51st IEEE Conference on Decision and Control
December 10-13, 2012. Maui, Hawaii, USA

Approximating the Likelihood of Historical Airline Actions to Evaluate Airline Delay Cost Functions

Michael Bloem[1], Haiyun Huang[2], and Nicholas Bambos[3]

Abstract— **Delay cost functions that quantify the cost of delay to airlines are essential to air traffic management research. Seventeen delay cost functions from previous research are evaluated with airline actions in Airspace Flow Programs. Airlines are assumed to solve a minimum cost perfect matching problem when matching flights to slots. Unobserved aspects of airline costs are accounted for by adding a noise term to the cost functions. The goal of this research is to find the cost function and corresponding noise parameters that maximize the likelihood of airline actions during 32 Airspace Flow Programs in the summer of 2006. A heuristic is developed that finds cost noise parameters that maximize an approximation of the log-likelihood of the airline actions. When applied to sample estimation problem instances generated by solving linear programming problems with known noise parameters, the heuristic can more accurately estimate noise parameters than a simple simulation-based approach. Validation efforts based on synthetic airline action data generated with known delay cost functions and noise parameters demonstrate that the heuristic is in many cases able to correctly identify as most likely the delay cost function that was in fact used to generate the synthetic data. However, the heuristic also under-estimates the magnitude of the cost noise variance on these estimation problem instances. Delay costs that are proportional to the length of delay, but with larger proportionality constants for flights bound for hub airports, maximize the approximation of the log-likelihood of the historical airline actions. The estimated standard deviations of the cost noise, expressed as a fraction of the average assignment cost for the historical matchings, ranged from 0.1 to 0.7 for cost functions that achieved relatively large approximate log-likelihoods.**

I. Introduction

Flight delays impose undesirable financial costs on airlines. These costs can be most accurately estimated by airlines because they alone are aware of most of the relevant factors, but airlines are reluctant to reveal their costs because doing so could be advantageous to their competitors. Air traffic management researchers are more likely to develop tools that help airlines and other stakeholders if they accurately understand airline behavior. This is true even for researchers designing collaborative mechanisms designed to elicit cost-related information from airlines and to allocate capacity accordingly. Without the benefit of knowing the actual financial cost of delay for airlines, researchers typically assume that airlines take actions to minimize a delay cost function that can be computed with publicly-available data. Hopefully, this function is strongly correlated with the financial costs that airlines are actually trying to reduce and therefore leads to accurate airline behavior models.

Aside from our own previous work [1], only one effort has been made to tune and validate these delay cost functions with records of airline actions. In her dissertation [2], 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 can make based on the cost of each choice. Xiong's research revealed many characteristics of airline delay costs, but it also has some limitations. Most 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, most air traffic management research assumes a separable delay cost model. Some airline decision support tools are also based on separable delay cost models [3]–[5]. 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, such as the minutes of delay multiplied by a weight related to the time-of-day or destination airport [6], [7].

The goal of this research is to find, from a set of delay cost functions proposed for separable delay cost models, the functions and corresponding additive cost noise parameters that maximize the likelihood of historical airline actions in Airspace Flow Programs. To this end, a heuristic is developed that finds cost noise parameters that maximize an approximation of the log-likelihood of the airline actions. In our previous work, this heuristic was applied to this problem but was not justified or validated [1]. In this paper, the heuristic is justified and some validation of its performance is presented.

The remainder of this paper is structured as follows. Section II provides background information about Airspace Flow Programs. The airline behavior model, maximum likelihood estimation problem, and two heuristics for this problem are defined in Section III. In Section IV, estimation problem instances with known noise parameters are used to investigate the validity of the heuristics. Airline delay cost functions are evaluated in Section V and conclusions are in Section VI.

[1]NASA Ames Research Center, Moffett Field, CA 94035-0001 michael.bloem@nasa.gov
[2]Science & Technology Corporation, Hampton, VA 23666-6413 and Delft University of Technology, Delft, the Netherlands h.huang@student.tudelft.nl
[3]Stanford University, Stanford, CA 94305 bambos@stanford.edu

II. Background

An Airspace Flow Program (AFP) is a mechanism used by the Federal Aviation Administration (FAA) to assign departure delays to aircraft in order to reduce demand for a region of airspace known as a "Flow Constrained Area" (FCA). AFPs are based on slots. A slot is the right to fly into an FCA during a specified period of time. The FAA assigns departure times to flights so that each flight arrives at the FCA approximately at the time of the slot to which it is assigned. Slots are allocated to airlines with an algorithm 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 the airline can adjust this assignment. Some airlines alter assignments of flights to slots in AFPs thousands of times each year. In this paper, records of these airline assignment actions are used to evaluate airline delay cost functions.

