Evaluating Delay Cost Functions with Airline Actions in Airspace Flow Programs

Michael Bloem

NASA Ames Research Center

Moffett Field, CA, USA

michael.bloem@nasa.gov

Haiyun Huang*

Science & Technology Corporation

Hampton, VA, USA and

Delft University of Technology

Delft, the Netherlands

h.huang@student.tudelft.nl

Abstract— **Air traffic management research and simulation use delay cost functions that attempt to quantify the cost of delay to airlines. Seventeen delay cost functions from previous research are evaluated with airline actions in Airspace Flow Programs. Airline actions from 34 days in the summer of 2006 were used to compute four metrics designed to quantify the consistency of the airline actions with each of the cost functions. Two of these metrics compare the cost of airline actions to the cost of the default first-scheduled-first-served actions. These metrics identify delay cost functions in which costs increase in discrete steps as delay increases as most consistent with airline actions. The other two metrics compare the cost of the airline actions to the minimum costs. These metrics identify delay costs that are proportional to the length of delay but with larger proportionality constants for flights bound for hub airports as most consistent with airline actions.**

Keywords— *Delay; Metrics; Airline Behavior Modeling; Collaborative Decision Making; Air Traffic Flow Management Slots; Airspace Flow Programs*

Nomenclature

$(\mathcal{F}, \mathcal{S}, \mathcal{M})$	Set of corresponding sets of flights, sets of slots, and matchings
α	Weighting parameter
$\beta(t_f, d)$	Time-of-Day Delay multiplier
δ	Length of slot time window in minutes
$\eta(e_f, d)$	Monetary Delay multiplier
$\gamma(f)$	Connection Delay multiplier
$\gamma'(f, a_f)$	Airline Connection Delay multiplier
$\hat{L}(\sigma^2)$	Approximate log-likelihood when the variance is σ^2 and the mean is assumed 0
$\rho(d)$	Step delay cost
$\sigma^{2\star}$	Estimate of the additive cost noise variance
$\tilde{R}_{\mathrm{FSFS}}$	Median of FSFS ratios
$\tilde{R}_{\mathrm{min}}$	Median of minimum ratios
ε	Additive cost noise
a_f	Airline operating flight f

$c(f, d)$	Delay cost function for flight f and d minutes of delay
d	Delay in minutes
$d(f, s)$	Delay in minutes resulting from assigning flight f to slot s
e_f	Aircraft type used for flight f
F	Set of flights
f	A flight
I	Improvement frequency
$J(F, S, M, c)$	The cost of matching the set of flights F to the set of slots S as specified by matching M when using cost function c
$J^\star(F, S, c)$	Minimum cost for a perfect matching of the flights F to the slots S when the cost function is c
M	A matching of flights to slots
n	Number of flights and slots
p_f	Number of passengers on flight f
R_{FSFS}	The FSFS ratio
R_{min}	The minimum ratio
S	Set of slots
s	A slot
t_f	Scheduled time of arrival for flight f
t_s	Slot time

I. Introduction

The cost of delaying a flight differs from flight to flight. This cost can be most accurately estimated by the airlines, but unfortunately airlines are reluctant to reveal their costs because doing so could be advantageous to their competitors. Having an accurate understanding of airline costs is important in air traffic management research. Airline cost functions impact the design of air traffic management concepts, help determine the value of new concepts, and can form the basis of airline behavior models used in simulations. Without the benefit of knowing actual airline delay costs, researchers must infer these costs.

Only one effort has been made to tune and validate delay cost functions with records of airline actions. In her disserta-

tion [1], Xiong used airline flight cancellation and slot usage data from Ground Delay Programs to tune parameters in discrete choice models of airline behavior. Discrete choice models assign a probability to every possible choice an airline has based on the cost of each choice. Xiong's research revealed many characteristics of airline delay costs, but there are some limitations of her work. Discrete choice models cannot easily handle cases with many discrete choices, preventing the use of some of the available data. More importantly, Xiong did not study any "separable" delay cost models. Separable delay cost models use a flight delay cost function to compute the cost of delaying each flight and assume that the total airline cost is the sum of the individual flight delay costs. While there are exceptions [2–8], most research uses a separable delay cost model [9–18]. Some airline decision support tools are also based on separable delay cost models [19–21]. Furthermore, Xiong's use of Ground Delay Program data limited her ability to investigate the difference in the cost of delay for hub-bound flights and other flights. Finally, while Xiong studied linear models with data-intensive variables related to airline revenues, her work did not consider some simple variables from previous research and airline decision-support tools, such as those in [20, 22–24].

The goal of this research is to determine the degree to which delay cost functions proposed for separable delay cost models are consistent with how airlines assign flights to slots in Airspace Flow Programs. None of these delay cost functions have been validated with airline actions. The inability of discrete choice models to handle cases with many choices would prevent the use of more than half of the available data for some airlines, so a new approach for validating cost functions is needed. Four new metrics are proposed for evaluating the extent to which cost functions are consistent with airline actions in Airspace Flow Programs. Two of the four metrics compare the costs incurred by airlines to the costs of default actions, while the other two compare to the minimum costs. One metric adds a noise term to delay cost functions to account for unobserved aspects of airline costs and then estimates the parameters of the noise with airline action data.

The remainder of this paper is structured as follows. Section II contains background information about Airspace Flow Programs. The model, metrics, and data used to evaluate cost functions are presented in Section III. Results are discussed in Section IV. The paper finishes with proposals for future work in Section V and conclusions in Section VI.

