Centralized Disaggregate Stochastic Allocation Models for Collaborative Trajectory Options Program (CTOP)

Guodong Zhu and Peng Wei
Iowa State University
Ames, IA, USA

Robert Hoffman and Bert Hackney
Metron Aviation
Dulles, VA, USA

***Abstract*— Collaborative Trajectory Options Program (CTOP) is a relatively new Traffic Management Initiative (TMI) which assigns delays and/or reroutes around one or more Flow Constrained Area (FCA)-based airspace constraints in order to balance demand and capacity. CTOP allows flight operators to submit a set of desired reroute options, called a Trajectory Options Set (TOS), for each flight. This paper aims to answer the following question: given TOSs and scenario-based capacity information, what is the best system performance we can achieve in terms of total route and delay costs? Three disaggregate stochastic programming models, including one static model and two dynamic models, are proposed in this paper, which provides three benchmarks for a Collaborative Decision Making (CDM)-compatible CTOP. If we further assume the travel time between constrained areas is the same for all flights, which is a mild condition, we can get three corresponding more aggregate models which have better computational efficiency. Important issues like equity, maximum delay, intra-airline cancellation, and substitution are discussed. A realistic CTOP use case has been created to test effects of TOS participation and to compare static versus dynamic planning techniques. These results can be coded in a decision support tool and can help air traffic managers understand, initiate, and perform post-analysis for CTOP programs.**

***Keywords*— CTOP; TOS; Stochastic Programming; Static Model; Dynamic Model; TMI**

NOMENCLATURE

Notation Used in All Six Models

N	Number of flights
N_i	Number of route options for flight i
c_{ij}	Cost of flight i taking route j
δ_{ij}	Binary indicator whether flight i takes route j
d_{ij}	Ground delay of flight i if taking route j
P	Number of PCAs, $k = 1, \cdots, P$
T	Number of time periods, $t = 1, \cdots, T$
Q	Number of scenarios
p_q	Probability of scenario q occurring
$a^k_{q,ij}$	Air delay of flight i taking route j before entering PCA k under scenario q
M^k_{tq}	Real capacity of PCA k in time period t under scenario q
$\tau^k_{q,i}$	Time period in which flight i crosses PCA k under scenario q
t^k_{ij}	Time period in which flight i taking route j is scheduled to cross PCA k
Ω_{ij}	Ordered set of indices of the PCAs which flight i taking route j crosses
Φ^k	Set of indices of the routes which are planned to cross PCA k
$B^k_{q,i,t}$	Binary indicator whether flight i crosses PCA k in time period t under scenario q
$B^k_{i,j,t}$	Binary indicator whether flight i taking route j will cross PCA k (first PCA on route j) in time period t
C	Set of ordered pairs of PCAs. $(k, k') \in C$ iff k is connected to k' in the directed graph of PCAs
$\Delta^{k,k'}$	Number of time periods to travel from PCA k to k', defined for all pairs $(k, k') \in C$
$P^k_{t,r}$	Planned direct demand at PCA k in time period t from flights with same path r
$L^k_{t,r,q}$	Number of flights with same path r that actually cross PCA k in time period t under scenario q
$A^{k,q}_{t,r}$	Number of flights with same path r taking air delay before entering PCA k
$\mathcal{M}$	Limit on delay a flight can take

Notation in Semi-Dynamic and Dynamic Models

t_s	Time period at which stage s begins
d_{qsij}	Ground delay of flight i which is scheduled to depart during stage s taking route j under scenario q
δ_{qsij}	Binary indicator whether flight i, scheduled to depart during stage s, will take route j under scenario q
B	Total number of branches in the scenario tree
N_b	Number of scenarios corresponding to branch $b \in \{1, \dots, B\}$
$B^k_{q,i,j,t}$	Binary indicator whether flight i taking route j will cross PCA k (first PCA on route j) in time period t under scenario q
$P^{k,q}_{t,r}$	Planned direct demand at PCA k in time period t from flights with same path r under scenario q