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 [8]. For example, 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. To make this problem tractable, a separable delay cost model will be utilized. 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.

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" 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.

Let F be the set of flights belonging to an airline in a matching, and let S be a set of all of the airline's slots in the matching. The number of flights and slots is n. Associated with each flight $f_i \in F$ is a scheduled time of arrival at the constrained resource t_{f_i} and associated with each individual slot $s_j \in S$ is a time t_{s_j} and a time window $[t_{s_j}, t_{s_j} + \delta]$ for some $\delta \geq 0$. A flight f_i can only be assigned to a slot s_j if it can arrive at the FCA before the end of the time window corresponding to the slot (that is, if $t_{f_i} \leq t_{s_j} + \delta$). There are historical assignments of flights to slots where $t_{f_i} > t_{s_j}$, so delay is computed as $d(f_i, s_j) = \max\{0, t_{s_j} - t_{f_i}\}$. 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 the perfect matching M selected by an airline. Matrix M is a square $n \times n$ 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 that computes a cost of $c(f_i, d)$ associated with delaying flight f_i by d minutes, the cost of a matching is $\sum_{i=1}^{n} \sum_{j=1}^{n} c(f_i, d(f_i, s_j)) M_{ij}$.

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 [8]. 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 f_i by assigning it to s_j is $c(f_i, d(f_i, s_j)) + \varepsilon_{ij}$. The deterministic part of the cost that can be computed using only observed publicly-available data is $c(f_i, d(f_i, s_j))$ and the stochastic part that accounts for unobserved factors known only to the airline is ε_{ij}. Assume that ε_{ij} is identically and independently distributed (iid) for all $i, j \in \{1, 2, \ldots, n\}$. This assumption is unlikely to be true in some cases because, for example, if delays for a particular $f_{i'}$ are costly in a way that is not accounted for by the deterministic part of the cost function, the corresponding additive cost noise random variables $\varepsilon_{i'j}$ $\forall j = 1, 2, \ldots, n$ may not be independent and may all have a positive mean that other ε_{ij} variables do not have. Although unlikely in some cases, this assumption is made because it enables the development of the heuristic described in sub-section III-C. Additionally, assume that the distribution of the ε_{ij} variables is Gaussian with mean μ and variance σ^2.

Under these assumptions, the minimum cost perfect matching problem faced by the airlines is

$$\begin{aligned} \text{minimize} \quad & \sum_{i=1}^{n} \sum_{j=1}^{n} (c(f_i, d(f_i, s_j)) + \varepsilon_{ij}) X_{ij} \\ \text{subject to} \quad & \sum_{j=1}^{n} X_{ij} = 1 \quad \forall i \in \{1, 2, \ldots, n\} \\ & \sum_{i=1}^{n} X_{ij} = 1 \quad \forall j \in \{1, 2, \ldots, n\} \\ & X_{ij} \in \{0, 1\} \quad \forall i, j \in \{1, 2, \ldots, n\}, \end{aligned} \tag{1}$$

where X_{ij} is 1 when f_i is assigned to s_j and 0 otherwise. Since ε_{ij} models factors impacting the cost that are known by the airline but unobserved by the public, the airlines are assumed to solve a deterministic optimization problem in which the realized ε_{ij} values for all $i, j \in \{1, 2, \ldots, n\}$ are revealed before the optimization.

B. Estimation Problem Statement

The problem under consideration in this paper is to find the delay cost functions and corresponding noise parameters that maximize the likelihood of airline actions in AFPs. More precisely, for each airline and for each candidate cost function c_k, we seek to find the noise parameters θ that maximize the likelihood $L_k(\theta) = \Pr(\mathcal{M}|\theta, \mathcal{F}, \mathcal{S})$. If each $(F, S, M) \in (\mathcal{F}, \mathcal{S}, \mathcal{M})$ is independent, the log-likelihood is

$$\ell_k(\theta) = \sum_{(F,S,M)\in(\mathcal{F},\mathcal{S},\mathcal{M})} \log \Pr(M|\theta, F, S). \qquad (2)$$

The delay cost function that maximizes the likelihood is the one for which θ values are found that achieve the largest value for $\ell_k(\theta)$ (and therefore also the largest $L_k(\theta)$).