II. Background

Airspace Flow Programs (AFPs) are a mechanism used by the Federal Aviation Administration (FAA) in the United States to assign departure delays to aircraft when demand for a region of airspace known as a "Flow Constrained Area" (FCA) exceeds capacity. This mechanism is based on the concept of slots. A slot is the right to fly into the FCA in a specified period of time. The FAA enforces slot ownership rights by assigning departure times to flights bound for an FCA so that each flight arrives at the FCA approximately at the time of the slot to which it is assigned. During AFPs, slots are allocated to airlines with

an algorithm referred to as "ration by schedule" (RBS) that is based on a first-scheduled-first-served (FSFS) principle. By default, each airline's flights are assigned to their allocated slots in a FSFS manner, but they can adjust this assignment as they see fit. Large airlines alter the assignment of their flights to their slots in AFPs thousands of times each year.

III. Method

A. Airline Behavior Model

Airlines have tools and procedures that allow them to make acceptable decisions during an AFP, but the problem they face during an AFP is complicated and difficult for researchers to model [25]. There are many possible ways for an airline to assign their flights to their slots. A flight can be canceled or routed out of the relevant FCA. The assignment of flights to slots is not just a one-time decision but can be changed repeatedly during the AFP. Uncertain factors impact the airline, such as if and how the FAA will alter the AFP parameters and when the AFP will end. Cancelations by other airlines can impact an airline's allocation of slots. Furthermore, the impact of delaying each flight is difficult to compute because passenger, luggage, crew, and aircraft connections mean that delaying one flight may impact several other flights. Mechanical or crew "time-out" issues can arise and further complicate matters. It is not obvious to researchers how airlines consider all of these factors when making decisions.

To make this problem tractable, a separable delay cost model will be assumed. More specifically, it is assumed that airlines attempt to minimize the sum of the delay costs associated with assigning each flight to each slot. This model assumes that airlines ignore the dynamic nature of the AFP, uncertainties, the possibility of canceling flights or routing them out of the FCA, and the behavior of other airlines, except when the flight delay cost function attempts to include such issues. More sophisticated models and solution methods have been developed that consider the possibility of canceling flights, non-separable cost functions, and other issues [3–7, 19]. These techniques may be more accurate, but the separable delay cost model allows for computationally simple evaluations of delay cost functions with AFP data.

If airlines minimize a separable delay cost, then the problem faced by airlines when assigning flights to slots is well-known and referred to as the "minimum cost perfect matching" or "assignment" problem. Given a set of flights and slots, a "matching" is a set of connections between flights and slots such that flights are matched to only one slot and vice versa. A "perfect matching" is any matching in which no flight or slot is left unmatched. Several algorithms can solve this problem efficiently, even for cases where there are hundreds or thousands of flights and slots. This is not necessarily the case for the more sophisticated problem models.

B. Notation

Before defining the four metrics used to evaluate the degree to which a delay cost function is consistent with historical

assignments of flights to slots in AFPs, some notation will be introduced. Let F be the set of flights belonging to an airline in a matching, and let S be the airline's slots in the matching. The number of flights and slots is n. Associated with each flight $f \in F$ is a scheduled time of arrival at the constrained resource t_f and associated with each slot $s \in S$ is a time t_s and a time window $[t_s, t_s + \delta]$ for some $\delta \geq 0$. A flight f can only be assigned to a slot s if $t_f \leq t_s + \delta$. The cost of assigning a flight f to a slot s is given by the cost function under consideration, which is a function of f and the delay associated with assigning f to s. There are historical assignments of flights to slots where $t_f > t_s$, so delay is computed as $d(f, s) = \max\{0, t_s - t_f\}$. This information makes up the data for the minimum cost perfect matching problem assumed to be solved by the airline.

The set of historical matchings of flights and slots by an airline is denoted by $(\mathcal{F}, \mathcal{S}, \mathcal{M})$. An element $(F, S, M) \in (\mathcal{F}, \mathcal{S}, \mathcal{M})$ contains the set of flights F and set of slots S associated with a matching M selected by an airline. The matching M is a perfect matching that assigns each $f \in F$ to exactly one $s \in S$. It is a square binary matrix with an entry for each possible assignment of a flight to a slot. Element M_{ij} is 1 if f_i is assigned to slot s_j and is 0 otherwise. For a cost function c_k, the cost of a matching is

$$J(F, S, M, c_k) = \sum_{f_i \in F, s_j \in S} c_k(f_i, d(f_i, s_j)) M_{ij}.$$

If a non-separable cost model were used, this equation could not be expressed as a sum of individual flight delay costs.

C. Metrics

The general approach underlying the metrics proposed here involves computing and comparing the airline cost, FSFS cost, and minimum cost corresponding to each airline matching. The computation of these total cost values for a particular flight-slot-matching triplet (F, S, M) is depicted in Fig. 1. Given this data and a particular c_k, the "airline cost" $J(F, S, M, c_k)$ can be computed. This cost incurred by the airline is compared to two other total costs values. The first of these is the "FSFS cost" produced by a FSFS matching M^{FSFS}: $J(F, S, M^{\mathrm{FSFS}}, c_k)$. The second of these is the "minimum cost" produced by any optimal matching $M^\star$ with cost function c_k: $J^\star(F, S, c_k)$.
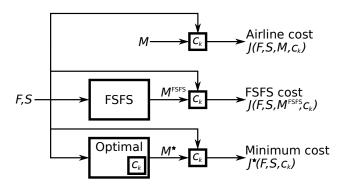

Figure 1: Calculation of cost values for a set of flights and slots.

Some example matchings are shown in Fig. 2. Fig. (a) is the default FSFS matching (M^{FSFS}), Fig. (b) is the airline-selected matching (M), Fig. (c) is a minimum cost perfect matching for some cost function A, and Fig. (d) is a minimum cost perfect matching for some cost function B. Fig. 3 shows example cost values incurred by these matchings.