Notation in Dynamic Models

δ_{qtij}	Binary indicator whether flight i, scheduled to depart in time period t, will take route j under scenario q
Y_{qit}	Binary indicator whether flight i is released for departure during time period t under scenario q
Dep_i	Original scheduled departure time for flight i
o_b, μ_b	Start and end nodes of branch b

I. INTRODUCTION

The goal of Air Traffic Flow Management (ATFM) is to alleviate projected demand-capacity imbalances at airports and in en route airspace through formulating and applying strategic Traffic Management Initiatives (TMIs). Two classical types of TMIs are Ground Delay Programs (GDP) and Airspace Flow Programs (AFP), which apply ground delay to flights bound for congested airspace which would otherwise experience costly and unsafe air delay. These traditional TMIs have several limitations, including the need for at least one TMI for each region of congested airspace, low common situational awareness among the Federal Aviation Administration (FAA) and airspace users, and low flexibility.

Designed to be a superset of the classical TMIs, a Collaborative Trajectory Options Program (CTOP) combines many features from its predecessors and brings two important new features: first, it can manage multiple constrained regions in an integrated way with a single program; second, it allows flight operators to submit a set of desired reroute options (called a Trajectory Options Set or TOS), which provides great flexibility and efficiency to the airspace users. In a TOS, each option is associated with a Relative Trajectory Cost (RTC) and some usage restrictions, such as time window of validity of the routing option. RTC is expressed in equivalent ground delay minutes, which encodes flight operators' conditional preferences among different route choices [1].

Maximizing airspace utilization and preserving equity among competing airspace users are two objectives of ATFM. In the current Collaborative Decision Making (CDM) paradigm, Ration-by-Schedule (RBS) is accepted as the standard principle for equitable resource allocation [2][3]. A major research question for TMI optimization is how to plan rates for airports or airspace Flow Constrained Areas (FCAs). "Rate" is the number of flights that will be admitted to the FCA in given time interval. Since there is inherent uncertainty in weather forecasts, and the demand can also be stochastic due to flight cancellation, drift, pop-up flights, and TOS submission in CTOP, we need to deal with a sequential Decision Making Under Uncertainty (DMU) problem. Various DMU frameworks have been explored by researchers, including Markov Decision Process (MDP) [4][5], Chance-Constrained Programming (CCP) [6][7][8], and Robust Optimization (RO) [9]. Simulation-based optimization has also been used to determine the GDP parameters under uncertainly [10]. The dominant approach in ATFM literature is stochastic programming, in which the capacity uncertainty is represented by a finite number of scenarios arranged in a scenario bush (two-stage case) or scenario tree (multistage case). Most of the literature on TMI optimization is focused on capacity uncertainty and the Single Airport Ground Holding Problem (SAGHP).

Two pioneering works on applying two-stage and multistage stochastic programming on SAGHP were done by Richetta et al. in the early 1990s [11][12]. The first stochastic model that conforms to the current CDM operating procedure, published by Ball et al. [13], is a two-stage high aggregate model that directly computes Planned Acceptance Rates (PARs) for a weather-impacted airport. It was later proved that under very mild conditions, the model in [11] can also generate CDM-compatible solutions [14]. In the aforementioned models, once a ground-delay decision is made, it cannot be revised, even though the flight is still on the ground and further ground-holding is possible. Mukherjee formulated a flight-level multistage model that allows a flight to take ground delays multiple times based on the latest capacity information and the scenario tree structure [15]. Importantly, his model gives the theoretical system cost lower bound for the scenario-based SAGHP optimization problem.