C. Linear Program Cost Approximate Maximum Likelihood Estimation

The Linear Program Cost Approximate Maximum Likelihood Estimation (LPCAMLE) heuristic attempts to find the maximum likelihood estimates of parameters of a Gaussian noise term that is added to Linear Program (LP) cost vectors. It uses a data set consisting of LP instance coefficient values and corresponding LP solutions, but not any samples from the noise distribution. It is based on an approximation of the likelihood motivated by LP sensitivity analysis referred to as the LP Sensitivity Analysis (LPSA) approximation.

Suppose an entity solves a set of N deterministic LPs of the form

$$\begin{array}{ll} \text{minimize} & (c_i + \epsilon_i)^T x_i \\ \text{subject to} & A_i x_i = b_i \\ & x_i \succeq 0, \end{array} \qquad (3)$$

where, before the LP instances are solved, each element of each ϵ_i vector is sampled independently from the same Gaussian distribution with mean μ and variance σ^2. Let $\theta = (\mu, \sigma^2)$ denote the parameters of this Gaussian distribution. The LP instance coefficient and solution data available when estimating θ is $\{x_i^\star, A_i, b_i, c_i\}_{i=1}^N$, where $x_i^\star$ is the solution of problem instance i selected by the entity under observation.

Let $J^\star(A_i, b_i, c_i, \epsilon_i)$ be the optimal value of (3) for a particular noise sample ϵ_i and for problem parameters A_i, b_i, and c_i. Similarly, let $x_i^{\star,\epsilon_i}$ be optimal primal variables for LP instance i and a particular noise sample ϵ_i.

LP sensitivity analysis can be used to approximate the optimal solution of problem (3) as ϵ_i changes [9]. The approximation of the solution is

$$J^\star(A_i, b_i, c_i, \epsilon_i) \approx J^\star(A_i, b_i, c_i^j, 0) + \epsilon_i^T x_i^{\star,0}. \qquad (4)$$

The likelihood of the solution data $x_i^\star$ for a single LP instance is, by definition, the probability that $x_i^\star$ achieves the minimum cost, given the noise distribution (parameterized by θ) and the LP instance parameters. This probability is

$$\Pr(x_i^\star|\theta, A_i, b_i, c_i) = \Pr\left((c_i + \epsilon_i)^T x_i^\star = J^\star(A_i, b_i, c_i, \epsilon_i)\middle|\, \theta\right).$$

Approximation (4) can be used to approximate the likelihood of the solution data for a single LP instance:

$$\Pr(x_i^\star|\theta, A_i, b_i, c_i) \approx \Pr\left((c_i + \epsilon_i)^T x_i^\star = J^\star(A_i, b_i, c_i, 0) + \epsilon_i^T x_i^{\star,0}\middle|\, \theta\right).$$

Some simplification leads to the LPSA approximation of the likelihood

$$\Pr(x_i^\star|\theta, A_i, b_i, c_i) \approx \Pr\left(\epsilon_i^T (x_i^{\star,0} - x_i^\star) = c_i^T (x_i^\star - x_i^{\star,0})\middle|\, \theta\right), \qquad (5)$$

which is simply the pdf of a Gaussian random variable with mean $\mathbf{1}^T(x_i^{\star,0} - x_i^\star)\mu$ and variance $\|x_i^{\star,0} - x_i^\star\|_2^2 \sigma^2$ evaluated at $c_i^T(x_i^\star - x_i^{\star,0})$.

LPCAMLE attempts to maximize the log-likelihood $\ell(\theta)$:

$$\ell(\theta) = \sum_{i=1}^N \log \Pr(x_i^\star|\theta, A_i, b_i, c_i). \qquad (6)$$

To create a tractable problem, LPCAMLE uses the LPSA approximation (5) to instead maximize the approximate log-likelihood

$$\hat{\ell}(\theta) = \sum_{i=1}^N \log \Pr\left(\epsilon_i^T (x_i^{\star,0} - x_i^\star) = c_i^T (x_i^\star - x_i^{\star,0})\middle|\, \theta\right). \qquad (7)$$