(a) FSFS matching (M^{FSFS})   (b) Airline matching (M)

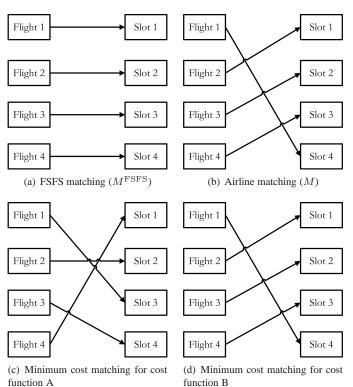(c) Minimum cost matching for cost function A   (d) Minimum cost matching for cost function B

Figure 2: Example matchings of a set of flights to a set of slots.
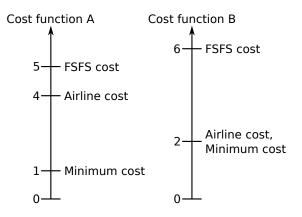

Figure 3: Example cost values.

One way to infer whether an airline was using a particular cost function is to see if the airline cost is lower than the default FSFS cost. In the example, the airline cost is slightly lower than the FSFS cost for function A, but more significantly lower for function B. Two metrics are based on comparing these two cost values. A second approach is to compare the airline cost to the minimum cost for each cost function. In the example, the

airline cost equals the minimum cost for cost function B but not for function A. The last two metrics are based on comparisons of the airline cost with the minimum cost for each cost function.

1) First-Scheduled-First-Served Ratio

The consistency of the matchings selected by airlines with a given cost function can be evaluated by comparing the airline cost to the corresponding FSFS cost for each set of flights and slots. The "first-scheduled-first-served ratio" metric for a set of flights and slots is the airline cost divided by the FSFS cost for a given cost function:

$$R_{\text{FSFS}} = \frac{\text{airline cost}}{\text{FSFS cost}}. \tag{1}$$

If the FSFS ratio is less than 1, then the airline cost is lower than the FSFS cost. FSFS ratio values larger than 1 mean that the airline could have incurred a lower cost with M^{FSFS}, the default matching. This would not happen if the airline was using c_k and solving a minimum cost perfect matching problem. Therefore, the distribution of FSFS ratio values in the data can be used to study consistency. The smaller the FSFS ratios are for a particular cost function, the more consistent the airline matchings are with the separable cost model assumption and that cost function. In the example, the FSFS ratio for the airline matching is $\frac{4}{5}$ for cost function A and $\frac{1}{3}$ for cost function B, indicating that this matching is more consistent with cost function B. Cost functions are ranked according to the median FSFS ratio over all the matchings ($\tilde{R}_{\text{FSFS}}$), with the 75$^{\text{th}}$ and 25$^{\text{th}}$ percentiles used as tiebreakers.

2) Improvement Frequency

A second metric of consistency is also based on a comparison with the FSFS cost and is referred to as the "improvement frequency." The more frequently that the airline costs are less than the FSFS costs, the more consistent airline actions are with that particular cost function. More precisely, the improvement frequency for cost function k is

$$I(c_k) = \frac{1}{N} \sum_{(\mathcal{F}, \mathcal{S}, \mathcal{M})} \mathbf{1}_{\{J(F, S, M, c_k) < J(F, S, M^{\text{FSFS}}, c_k)\}}, \tag{2}$$

where N is the number of flight-slot-matching triplets (F, S, M) in $(\mathcal{F}, \mathcal{S}, \mathcal{M})$ and $\mathbf{1}_{\{x\}}$ evaluates to 1 if some condition x is true and to 0 if x is false. The closer $I(c_k)$ is to 1, the more consistent an airline's actions are with c_k and the separable cost model assumption.

3) Minimum Ratio

The consistency of the matchings selected by airlines can also be evaluated by comparing the airline costs to the minimum costs. The "minimum ratio" for a set of flights and slots and a particular cost function is the ratio of the airline cost to the minimum cost. Given F, S, M, and c_k, the minimum ratio is

$$R_{\text{min}} = \frac{\text{airline cost}}{\text{minimum cost}}. \tag{3}$$

A low R_{min} value indicates that the airline cost is closer to the minimum cost for a particular set of flights and slots and a particular cost function. A value of 1 is the lowest possible value and it indicates that the matching M is a minimum cost perfect matching. In the example, the minimum ratio for M with cost function A is 4, but the minimum ratio with cost function B is 1, again indicating a greater consistency of M with cost function B. The overall consistency will be evaluated by studying the distribution the minimum ratio values in the data for each airline and for each delay cost function. The more frequently the values are close to 1 for a particular cost function, the more consistent the airline matchings are with the separable cost model assumption and that cost function. Again, the median of the minimum ratio values for all the matchings ($\tilde{R}_{\text{min}}$) will be used to rank cost functions, with the 75$^{\text{th}}$ and 25$^{\text{th}}$ percentiles used as tiebreakers.

4) Approximate Log-Likelihood

Even if airlines do minimize a separable delay cost when matching flights and slots, it is unlikely that their delay cost function can be computed exactly from publicly available data [25]. One way to handle this issue is to add a noise term to the delay cost function to account for unobserved factors that impact the cost. Assume that the actual cost of delaying a flight f by d minutes is $c(f, d) + \varepsilon$. The deterministic part of the cost that can be computed from available data is $c(f, d)$ and the stochastic part that accounts for unobserved factors is ε. Assume that for each assignment of a flight to a slot, ε is identically and independently distributed (iid). In reality this assumption is unlikely because if a particular flight is costly in a way that the deterministic part of the cost function does not account for, it will have a positive additive cost noise for each possible slot assignment. Additionally, assume that ε is normally distributed with mean μ and variance σ^2. If ε is assumed to be zero-mean, then the assumptions will be referred to as the zero-mean iid normal additive noise assumptions.