This paper presents two major extensions to the classical SAGHP stochastic programming models. First, we generalize the single constrained resource planning models to the case of having multiple constrained resources; second, we generalize single route delay planning models to the case in which a flight can have multiple rerouting options apart from taking delays. Using the dominant DMU framework in ATFM, we have solved a very general and fundamental research problem, whose result is not only meaningful to CTOP research in the U.S., but also is helpful to researchers in other regions. In this work, six disaggregate stochastic programming models are formulated, including three fully disaggregate and three partially disaggregate models. The models are disaggregate because the route cost (RTC) is different for each flight. Some models are partially disaggregate because once the routing decision for each flight has been made, we can group flights by the congested regions in which they travel, in order to reduce the number of decision variables and constraints. Similar to the theoretical value of Mukherjee's work for the SAGHP [15], our dynamic models can give the theoretical lower bound for the very general multiple constrained resources multiple route options rerouting, ground and air hold problem.

There are five companion CTOP planning papers to this work: in [16], we proposed a deterministic mixed-integer linear programming model to optimally allocate route and slots to flights, and demonstrated the benefit of intra-airline cancellation and substitution. The models in [16] can be viewed as special cases of models in this paper where capacity information is known perfectly. In [17], we presented a highly aggregate CDM-compatible two-stage stochastic model which can generate rates for the FCAs; in [18], we identified and addressed some of the problems in [17], and proposed a multistage model for FCA rate planning; in [19], we pointed out why CTOP rate planning is essentially a multi-commodity flow problem, given the correct boundary conditions, and formulated three highly computationally efficient aggregate stochastic models to plan delays for flights traversing multiple congested regions; in [20], we studied the impact of demand uncertainty on TMI optimization, introduced a heuristic saturation technique, and discussed its important role in GDP and CTOP rate planning.

II. Some Concepts and Model Assumptions

A key concept used in this paper is a Potentially Constrained Area (PCA), which is the physical airspace region or resource in which demand may exceed capacity and whose future capacity evolutions are represented by a finite set of scenarios. This is different from a more familiar concept of FCA, which serves like a valve to control the traffic flows into a region. In this paper, we will not have FCAs in our models since we will control individual flights rather than setting limits for the amount of traffic into a certain airspace.

A related concept is the PCA network, which refers to a directed graph that links the PCAs and models the potential movement of traffic between them (Figure 4).

Here are the key assumptions used in this paper:

- The flights captured by CTOP are not controlled by other TMIs at the same time; in other words, we do not consider the TMI interaction.
- The assigned route information, the topology of the PCA network, unimpeded PCA entry times, CTOP window, and scenario-based PCA capacity information are all available. These are our model inputs.
- All flights are required to exit the PCA network by the end of the planning horizon. This boundary condition ensures that results from different models can be fairly compared.
- Flight separation will be enforced by the Air Traffic Control (ATC) controllers. We need to make sure the number of flights traversing the PCA in each time period is no larger than its physical capacity; therefore the model is a strategic interval-based model. The bin size used in this paper is 15 minutes. We can only delay a flight by integer multiples of the size of a time bin.
- We can proactively ground-hold and air-hold flights. Therefore, models are centralized and will be used to achieve the theoretical system cost lower bound.

III. Two-stage Static Stochastic Model

In this section, we will introduce the two-stage static model. In a two-stage model, we will make the first stage decision at the beginning of a CTOP: determining both the route and ground delay for each flight. At the second stage, we will take the recourse action and determine the air delays needed to pass the constrained areas.

A. Fully Disaggregate Version

In a fully disaggregate model, we will determine route, the amount of ground delay and air delay for each flight.

The objective function minimizes the total route cost, ground delay, and expected air delay costs. α and β are the weighting coefficients. Since the size of a time interval is 15 minutes, $d_{ij} = 1$ means flight i will need to take 15 minutes of ground delay before flying route j.

$$\min_{\delta,d,a} \sum_{i=1}^{N} \sum_{j=1}^{N_i} (c_{ij}\delta_{ij} + \alpha d_{ij}) + \beta \sum_{q=1}^{Q} p_q \sum_{i=1}^{N} \sum_{j=1}^{N_i} \sum_{k\in\Omega_{ij}} a^k_{q,ij} \quad (1)$$