A maximization problem with objective function $\hat{\ell}(\theta)$ can be solved analytically using the first-order necessary conditions and the pdf of Gaussian random variables. The resulting LPCAMLE estimates $\theta^\star = (\mu^\star, \sigma^{2\star})$ are

$$\mu^\star = \frac{\sum_{i=1}^N p_i v_i / q_i}{\sum_{i=1}^N p_i^2 / q_i} \qquad (8)$$

and

$$\sigma^{2\star} = \frac{1}{N} \sum_{i=1}^N \frac{(v_i - p_i \mu^\star)^2}{q_i}, \qquad (9)$$

where $p_i = \mathbf{1}^T(x_i^{\star,0} - x_i^\star)$, $q_i = \|x_i^{\star,0} - x_i^\star\|_2^2$, and $v_i = c_i^T(x_i^\star - x_i^{\star,0})$. If $p_i = q_i = 1$ for all i and the v_i are viewed as the sample data, equations (8) and (9) simplify to the standard maximum likelihood estimates of the parameters of a normal distribution. If $x_i^\star = x_i^{\star,0}$, then $q_i = 0$ and the terms in the sums in eqs. (8) and (9) are invalid. This situation corresponds to evaluating the pdf of a Gaussian random variable with a variance of zero in the LPSA approximation (5). It is not clear how to handle this situation; such instances are skipped for now.

To apply LPCAMLE to the estimation problem in sub-section III-B, problem (1) must be posed as an LP. If the appropriate vectors c_i, b_i, x_i, and ϵ_i and matrix A_i are constructed, the LP relaxation of the integer program (1) is indeed identical to the LP (3). Each of the n^2 elements of the vector ϵ_i correspond to one ε_{ij}. Due to the total unimodularity of the appropriate A_i matrix and the integrality of the appropriate b_i vector, this LP is guaranteed to produce the same minimum cost value as the corresponding integer program (1), even though the LP solution may not be integer [10]. Therefore, the LPCAMLE heuristic can be applied in an attempt to solve the estimation problem posed in sub-section III-B.

D. Simulation Approximate Likelihood Estimation

The likelihood can also be approximated using simulations [11]. This can be accomplished by selecting a problem instance i_b randomly from the N matchings in $(\mathcal{F}, \mathcal{S}, \mathcal{M})$, generating a corresponding noise sample ϵ_b from a distribution with parameters θ, and then solving the corresponding LP problem instance (3) to find the minimum cost value $c_{i_b}^T x_{i_b}^{\star,\epsilon_b}$ for the airline matching problem (1) with noise sample ϵ_b. If this is done B times, then the simulation approximation of the likelihood is

$$\Pr(\mathcal{M}|\theta,\mathcal{F},\mathcal{S}) \approx \frac{1}{B}\sum_{b=1}^{B}\mathbb{1}\left\{\frac{(c_{i_b}+\epsilon_b)^T x_{i_b}^\star - (c_{i_b}+\epsilon_b)^T x_{i_b}^{\star,\epsilon_b}}{(c_{i_b}+\epsilon_b)^T x_{i_b}^{\star,\epsilon_b}} \leq \xi\right\}, \tag{10}$$

where ξ is a parameter specifying how close the cost induced by the historical solution data $x_{i_b}^\star$ must be to be to the optimal cost to be considered an optimal solution for this problem instance with sampled noise ϵ_b. Here $\mathbb{1}\{a\}$ equals 1 if the condition a is true and equals 0 if a is false.

In this case, a set $\Theta = \{\theta_k\}_{k=1}^K$ of candidate values for θ must be defined. Then, the maximum likelihood $\theta^\star$ is computed as

$$\theta^\star \in \arg\max_{k=1,\ldots,K} \Pr(\mathcal{M}|\theta_k,\mathcal{F},\mathcal{S}), \tag{11}$$

where approximation (10) is used to approximate $\Pr(\mathcal{M}|\theta_k,\mathcal{F},\mathcal{S})$.

IV. Validation

A. Implementation Notes

The heuristics were implemented and tested in Matlab. LP instances were solved with CVX, a package for specifying and solving convex programs in Matlab [12]. The set Θ contained possible values for σ^2 of $\{1, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100\}$. The mean was 0 for each $\theta \in \Theta$. The LPCAMLE heuristic implemented here assumes that the noise mean is zero by setting $\mu^\star = 0$ in eq. (9) when computing $\sigma^{2,\star}$. When approximating likelihoods with simulation as in eq. (10), B was set to 1000 and ξ was set to 0.01. These values were selected because they performed relatively well on the problem instances under investigation.