The fourth metric proposed in this research is based on a heuristic for approximating the variance of the zero-mean iid noise that maximizes the likelihood that an airline was using a given delay cost function. This heuristic is referred to as "Linear Program Cost Approximate Maximum Likelihood Estimation" (LPCAMLE), and it is based on linear programming sensitivity theory and maximum likelihood estimation [26]. The derivation of this heuristic is not included here. The metric is the approximate log-likelihood $\hat{L}(\sigma^{2\star})$ that an airline was using a given cost function with zero-mean normal delay cost noise and an LPCAMLE-estimated normalized variance $\sigma^{2\star}$. The zero-mean normalized variance estimate is

$$\sigma^{2\star} = \frac{1}{N} \sum_{(\mathcal{F}, \mathcal{S}, \mathcal{M})} \left(\frac{J(F, S, M - M^{\star}, c_k / \bar{c}_k)}{\|M - M^{\star}\|_2} \right)^2, \tag{4}$$

where here $\| \cdot \|_2$ is the entry-wise matrix 2-norm and $c_k / \bar{c}_k$ is the cost function k normalized by the mean flight delay cost for this cost function in the operational data. The approximate

log-likelihood is then computed as

$$\hat{L}(\sigma^{2\star}) = \sum_{(\mathcal{F}, \mathcal{S}, \mathcal{M})} \log g_{\sigma^{2\star}} \left(\frac{J(F, S, M - M^{\star}, c_k/\bar{c}_k)}{\|M - M^{\star}\|_2} \right),$$

(5)

where $g_{\sigma^2}(x)$ is a probability density function for a normal random variable with mean zero and variance σ^2 evaluated at x.

The standard deviation estimate normalized by the mean flight delay cost in the airline matchings ($\sigma^{\star}$) will also be reported. This quantity gives an idea of the relative magnitude of the deterministic and stochastic portions of the delay cost functions.

Validation efforts indicate that if a particular delay cost function is used to select matchings, then it achieves larger approximate log-likelihood values than other candidate functions. The main case in which the approximate log-likelihood does not work is when a cost function produces the same total cost of a matching for many or all possible matchings. Validation also suggests that the LPCAMLE estimates may not have stabilized until after 200 matching data points, so results for airlines with fewer than 200 matchings will not be presented.

D. Delay Cost Functions

A set of candidate cost functions are evaluated. The functions depend on publicly available (or at least approximable) characteristics of the flight f and the minutes that the flight is delayed d. Characteristics of the flight f include the airline operating the flight (a_f), the scheduled time of arrival (t_f), the number of passengers on the flight (p_f), and the aircraft type used for the flight (e_f). The cost functions that will be evaluated are documented in Table I.

The US Department of Transportation considers a flight delayed when it arrives 15 or more minutes after its scheduled arrival time, and it reports "on-time performance" data that may impact customer perception of airlines. Airlines attempt to reduce the number of flights that are counted as delayed [5, 7, 25]. Therefore, the first cost function is equal to 1 if a flight will be counted as delayed and 0 otherwise.

The second cost function is the minutes of delay multiplied by the number of passengers on the flight. This cost function has been used in an airline decision-support tool [19], and it is related to the number of passengers that will miss a connection when a flight is delayed by some amount. Cost functions 3 and 4 are the squared delay and squared passenger delay, respectively. Functions of this form are proposed in [12]; they provide a simple means of capturing increasing marginal delay costs due to missed connections.

The Time-of-Day Delay (cost function 5) is the minutes of delay multiplied by a multiplier that is a function of the scheduled time of arrival and the minutes of delay. This function is based on [22], in which an airline schedule was analyzed to quantify how the magnitude and time of day of a delay impact airline delay costs. It has also been used in airline decision-support tools [27]. The form of the multiplier $\beta(t_f, d)$ can be found in [22].

Cost function 6 is referred to as Connection Delay and it was proposed in [23] and [24]. It attempts to capture the fact that delaying flights bound for hub airports is especially costly because these flights are likely to involve passengers, crews, and aircraft that need to connect to other flights. The cost is computed as the minutes of delay times a multiplier $\gamma(f)$ that is 2 for flights bound for high connection rate airports (also known as hubs), 1.5 for flights bound for medium connection rate airports, and 1 for all other flights. The classification of airports into these categories is specified in [24]. An airline-specific version of this cost function was also developed and is referred to as the Airline Connection Delay (cost function 7). In this function, the multiplier $\gamma'(f, a_f)$ is a function of the airline: the high connection rate and medium connection rate airports vary from airline to airline.

Previous research has attempted to calculate the monetary cost of delay in Europe [28]. This work has been adapted for the US market [29]. Cost function 8 is the Monetary Delay, and it is an implementation of the function in [29]. This cost function is also computed by multiplying the minutes of delay by a multiplier $\eta(e_f, d)$ that is a sum of per-minute fuel, crew, maintenance, passenger, and other costs.

Cost function 9 is referred to as the Step Function because it generates costs that increase in discrete "steps" as various delay thresholds are exceeded. The motivation for this form is that, for example, a delay of less than 60 minutes is assumed to provide sufficient time for passengers to make connections but a delay greater than 60 minutes does not. A function of this form has been used in airline decision-support tools [7, 20].

The remaining eight cost functions include two or more of these first nine cost functions. Many factors impact airline actions [25], and these functions involve more factors than any of the first nine functions do on their own. Cost functions 10–15 involve a product of two or more multipliers and delay. For example, the product of the Time-of-Day Delay multiplier, the Monetary Delay multiplier, and the delay is cost function 14, which has been used in previous research [18, 30]. Cost functions 16 and 17 are convex combinations of two other cost functions (which means that $\alpha \in [0, 1]$). The α_{16} and α_{17} parameters in these cost functions were tuned by hand; the best performance was observed when they were both set to $\frac{13}{14}$.

E. Data

The historical matchings used in this study are recorded in 34 days of Expected Departure Clearance Time (EDCT) log files from June–August 2006. The files contain information about airline actions during GDPs and AFPs throughout the National Airspace System [31, 32]. "Simplified Substitution" messages in these files specify sets of flights, sets of slots, and the corresponding airline-selected matching.

