

Ground Delay Program Decision-making using Multiple Criteria: A Single Airport Case

Yi Liu

Department of Civil and Environmental Engineering
Nextor
Berkeley, CA
liuyisha@berkeley.edu

Mark Hansen

Department of Civil and Environmental Engineering
Nextor
Berkeley, CA
mhansen@ce.berkeley.edu

Abstract—In this paper, we develop Ground Delay Program (GDP) models using continuum approximation. Both early GDP cancellations and GDP extensions are considered in the models. We then identify and define four performance criteria for GDP: capacity utilization, predictability, efficiency and equity. Using the proposed GDP models, we represent the trade-offs between the performance goals and relate these to the GDP decisions on planned clearance time, scope and early cancellation policy. Each flight operator may prefer a different point on the performance trade-off curves and correspondingly opt for different GDP plans. The decision-making process is formed as a utility optimization problem in our work. Specifically, we employ a linear utility function to illustrate how the trade-off curves could be used by flight operators to select their preferred GDP decisions. The research would lead to improved GDP decision-making, in which traffic managers and flight operators can make informed trade-offs based on their assessment of the importance of different performance criteria.

Keywords—ground delay program; performance metrics; performance trade-offs; user optimization

I. INTRODUCTION

On the day-of-operation, airport capacity varies and is often reduced due to poor weather, traffic congestion, or other factors. Ground Delay Programs (GDPs) are usually implemented in this case in order to balance the arrivals with the reduced capacity at the destination airports [1]. This is accomplished by delaying takeoffs at the departure airports. As a result, GDPs transfer expensive and unsafe airborne delays to cheaper and safer ground delays. In 2011, 1073 GDPs were issued at 50 airports. As one of the most common Traffic Management Initiatives (TMI), GDPs are essential to keeping the air traffic efficient and safe. However, the current GDP planning process is ad-hoc and subjective in several respects.

First, different managers may create different plans for the same situation. Depending on their temperament and experience, a manager may set higher or lower capacity rates, and shorter or longer program durations. Clear evaluation criteria to assist managers in designing GDPs are needed. Although TMI performance categories are described in the literature [2; 3; 4], associated performance criteria and day-of-

operation performance metrics are not defined for GDPs except equity metrics [5; 6; 7].

Second, flight operators influence the GDP decisions through planning telecons with the traffic managers, and frequent interaction with the command center personnel. The inputs from the flight operators focus on the decisions on the GDP parameters, and not the underlying performance objectives. It is unclear to both the Federal Aviation Administration (FAA) and flight operators how the performance of the program will be influenced by the GDP decisions. A mechanism linking the GDP performance metrics to its decision variables is missing.

Third, in the vast majority of the existing studies dealing with GDP decision-making, there is a sole objective--minimizing the delay cost [6; 8; 9; 10; 11; 12]. A few efforts have also been made to investigate trade-offs between performance goals, in particular with respect to equity and efficiency [6; 13]. However, little effort has been made to consider other goals, such as predictability and throughput. We therefore lack the ability to evaluate GDP performance using multiple criteria. While delay is an adequate measure of operational effectiveness in some instances, it does not present a complete picture of the many aspects of performance that determine the quality and level of service that Air Traffic Control (ATC) users receive [2]. Different flight operators may have different preferences on performance goals. For instance, low-cost carriers may consider efficiency more important, whereas cargo airlines may consider predictability more valuable. Therefore, an improved decision-making process should be able to measure various dimensions of the GDP performance.

In this paper, we propose to address these issues by developing GDP performance criteria and assessing the trade-offs between multiple performance goals in a manner that can inform air traffic management decision-making. We identify four performance criteria for GDPs: capacity utilization, predictability, efficiency, and equity; and specify performance metrics for them. We also build performance trade-off curves between the criteria and associated metrics using GDP models based on continuum approximations, and illustrate how these curves could be used to assist in GDP decision-making processes when the objective is a linear function of the goal

metrics. The research forms a basis for assessing the performance of GDPs using multiple criteria and will ultimately lead to improved GDP decision-making, in which traffic managers and flight operators can make informed trade-offs based on their assessment of the importance of different performance criteria.

The remainder of the paper is organized as follows. Section II introduces our GDP models. Section III specifies GDP performance criteria and the associated performance metrics. Section IV presents the performance trade-offs and discusses how the trade-off curves can assist in the GDP decision-making process. Section V summarizes the paper and discusses future research.

II. GROUND DELAY PROGRAM MODELS

In this study, we model GDP performance by using the continuum approach based on queueing diagrams [14].

When the Airport Acceptance Rate (AAR) at the airport is lowered by bad weather, a GDP will be implemented to balance the demand with the reduced capacity. There are three decision variables in the GDP design: the duration of the program, planned airport acceptance rates, and the scope of the area that is subject to GDP. If the duration of the poor weather is underestimated, then there will be airborne delay since more arrivals will be released than the destination airport can accommodate. Airborne delay could also occur if the AAR is overestimated. On the contrary, capacity may be underutilized if the estimate is too conservative. Compared to the duration of the program, planned airport acceptance levels are more predictable. For example, at SFO, fog could preclude simultaneous arrival operations on its closely spaced parallel runways, which reduces the arrival capacity from 60 to 30 flights per hour [15]. However the duration of the fog is hard to predict due to large uncertainty in the weather forecast. In this research, we investigate how the uncertainty in the duration of the program will affect the GDP performance and assume no uncertainty in airport acceptance levels. Since the GDP start time is closer to the program file time, conceptually more uncertainty exists in the prediction of the GDP end time. In the current models, we assume the program start time is fixed. Therefore, errors in predicting the GDP end time would account for all the uncertainty in the program duration. The third decision variable is scope. When a GDP is called, the FAA exempts flights from the program by limiting the scope of the GDP to a geographical area surrounding the destination airport [13]. With a small scope, more flights will be exempted from the GDP, but longer delays will be imposed on the affected flights and thus reduce equity. As discussed later, the decision on the scope has substantial impacts on GDP performance.