For each flight, one and only one route should be chosen:

$$\sum_{j=1}^{N_i} \delta_{ij} = 1, \text{ for } i = 1, \cdots, N \quad (2)$$

Only if we choose route j for flight i, can d_{ij} or $a^k_{q,ij}$ be nonzero. $\mathcal{M}$ is a parameter and can be flight or even route specific. It puts limit on the maximum delay flight i can take.

$$0 \le d_{ij} + \sum_{k\in\Omega_{ij}} \sum_{q=1}^{Q} a^k_{q,ij} \le \delta_{ij}\mathcal{M} \quad \forall i, j \quad (3)$$

The ETA of flight i for PCA k along route j under scenario q equals to the nominal PCA entry time plus the ground delay taken on the ground and the amount of air delay taken before PCA k and all upstream PCAs:

$$\tau^k_{q,i} = \sum_{\substack{j=1\ldots N_i \\ (i,j)\in\Phi^k}} (t^k_{ij}\delta_{ij} + d_{ij} + \sum_{k\in\Omega_{ij};\mathrm{id(k)}\le\mathrm{id}(k)} a^{\mathrm{k}}_{q,ij}) \quad \forall i, k, q \quad (4)$$

Route j may sequentially cross several PCAs. Here id(k) is the sequence number of PCA k on that route.

Auxiliary binary variable $B^k_{q,i,t}$ connects the flight arrival time and flight count. A flight can only arrive at a PCA once:

$$\sum_{t=1}^{T} B^k_{q,i,t} \le 1 \quad \forall i, k, q \quad (5)$$

If $\tau^k_{q,i} = t'$, then $B^k_{q,i,t'} = 1$:

$$\sum_{t=1}^{T} t B^k_{q,i,t} = \tau^k_{q,i} \quad (6)$$

Note that if the route j assigned to flight i does not pass PCA k, then $\tau^k_{q,i} = 0$ and thus $B^k_{q,i,t} = 0$ for all t.

Finally, we have the set of physical capacity constraints:

$$\sum_{i=1}^{N} B^k_{q,i,t} \le M^k_{tq} \quad \forall\ t, k, q \quad (7)$$

It is easy to see that if a flight takes the route-out option and does not pass any PCAs, then $B^k_{q,i,t} = 0\ \forall q, k, t$ and will not take any slot.

In a fully disaggregate model, we track the motion of each individual flight as it moves in time, hence it is a Lagrangian type of model. After solving the model, all the information for every flight will be known. Fully disaggregate model provides more direct control, but it interferes with current traffic flow management practices, such as accepted practices in the FAA rationing logic for allocating arrival slots to flights

B. Partially Disaggregate Version

If we assume the travel time between two consecutive PCAs is the same for all flights, which is a rather mild assumption, we can aggregate flights by paths and get a more efficient formulation. A path is defined by a sequence of PCAs that flights traverse. For example, in Figure 4, PCA1 $\rightarrow$ PCA_EWR is one path, and PCA $\rightarrow$ PCA_EWR is another path. A path is different from a route, which starts and ends at airports and is composed of waypoints. Different routes can have the same path, if they cross the same PCAs.

Again in this first stage we need to select route and amount of ground delay for each flight:

$$\sum_{j=1}^{N_i} \delta_{ij} = 1, \text{ for } i = 1, \cdots, N \\ 0 \le d_{ij} \le \delta_{ij}\mathcal{M} \quad \forall i, j \tag{8}$$

Once the route and ground delay have been determined for flight i, we want to know when it will reach the first PCA k on the assigned route j. The binary variable $B^k_{i,j,t}$ satisfies:

$$\sum_{t=t^k_{ij}}^{T} tB^k_{i,j,t} = t^k_{ij}\delta_{ij} + d_{ij} \tag{9}$$

$$\sum_{\substack{j=1\ldots N_i \\ (i,j)\in\Phi^k}} \sum_{t=t^k_{ij}}^{T} B^k_{i,j,t} \le 1 \quad \forall i, k \tag{10}$$

Constraint (10) means flight i will only pass PCA k once.