These messages contain enough information to define the minimum cost perfect matching problems that it is assumed that the airline solved to select the specified matching. However, some assumptions were made in processing the EDCT log file data. For example, the scheduled time of arrival t_f for a flight

Number	Name	Function
1	On-time Performance	$c_1(f,d) = \mathbf{1}_{\{d>15\}}$
2	Passenger Delay	$c_2(f,d) = p_f d$
3	Squared Delay	$c_3(f,d) = d^2$
4	Squared Passenger Delay	$c_4(f,d) = (p_f d)^2$
5	Time-of-Day Delay	$c_5(f,d) = \beta(t_f,d)d$
6	Connection Delay	$c_6(f,d) = \gamma(f)d$
7	Airline Connection Delay	$c_7(f,d) = \gamma'(f,a_f)d$
8	Monetary Delay	$c_8(f,d) = \eta(e_f,d)d$
9	Step Function	$c_9(f,d) = \rho(d)$
10	Time-of-Day Connection Delay	$c_{10}(f,d) = \beta(t_f,d)\gamma(f)d$
11	Time-of-Day Passenger Delay	$c_{11}(f,d) = \beta(t_f,d)p_f d$
12	Connection Passenger Delay	$c_{12}(f,d) = \gamma(f)p_f d$
13	Time-of-Day Connection Passenger Delay	$c_{13}(f,d) = \beta(t_f,d)\gamma(f)p_f d$
14	Time-of-Day Monetary Delay	$c_{14}(f,d) = \beta(t_f,d)\eta(e_f,d)d$
15	Connection Monetary Delay	$c_{15}(f,d) = \gamma(f)\eta(e_f,d)d$
16	Connection and Monetary Combination Delay	$c_{16}(f,d) = \alpha_{16}c_6(f,d) + (1-\alpha_{16})c_8(f,d)$
17	Airline Connection and Monetary Combination Delay	$c_{17}(f,d) = \alpha_{17}c_7(f,d) + (1-\alpha_{17})c_8(f,d)$

was set equal to the EENTRY field in the EDCT log files, but EENTRY is not actually the scheduled time of arrival but rather an estimate of the earliest time the flight can arrive at the FCA. Furthermore, when airlines choose to keep the default FSFS assignment of flights to slots it is not recorded in the EDCT file. EDCT log files give an incomplete picture of how each airline used slots, which impacts the analysis presented here.

Other sources of data that were used in computing the cost function values were Aircraft Situational Display to Industry (ASDI) files, OAG data about the number of seats on various aircraft types [33], and an average load factor computed by GRA [34].

At any time during an AFP, airlines can submit Simplified Substitution messages to the FAA that specify sets of flights and slots and a matching. Some airlines specify matchings frequently while others do so relatively rarely. A histogram of the number of matching messages for the 18 airlines in the data set is shown in Fig. 4. Larger numbers of matchings will lead to more meaningful results. There were Simplified Substitution messages specifying matchings for 18 airlines, but 11 of those airlines submitted matching messages less than 100 times in the 34 days in the data set.

Each Simplified Substitution message can specify as many flights and slots as the airline would like to match. Some airlines match many flights and slots, but more frequently only a few flights and slots are matched. Histograms of the number of flights and slots in the matchings submitted by two airlines are presented in Fig. 5. Matchings with more flights and slots reveal more about airline preferences than matchings with just a few flights and slots because airlines only have a few choices when there are only a few flights and slots. For both airlines, the majority of matchings involve less than 10 flights and slots. However, airline E has some matchings with more than 100
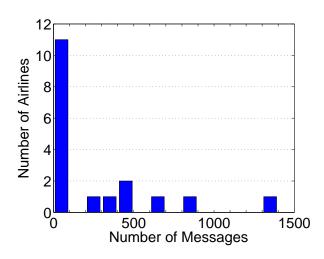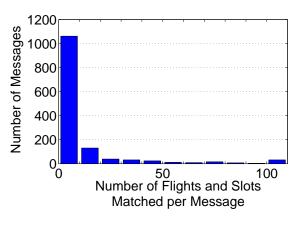

Figure 4: Histogram of the number of matching messages for each airline in the data set.

flights and slots. The largest matching for airline G contains less than 30 flights and slots.

IV. RESULTS

The four metrics described in sub-section III.C were computed with the data described in sub-section III.E. The cost functions that are most consistent with the airline actions in the data are presented here.

Validation efforts suggest that at least 200 matchings are needed for one of the metrics. There are seven airlines with more than 200 matchings for which results will be presented. The number of matchings and median number of flights and slots in the matchings for these airlines are presented in Table II.
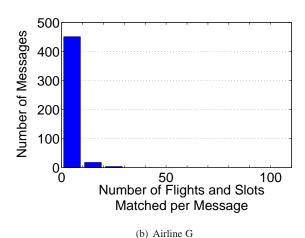
(a) Airline E

(b) Airline G

Figure 5: Histograms of the number of flights and slots for the matchings of two airlines. The bar furthest to the right counts all entries with more than 100 flights and slots.

TABLE II: AIRLINE MATCHING CHARACTERISTICS

Airline	Matchings	Median Number of Flights and Slots
A	834	4
B	410	12.5
C	293	9
D	302	3
E	1368	4
F	618	3
G	473	2

A. FSFS Ratio

The cost functions with the lowest median FSFS ratio ($\tilde{R}_{\text{FSFS}}$) for each airline are shown in Table III. The column labeled "1st" contains the number of the cost function from Table I with the lowest median FSFS ratio, etc. Asterisks designate cost functions that are tied even after using the tiebreakers. The $\tilde{R}_{\text{FSFS}}$ values are often equal to 1 because the airline-selected matchings often achieve the same cost as the FSFS matching, particularly when the matchings are small.