The queueing diagram of the arrival traffic at a GDP airport is shown in Fig. 1. The brown solid line represents the scheduled cumulative demand curve, which is linear based on the assumption of a constant schedule demand rate λ . This curve determines the Original Time of Arrival (OTA) of flights in the GDP. The piece-wise blue dash line represents the planned cumulative arrival curve under GDP, which is the basis for allocating Controlled Time of Arrival (CTA) for the

delayed flights. In the plan, a low AAR level, C_L , is expected between time zero and T . At time T , the AAR recovers and the airport is able to land arrivals at rate, C_H . The virtual queue on the ground first grows, and then the queue starts to clear at time T . The queue vanishes in the system at time, T_2 , when the planned cumulative arrival curve intersects the schedule cumulative demand curve and overlaps with it afterwards.

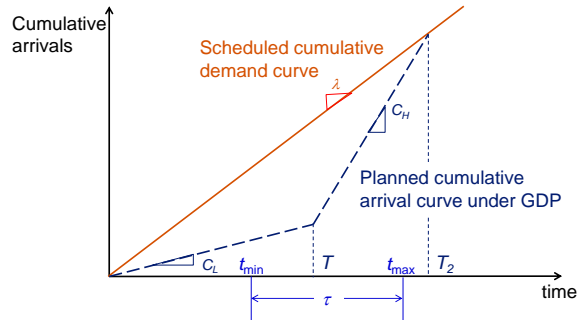


Figure 1. Queueing diagram of the arrival traffic at the GDP affected airport

With this set-up, we can express the duration of the program as

$$T_2 = (C_H - C_L) \cdot T / (C_H - \lambda) \quad (1)$$

where, T is the time when the higher AAR level is available, which could be the expected weather clearance time if the GDP is called due to poor weather. Given the monotonic linear relationship between T and T_2 , we will consider T as a GDP decision instead of T_2 in the rest of the paper.

Due to errors in prediction, the actual time when the weather clears, τ , may be different from the planned clearance time, T . After decisions are made on T , there are two possible scenarios: early weather clearance, $\tau < T$, and late weather clearance, $\tau > T$. These two scenarios are modeled in Sections A and B respectively. To increase tractability in the process, we assume that the actual clearance time is uniformly distributed between t_{min} and t_{max} . Consequently, the expected clearance time T will be a value between the same bounds.

A. GDP Models with Early Clearance

In the case of early clearance, there is unexpected extra capacity, $C_H - C_L$, between τ and T . We could choose not to revise the program and we will have enough capacity to land the planned arrivals as in the original GDP. Alternatively, we could revise the program by taking advantage of the extra capacity. In practice, traffic managers usually cancel the GDP early in this case to maximize the arrival throughput, which is counted as the total number of arrivals at any time. Revision will certainly improve capacity utilization and efficiency, but probably reduce predictability, which is evaluated as the consistency with the original GDP plan. There is widespread consensus in the community that predictability is also important. Therefore, in this research we consider early cancellation as an option, but also allow the option of not revising the GDP. The queueing diagrams for the arrivals are shown in Fig. 2 for both cases. Unnecessary delay, which is

defined as the difference between the realized delay and the minimum delay if we had perfect information about when high capacity would be available when designing the GDP, is highlighted by the shaded area. In the ideal case, delay would vanish in the system at time τ_2 .

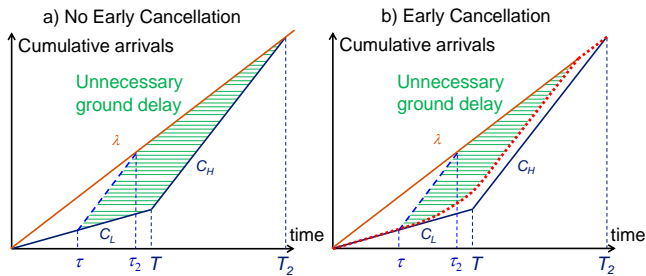


Figure 2. Queuing diagrams for arrivals, early clearance case

As seen in Plot a), if we don't consider GDP cancellation, then the realized cumulative arrival curve will overlap with the planned cumulative arrival curve and realized delay will be equal to the planned delay. The program is therefore very predictable.

On the contrary, if we cancel the GDP early to utilize the unexpected increase in capacity, then we could reduce unnecessary delay in the system as shown in Plot b). In the case of early cancellation, if we simply terminate the GDP by releasing flights at the earliest possible takeoff time, some flights may encounter airborne delay before landing since we do not have infinite arrival capacity. To avoid incurring airborne delay, we need to release flights at a rate commensurate with the amount of capacity we have. In our model, we assume that GDP is revised at time τ and the revised time slots are assigned instantaneously. The model with early cancellation is considerably more complicated because we need to revise the cumulative arrival curve, which will serve as the basis of the new CTA. The revised cumulative curve is jointly determined by the available capacity and the available cumulative arrival demand. Since we are considering early cancellation, the available capacity is C_L before τ and C_H afterwards. The available demands are obtained by releasing flights at their earliest possible takeoff time. At the time when capacity actually increases at τ , flights heading to the congested airport are either:

- Type I: these flights have already departed. They are scheduled to depart before τ and actually have departed before τ under the original GDP. Assigned ground delay in the original GDP has fully occurred for Type I flights and they will therefore arrive at their allocated CTA. The available cumulative demand curve is then the same as the planned cumulative arrival curve under the original GDP.