We can pre-process all the candidate routes for all flights and categorize them according to paths. The planned direct demand at PCA k from flights with the same path r is:

$$P^k_{t,r} = \sum_{i=1}^{N} \sum_{j\in\text{Path } r} B^t_{k,i,j} \quad t = 1, \cdots, T \tag{11}$$

In (11), k is the first PCA on path r, $k = r_1$. The rest of constraints are similar to [19] section III.

$$L^k_{t,r,q} = \begin{cases} \text{if } k = r_1 & P^k_{t,r} - (A^k_{t,r,q} - A^k_{t-1,r,q}) \\ \text{else} & \text{UpPCA}^k_{t,r,q} - (A^k_{t,r,q} - A^k_{t-1,r,q}) \end{cases} \\ \text{UpPCA}^k_{t,r,q} = L^{k'}_{t-\Delta^{k',k},r,q} \quad (k', k) \in r \\ M^k_{t,q} \ge \sum_r L^k_{t,r,q} \\ P^k_{t,r}, L^k_{t,r,q} \ge 0 \tag{12}$$

We require all flights to land or exit the PCA network at the end of the planning horizon:

$$\sum_{t=1}^{T} P^{k=r_1}_{t,r} = \sum_{t=1}^{T} L^{k=r_{-1}}_{t,r,q} \quad \forall r, q \tag{13}$$

The left-hand side of (13) is the total demand of commodity r (flights that fly path r) which enter the PCA system through the first PCA on path r. The right-hand side is the cumulative amount of commodity r which exits the PCA system via the last PCA on path r.

The objective function now becomes:

$$\min_{\delta,d,a} \sum_{i=1}^{N} \sum_{j=1}^{N_i} (c_{ij}\delta_{ij} + \alpha d_{ij}) + \beta \sum_{q=1}^{Q} p_q \sum_{k\in\text{PCAs}} \sum_{t=1}^{T} \sum_r A^k_{t,r,q} \tag{14}$$

We call the second model a partially disaggregate model because, when assigning routes and ground delays to flights, we consider each flight individually, but once the flights enter the PCAs we treat them as aggregate traffic flows. The second model can also be viewed as an Lagrangian-Eulerian type model, because we will first make decisions for each flight, then focus on observing traffic flows at specific locations, i.e. PCAs, in the airspace. After solving the model, we do not know how much air delay each flight will take. This information will only be known after we run the resource allocation algorithm.

IV. Multistage Semi-Dynamic Stochastic Model

A shortcoming of the two-stage model is that it does not take advantage of the structure information of a scenario tree. In this section, we will formulate a Richetta's type multistage model to overcome this limitation [12]. Similar to Richetta's SAGHP model, in CTOP we will determine the route and ground delay at a flight's original scheduled departure time, not when CTOP is just proposed. Richetta's multistage model is often called a semi-dynamic model because, in contrast to Mukherjee's dynamic model, in Richetta's model the ground delay, once assigned, cannot be revised. This also holds true for our model: a flight can be only ground-delayed once and assigned a route once. As a result, Richetta's type model is less efficient in terms of system delay cost, but has a higher predictability about departure time and route assignment than Mukherjee's type model. The essence of Richetta's type model is that, by delaying the time of assigning ground delays and/or reroutes, we can take advantage of more weather information and make better decisions.

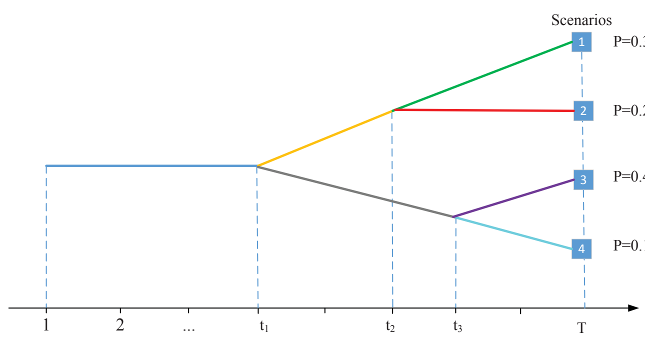

Fig. 1. A scenario tree with four scenarios [15]

In this model, we will use the concept of stage. A stage can comprise several time periods, at which we have the same weather information. For example in Figure 1, there are four stages, and dotted vertical lines indicate the starting times of each stage. Because we have multiple PCAs in a CTOP, the branching point in a scenario tree means we have new weather information for at least one PCA. This time the route and ground delay decisions are also scenario-dependent.

