

Causal Demand Modelling for Applications in En Route Air Traffic Management

Ivan Tereshchenko

Department of Civil and Environmental Engineering,
University of California, Berkeley
Berkeley CA, USA

Mark Hansen

Department of Civil and Environmental Engineering
University of California, Berkeley
Berkeley, CA, USA

Abstract—Increasing air traffic volume makes en route Traffic Management Initiatives (TMIs) more important than ever before. The effective execution of en route TMIs depends on accurate predictions of airspace demand. Precise forecasts of airspace demand require causal models of route choice. Previous research shows that obtaining such models is extremely difficult, due to the complex nature of the airspace system. In this paper, we test three methods for making causal estimates of route utility in the context of two en route TMIs – the Airspace Flow Program (AFP) and Collaborative Trajectory Options Program (CTOP). The testing was done using simulated TMI data. We show that statistical models of the behavior of individual flights produce biased estimates of route utility. Models based on changes in aggregate delay produce better estimates; however, such models are harder to implement in practice. Finally, CTOP offers data structures that allow us to achieve higher quality airspace demand predictions.

Keywords—Air Traffic Management; AFP; CTOP; Discrete Choice Modelling; Demand Prediction; Queueing.

I. INTRODUCTION

A. Background

1) En Route Traffic Management Initiatives

En route traffic management initiatives are programs conducted by air traffic management authorities in order to balance airspace capacity and demand during instances of heavy traffic volume and severe weather. For instance, thunderstorms might make some airspace sectors unavailable for entry, which reduces the capacity of the airspace. TMIs assume control over aircraft that were scheduled to cross impacted sectors of the airspace and determine these aircrafts' departure times and routes in a way that balances the capacity of the airspace with scheduled demand.

The focus of this paper is to determine if the existing statistical tools allow us to reliably forecast airspace demand during en route traffic management initiatives (TMIs), such as the Airspace Flow Program (AFP) and Collaborative Trajectory Options Program (CTOP). AFP is currently one of the most widely used en route TMIs in the US. AFP works in a similar

way to a Ground Delay Program (GDP): it assigns departure delays to aircrafts whose original routes would penetrate a certain section of the airspace (Flow Constrained Area or FCA) and allows airlines to reroute their flights out of the AFP in order to reduce delays. CTOP is a further development of AFP that is intended to replace it in the future. In CTOP, flights communicate their route preferences to the ATM through so-called Trajectory Option Sets (TOS). In a TOS, for each potential route the aircraft operators assign the so-called Relative Trajectory Cost (RTC) - measure of cost of airborne delay for a given route measured in terms of departure ground delay.

2) Problem Statement

One of the recurring issues in en route TMIs is spillover effects. Some TMI-controlled flights receive a departure delay, while others reroute around the constrained area. Large departure delays result in delay propagation throughout the system. The purpose of reroutes is to reduce delay and delay propagation. However, rerouted flights may be funneled into other regions of the airspace, which may require the creation of additional TMIs. Knowing which flights will be delayed and by how much, and which reroutes other flights will take, would help the Air Traffic Management to better plan for TMI extension, staffing, and other planning activities. For example, this information can be incorporated into stochastic optimization models for capacity rate setting under uncertainty, such as Enhanced Stochastic Optimization Model (ESOM) [1][4]. This issue will become more significant as air traffic grows, since congested airspace will have less capacity available for AFP and CTOP reroutes.

In other words, we would like to be able to predict airspace demand to a high degree of accuracy. There are several factors that will likely make this challenging. First, airlines do not easily disclose their decision-making policies to aviation authorities, since these policies contain proprietary information that they fear might be used against them by authorities or competitors. Second, en route TMIs that control hundreds of flights at the same time are relatively rare (< 50 instances per year in the United States), which limits the amount of data available to train machine learning models that can forecast precise actions that

aircraft operators may take. As a result, we would like to develop a structured model that underlines the decision-making process of aircraft operators and allows us to forecast flight-level behavior based on the flights' observable attributes, such as origin-destination pair, departure time, assigned ground delay and flight time, and route choice decisions that were made in the past.

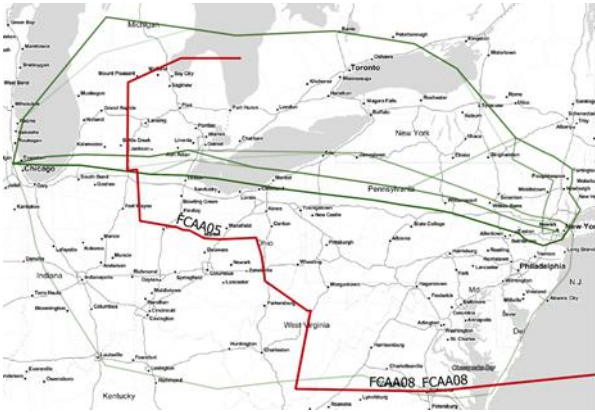


Figure 1. Illustration of Airspace Flow Program with two FCAs (red lines), flights that chose shorter paths and rerouted to the north and to the south of FCA (green lines).

3) Structured Demand Model

We have developed such a structured model in a previous paper [6] where we attempted to estimate a statistical model of decision-making in the AFP setting. In that paper we assumed that for a given flight each potential route has utility (or disutility) associated with it. The utility of a route is a linear function (weighted sum) of the route's airborne delay and ground delay:

$$U_i = \beta_1 * Delay + \beta_2 * Flight Time + \epsilon \quad (1)$$

In this equation U_i is the random utility term of a given route i , β_1 and β_2 are utility weights associated with AFP delay and a route's flight time respectively, and ϵ is the unobserved utility term. This approach has been used in multiple other studies that are related to modelling of cancellation behavior during a Ground Delay Program (GDP) [5][7][8].