- Type II: these flights are being held on the ground at τ . Type II flights would have already departed in the original schedule but are waiting on the ground at time τ due to the initiation of the original GDP. These flights can, in principle, depart immediately if capacity permits. If these flights are all released at time τ , the cumulative arrival curve after this action is then the available cumulative demand curve for this type.
- Type III: these flights are scheduled to depart after τ . Ground delay assigned in the original GDP has not occurred yet for flights of this type. Therefore, there would be no delay for these flights if they were allowed to take off as scheduled. Assuming they depart as scheduled, they will arrive earlier than the time slots assigned to them under the original GDP. The available cumulative arrival demand curve is the same as the scheduled cumulative arrival curve.

The total available cumulative arrival demand after revision is calculated as the sum of the available cumulative demands of each type. The difference between the available demand and the planned cumulative arrival demand curve in the original GDP reflects the effect of GDP revision, which is affected by the range of flight time. At this stage, we assume all the flights that are heading to the affected airport will be subject to the GDP. The maximum flight time of these flights is denoted as F_{max} . When F_{max} is small, the delayed flights are concentrated in the vicinity of the affected airport and they could arrive at the airport earlier under revision, which enables the airport to utilize the expected extra capacity efficiently. If the maximum flight time is increased, delay would be absorbed by more flights but the utilization of the unexpected extra capacity would be less efficient. In this analysis, the flight time is assumed to follow a uniform distribution between the minimum flight time F_{min} and F_{max} .

B. GDP Extension Models

In the case of late clearance, GDP extension is assumed to further transfer airborne delay to ground delay if possible. In our model, we also assume that the actual clearance time is known at time T . Extension is realized by giving priority to flights in the air and further holding flights on the ground if necessary. Similar to the case of early clearance, we can also categorize flights into three groups: Types I, II and III. However, the critical time used to define the groups will be T instead of τ as in the previous case. Type I flights have taken off at time T , whereas Types II and III flights are still on the ground. After T , all the capacity will be dedicated to land Type I flights first since they are the flights in the air. Types II/III flights will be held until there are available arrival slots. It should be noticed that extension can transform airborne delay to ground delay, but it is not able to reduce the amount of delay due to limited capacity. Four cases would happen for GDP extension, as illustrated in Fig. 3.

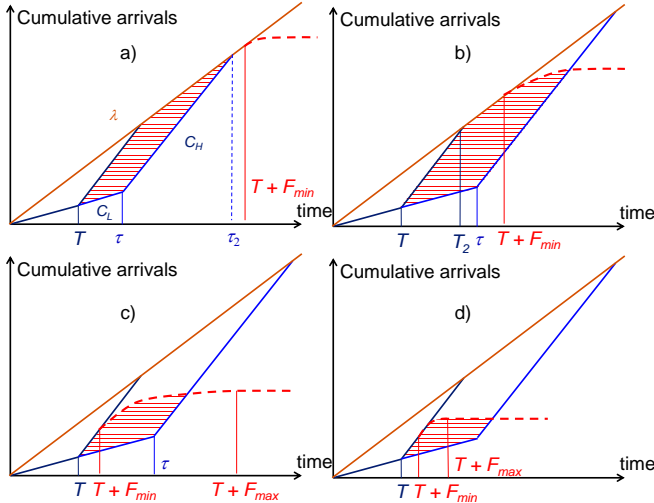


Figure 3. GDP extension models

In each plot, the red dash curve represents the cumulative arrival demand for Type I flights which have taken off before time T . It overlaps with the planned cumulative arrival curve as in the original GDP until $T+F_{min}$ when arrival capacity is also planned to serve flights taking off after T , and the cumulative arrival curve for Type I flights thus deviates from the planned cumulative arrival curve afterwards. The piece-wise blue linear line with the slope jump at τ is the actual cumulative arrival curve. Therefore, the shaded area in the plot, between the cumulative arrival demand curve for Type I flights and the actual cumulative arrival curve, represents the amount of unavoidable airborne delay, which is absorbed by Type I flights since the other flights will be further delayed on the ground and only be released when slots are available.

For all the cases, delay vanishes in the system at time τ_2 , and the trapezoid between the planned cumulative arrival curve and the actual cumulative arrival curve is the total amount of delay. Again, GDP extension is not reducing the total amount of delay. However, the amount of airborne delay could vary. In Plot a), if a flight takes off after T , then it will arrive after τ_2 since $T+F_{min}$ is later than τ_2 . All the AARs before τ_2 in the original GDP are planned for Type I flights which have taken off at time T . As a result, all the unexpected delay is realized as airborne delay. This is a special case of the GDP extension model where no action is needed. In the other three plots, GDP extension should be considered and part of airborne delay could be effectively transferred to ground delay. The amount of airborne delay depends on the relative magnitudes of $T+F_{min}$, $T+F_{max}$, the planned delay clearance time T_2 , and the actual capacity recovery time τ . Plots b), c) and d) graphically demonstrate the possible scenarios. We have developed analytical expressions for the various curves depicted in the above queuing diagrams. The expressions are somewhat complicated and dependant on inequality conditions. We will not present these formulas or their algebraic derivations here.