A. Fully Disaggregate Version

The objective function minimizes the expected route, ground delay, and air delay costs:

$$\min \quad \sum_{q=1}^{Q} p_q \sum_{i=1}^{N} \{ \sum_{j=1}^{N_i} (c_{ij}\delta_{qsij} + \alpha d_{qsij}) + \beta \sum_{j=1}^{N_i} \sum_{k\in\Omega_{ij}} a^k_{q,ij} \} \tag{15}$$

Note that the original departure stage number s is uniquely determined by the flight ID i. One and only one route should be chosen for each flight i:

$$\sum_{j=1}^{N_i} \delta_{qsij} = 1, \text{ for } i = 1, \cdots, N, \text{ for } q = 1, \ldots, Q \quad (16)$$

Only if we choose route j for flight i can d_{qsij} and $a^k_{q,ij}$ be nonzero.

$$0 \leq d_{qsij} + \sum_{k \in \Omega_{ij}} a^k_{q,ij} \leq \delta_{qsij} \mathcal{M} \quad (17)$$

The ETA of flight i at PCA k along route j under scenario q is:

$$\tau^k_{q,i} = \sum_{\substack{j=1\ldots N_i \\ (i,j)\in\Phi^k}} (t^k_{ij}\delta_{qsij} + d_{qsij} + \sum_{k\in\Omega_{ij};\mathrm{id(k)}\leq \mathrm{id}(k)} a^{\mathrm{k}}_{q,ij}) \quad (18)$$

The capacity constraints are the same as in static model (5)-(7).

The most important set of constraints in a multistage model are nonanticipativity constraints, which ensure that decisions made at time t are solely based on the information available at t [21].

$$\begin{aligned} \delta_{q^b_1 sij} &= \cdots = \delta_{q^b_{N_b} sij} \\ d_{q^b_1 sij} &= \cdots = d_{q^b_{N_b} sij} \end{aligned} \quad (19)$$

The constraints mean if a set of scenarios are on the same branch, we should take exactly the same actions with respect to the set of scenarios. The branch(es) information is determined by stage s, which is in turn determined by a flight's original scheduled departure time. For example, if flight i is scheduled to depart at time period $t_1 + 1$, then we have:

$$\begin{aligned} \delta_{12ij} &= \delta_{22ij} \quad d_{12ij} = d_{22ij} \\ \delta_{32ij} &= \delta_{42ij} \quad d_{32ij} = d_{42ij} \end{aligned}$$

The major improvement of this model over the two-stage static model is that the model explicitly takes into account the updated capacity information.

B. Partially Disaggregate Version

In a multistage model, binary variable $B^k_{q,i,j,t}$ becomes scenario-dependent and now satisfies:

$$\begin{aligned} &\sum_{t=t^k_{ij}}^{T} B^k_{q,i,j,t} \leq 1 \\ &\sum_{t=t^k_{ij}}^{T} tB^k_{q,i,j,t} = t^k_{ij}\delta_{qsij} + d_{qsij} \end{aligned} \quad (20)$$

The direct demand at PCA k with same path r is also scenario-dependent now:

$$P^{k,q}_{t,r} = \sum_{i=1}^{N} \sum_{j\in \text{Path } r} B^k_{q,i,j,t} \quad (21)$$

The other constraints are the same as in [19], section IV; we omit them here. The objective function is:

$$\begin{aligned} \min \quad &\sum_{q=1}^{Q} p_q \{ \sum_{i=1}^{N} \sum_{j=1}^{N_i} (c_{ij}\delta_{qsij} + \alpha d_{qsij}) + \\ &\beta \sum_{k\in \text{PCAs}} \sum_{t=1}^{T} \sum_{r} A^k_{t,r,q} \} \end{aligned} \quad (22)$$

V. Multistage Dynamic Stochastic Model

In this section, we will introduce Mukherjee's type multistage model, also known as the truly dynamic model, for CTOP. The idea of this model is that when making ground delay/rerouting decisions, we will consider the fact that a flight may be further ground delayed/rerouted later on, in other words "plan to replan."