Each flight faces two possible route options: to stay on the shortest path route and receive a significant departure delay or to reroute out of AFP and receive no ground delay in exchange for some airborne delay associated with flying a longer route. By using observed choices of routes that flights took during AFP we can estimate the weights that are associated with airborne delay and ground delay in a utility function. This statistical model is called Random Utility Model (RUM), because it treats the utility as a sum of deterministic (observed) and stochastic (unobserved) components.

The ratio of the estimated weights for airborne and ground delays is called the Delay Cost Ratio (DCR). DCR is the interpretable component of the model, since it represents the tradeoff between the two types of aircraft delay. Most studies assume DCR to vary between 2 and 3 [2][3]. For example, this means that airlines should be willing to take reroutes that require 10 minutes of extra flight time in order to avoid 20 to 30 minutes of ground departure delay. In the previous paper we found that,

according to the statistical estimates, DCR are anywhere between 4.5 and several hundred. We proposed several possible explanations for this behavior, such as boundedly rational behavior in the presence of uncertainty and the imprecise nature of the data. We also could not discount the possibility that airlines do have very high DCR values. Finally, we hypothesized that airlines' coordination of actions of flights in the AFP queue reduces the apparent sensitivity of flights to ground delay and distorts our estimates. In this paper we will focus on this hypothesis.

4) Goal of the paper

The goal of this paper is to develop and test several DCR estimation techniques. We test these techniques on simulated data from stylized models of airline behavior in the context of AFPs or CTOPs. The models do not reflect all complex interaction that are present in the real air transportation system, but only some of them, such as the interaction of flights in the AFP or CTOP queue. We also assume that the behavior of airlines is very simple: the cost function is linear with respect to ground delay and flight time, and the airline response to the TMI follows a very simple optimization model. The reason we do this is that the demand modelling techniques that work on simple simulated data have a chance to perform well in real-world data. However, if these techniques fail on simplified data, they will also fail on real AFP and CTOP data. We can assess the accuracy of our techniques, because we set all simulation parameters and can compare the outputs of our models to assumed ground-truth values.

In the first part of the paper we test two possible ways of estimating DCR from simulated observational data. First, we use random utility model and simulated flight-level route choice data. Second, we develop and implement a simple method that uses aggregate delay data to calculate the upper bound for DCR values. In this part of the paper we are aiming to test the hypothesis that the coordination between flights in an AFP queue does not allow us to reliably estimate cost model parameters even for the simplest cost models. Then we proceed to create a route cost estimation technique that can be applied to forecast demand in the context of CTOP. Finally, we discuss the results and provide several consequences that are relevant to the practice of Air Traffic Management.

II. ESTIMATING RELATIVE TRAJECTORY COST

The purpose of this paper is to explore in depth one of the possible explanations of previous results. We hypothesized that the unexpectedly small sensitivity of route utility to assigned ground delay is the result of airlines' optimization of aggregate outcomes using the components of Collaborative Decision Making (CDM). When flight operators cancel or reroute flights out of AFP, they do not lose their arrival slots. Instead, they move some of the other flights up in the queue in order to reduce overall delays in a TMI. When we observe the route choices of flights it appears to us as a lower sensitivity of route utility to departure delay.

In this paper we conduct computational experiments, in which we simulate flight schedules in the context of an AFP, assign departure delays to each flight in the queue, and then allow airlines to choose reroutes and flight substitutions using a

simple optimization model. We then use the optimal solutions to airlines' problems as data inputs into random utility models in order to estimate DCR values that were used by airlines as inputs in their optimization programs. After this, we develop a simple aggregate method for estimating the upper bound of DCR. Finally, we propose a method for demand prediction that can be implemented in the next generation of en route TMIs such as CTOP.

A. Flight-level Simulation

1) General Description

The first step is to simulate notional AFP data that we can use for model estimation and validation. This simulated data contains information about flow constraints, airspace capacity, reroutes, schedule arrival time at flow constraints, reroute costs, and AFP-assigned delays. The flight-level AFP simulation constrains several components: simulation of notional flight schedules, FCAs, direct routes, reroutes and RTC values, and generation of airline response using a simple optimization model.

2) Simulating Schedules and RTC

The first step was to simulate airline schedules, FCAs, reroutes, and RTC values for every reroute. In order to have more control over the analysis and to get more general results we did not use real airline schedules in the simulation. Instead, we randomly sampled origin and destination coordinates from two, two-dimensional Gaussian distributions to generate a list of flights. In a similar vein we generated two end-points for a linear FCA that lies roughly in the middle of the analysis region. For each flight we computed shortest path flight times assuming that aircraft speed is normally distributed across flights with a mean of 450 knots. Every flight was assigned an FCA arrival time. The arrival times were uniformly distributed across the analysis period of several hours.

Then we generated simplified reroute trajectories that circumvent the FCA via the shortest available path. The additional flight time for a reroute is therefore the difference between the reroute flight time and the shortest path flight time. After calculating reroutes, we computed RTC values for the reroute. We assumed that the RTC of a reroute is equal to the additional flight time for the reroute multiplied by DCR, which we assume to be constant for every flight in a given simulation. In three simulation runs we assumed that DCR is equal to 2, 3 and 5 (see more detail below).

The next step is to conduct AFP delay assignment. The duration of the AFP was set to 4 hours—the average real-world AFP duration. The total number of flights controlled by AFP in every scenario is equal to 600 flights, approaching the real-world value for the United States. The capacity rate of AFP was set to 75% of demand in order to generate delays that approximate real-world levels. The delays were assigned to the simulated flights using the Ration-By-Schedule rule applied to the FCA arrival times calculated above, without exemptions. Reroutes were assumed to have a required delay of zero minutes.

As a result, we have simulated notional airline schedules, complete with shortest path routes and reroutes, RTC for reroutes, and required AFP delay values. Note that the analysis is simplified—we do not embed simulated airline schedules into

broader airline hub-and-spoke networks. The purpose of this is to isolate the effect AFP structure on observations of revealed airline preferences and costs.