C. Impact of GDP Scope

So far, we have assumed that all the flights heading to the affected airport with arrival time in the constrained period are involved in the GDP. In practice, only flights within a certain region will be subject to the GDP. Flights from outside the scope region will be exempt from the program. As mentioned, scope is an important design parameter of the GDP. In this analysis, it is reflected by F_{scope} , the maximum flight time of the GDP affected flights. Flights with flight time between F_{scope} and F_{max} will be exempt from the program. The demand rate of the exempted flights is denoted by λ_e . By assuming a uniform distribution for flight time, we obtain

$$\lambda_e = (F_{max} - F_{scope}) \cdot \lambda / (F_{max} - F_{min}) \quad (2)$$

All the delay will be absorbed by the non-exempted flights whereas the exempted flights will arrive at the airport on time. The queuing diagrams of the GDP arrivals for the non-exemption case and the exemption case are shown in Fig. 4. The non-exemption case is represented with dashed lines and the case with exempted flights is represented with solid lines. Compared to the non-exemption case, both the demand rates and the capacity rates in the exemption case are reduced by λ_e . The delay clearance time in the exemption case is denoted as $T_{2,e}$. It is found that $T_{2,e}$ is equal to T_2 , when delay clears in the non-exemption case. It is also easy to prove that the planned system delays in the two cases are the same. With a small F_{scope} , the GDP affected flights will locate in the vicinity of the airport and GDP revision will be more efficient.

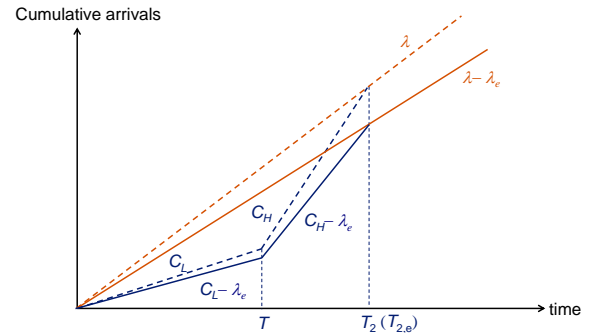


Figure 4. Queuing diagrams of GDP affected arrivals, non-exemption and exemption cases

III. PERFORMANCE METRICS

In this section, we will define our four performance criteria and quantify them with the proposed GDP models.

A. Capacity Utilization

This metric is specified to measure how fully we used our capacity. It is defined as the ratio of throughputs:

$$\alpha_c(\tau, T) = N_R / N_I \quad (3)$$

where, N_I is the ideal throughput, the total number of arrivals, under perfect information at the time when queue clears, τ_2 ; N_R is the realized throughput at this time. These values are shown in Fig. 5, for the cases of early clearance and

late clearance respectively. As we see in Plot a), the realized throughput is less than the idealized throughput at τ_2 . Therefore, capacity utilization is less than 1 in the case of early clearance. However, N_R is increased if we consider early GDP cancellation, which benefits capacity utilization. In the case of late clearance, the ideal throughput is the same as the realized throughput since delay could only clear at time τ_2 even if we had perfect information at the beginning. As a result, capacity utilization is equal to 1.

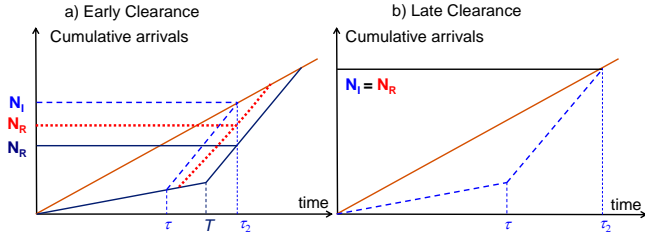


Figure 5. Ideal throughput and realized throughput

Note: The red dotted line represents the cumulative throughput curve with early cancellation

B. Predictability

In principle, predictability is defined to measure the accuracy of the information available in advance. In the design of a GDP, a certain amount of delay is planned. Flight operators will be informed of the expected delays before the GDP is implemented. The realized delay will usually be different from the planned delay due to error in prediction. Predictability is then identified to measure how different the realized delay is from the planned delay:

$$\alpha_p(\tau, T) = \min(D_P, D_R) / \max(D_P, D_R) \quad (4)$$

where, D_P is flight delay planned at the beginning of the GDP; D_R is total realized flight delay. Defined as this, predictability reflects how consistent the program is with the original plan.

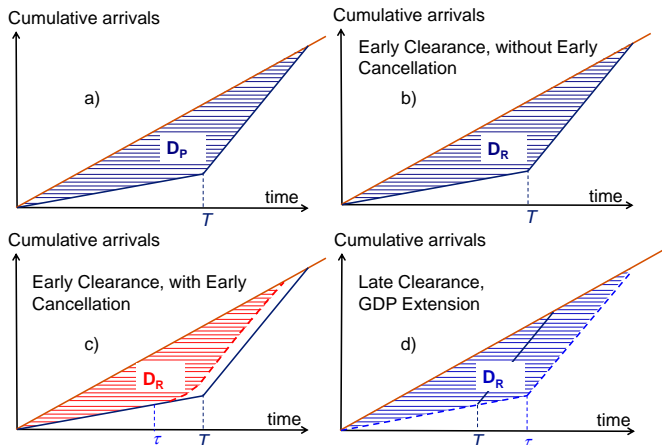


Figure 6. Planned delay and realized delay

As shown in Fig. 6, D_P is determined by the planned weather clearance time at the beginning of the GDP and does not change with the actual clearance time. On the contrary, D_R depends on when the weather will clear and whether we choose to revise the GDP or not. In the case of early clearance, D_R will be equal to D_P if we choose not to revise the GDP. Flight operators could run their operations relying on the CTA allocated in the original GDP and no further adjustment will be needed. If the GDP is revised, as seen in the bottom-left plot, we will be able to save delay in the system and realized delay will be less than the planned delay. In the case of late weather clearance, as in the bottom-right plot, realized delay will be larger than the planned delay due to unexpected late capacity recovery.