A. Fully Disaggregate Version

The first set of constraints enforce that in any scenario q, a flight can only depart once and only one route can be chosen:

$$\sum_{t=1}^{T} \sum_{j=1}^{N_i} \delta_{qtij} = 1 \quad \forall i, q \quad (23)$$

and the route is chosen at the actual departure time:

$$\sum_{j=1}^{N_i} \delta_{qtij} = Y_{qit} \quad \forall i, t = \text{Dep}_i, \ldots, T, q \quad (24)$$

Note that the t in δ_{qtij} is no longer solely determined by i. The key difference between Mukherjee's type model and Richetta's type model is that we explicitly use the actual departure time as a decision variable, since we will impose the nonanticipativity constraints at the actual departure time. For every branch in the scenario tree we have the following nonanticipativity constraints:

$$\delta_{q^b_1 tij} = \cdots = \delta_{q^b_{N_b} tij} \quad \forall i, b, j, o_b \leq t \leq \mu_b \quad (25)$$

Because of constraint (24), (25) also has the following effect:

$$Y_{q^b_1 it} = \cdots = Y_{q^b_{N_b} it} \quad \forall i, b, o_b \leq t \leq \mu_b \quad (26)$$

Constraint (26) is what we have seen in the GDP problem [15]. The ground delay for flight i under scenario q is:

$$\sum_{t=1}^{T} (t - \text{Dep}_i) Y_{qit} \quad (27)$$

Again, only if we choose route j for flight i can air delay $a^k_{q,ij}$ be nonzero.

$$0 \leq \sum_{k\in\Omega_{ij}} a^k_{q,ij} \leq \mathcal{M} \sum_{t=1}^{T} \delta_{qtij} \quad \forall i, j, q \quad (28)$$

The ETA of flight i at PCA k along route j under scenario q is:

$$\tau^k_{q,i} = \sum_{\substack{j=1\ldots N_i \\ (i,j)\in\Phi^k}} (t^k_{ij} \sum_{t=1}^{T} \delta_{qtij} + \sum_{k\in\Omega_{ij}} a^k_{q,ij}) + \sum_{t=1}^{T} (t - \text{Dep}_i) Y_{qit} \quad (29)$$

The capacity constraints are the same as in static model (5)-(7). Finally, the objective function is:

$$\min \quad \sum_{q=1}^{Q} p_q \sum_{i=1}^{N} \{\alpha \sum_{t=1}^{T} (t - \text{Dep}_i) Y_{qit} + \sum_{t=1}^{T} \sum_{j=1}^{N_i} c_{ij} \delta_{qtij} + \beta \sum_{j=1}^{N_i} \sum_{k \in \Omega_{ij}} a^k_{q,ij}\} \tag{30}$$

B. Partially Disaggregate Version

Binary variable $B^k_{q,i,j,t}$ satisfies:

$$\begin{aligned} &\sum_{t=t^k_{ij}}^{T} B^k_{q,i,j,t} \leq 1 \\ &\sum_{t=t^k_{ij}}^{T} t B^k_{q,i,j,t} = t^k_{ij} \sum_{t=1}^{T} \delta_{qtij} + \sum_{t=1}^{T} (t - \text{Dep}_i) Y_{qit} \\ &P^{k,q}_{t,r} = \sum_{i=1}^{N} \sum_{j \in \text{Path } r} B^k_{q,i,j,t} \end{aligned} \tag{31}$$