B. Airline Response Model

The next step in the AFP simulation process is airline response. Our goal is to demonstrate that the ability of airlines to respond to AFP delay assignments hinders our ability to accurately assess airlines' sensitivity to changes in delay.

The airline response model is a simple integer programming model program in which the airline may reroute flights and reallocate their vacated slots to other flights.

$$\min \sum_i \sum_t [a_t x_{it} - s_i(1 - y_i) + r_i * y_i] \quad (2)$$

s. t.

$$1.) \sum_t a_t x_{it} \geq s_i(1 - y_i) \quad \forall i$$

$$2.) \sum_t a_t x_{it} - s_i(1 - y_i) \leq T_{max} \quad \forall i$$

$$3.) \sum_t x_{it} + y_i = 1 \quad \forall i$$

$$4.) \sum_i x_{it} \leq 1 \quad \forall t$$

$$x_{it}, y_i \in B$$

Where a_t is the scheduled FCA arrival time of slot t , x_{it} indicates whether slot t is assigned to flight i , y_i indicates reroute for flight i , s_i is scheduled arrival time at FCA for flight i , and r_i is the RTC for a reroute for the flight. The objective is to minimize the total cost of TMI for an airline. The cost has two components: delay cost in minutes and reroute cost in ground delay unit equivalents. The cost of a reroute is equal to the RTC value of reroute for each flight. The first set of constraints ensures that the flight is not moved to the slot preceding its scheduled arrival time. The second set of constraints ensures that the flight does not get rescheduled to a much later slot. The third set of constraints makes sure that exactly one action is chosen for each flight. Finally, the fourth set of constraints ensures that each slot is used at most once.

In other words, airlines are allowed to reroute and substitute flights. When a flight is rerouted, its AFP slot is not lost but instead can be filled in by another flight, thereby reducing delay for a specific flight and the carrier overall. This was done to isolate the effect of rerouting. Adding the ability to cancel complicates the analysis, but the results stay qualitatively the same.

There is one optimization problem per airline. One output of the airline response model for each airline mathematical program is a binary decision vector that has the same dimension as the number of flights, n , controlled by the AFP, which indicates which flights are rerouted. A second output is an n by n vector specifying the assignment of slots to flights. After simulating multiple scenarios, the next step is to construct a RUM-model that attempts to predict whether airlines will reroute a particular flight. The binary reroute decision vector will be the Y-variable, while flight time and AFP delay values

created in the simulation constitute the X-variables. The ratio of the flight time coefficient to delay coefficient from the RUM-model is our estimate of DCR. In other words, our goal is to use the random utility model to recover the value of DCR and compare it to the ground-truth RTC that was assumed in the simulation used to generate the data.

C. Simulation scenarios

We would like to see the effect that airline-level aggregate optimization has on DCR estimates from binary logit model reroute decisions. In order to do this, we generate multiple simulations in which we vary only one parameter – the number of flights controlled by AFP per airline. The higher the number of flights (slots) for an airline, the higher the flexibility to reroute and substitute flights. In different simulation runs we set the number of flights per airline to values between 600 and 1. We expected that the RUM-estimates of RTC will be different for different simulation runs. We generate 10 simulation runs in order to obtain a sample of 6,000 flights.

We conduct this simulation process three times, each time setting ground-truth DCR value to a different level: 2, 3, and 5.

D. Estimates of Random Utility Model

1) Model Description

Finally, we input all simulated datasets into the Random Utility Model. We chose the simplest kind of RUM-model: binary logit. Other types of models, such as mixed logit, produce different estimates, but do not qualitatively change the conclusions.

The dependent variable is a binary decision variable that takes the value of zero if the flight does not reroute, and the value of one if the flight reroutes. There are two independent variables - flight time and required delay. Required delay for the reroute option is assumed to be zero, in line with AFP design. We model the probability that the flight operator chose one of the options:

$$\Pr(Y_i = 0 \text{ or } 1) = \frac{e^{x_i\beta}}{\sum e^{x_j\beta}} \quad (3)$$

Here, i is the set of routes available to a given flight, which consists of two options – a shortest path option and a reroute. $x\beta$ is the deterministic component of random utility:

$$V_i = \beta_0 + \beta_1 * Delay_i + \beta_2 * Flight Time_i \quad (4)$$

As a result of model estimation, we obtain three regression coefficients: regression intercept, flight time and delay coefficients. The flight time and delay coefficients show the change in route's utility as flight time and delay on the route increase by one minute. The ratio of the flight time and delay coefficients is the DCR value, which shows the amount of reduction in ground delay that is required to offset the increase in flight time equal to 1 minute.

2) Summary of Results

DCR values were computed for each simulation run. Binary logit DCR estimates vary by the number of slots per airline in the AFP and by the assumed ground-truth value of DCR specified for the simulation run. Figure 2 summarizes results. The x-axis is the number of flights (slots) available per airline in AFP. The

y-axis is the regression estimate of DCR for each simulation run. Each line indicates a separate set of simulations with different assumed values of DCR. For ease of interpretation, the plot is logarithmic.

While the assumed ground truth values of DCR are equal to 2, 3 and 5 in the three sets of simulations, the estimated values are very different. For simulations with a very large number of AFP-controlled flights per airline the estimate of DCR is close to 100—two orders of magnitude higher than the ground-truth value.