C. Efficiency

A primary motivation for the GDP is that as long as delay is unavoidable, it is cheaper and safer for flights to absorb delay on the ground before takeoff, rather than in the air. The efficiency metric is defined to examine how efficient the program is in transferring airborne delay to ground delay. For this metric, we distinguish the cost of airborne delay from the cost of ground delay and assume the cost ratio is $\beta (>1)$. In other words, 1-minute of airborne delay is equivalent to β -minute of ground delay. The efficiency metric is expressed as

$$\alpha_e(\tau, T) = C_I / C_R \quad (5)$$

where, C_I is the minimum cost that would be incurred if perfect information were available about when the capacity will increase; C_R is the total realized cost at the end of the program. The costs are illustrated in Fig. 7.

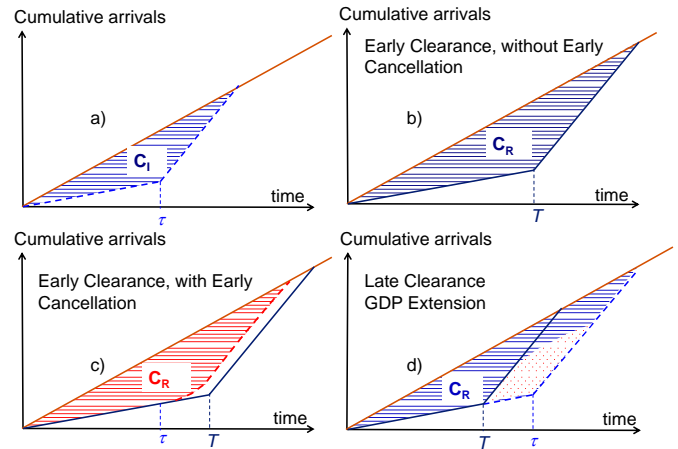


Figure 7. Queuing diagrams of GDP affected arrivals, non-exemption and exemption cases

As in plot a), C_I is determined by the actual clearance time τ and is always in the form of ground delay. There are three different cases for the realized delay cost, as in plots b), c) and d). Airborne delay occurs only in the case of GDP extension, which is highlighted by the dotted area in plot d).

D. Equity

As discussed earlier, when there are flights exempted from the GDP, planned delay clearance time and total planned delay are still the same as the case without exemption. However, with more flights exempt from the program, more delays are allocated to the non-exempt flights while the exempt flights are essentially ‘free-riders’. This raises the equity issue in the design of GDPs: how much of the demand should be exempt from the program. In practice, the FAA exempts flights from a GDP by limiting the scope of the GDP to a geographical area surrounding the destination airport [13]. A flight operator, whose flights are mostly long-haul, will prefer a smaller scope so that more of its flights can arrive on time. On the contrary, flight operators with more short-haul flights may prefer a larger scope, in which case total delay will be absorbed by more flights and delay per flight will be reduced. Preserving equity among competing flight operators is an important goal of the FAA. In this study, the equity metric is defined by comparing the maximum planned flight delay to the maximum planned flight delay when no flights are exempted from the GDP [5; 6; 7]. Different from the other performance metrics, equity will only be measured when the GDP is planned and its value will not be updated upon a GDP revision.

As shown in Fig. 8, the maximum planned flight delay among all the affected flights, denoted as $d_{p,\max}^e$, is encountered by the flight that is planned to arrive at T . In the non-exemption case, the total system delay is absorbed by all the flights and thus the maximum planned flight delay is minimized.

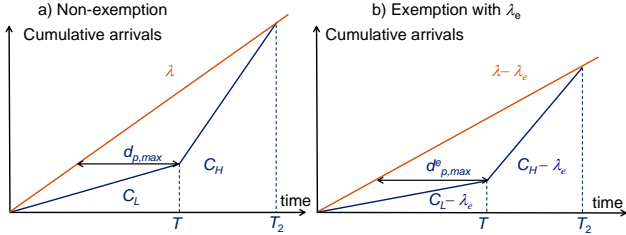


Figure 8. Maximum planned delays, non-exemption and exemption cases

Following the proposed models, the maximum planned delay can be expressed as

$$d_{p,\max}^e = (\lambda - C_L) \cdot T / (\lambda - \lambda_e) \quad (6)$$

and the performance metrics of equity could then be written as:

$$\alpha_f = \frac{\min(d_{p,\max}^e)}{d_{p,\max}^e} = \frac{\lambda - \lambda_e}{\lambda} = \frac{F_{scope} - F_{\min}}{F_{\max} - F_{\min}} \quad (7)$$

As seen in the formula, the equity of a GDP is independent of the decision on the clearance time and only affected by the scope. Specifically, equity is an increasing function of F_{scope} .

E. Summary

We constructed performance metrics for capacity utilization, predictability, efficiency and equity in the previous sections. All the metrics are dimensionless, and between 0 and 1. The expected value of the equity metric is determined once we select the scope of the GDP and it is independent of the prediction errors. However, the values of the other three metrics depend on τ , the real weather clearance time. Since GDP decisions are made before the real clearance time is known, the program performance should be assessed using the expected values of the defined metrics. In our analysis, we integrate the values of the three metrics over τ to get the expected values of the performance:

$$\begin{aligned} \alpha(T) &= E[\alpha(\tau, T)] = \int_{t_{\min}}^{t_{\max}} \alpha(\tau, T) \cdot f(\tau) d\tau \\ &= \frac{1}{t_{\max} - t_{\min}} \cdot \left[\int_{t_{\min}}^T \alpha(\tau < T | T) d\tau + \int_T^{t_{\max}} \alpha(\tau \geq T | T) d\tau \right] \end{aligned} \quad (8)$$

where, α is any of the three metrics: capacity utilization, predictability and efficiency. The expressions for the metrics under the cases of early clearance and late clearance are very complicated and details are not presented in this paper. Instead, below we will show the characterized relationship between performance and GDP decisions and the performance trade-offs between competing goals in the developed models.