When comparing outcomes of the LPCAMLE heuristic across different cost functions in sub-sections IV-C and V, the assignment costs must be normalized to have roughly the same magnitude to allow for fair comparisons of variance estimates and approximate log-likelihoods. To allow for such comparisons, the assignment cost for each cost function will be normalized by the average observed assignment cost for the cost function ($\bar{c}$).

B. Comparison of Heuristics

Before using LPCAMLE to evaluate airline delay cost functions, some sample estimation problem instances based on generic LPs will be specified and solved to demonstrate the behavior of the LPCAMLE and simulation heuristics.

For these sample estimation problem instances, $A_i = A$, $b_i = b$, and $c_i = c$ for all of the $N = 1000$ solved LP problem instances in the data set. The noise parameters are $\theta = (0, 25)$. The size of A is 100×50. Each element in A was generated by sampling from a Gaussian distribution with mean zero and standard deviation 100. Each element in the cost vector c is sampled from a uniform distribution on $[0, 100]$. The b vector is computed by post-multiplying A by a vector of ones. For this portion of the validation, the heuristics used the true cost vector c when estimating the variance.

TABLE I
Comparison of Heuristics on Sample Problem Instances
($\sigma^2 = 25$)

Seed	LPCAMLE $\sigma^{2,\star}$	Simulation $\sigma^{2,\star}$
0	26.8	1
1	25.9	1
2	24.8	1
3	33.6	1
4	68.1	1

Different LP coefficients and corresponding LP solution data sets for different estimation problem instances are generated by initializing a random number generator with different seed values. Then, the LPCAMLE and simulation heuristics are executed with the generated data sets to estimate the maximum likelihood estimates for θ.

Table I shows the results of the two heuristics. For three of the five estimation problem instances (produced using random number generator seeds 0-2), the LPCAMLE variance estimate $\sigma^{2,\star}$ is within 8% of the actual variance $\sigma^2 = 25$. For one of the other cases the estimate is off by almost 35% and in the final case it is off by 172%. It seems that characteristics of the problem coefficients A, b, and c can impact the quality of the LPCAMLE estimates. The heuristic based on the simulation approximate likelihood, on the other hand, estimates a variance value of 1 for each of the problem instances. This poor performance was not analyzed in detail, but it might be related to small cost noise values failing to perturb the problem cost vector sufficiently to drive the solution to an $x_i^{\star,\epsilon}$ that is not equal to the zero-noise solution $x_i^{\star,0}$. Furthermore, the simulation-based heuristic requires almost 1000 seconds of computation time while the LPCAMLE heuristic completes in about 0.5 seconds. Due to its poor performance, no further results from the simulation-based heuristic are presented.

C. LPCAMLE Validation with Synthetic Matching Data

The performance of LPCAMLE was investigated with some synthetic airline matching data generated by solving the airline matching problem (1) with known delay cost functions and noise sampled from known distributions. This synthetic data was based on the historical matchings of two airlines referred to as airline E and airline G. There were 1368 matchings for airline E and 473 matchings for airline G.

The delay cost functions will be explained in subsection V-A. Six of the delay cost functions that perform relatively well on the historical matching data are investigated in this validation work. They are cost functions 2, 5–8, and 16 (see Table III).

The standard deviation of the noise (σ) was varied in this validation work. It was assigned four values ranging from $\bar{c}$ down to $\bar{c}/10$. For each of two airline data sets, six delay cost functions, and four zero-mean additive cost noise standard

TABLE II

CANDIDATE COST FUNCTIONS WITH LARGEST APPROXIMATE LOG-LIKELIHOOD

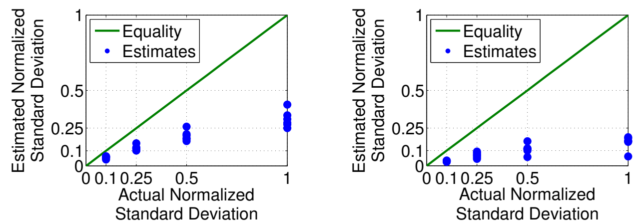
Generating Cost Function	**Airline E**				**Airline G**			
	$\frac{\sigma}{\bar{c}} = 0.1$	$\frac{\sigma}{\bar{c}} = 0.25$	$\frac{\sigma}{\bar{c}} = 0.5$	$\frac{\sigma}{\bar{c}} = 1.0$	$\frac{\sigma}{\bar{c}} = 0.1$	$\frac{\sigma}{\bar{c}} = 0.25$	$\frac{\sigma}{\bar{c}} = 0.5$	$\frac{\sigma}{\bar{c}} = 1.0$
2	2	2	2	2	2	2	2	2
5	5	5	5	5	2	2	2	2
6	6	6	6	6	6	2	2	2
7	7	7	7	7	7	2	2	2
8	8	8	8	16	2	2	2	2
16	16	16	16	16	2	2	2	2