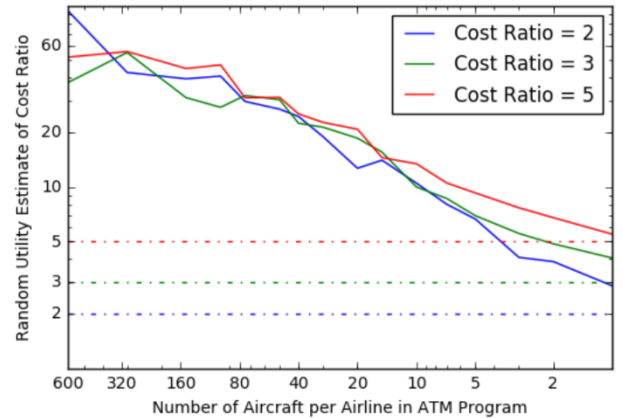


Figure 2. Summary of random utility model estimates for three simulation runs with varying cost ratios.

As the number of flights per airline in the AFP decreases, DCR estimates slowly approach the ground truth value. For simulations with one AFP slot per airline—implying TMI actions for each flight are made independently—the estimated DCR is only slightly higher than the ground truth values. Statistical methods, such as binary logit, are best-fitted to deal with such cases, since the primary assumption in most statistical methods is independence between observations.

The results suggest that DCR estimates based on real-life data that we obtained in preceding research were not an accident or the result of poor data. The analysis of simulated data shows that simple intra-airline coordination of actions that is allowed within the CDM framework results in an inability to estimate the true sensitivity of flights to changes in TMI delay.

3) Causal Inference Interpretation

The fact that a particular model failed to recover true route cost parameters does not necessarily mean that another econometric model cannot succeed. However, there is a reason why statistical models that treat flights as individual observations will not be able to estimate the true cost parameters. The results that we obtained can be explained within the framework of causal inference that employs systems of simultaneous equations that was proposed by Judea Pearl in 1995[5]. It has been shown that in order to recover the parameters of simultaneous equations, the assignment model has to be represented by Directed Acyclical Graph, or DAG. The following is an example of DAG:

- 1) Ground delay and flight time influence utilities of routes.

2) Utilities of routes with unobserved attributes described by a probability distribution with known moments influence route choice.

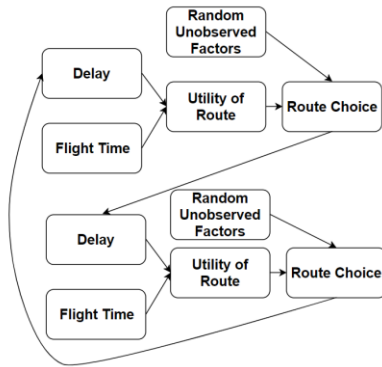


Figure 3. Directed Cyclical Graph that represents the route choice decision-making model. Boxes represent variables, arrows represent influences that these variables make on other variables.

In this case we would be able to get regression coefficients for ground delay and flight time and compute a DCR that matches the DCR that we assumed in order to generate our simulated data. This is the case when the carrier has one flight controlled by TMI and may choose to reroute it. As we have seen, in this case we can recover DCR estimates fairly well. However, when there is more than one flight, we are dealing with a Directed Cyclical Graph model:

- 1) Ground delay and flight time for flight 1 influence choice for flight 1.
- 2) Choice for flight 1 influences ground delay for flight 2.
- 3) Ground delay for flight 2 influences choice for flight 2.
- 4) Choice for flight 2 influences delay for flight 1. This closes the cycle.

As we add more and more flights, we add more and more cycles, which makes DCR estimates deviate more and more from the true values. The deviation happens due to the ability of airlines to optimize their queues. Rerouting only a handful of flights with the lowest RTC values can reduce departure delay and make rerouting other flights unnecessary.

E. Aggregate Approach

1) Method Description

As the previous sections shows, the inability to precisely estimate relative trajectory costs is related to the ability of airlines to use rerouting and slot substitution to optimize their schedules. Statistical models are poorly suited for this type of problem, since all flights are treated as independent observations. However, there might be a way to estimate DCR using more traditional deterministic queuing models.

Figure 4 is a queuing diagram for a hypothetical single-FCA en route TMI that fits the description given in the previous section. For a certain period, the capacity of the FCA is below the scheduled demand. In order to balance the demand, ground delay is assigned to flights. Airlines, however, can "extract" flights from the TMI via rerouting and cancellation. On the

queuing diagram this amounts to reducing the slope of the demand curve to match the capacity curve. To reduce ground delay to zero, airlines need to reroute N_{ex} flights—the maximum excess accumulation of flights in the system. However, even rerouting a small portion of flights leads to a large percentage reduction in departure delay. The delay is represented by the triangular area between two curves. If the airline reroutes N_{ex} flights this area will be reduced to zero.

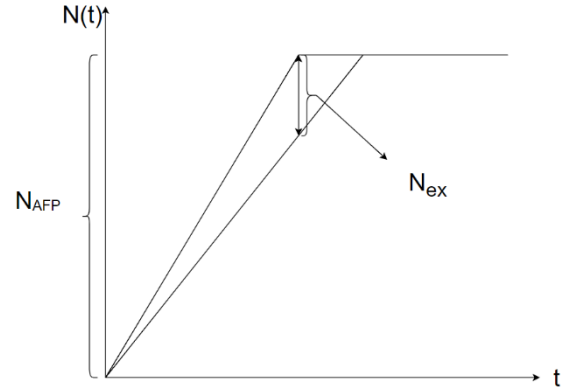


Figure 4. Single-FCA AFP queueing diagram. N_{AFP} is the set of flights controlled by AFP. N_{ex} is the excess accumulation.

As a result, we can write down the following inequality:

$$\Delta D_i * c_{del} \geq \Delta FT_i * c_{del} * DCR \quad (5)$$

Here, ΔD is the reduction in delay that results from reroutes, c_{del} is the cost of one unit of ground delay, ΔFT is the total additional flight time for all rerouted flights. On the righthand side, we decompose the cost of one unit of flight time as $c_{del} * DCR$, recalling that DCR is conventionally defined as the ratio of the cost of one unit of airborne delay and one unit of ground delay. Index i stands for airline. We can rewrite the inequality above to obtain the following upper bound on DCR:

$$DCR \leq DCR_{High}^i = \frac{\Delta D}{\Delta FT} \quad (6)$$

This expression can be used to compute the upper bound on DCR using the simulated AFP data. We can do this by computing the total reduction in ground delay for all carriers and dividing it by the sum of additional flight times on all rerouted flights.