IV. PERFORMANCE TRADE-OFFS AND USER OPTIMIZATION

In this section, through a numerical example, we will first present the influence of the GDP decisions on the performance expectations and the trade-offs among multiple performance goals. Then, we will illustrate how the research results could assist decision-making in the design of GDP under capacity uncertainty.

The set of parameter values in the example is shown in Table 1. AAR values are chosen based on the airport capacity benchmark report by the FAA [16]. The lower and upper bounds for τ are estimated after reviewing the air traffic control system command center advisories database, which is available on the FAA website. The cost ratio of airborne delay to ground delay is set as 2 [6].

TABLE I. PARAMETER VALUES IN THE NUMERICAL EXAMPLE

Parameter	Notation	Values	Unit
Scheduled demand rate	λ	60	Arrival per hour
High airport acceptance rate	C_H	80	Arrival per hour
Low airport acceptance rate	C_L	40	Arrival per hour
Lower bound for τ	t_{\min}	2	Hour
Upper bound for τ	t_{\max}	6	Hour
Minimum flight time	F_{\min}	0.5	Hour
Maximum flight time	F_{\max}	7	Hour
Cost ratio	β	2	-

A. Performance Metrics and Their Trade-offs

There are three decisions in the design process of a GDP besides the fixed AARs: scope of the GDP, planned clearance time, and early cancellation policy. As discussed before, equity is determined by the GDP scope only, and it increases monotonically with scope. Expectations of the other three performance metrics also depend on the assumed clearance time and the policy on early cancellation. The influence of the decisions on the three performance metrics is shown in Fig. 9, where the left three plots demonstrate the non-exemption case and the right plots demonstrate the exemption case, with half of the demand exempted from the GDP.

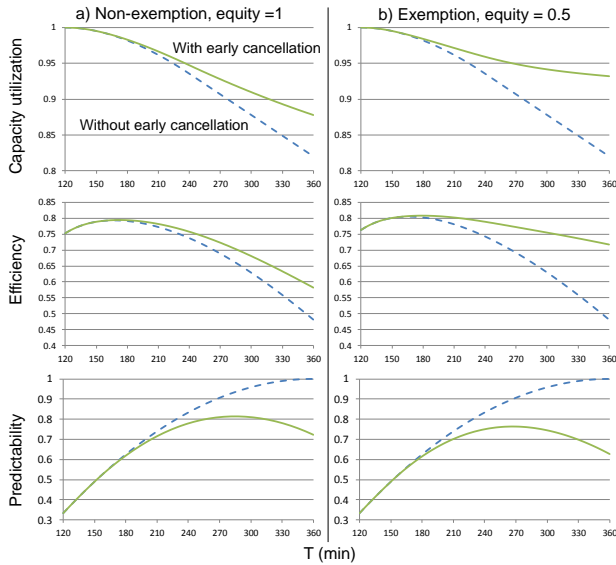


Figure 9. Values of performance metrics as functions of T , with and without early cancellation, with and without GDP exemption

When no flight is exempted from the GDP, equity is equal to 1 as for the left three plots. The blue dash curves depict the relationship between the performance metrics and the planned clearance time without considering early GDP cancellation. In this case, capacity utilization decreases with the planned clearance time because there is a larger chance of early clearance with larger T and part of the high capacity cannot be utilized. As T increases, efficiency first increases because of the reduced chance of expensive airborne delay. After a certain point, efficiency will decrease with T because it is very likely that realized ground delay is much larger than it could be if we had perfect information. Predictability increases with T without early GDP cancellation. With a larger T , it is very likely that capacity will recover earlier than planned, without early cancellation, and the realized delay will be the same as the planned delay so that predictability approaches 1. The green solid curves represent the relationship between the performance metrics and the planned clearance time with considering early GDP cancellation. Compared to before, early GDP cancellation allows us to take advantage of the unexpected high capacity and reduce delay in the system in the case of early weather clearance. This benefits capacity utilization and efficiency. On the contrary, predictability degrades when we permit early

GDP cancellation. Basically, the more adaptable we make the GDP, the less predictability we have. The impact of early GDP cancellation is more obvious with a larger T for all three metrics, with the difference between the two curves increasing with T .

The conclusions above all also hold true when there is exemption, as shown in the right plots. By comparing the green solid curves in the right plots to those in the left plots, we find that exemption improve the performance of capacity utilization and efficiency, but reduce predictability in the system. By exempting long-haul flights from the GDP, the delayed flights are concentrated in the vicinity of the affected airport and they could arrive at the airport earlier if there is early cancellation, which enables the airport to utilize the unexpected extra capacity earlier. This benefits capacity and efficiency, but reduces predictability. In the case of GDP extension, more flights will still be on the ground if they are closer to the affected airport. As a result, more airborne delay could be transferred to ground delay, which further increases efficiency. By comparing the blue dash curves, we find that the plots for capacity and predictability are the same regardless of exemption rate whereas efficiency is slightly improved with exemption. The efficiency gain from reduced scope is because the GDP extension can shift more airborne delay to ground delay.