TABLE III: COST FUNCTIONS WITH LOWEST MEDIAN FSFS RATIO

Airline	1st	$\tilde{R}_{\text{FSFS}}$	2nd	$\tilde{R}_{\text{FSFS}}$	3rd	$\tilde{R}_{\text{FSFS}}$
A	1	1.000	9	1.000	6	1.000
B	9	0.978	1	1.000	2	1.006
C	9	1.000	1	1.000	16	1.013
D	9	1.000	1	1.000	7	1.000
E	1	1.000	9	1.000	12	1.000
F	9	1.000	1*	1.000	6*	1.000
G	1*	1.000	2*	1.000	9*	1.000

For every airline, cost functions 1 (On-time Performance) and 9 (Step Function) achieved or tied for the first- and second-lowest $\tilde{R}_{\text{FSFS}}$ values (after ties were broken). These cost functions are similar in that they both produce costs that increase in discrete steps as delays increase. Other cost functions that are among the top three most consistent cost functions with the actions of some airline according to the FSFS ratio are cost functions 2 (Passenger Delay), 6 (Connection Delay), 16 (Connection and Monetary Combination Delay), and 12 (Connection Passenger Delay).

B. Improvement Frequency

The cost functions with the largest improvement frequency for each airline are shown in Table IV. Asterisks designate cost functions with the same I value. Among these cost functions the improvement frequency ranges from 0.188 to 0.522. These values are low on a scale from 0 to 1, partially because I does not count the many instances in which the airline cost equals the FSFS cost. Relatively low values for I do not necessarily mean that the separable cost model assumption is invalid. They could be the result of not using appropriate delay cost functions or the many cases where the airline cost equals the FSFS cost.

TABLE IV: COST FUNCTIONS WITH LARGEST IMPROVEMENT FREQUENCY

Airline	1st	I	2nd	I	3rd	I
A	1	0.363	9	0.357	12	0.213
B	9	0.522	1	0.463	12	0.322
C	9	0.352	1	0.317	8*, 12*	0.188
D	9	0.457	12	0.424	16	0.391
E	12	0.477	2	0.438	9	0.392
F	9	0.354	15	0.291	16	0.283
G	9	0.241	1	0.197	8	0.152

As was the case when cost functions were evaluated with $\tilde{R}_{\text{FSFS}}$, the I values indicate that cost functions 1 (On-time Performance) and 9 (Step Function) are often most consistent with airline actions. Based on I values, cost function 12 (Connection Passenger Delay) is among the top three most consistent cost functions for all but two of the airlines. Cost functions 2 (Passenger Delay), 8 (Monetary Delay), and 16 (Connection and Monetary Combination Delay) also place in the top three for at least one airline according to I values.

C. Minimum Ratio

The cost functions with the smallest $\tilde{R}_{\min}$ values for each airline are shown in Table V. Again, ties are broken with the 75th and 25th percentiles. Among these cost functions, it is common for the $\tilde{R}_{\min}$ value to be 1, indicating that airline-selected matchings frequently achieve minimum costs.

TABLE V: COST FUNCTIONS WITH LOWEST MEDIAN MINIMUM RATIO

Airline	1st	$\tilde{R}_{\min}$	2nd	$\tilde{R}_{\min}$	3rd	$\tilde{R}_{\min}$
A	1	1.000	7	1.028	6	1.055
B	7	1.085	6	1.105	5	1.115
C	6	1.095	5	1.104	16	1.110
D	7	1.000	9	1.000	1	1.000
E	6	1.000	9	1.000	1	1.000
F	1	1.000	6	1.000	7	1.000
G	2	1.000	17	1.000	7	1.000

According to the minimum ratio, cost functions 6 (Connection Delay) or 7 (Airline Connection Delay) are most consistent with the matchings of more than half of the airlines. At least one of these two similar cost functions places in the top three for every airline. Cost functions that are consistent with airline actions according to both the minimum ratio and the previous two metrics are cost functions 1 (On-time Performance), 9 (Step Function), 16 (Connection and Monetary Combination Delay), and 2 (Passenger Delay). Like cost functions 6 and 7, cost functions 5 (Time-of-Day Delay) and 17 (Airline Connection and Monetary Combination Delay) are identified as consistent with the actions of some airlines by the minimum ratio, but not by the FSFS ratio or I. These differences occur because the FSFS ratio and I evaluate consistency based on the FSFS cost while the minimum ratio uses the minimum cost.

D. Approximate Log-Likelihood

The cost functions with the largest $\hat{L}(\sigma^{2\star})$ for each airline are shown in Table VI. The approximate log-likelihood values can be used to see the relative performance of the cost functions for each airline but cannot be compared across airlines because each airline has a different number of matchings in the data.

The corresponding estimates of the standard deviation of the additive cost noise normalized by the average cost per assignment are also in this table. These $\sigma^\star$ values do not indicate the consistency of the airline matchings with a cost function and the separable cost model and zero-mean additive cost noise assumptions. However, they do indicate the relative magnitudes of the observed and unobserved aspects of flight delay costs. Smaller $\sigma^\star$ values indicate that the the airline matchings are best explained with additive cost noise values that are relatively small compared to the deterministic part of the cost functions. Most are between 0.1 and 0.7, but validation work suggests that these are likely under-estimates.

As was suggested by the minimum ratio, the closely-related cost functions 6 (Connection Delay) and 7 (Airline Connection Delay) are most consistent with the matchings of most of the airlines. Cost functions 16 (Connection and Monetary Combination Delay) and 17 (Airline Connection and Monetary Combination Delay) are other similar cost functions that achieve one of the top three largest $\hat{L}(\sigma^{2\star})$ values for most of the airlines. Other cost functions that achieve top-three $\hat{L}(\sigma^{2\star})$ values for at least one airline are 1 (On-time Performance), 5 (Time-of-Day Delay), 7 (Airline Connection Delay), 8 (Monetary Delay), and 2 (Passenger Delay).

V. FUTURE WORK

This work could be immediately improved by using Aggregate Demand List (ADL) files rather than EDCT log files. ADL files more accurately capture what actions airlines took during AFPs than EDCT log files [35]. Another immediate extension would be to analyze GDP data as well as AFP data. With a small change to the minimum cost perfect matching problem, cancellations and route-outs could also be studied with the four metrics proposed here. Finally, more cost functions could be analyzed, particularly combinations of existing cost functions.