We use the simulated data from before to compute this upper bound on DCR and compare it with the ground truth value and RUM estimates.

In some cases, especially when the number of slots per airline is small, we cannot compute the upper bound of DCR, since in these cases no flights reroute. Instead, we can compute the lower bound. We can write down the following inequality:

$$TD * c_{del} \leq \Delta FT * c_{del} * DCR \quad (7)$$

TD stands for total delay. The inequality is saying that, if no flights rerouted, the total cost of delay should be lower than the total cost of airborne delay required to reroute out of AFP. We can rewrite the inequality to obtain the lower bound of DCR:

$$DCR_{Low}^i = \frac{TD}{\Delta FT} \geq DCR \quad (8)$$

The estimates of the upper bound and the lower for each airline might be highly variable due to the fact that some airlines will

be able to achieve a higher reduction in delay from reroutes due to chance. Assuming that all airlines have the same DCR, we can bring our estimates closer to the ground-truth value:

$$DCR_{High} = \min\{DCR_{High}^i\} \quad (9)$$

$$DCR_{Low} = \max\{DCR_{Low}^i\} \quad (10)$$

The first expression says that the best estimate of the upper bound of DCR is the smallest upper bound estimate across all airlines. The second expression states that the best lower bound estimate is the maximum lower bound.

2) Summary of Results

Figure 5 summarizes the estimates of DCR as a function of the number of slots per airline in AFP. The estimates of the aggregate model are much closer to the true parameter values. Moreover, these estimates converge to the true value faster than the random utility estimates.

In the second figure we summarize the estimates for the lower bound of DCR. These values were calculated using observations, where none of the flights in a particular airline rerouted. Such cases are most likely for airlines with a small number of slots. We obtained the estimates for airlines that have 10 or fewer slots per AFP. The lower bound estimates converge to the true parameter value.

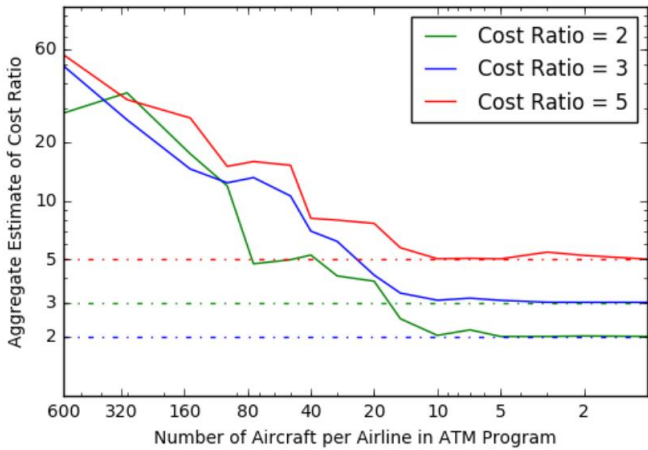


Figure 5. Summary of aggregate model estimates – upper bound.

Although the aggregate approach to DCR estimation allowed us to come closer to estimating the cost parameter for the simple model of reroute cost, the implementation of this approach in practice is very complicated. First, airlines are allowed to cancel flights to reduce departure delay, which makes estimation of reroute costs more challenging. Additionally, there are multiple sources of delay that interfere with AFP delay. The result of this is the inability to estimate the cost ratio in the majority of cases.

For the case of airlines with a large number of slots, the second estimation method fails to recover RTC estimates. We could precisely estimate the DCR parameter if the rerouting cost were exactly equal to the reduction in the cost of delay to the remaining flights. However, typically the reduction in ground delay cost is much greater than the increase in the cost of airborne delay. In most cases this makes our estimates of ΔRTC much larger than the true value.

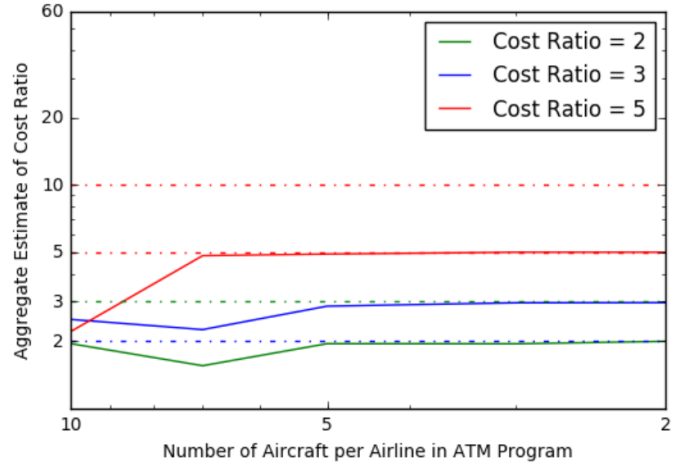


Figure 6. Summary of aggregate model estimates – lower bound.

F. Real-life AFP Dynamics

Even though our methods may not precisely estimate reroute costs, we can draw several conclusions from our observations, which can be used to make qualitative and quantitative predictions about the behavior of airlines. Optimization of airline schedules in response to a TMI has relatively simple consequences: instead of rerouting flights with high delays, airlines may want to spread reroutes uniformly over schedules. This is done in order to match the slope of the arrival curve of the queueing diagram with the slope of the departure curve. Rerouting in this manner maximizes its benefits on airline on-time performance. This leads to a simple approximate algorithm for determining which flights to reroute:

Step 1: Calculate the number of flights that need to be rerouted using the queueing diagram.

Step 2: Select a subset of flights with the shortest, least costly reroutes.