Performance trade-off curves are shown in Fig. 10. Movement toward the right along these curves is associated with earlier planned clearance times. Equity is equal to 1 for the left two plots. The bottom-left plot presents the trade-offs between efficiency and capacity utilization. The dashed blue line is for the case without early cancellation and the solid green line is for the case with early cancellation. Both plots have internal optima, and the points located on the left of the internal peaks are inferior because we can increase efficiency and capacity utilization simultaneously by decreasing T . On the right of the peaks, the line for early cancellation is above that for no early cancellation. Therefore, if only efficiency and capacity utilization are concerned, then we will always choose to terminate the GDP earlier if possible and we tend to pick an earlier planned clearance time. The situation changes when predictability is also taken in to account. Comparing the two dashed plots on the left, we see that a choice of larger T degrades efficiency and capacity utilization but benefits predictability. As a result, flight operators who value predictability more may prefer a larger T . Additionally, early cancellation is not necessarily better when predictability is important. For instance, at capacity utilization equal to, say, 0.9, both predictability and efficiency are higher if we choose not to terminate the GDP earlier. It should be pointed out that, for capacity utilization to be equal, the planned clearance time of the case with no early cancellation must be smaller than that of the early cancellation case here. The trade-off relationship is similar when equity is reduced to 0.5, as shown in the right plots. However, the differences between the early cancellation and no early cancellation cases are more pronounced, since early cancellation is more effective with long-haul flights exempted. Moreover, the performance of efficiency and capacity in the case of early cancellation are improved in general with exemption, as the possible metric values are

spread at higher levels, comparing the two green curves in the plots at the bottom of Fig. 10. This performance improvement is at the expense of equity.

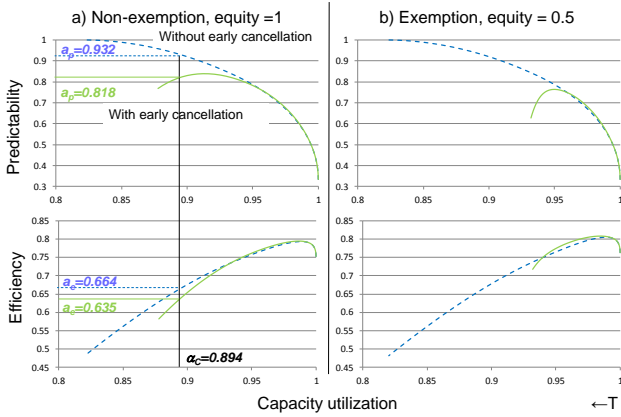


Figure 10. Performance trade-off curves, with T increase from the right to the left

B. User Optimization

Different flight operators may have different preferences regarding performance goals. Each flight operator may prefer a different point on the trade-off curves, and correspondingly opt for different GDP plans. Rationally, each user will prefer the point that maximizes their utility. The decision-making process could then be formed as a utility optimization problem. Here, we assume that the FAA will predetermine the level of equity and the flight operators will only consider the other three metrics in the utility function. We use linear utility functions to illustrate how the trade-off curves could be used by flight operators to select their preferred GDP decisions. Using the same set up, optimal solutions can always be found for users with concave utility functions.

The constrained optimization problem is shown as the following. Given the value of the equity metric by the FAA, each system user will choose a set of the performance vectors, $[\alpha_c, \alpha_p, \alpha_e]$ to

$$\text{maximize: } U(\alpha_c, \alpha_p, \alpha_e) = C_c \cdot \alpha_c + C_p \cdot \alpha_p + C_e \cdot \alpha_e$$

$$\text{subject to: } F_{\text{early_cancellation}}(\alpha_c, \alpha_p, \alpha_e) = 0$$

$$\text{or } F_{\text{no_early_cancellation}}(\alpha_c, \alpha_p, \alpha_e) = 0$$

where, the constraints are limiting the feasible region of the performance vectors to the points on the performance trade-off curves. Implied ideal plans of three users with different preferences on performance goals are compared in Table 2. User 1 is concerned most with capacity utilization and weighs predictability and efficiency equally. Predictability has double the importance to User 2 compared to the other metrics, and is an even more importance goal to User 3. If User 1 is the only user in the system, and FAA wants to maximize equity by exempting no flights, then the GDP should be planned for 3.8 hours and early cancellation should be permitted. If half of the demand is exempt, then the planned duration of the GDP could

be slightly longer. T increases with the exemption rate under early cancellation because early cancellation can recoup more throughput if more short-haul flights are involved in the program. With predictability as the most critical performance goal, both User 2 and User 3 choose not to revise the program in the case of early clearance. The preferred planned clearance times of the two users are not affected by the GDP scope, because scope affects performance only if early cancellation is permitted. User 3, with his stronger emphasis on predictability, prefers a very conservative approach with the planned clearance time approaching the maximum duration of the low capacity period.

TABLE II. PREFERRED GDP DECISIONS BY DIFFERENT SYSTEM USERS

User	Weights			Equity = 1		Equity = 0.5	
	C_c	C_p	C_e	$D1^a$	$D2^b$	$D1$	$D2$
1	0.5	0.25	0.25	3.8	Yes	3.88	Yes
2	0.25	0.5	0.25	4.88	No	4.88	No
3	0	0.75	0.25	5.4	No	5.4	No

a. Decision 1: planned clearance time, T (hour); b. Decision 2: early cancellation or not?

V. CONCLUSIONS

In this paper, we have modeled the GDPs for a single airport case using continuum approximation. Both early weather clearance and late weather clearance are considered. In the case of early clearance, we consider early GDP cancellation as an option, but also allow the option of not revising the GDP. In the case of late clearance, GDP extension is assumed to further transfer some airborne delay to ground delay. When extending the GDP, landing priority is given to the en-route flights and flights that have not taken off yet are further delayed on the ground if necessary.

Based on the proposed models, we have identified performance metrics in the GDP design for four performance goals: capacity utilization, predictability, efficiency, and equity. All the metrics are dimensionless and have values between 0 and 1. Realized values of the performance metrics require the realized value of τ , the actual time when the AAR increases. Since the actual values cannot be known at the time of decision, expected values of the metrics are derived to support decision-making. Equity is determined by the GDP scope only, and it increases monotonically with scope. Expectations of the other three performance metrics also depend on the assumed clearance time and the policy on early cancellation. Without considering early GDP cancellation, capacity and predictability vary with T monotonically, whereas there is an intermediate peak for efficiency. Early GDP cancellation saves delay in the system and enables the delay to clear earlier. While this adaptability benefits capacity and efficiency, it degrades predictability.

