traffic control,” *Transportation Science*, 1993.
- [11] O. Richetta, “Optimal algorithms and a remarkably efficient heuristic for the ground-holding problem in air traffic control,” *Operations Research*, 1995.
- [12] M. Terrab and A. R. Odoni, “Strategic flow management for air traffic control,” *Operations Research*, 1993.
- [13] M. O. Ball, R. L. Hoffman, A. R. Odoni, and R. Rifkin, “The static stochastic ground holding problem with aggregate demands,” accepted for publication by *Operations Research*.
- [14] Volpe National Transportation Systems Center, “Enhanced Traffic Management System (ETMS): Functional description,” Tech. Rep., United States Department of Transportation, 1999.
- [15] M. O. Ball, R. L. Hoffman, W. D. Hall, and A. Muharremoglu, “Collaborative decision making in air traffic management: A preliminary assessment,” Tech. Rep., NEXTOR, 1998.