Some delay cost functions can achieve similar or identical total delay costs for many possible matchings while other functions will achieve similar or identical total delay costs for few or none of the possible matchings. This may bias the results presented here and should be addressed explicitly in future work.

Even if this work were extended to consider cancellations and route-outs, the assumption that airlines minimize a separable cost leads to a simple model of their behavior in AFPs. More non-separable cost functions should be evaluated with airline action data. The uncertain dynamics of AFPs also may play an important role in airline decisions, and this should be studied.

VI. CONCLUSIONS

Valid models of airline behavior are essential for meaningful air traffic management research. In this paper, airline actions in Airspace Flow Programs were used to evaluate several proposed flight delay cost functions used in separable airline cost models. Two different classes of cost functions were identified as most consistent with airline actions because two different classes of metrics for evaluating consistency were used. When the consistency of an airline's matchings with a cost function is evaluated by comparing the costs achieved by the airline matchings with the costs of the default first-scheduled-first-served matchings, cost functions 1 (On-time Performance) and 9 (Step Function) are most consistent with the matchings of most airlines. These functions produce costs that increase in discrete steps as delay thresholds are exceeded. When the consistency of an airline's matchings with a cost function is evaluated by comparing the costs achieved the airline matchings to the minimum costs, cost functions 6, 7, 16, and 17, all of which are closely related to Connection Delay, are most consistent with the matchings of most airlines. These cost functions produce costs that are proportional to the length of the delay but with proportionality constants that are larger for flights bound to hub airports.

Finally, the linear programming cost approximate maximum likelihood method estimates the standard deviation of a

TABLE VI: Cost Functions with Largest Approximate Log-Likelihood

Airline	1st	$\hat{L}(\sigma^{2\star})$	$\sigma^\star$	2nd	$\hat{L}(\sigma^{2\star})$	$\sigma^\star$	3rd	$\hat{L}(\sigma^{2\star})$	$\sigma^\star$
A	7	−1159	0.487	6	−1252	0.539	1	−1318	1.216
B	7	−773.5	0.454	5	−858.3	0.509	17	−873.8	0.525
C	6	−591.3	0.587	16	−603.6	0.615	7	−612.7	0.623
D	7	12.97	0.037	17	−150.4	0.133	16	−281.6	0.246
E	6	−1627	0.369	16	−1715	0.336	17	−1785	0.351
F	16	86.98	0.064	17	−9.009	0.083	6	−48.54	0.050
G	2	−0.366	0.051	17	−239.5	0.164	8	−268.6	0.185

noise term that was added to cost functions to account for unobserved aspects of airline costs. The standard deviation values, expressed as a fraction of the average assignment cost for the historical matchings, ranged from 0.1 to 0.7 for cost functions with relatively large approximate log-likelihoods.

References

[1] J. Xiong, *Revealed Preference of Airlines' Behavior under Air Traffic Management Initiatives*. PhD thesis, University of California, Berkeley, 2010.

[2] J. W. Bono and D. H. Wolpert, "Analyzing policy risk and accounting for strategy: Auctions in the National Airspace System," Working paper 2010-04, American University, Department of Economics, February 2010.

[3] G. C. Carr, H. Erzberger, and F. Neuman, "Airline arrival prioritization in sequencing and scheduling," in *Proc. of USA/Europe Air Traffic Management Research & Development Seminar*, (Orlando, FL), December 1998.

[4] S. Luo and G. Yu, "Airline schedule perturbation problem: Landing and takeoff with nonsplittable resource for the Ground Delay Program," in *Operations Research in the Airline Industry* (G. Yu, ed.), ch. 14, pp. 404–432, Springer, 1998.

[5] H. Ergin, "Lessons learned in implementing a slot optimizer during Ground Delay Programs," in *Proc. of Airline Group of the International Federation of Operational Research Societies (AGIFORS) Airline Operations Meeting*, 2009.

[6] K. Abdelghany, A. Abdelghany, and T. Niznik, "Managing severe Airspace Flow Programs: The airlines' side of the problem," in *Proc. of Airline Group of the International Federation of Operational Research Societies (AGIFORS) Airline Operations Meeting*, (Denver, CO), 2007.

[7] L. Gumireddy and I. Ince, "Optimization for heirarchical objectives during Ground Delay Programs," in *Proc. of Airline Group of the International Federation of Operational Research Societies (AGIFORS) Airline Operations Meeting*, (Rome, Italy), 2002.

[8] H. Balakrishnan, "Techniques for reallocating airport resources during adverse weather," in *Proc. of IEEE Conference on Decision and Control*, (New Orleans, LA), December 2007.

[9] M. E. Berge, C. A. Hopperstad, and Áslaug Haraldsdóttir, "Airline schedule recovery in collaborative flow management with airport and airspace capacity constraints," in *Proc. of USA/Europe Air Traffic Management Research & Development Seminar*, (Budapest, Hungary), June 2003.

[10] T. Vossen and M. Ball, "Optimization and mediated bartering models for Ground Delay Programs," *Naval Research Logistics*, vol. 53, pp. 75–90, 2006.

[11] T. W. M. Vossen and M. O. Ball, "Slot trading opportunities in collaborative Ground Delay Programs," *Transportation Science*, 2005.

[12] M. Ball, G. Dahl, and T. Vossen, "Matchings in connection with Ground Delay Program planning," *Networks*, vol. 53, no. 3, pp. 293–306, 2009.

[13] A. Ranieri and L. Castelli, "A market mechanism to assign air traffic flow management slots," in *Proc. of USA/Europe Air Traffic Management Research & Development Seminar*, (Napa, CA), June 2009.