Using the expectations of the performance goals, we further construct trade-off curves between the performance goals and relate these to the GDP decisions on planned clearance time, scope, and early cancellation policy. It is found that when only efficiency and capacity utilization are concerned, we will always choose to terminate the GDP earlier if possible and we

tend to pick a small planned clearance time. The situation changes when predictability is also taken in to account. A choice of larger T or larger scope degrades efficiency and capacity utilization but benefits predictability. Additionally, early cancellation is not necessarily better when predictability is highly valued. The conclusions are true at different levels of equity.

With the same set-up, we have also investigated the utility optimization in the GDP with a linear utility function. Equity is predetermined by the FAA, whereas flight operators state their preferences on the other three performance metrics by providing the relative weights they would attach to the performance goals. Optimal values of T are obtained by maximizing the user utilities assuming the actual clearance time following a uniform distribution. The results indicate that different flight operators will opt for different GDP plans according to their assessment of the importance of different performance criteria.

Decisions on GDPs and other traffic management initiatives are made every day, and without explicit considerations of the performance trade-offs involved. In its own performance metrics, it appears that the FAA places considerable emphasis on capacity utilization. For example, one of the most widely tracked metrics is the ratio of operations to called rates. It is interesting in this regard that, as shown in Fig. 10, substantial increases in predictability, which many flight operators claim to highly value, can be obtained with modest reductions in throughput. This is but one example in which understanding the trade-offs involved, and finding a way to elicit input from the flight operator community on how to balance competing goals may change current practice. Work on how to take conflicting inputs from different flight operators and arrive at an acceptable compromising plan is currently underway.

REFERENCES

[1] Xiong, J., "Revealed preference of airline's behavior under air traffic management initiatives", PhD Thesis, University of California, Berkeley, California, 2010.

[2] Bolczak, C.N., Hoffman, J.H., Jensen, A.J., Trigeiro, W.W., National airspace system performance measurement: overview, MITRE Technical Report MTR 97W0000035, Center for Advanced Aviation System Development, McLean, Virginia, 1997.

[3] Bradford, S., Knorr, D., Liang, D., "Performance measures for future architecture", Proceedings of 3rd USA/Europe Air Traffic Management R&D Seminar, Napoli, Italy, 2000.

[4] Sridhar, B., Grabbe, S.R., Mukherjee, A., "Modeling and optimization in traffic flow management", Proceedings of the IEEE 96(12), 2008, pp. 2060-2080.

[5] Hoffman R., Ball, M.O, Mukherjee, A., "Ration-by-Distance with equity guarantees: A new approach to Ground Delay Program planning and control", the 7th ATM R&D Seminar, Barcelona, Spain, 2007.

[6] Mukherjee, A., Hansen, M., "A dynamic stochastic model for the single airport ground holding problem", Transportation Science 41(4), 2007, pp. 444-456.

[7] Ball, M.O., Hoffman, R., Mukherjee, A., "Ground delay program planning under uncertainty based on the ration-by-distance principle", Transportation Science 44 (1), 2010, pp. 1-14.

[8] Odoni, A.R., "The flow management problem in air traffic control". In: Odoni, A.R., Bianco, L., Szego, G. (Eds.), Flow Control of Congested Networks, Springer-Verlag, Berlin, 1987, pp. 269-288.

[9] Hoffman, R., Ball, M.O., "A comparison of formulations for the single-airport ground-holding problem with banking constraints", Operations Research, 48 (4), 2000, pp. 578-590.

[10] Ball, M.O., Hoffman, R., Odoni, A., Rifkin, R., "A stochastic integer program with dual network structure and its application to the ground-holding problem", Operations Research 51(1), 2003, pp. 167-171.

[11] Liu, P.B., Hansen, M., Mukherjee, A., "Scenario-based air traffic flow management: from theory to practice", Transportation Research Part B 42, 2008, pp. 685-702

[12] Mukherjee, A., Hansen, M., Grabbe, S., "Ground delay program planning under uncertainty in airport capacity", AIAA Guidance, Navigation, and Control Conference Chicago, 2009.

[13] Ball, M.O., Lulli, G., "Ground delay program: optimizing over the included flight set based on distance", Air Traffic Control Quarterly 12, 2004, pp. 1-25.

[14] Daganzo, C.F., Fundamentals of Transportation and Traffic Operations, Elsevier Science Inc., New York, pp. 1997.

[15] Cook, L.S., Wood, B., "A model for determining ground delay program parameters using a probabilistic forecast of stratus clearing", 8th USA/Europe Air Traffic Management R&D Seminar, 2009.

[16] FAA, Airport Capacity Benchmarking Report 2004. Washington, D.C., 2004.

AUTHOR BIOGRAPHY

Yi Liu is a PhD candidate in the Department of Civil and Environmental Engineering at the University of California, Berkeley working under Dr. Mark Hansen. She has an undergraduate degree in Hydraulic and hydropower engineering from Tsinghua University in China, a Masters in Geotechnical Engineering from Hong Kong University of Science and Technology. Her main research interest is in air traffic management and surface operations management.

Mark Hansen is a Professor of Civil and Environmental Engineering at the University of California, Berkeley. He graduated from Yale with a Bachelor's degree in Physics and Philosophy in 1980, and has a PhD in Engineering Science and a Masters in City and Regional Planning from UC Berkeley. Dr. Hansen worked as a physicist at the Environmental Protection Agency, prior to graduate school. His main research interest is in aviation system performance analysis, air traffic flow management, air transport economics and aviation safety. Professor Hansen is the Berkeley co-director of the National Center of Excellence in Aviation Operations Research.