deviation levels, a set of synthetic minimum cost perfect matchings were generated for the flights and slots in the airline matchings. Then, for each of these synthetic sets of matchings, the LPCAMLE heuristic was applied with each of the six cost functions used as the candidate cost function. The corresponding additive cost noise variance estimates and approximate log-likelihood values were recorded in each case. If LPCAMLE worked perfectly, it would produce exact estimates of the additive cost noise standard deviation and the cost function used to generate the synthetic data would always achieve the largest approximate log-likelihood.

Table II shows which of the candidate cost functions achieved the top-three largest approximate log-likelihood values in each synthetic scenario for airline E and airline G. For airline E, the generating cost function achieved the largest approximate log-likelihood in 23 of the 24 instances.

For the airline G, cost function 2 (Passenger Delay) often achieves the largest approximate log-likelihood. Airline G used the same aircraft type for almost every flight. Also, an annual average load factor was used for all flights. Therefore, cost function 2 fails to differentiate between matchings and almost any matching achieves the minimum cost. This causes LPCAMLE to compute a relatively large $\hat{\ell}(\sigma^{2\star})$ for cost function 2. In this sort of situation it may be wise to remove such cost functions from consideration. If the deterministic part of an airline's cost function fails to differentiate between most matchings, its optimal matchings would be determined almost entirely by realizations of the cost noise terms. These are assumed to depend *only* on factors *not* observed by the public, suggesting that its matchings do not depend on publicly-observed factors, which seems unlikely. If cost function 2 were removed, then there would only be four instances when the generating cost function did not achieve the largest $\hat{\ell}(\sigma^{2\star})$.

In Fig. 1, the estimates of the normalized standard deviation ($\sigma^\star/\bar{c}$) are plotted as a function of the actual normalized standard deviations ($\sigma/\bar{c}$) used to generate the synthetic matchings. For both airlines, the estimates are severe under-estimates of the actual standard deviation values. It is not clear why this is the case.



(a) Airline E (b) Airline G

Fig. 1. Normalized standard deviation estimates for synthetic data generated with four normalized actual standard deviation values.

V. EVALUATING AIRLINE DELAY COST FUNCTIONS

A. Delay Cost Functions

A set of candidate cost functions are evaluated. The functions depend on publicly available characteristics of a flight f_i and the minutes that the flight is delayed d. Ref. [1] explains the 17 evaluated cost functions in detail. Those that achieve one of the three largest $\hat{\ell}(\sigma^{2\star})$ for an airline are listed in Table III.

TABLE III

DELAY COST FUNCTIONS

Number	**Name**
1	On-time Performance
2	Passenger Delay
5	Time-of-Day Delay
6	Connection Delay
7	Airline Connection Delay
8	Monetary Delay
16	Connection and Monetary Combination Delay
17	Airline Connection and Monetary Combination Delay

B. Data

The historical matchings used in this study are from 32 unique AFPs and are recorded in 34 days of Expected Departure Clearance Time (EDCT) log files from June–August 2006. "Simplified Substitution" (SS) messages in these files specify a set of flights, a set of slots, and the corresponding airline-selected matching. If a few assumptions are made (see Ref. [1]), SS messages contain enough information to define the minimum cost perfect matching problems (1).

At any time during an AFP, airlines can submit SS messages to the FAA. Some airlines specify matchings frequently while others do so relatively rarely. There were SS messages specifying matchings for 18 airlines, but 11 of those airlines submitted messages less than 200 times and were therefore not investigated. Validation efforts not reported here due to space constraints suggest that at least 200 matchings are needed for one of the metrics used in this paper, so results for these 11 airlines are not presented. Also, each SS message can specify as many flights and slots as the airline would like to match. For most airlines, the majority of matchings involve less than 10 flights and slots. However, some matchings contain more than 100 flights and slots.