Step 3: Reroute flights roughly uniformly across the duration of AFP

We can demonstrate this behavior on the real-life AFP data. On the figure below you can see an instance of AFP that constrained a substantial number of rerouted flights. This plot uses data from an instance of AFP that occurred on July 19, 2012. The x-axis is airport arrival delay associated with choosing the shortest path flight option. The y-axis is airport arrival delay for rerouted flights, assuming a route that we constructed that avoids the FCA region. The red points are flights that remained in AFP; the blue points are rerouted flights. The airlines choose an acceptable arrival delay threshold, below which flight are allowed to reroute. In other words, in the case of this AFP, the airlines chose to reroute only those flights that will certainly arrive at their origin on time. Among the flights with such reroutes, the airline distributed rerouted flights roughly evenly across the schedule. We know this, because higher arrival delay corresponds to higher departure delay, and higher departure delay corresponds to a lower place in the AFP queue. As can be seen, the rerouted flights span the entire range

of possible arrival delay values, and generally have among the lowest arrival delays associated with choosing the rerouting option.

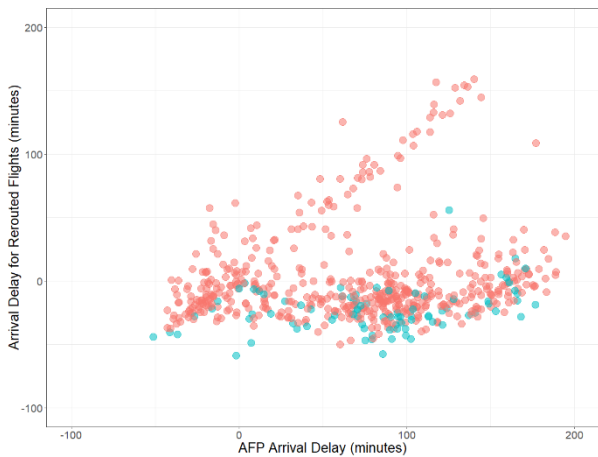
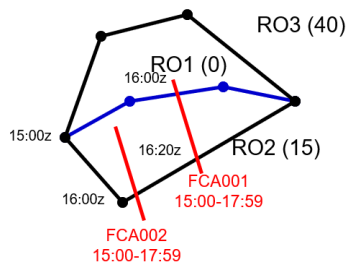


Figure 7. AFP that occurred on July 19, 2012 at FCA A05. Blue dots are rerouted flights, red dots – flights that remained in AFP.

G. CTOP Approach

1) Method Description

There is a third approach to modelling reroute cost, which can potentially give us more detailed models of demand. It will be available to us once CTOP TMI or similar TMIs outside of the US are fully implemented. This section focuses on how we can use our knowledge of the internal resource allocation of CTOP to estimate bounds on DCR.



ERTD	RO	RTC	RMNT	TVST	TVET	Delay Required	Adj. Cost
15:00	RO1	0	-	-	-	45	45+0=45
15:00	RO2	15	-	-	-	30	30+15=45
15:00	RO3	40	-	-	-	0	0+40=40

Figure 8. Illustration of TOS with three route options that traverse two FCAs. Note that required delay for each route option and adjusted cost are in red. RO3 route has the lowest adjusted cost and is assigned to this flight with 0 minutes of ground delay. Source: Metron Aviation.

CTOP combines rerouting and delay assignment in one TMI. At the start of CTOP, airlines submit a trajectory option set (TOS) for each participating flight. A TOS contains a list of potential flights along with RTC values for each route, where the shortest path route has an RTC of zero. The resource allocation algorithm loops through the list in pre-specified order and computes required departure delay for each TOS route. It then adds RTC values for each route with required delay values to compute the adjusted route cost:

$$Adj. Cost of Route i = Required Delay + RTC_{i1}$$

Figure 8 is an example of CTOP with two FCAs. The flight that is controlled by this CTOP submitted a TOS with three route options. They have a cost of 0, 15, and 40 minutes ground delay equivalent respectively. The CTOP resource allocation algorithm calculated that routes 1 and 2 require 45 and 30 minutes of ground delay. The route with the lowest adjusted cost is route 3, since it requires no ground delay and has an RTC of 40, which results in an adjusted cost of 40.

TOS option data will not be accessible to air traffic managers since it contains proprietary information that is relevant to airlines' internal operations. However, we can use the resource allocation algorithm to our advantage to learn DCR values.

When a flight is rerouted by CTOP we know that its adjusted cost is less than or equal to the required delay for the shortest path option, because otherwise the reroute would not be assigned:

$$AC_{RR} \leq RD_{Sh} \quad (12)$$

Or alternatively:

$$RD_{RR} + RTC_{RR} \leq RD_{Sh} \quad (13)$$

$$RTC_{RR} = \Delta FT * DCR \quad (14)$$

Where RD is required delay, AC is adjusted cost, ΔFT is additional flight time of reroute relative to shortest path option, RR is reroute, and Sh stands for shortest path.

By subtracting the required delay for the reroute from the required delay for the shortest path option we can compute the upper bound on RTC for the reroute. If we divide the RTC estimate by the airborne delay for the reroute we can get the estimate of DCR.

$$DCR \leq \frac{Req. Del. Reroute - Req. Del. Shortest Path}{\Delta Flight Time} \quad (15)$$

This assumes, as we do in previous sections, that the relative trajectory cost of the reroute is the linear function of its flight time:

$$RTC_{RR} = \Delta FT_{RR} * DCR \quad (16)$$

The resulting estimate of DCR computed for each rerouted flight is the upper bound estimate—DCR could be higher, but not lower. For every 1,000 CTOP flights we expect to see 50-100 rerouted flights, for which we can compute upper bound estimates of the cost ratio. Note, however, that this method only allows us to estimate DCR bounds for flights that a) complied with CTOP and b) received a reroute by CTOP. The flights that complied with CTOP are defined as flights that submitted a TOS with at least two alternative routes that differ from each other by RTC and resource demands. Without having additional information about whether a flight complied with CTOP or not we cannot distinguish the flights with one TOS option and the flights that submitted multiple routes but were assigned the shortest route by CTOP. As a result, DCR upper bounds will only be certain for a relatively small subset of flights, since in most CTOP instances only 5-10% of flights will receive reroutes.