[14] Y. Zhang and M. Hansen, "Regional GDP – extending Ground Delay Programs to regional airport systems," in *Proc. of USA/Europe Air Traffic Management Research & Development Seminar*, (Napa, CA), June 2009.

[15] J. Henderson, H. Idris, S. Ferguson, J. Krozel, and R. Kicinger, "User and service provider collaboration on flight route and delay under uncertainty," in *Proc. of AIAA Guidance, Navigation, and Control Conference*, (Toronto, Canada), August 2010.

[16] J. Rios, K. Sheth, and S. Gutierrez-Nolasco, "Incorporating user preferences within an optimal traffic flow management framework," in *Proc. of AIAA Guidance, Navigation, and Control Conference*, (Toronto, Canada), August 2010.

[17] S. Gutierrez-Nolasco and K. S. Sheth, "Analysis of factors for incorporating users preferences in air traffic management: A users' perspective," in *Proc. of AIAA Aviation Technology, Integration, and Operations Conference*, (Fort Worth, TX), September 2010.

[18] H. Gao, G. Hunter, F. Berardino, and K. Hoffman, "Development and evaluation of market-based traffic flow management concepts," in *Proc. of AIAA Aviation Technology, Integration, and Operations Conference*, (Fort Worth, TX), September 2010.

[19] A. Vasquez-Marquez, "American Airlines Arrival Slot Allocation System (ASAS)," *Interfaces*, vol. 21, January-February 1991.

[20] M. Brennan, "Simplified Substitutions – enhancements to substitution rules and procedures during Ground Delay Programs," in *Proc. of Airline Group of the International Federation of Operational Research Societies (AGIFORS) Airline Operations Meeting*, (Ocho Rios, Jamaica), May 2001.

[21] T. J. Niznik, "Optimizing the airline response to Ground Delay Programs," in *Proc. of Airline Group of the International Federation of Operational Research Societies (AGIFORS) Airline Operations Meeting*, (Ocho Rios, Jamaica), May 2001.

[22] R. Beatty, R. Hsu, L. Berry, and J. Rome, "Preliminary evaluation of flight delay propagation through an airline schedule," in *Proc. of USA/Europe Air Traffic Management Research & Development Seminar*, (Orlando, FL), October 1998.

[23] H. D. Sherali, R. W. Staats, and A. A. Trani, "An airspace-planning and collaborative decision-making model: Part II – cost model, data considerations, and computations," *Transportation Science*, vol. 40, pp. 147–164, May 2006.

[24] R. W. Staats, *An Airspace Planning and Collaborative Decision Making Model Under Safety, Workload, and Equity Considerations*. PhD thesis, Virginia Polytechnic Institute and State University, April 2003.

[25] H. Idris, A. Evans, R. Vivona, J. Krozel, and K. Bilimoria, "Field observations of interactions between traffic flow management and airline operations," in *Proc. of AIAA Aviation Technology, Integration and Operations Conference*, (Wichita, KS), September 2006.

[26] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, UK: Cambridge University Press, 2004.

[27] D. Hoitomt, D. Kraay, and B. Tang, "Flight sequencing under the FAA's Simplified Substitution rules," in *Proc.* *of Airline Group of the International Federation of Operational Research Societies (AGIFORS) Airline Operations Meeting*, (Istanbul, Turkey), April 1999.

[28] A. Cook, G. Tanner, and S. Anderson, "Evaluating the true cost to airlines of one minute of airborne or ground delay," Final Report, Performance Review Commission, Eurocontrol, April 2004.

[29] A. Q. Kara, J. Ferguson, K. Hoffman, and L. Sherry, "Estimating domestic US airline cost of delay based on European model," in *Proc. of International Conference on Research in Air Transportation*, (Budapest, Hungary), June 2010.

[30] H. Gao and G. Hunter, "Evaluation of user gaming strategies in the future National Airspace System," in *Proc. of AIAA Aviation Technology, Integration, and Operations Conference*, (Fort Worth, TX), September 2010.

[31] K. Howard and B. Sharick, ATMS/ETMS requirements, Federal Aviation Administration, `http://cdm.fly.faa.gov/ad/CDM-GDP_specs.htm`, March 2003.

[32] "Interface control document for substitutions during Ground Delay Programs, Ground Stops, and Airspace Flow Programs," interface control document, Federal Aviation Administration, `http://www.fly.faa.gov/Products/NASDOCS/nasdocs.jsp`, November 2006.

[33] OAG, "Aircraft information." `http://www.oag.com/NorthAmerica/airlineandairport/aircraftstatistics.asp`, 2010.

[34] GRA, Incorporated, "Economic values for FAA investment and regulatory decisions, a guide," Tech. Rep. Contract No. DTFA 01-02-C00200, FAA Office of Aviation Policy and Plans, Washington, DC, October 2007.

[35] K. Howard and M. Lehky, "Aggregate Demand List (ADL)/FSM broadcast data formats - version 10 revision 4." John A. Volpe National Transportation Systems Center Memorandum, `http://www.fly.faa.gov/Products/NASDOCS/ADL_Ver10_Rev4.pdf`, November 2005.

AUTHOR BIOGRAPHIES

Michael Bloem earned a B.S.E. degree with majors in electrical and computer engineering and economics from Calvin College in 2004 and an M.S. degree in electrical and computer engineering from the University of Illinois at Urbana-Champaign in 2007. He researches dynamic airspace configuration and traffic flow management in the Systems Modeling and Optimization branch of the Aviation Systems Division at NASA Ames Research Center, Moffett Field, CA. Mr. Bloem is a member of the AIAA.

Haiyun Huang is currently studying at the faculty of Aerospace Engineering at Delft University of Technology in the Netherlands, with a focus on Maintenance, Repair and Overhaul, Flight Path Operations and Air Traffic Management. He will receive his B.Sc. degree in Aerospace Engineering in the summer of 2011 and his M.Sc. degree in Aerospace Engineering will follow at the end of 2011.