C. Results

The cost functions with the three largest $\hat{\ell}(\sigma^{2\star})$ values for each airline, along with corresponding estimates of $\sigma^\star/\bar{c}$, are shown in Table IV. See Table III for the names of the cost functions and Ref. [1] for a detailed description of each. Smaller $\sigma^\star/\bar{c}$ values indicate that the the airline matchings are more likely 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 results in sub-section IV-C suggest that these estimates are probably low.

TABLE IV

COST FUNCTIONS WITH LARGEST APPROXIMATE LOG-LIKELIHOOD

Airline	1^{st}	$\frac{\sigma^\star}{\bar{c}}$	2^{nd}	$\frac{\sigma^\star}{\bar{c}}$	3^{rd}	$\frac{\sigma^\star}{\bar{c}}$
A	7	0.487	6	0.539	1	1.216
B	7	0.454	5	0.509	17	0.525
C	6	0.587	16	0.615	7	0.623
D	7	0.037	17	0.133	16	0.246
E	6	0.369	16	0.336	17	0.351
F	16	0.064	17	0.083	6	0.050
G	2	0.051	17	0.164	8	0.185

Cost functions 6, 7, 16, and 17, which are all based on Connection Delay, achieve relatively large approximate log-likelihood values for most airlines. These functions attempt to capture the fact that delaying flights bound to hub airports is especially costly because such 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 that is 2 for flights bound for high-connection-rate airports (hubs), 1.5 for flights bound for medium-connection-rate airports, and 1 for all other flights [7].

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. It is assumed that airlines solve a minimum cost perfect matching problem when matching flights to slots. Unobserved aspects of airline costs are accounted for by adding a Gaussian noise term to the cost functions. A heuristic was developed to find cost functions and corresponding noise parameters that maximize an approximation of the log-likelihood of airline actions during 32 Airspace Flow Programs in the summer of 2006. When applied to estimation problem instances based on linear programming problem coefficients and solution data generated with known noise parameters, the heuristic can more accurately estimate noise parameters than a simple simulation-based approach. Validation efforts based on synthetic airline action data generated with known delay cost functions and noise parameters demonstrates that the heuristic is in many cases able to correctly identify as most likely the delay cost function that was used to generate the synthetic data. However, the heuristic also under-estimates the magnitude of the cost noise variance for these problem instances. Costs that are proportional to the length of the delay, but with proportionality constants that are larger for flights bound to hub airports, maximize the approximation of the log-likelihood of the historical matchings of most airlines. Finally, the corresponding estimates of the standard deviations of the cost noise terms, expressed as a fraction of the average assignment cost for the historical matchings, ranged from 0.1 to 0.7.

ACKNOWLEDGMENT

The authors are grateful to Dr. Amar Choudry of STC for facilitating Mr. Huang's internship at NASA during the fall of 2010, which was supported by STC/STIEP/IRAD 6700-112.

REFERENCES

[1] M. Bloem and H. Huang, "Evaluating delay cost functions with airline actions in Airspace Flow Programs," in *Proc. of USA/Europe Air Traffic Management Research & Development Seminar*, Berlin, Germany, June 2011. [Online]. Available: http://www.atmseminar.org

[2] J. Xiong, "Revealed preference of airlines' behavior under air traffic management initiatives," Ph.D. dissertation, University of California, Berkeley, 2010.

[3] A. Vasquez-Marquez, "American Airlines Arrival Slot Allocation System (ASAS)," *Interfaces*, vol. 21, no. 1, January-February 1991.

[4] M. Brennan, "Simplified Substitutions – enhancements to substitution rules and procedures during Ground Delay Programs," in *Proc. of AGIFORS Airline Operations Meeting*, Ocho Rios, Jamaica, May 2001.

[5] T. J. Niznik, "Optimizing the airline response to Ground Delay Programs," in *Proc. of AGIFORS Airline Operations Meeting*, Ocho Rios, Jamaica, May 2001.

[6] 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.

[7] 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, no. 2, pp. 147–164, May 2006.

[8] 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.

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

[10] M. X. Goemans, "Lecture notes on bipartite matching," February 2009. [Online]. Available: http://math.mit.edu/~goemans/18433S09/matching-notes.pdf

[11] P. Marjoram, J. Molitor, V. Plagnol, and S. Tavaré, "Markov chain Monte Carlo without likelihoods," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 100, no. 26, pp. 15 324—15 328, December 2003.

[12] M. Grant and S. Boyd, "CVX: Matlab software for disciplined convex programming," Web page and software, http://cvxr.com/cvx, January 2009.