2) Summary of Results

Our task was to test how well we can estimate DCR values of individual flights using the methods outlined about. To achieve this, we simulated an instance of CTOP that included approximately 10,000 flights across the US NAS, controlled by

5 FCAs (see figure). Flight schedules, Trajectory Option Sets and FCA capacity rates were generated using a CTOP planning tool designed by Metron Aviation. This tool generates a TOS for each route that consists of routes that have been observed in previous flights. Some flights were allowed to submit a single route option in the TOS (shortest path option). We call such flights non-compliant. The RTC values for each TOS route were computed as the additional flight time for that route relative to the shortest path flight time, multiplied by DCR, which in the case of the simulation was set to be equal to 2 for all CTOP-compliant flights. Next, we would like to see how reliably we can estimate DCR values for flights using the method outlined above. If successful, this method will give us a method that can be used to forecast changes in airspace demand in future CTOPs, assuming that the RTC generation policy remains the same.

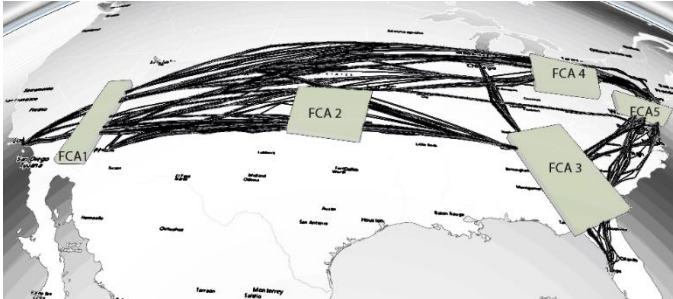


Figure 9. Simulated CTOP FCAs and flight tracks of CTOP-controlled flights. Source: Metron Aviation.

After obtaining the list of CTOP-controlled flights, we conducted CTOP resource allocation and assigned routes and delays to all flights. Due to low capacity rates that were specified in the simulation, approximately 2,500 flights were rerouted away from shortest path routes by CTOP. For this set of flights, we observe a set of two route options—the initial flight plan (shortest path) and the final flight plan (CTOP reroute). Using the CTOP resource allocation algorithm, we computed the required delay for the shortest path route option.

Figure 10 summarizes DCR upper bound estimates by various airlines. The estimates are summarized as empirical cumulative distribution functions grouped by air carrier. The true value of DCR was assumed to be equal to 2 for all flights. The estimates of DCR vary significantly across flights. While estimates for many flights are close to 2, for others the estimates are as high as several hundred. However, as in the previous method, assuming that all flights within a certain airline have the same DCR and assuming linear route cost functions, it suffices to find the lowest upper bound estimate of DCR. In all 4 cases displayed above the DCR estimate is equal to 2. In general, given large enough sample, the lowest upper bound DCR estimate will be equal to true DCR, assuming all flights within a certain group of flights have the same DCR.

This approach works when the RTC function is linear with respect to flight time and when there is no significant heterogeneity of ΔRTC values within the airline. In practice, this is unlikely to be the case. One future research direction is to modify the technique described above to predict route costs for these cases of non-linearity and heterogeneity.

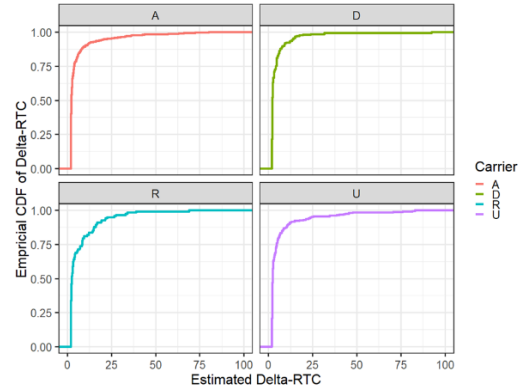


Figure 10. Empirical cumulative distribution functions of estimates of ΔRTC and the estimates of first moments.

These results show us that although this method works better than random utility models and aggregate models estimated on AFP data, since we have more information about flights' preferences, the CTOP-based estimates are not completely certain. The average estimate of DCR is far from the ground-truth value that was assumed to simulate the data. Using the minimum upper bound as DCR brings us much closer to the correct estimate. In practical settings, however, RTCs are unlikely to be a simple linear function of extra flight time, let alone one that has the same constant of proportionality across all flights. In our simulated example, the least biased DCR estimate is based on a single flight that yields the lowest upper bound estimate. We cannot predict how well this would approximate the actions of the rest of the flights in the system. As a result, the efficacy of this demand prediction model can only be tested ex post by comparing the outcomes of future CTOPs with our forecasts. This makes it necessary to constantly monitor the performance of such demand prediction models in order to avoid potentially dangerous forecasting errors.

III. DISCUSSION AND CONCLUSION

A. Summary of Route Cost Estimation Approaches

In this section we would like to summarize the results of the three demand models that were presented in the previous section: the random utility model and the aggregate model that used simulated AFP data, and the CTOP-based DCR estimation model. All three approaches share several characteristics. We simulated flight schedules, assumed that the relative trajectory cost of a route for any given flight is equal to the product of additional flight time relative to the shortest path flight time and the so-called DCR—the incremental change in relative trajectory cost as a result of 1 unit increase in flight time for a route. We then simulated several instances of TMIs that are based on real existing or planned TMIs: AFP and CTOP. For all three approaches we assigned arrival slots to flights that were included in the simulated schedules. Simulated AFP flights were allowed to respond to AFP assignment via reroutes and substitutions. Simulated CTOP flights, instead of responding to CTOP, supplied the air traffic management with information about their preferences—trajectory option sets—before assigning routes.

In all three cases we used airline response and resource assignment data to try to statistically estimate DCR values. These values, if estimated successfully, could be used to provide a detailed forecast of demand distribution over the airspace in en route TMIs. This would give planners an opportunity to see where potential new bottlenecks might occur and focus on preventing them.

In the simulated data, we made the assumptions as simple as possible. The route utility function is linear and depends on one parameter of interest—DCR. This was done to see if the tools at our disposal can reliably extract information from the simplest type of route cost models. If our models fail to recover parameters from simple data, they will fail at recovering parameters from more complicated data. Moreover, such models might be of little use in more realistic operational settings. In real-world TMIs, all inputs in the decision-making process, such as demand, capacity, flight time, and weather are uncertain. Such uncertainty may alter airline behavior. Cost models depend on network structure, fleet availability, load factors, and crew availability, making route cost a potentially non-linear stochastic function of flight time. Also, in TMIs such as AFP and CTOP airlines are not only allowed to reroute flights out of TMI, but can also cancel them, among other potential actions not taken into account in our simulations. These and other factors make the real NAS significantly more complex than the simulations presented in this paper. If our demand estimation models perform poorly in a simplified simulated setting, they will perform equally or worse in the real world.

B. Implications and Future Work

There are several implications of the inability to reliably estimate cost models that underly airlines' flight routing decisions. First, it is difficult to make forecasts of airspace demand that are more accurate than the simplest baseline demand models. As we have shown, even under idealized settings the statistical methods are not causal. Due to significant estimation bias, we will systematically underestimate the probability of reroutes for flights with low delays and overestimate the probability of reroutes of highly delayed flights. Flight cancellations, weather uncertainty, and other factors add additional layers of complexity that make it even harder to generate demand predictions.

Possibly, simple queuing models give us all the information we need to make demand predictions. The queuing diagram can tell how many flights need to reroute in order to reduce delays to a manageable level. Depending on the geometry of FCA, rerouting flights can spillover to one or both ends of the linear FCA. As a result, we can predict the changes in demand on the edges of FCAs relatively well, while potential bottlenecks that may arise in any other part of the airspace system may remain obscured. The development of such queuing models should be the focus of future research in TMI demand prediction.

One of the ways to solve the problem of demand forecasting in en route TMIs is to reframe the problem. Focusing on demand prediction is difficult and potentially harmful, since erroneous forecasts may lead to erroneous planning decisions. For example, if a surge in demand occurs at a certain airspace sector, while the forecasting model predicted a low probability of this

event, the sector control room may not be adequately staffed in order to safely conduct operations. A way to avoid this is to focus on predicting exposure to potential adverse events. For example, we may focus on finding sectors, where unanticipated changes in demand may endanger the users of the airspace system. Such problems constitute another line of future research, which might be easier to solve and may offer larger long-term benefits.

Traffic management initiatives such as CTOP are well suited to deal with such systemic fragilities. If a potentially fragile sector appears in the NAS, additional FCAs can be set up to meter demand to these areas. The spillover traffic along the edges of FCAs can also be effectively controlled by CTOP. Finally, as we have seen, the estimates of demand parameters for CTOP are much closer to reality due to the structure of the program. As a result, demand forecasts for CTOP can be made with greater accuracy than for programs like AFP.

Finally, CTOP provides us with data structures that allow us to better predict changes in the airspace demand. The CTOP-based models perform better than random-utility and aggregate models. This means that we can achieve a higher level of predictability in the system. In future work we will focus on the sensitivity of our estimates to non-linear and heterogeneous route costs. Additionally, we will look at the problem of airline responses to CTOP assignments, which will make a significant impact on distribution of demand in the airspace.

ACKNOWLEDGMENT

We would like to thank Robert Hoffman, Bert Hackney, and Rafal Kicingier of Metron Aviation for providing us with simulated CTOP data and valuable input into the paper.

REFERENCES

- [1] M.O. Ball, R. Hoffman, A.R. Odoni, R. Rifkin. "A stochastic integer program with dual network structure and its application to the ground-holding problem." *Operations research* 51.1, 2003: pp. 167-171.
- [2] A. Cook, and G. Tanner. "European Airline Delay Cost Reference Values," 2015.
- [3] R.Hoffman, B. Hackney, R. Kicingier, M.O. Ball, G. Zhu. "Computational Methods for Flight Routing Costs in Collaborative Trajectory Options Programs." 2018 Aviation Technology, Integration, and Operations Conference. 2018
- [4] R.Hoffman, B. Hackney, P. Wei, G. Zhu. "Enhanced stochastic optimization model (ESOM) for setting flow rates in collaborative trajectory options programs (CTOP)." 2018 Aviation Technology, Integration, and Operations Conference. 2018.
- [5] J. Pearl. "Causal diagrams for empirical research." *Biometrika* 82.4 (1995): 669-688.
- [6] M. Seelhorst, and M. Hansen. "Flight cancellation behavior and delay savings," 2012.
- [7] I. Tereshchenko, and M. Hansen. "Relative Trajectory Cost Estimation for CTOP Applications Using Multivariate Nonparametric Finite Mixture Logit." *International Conference for Research in Air Transportation*. 2018.
- [8] J. Xiong, and M. Hansen. "Modelling airline flight cancellation decisions," *Transportation Research Part E: Logistics and Transportation Review*, vol. 56, 2013, pp. 64-80.
- [9] J. Xiong, and M. Hansen, "Airlines' Flight Cancellation and Delay Trade-off Analysis Through Slot Credit Substitution Process," *TRB Annual Meeting*, January, 2014, paper 14-2279_A.