The other constraints are same as in [19], section IV, omitted here. The objective function is:

$$\min \quad \sum_{q=1}^{Q} p_q \{ \sum_{i=1}^{N} (\alpha \sum_{t=1}^{T} (t - \text{Dep}_i) Y_{qit} + \sum_{t=1}^{T} \sum_{j=1}^{N_i} c_{ij} \delta_{qtij}) + \beta \sum_{k \in \text{PCAs}} \sum_{t=1}^{T} \sum_{r} A^k_{t,r,q} \} \tag{32}$$

VI. Additional Modeling Considerations

A. Equity Issue

The current formulations only consider the efficiency issue. As in [16], we can easily add an equity term to the objective function to make efficiency-equity tradeoffs.

B. Limit on the Delays

It is straightforward to add the maximum ground and/or air delay limits, a concern for airlines, on each flight for fully disaggregate models. For example, for the multistage dynamic model we can have:

$$\sum_{t=1}^{T} (t - \text{Dep}_i) Y_{qit} + \sum_{\substack{j=1 \ldots N_i \\ (i,j) \in \Phi^k}} \sum_{k \in \Omega_{ij}} a^k_{q,ij} \leq \mathcal{M}_{ij} \quad \forall q \tag{33}$$

We may also want to restrict the maximum number of flights taking air delay before a PCA, i.e. the length of the queue before a constrained resource, which air traffic controllers care about. However, it is not easy to impose such a constraint for fully disaggregate models.

On the other hand, for partially disaggregate models it is easy to add the constraint for queue size, since we explicitly model the "inventory" of airborne flights at each PCA:

$$\sum_{r} A^{k,q}_{t,r} \leq \text{\# of Flights Limit} \quad \forall k, t, q \tag{34}$$

There is no way to restrict the maximum delay for each flight in the partially disaggregate model because we do not know the airline delay information at the flight level.

C. TOS Route Restrictions

In the proposed models, we ignore many practical restrictions including TOS route restrictions. There are three optional requirements for each route that can be provided by the flight operator: Required Minimum Notification Time (RMNT), which allows for needed preparation time, such as adding fuel; Trajectory Valid Start Time (TVST) and Trajectory Valid End Time (TVET), which are the earliest and latest acceptable takeoff times for that TOS option, respectively.

From the current time, flight's scheduled departure time and RMNT, we can directly add the minimum ground delay needed (MGD_{ij}) for flight i to take route j:

$$\begin{aligned} \text{two-stage}: &\quad d_{ij} \geq \delta_{ij} \text{MGD}_{ij} \\ \text{semi-dynamic}: &\quad d_{qsij} \geq \delta_{qsij} \text{MGD}_{ij} \quad \forall q \\ \text{dynamic}: &\quad \sum_{t=1}^{T} (t - \text{Dep}_i) Y_{qit} \geq \sum_{t=1}^{T} \delta_{qtij} \text{MGD}_{ij} \quad \forall q \end{aligned} \tag{35}$$

Similarly, TVET and TVST also impose upper and lower constraints on the required ground delay time.

D. Intra-airline Cancellation and Substitutions

Like Mukherjee et al. did for the GDP problem [15], we can formulate an optimization model for CTOP to enable collaborative decision making by allowing airlines to execute scenario-contingent cancellations and substitutions. The two key differences here are that we have multiple PCAs and that we need to consider the air delays before reaching PCAs. Define

- $v^t_{a,q,k}$: Number of slots owned by airline a for PCA k during time period t under scenario q
- $c(i, \lambda, \mu)$: Cost for flight i to take λ time periods of ground delay and μ time periods of air delay
- $B^q_{i,\lambda,\mu}$: Binary indicator whether flight i takes λ time periods of ground delay and μ time periods of air delay under scenario q

$v^t_{a,q,k}$ can be determined from the solutions of stochastic models. $c(i, \tau, \mu)$ is flight-specific and can be nonlinear to capture the downstream missed connection effect of crews, passengers, and the airframe. The optimization model remains linear due to the introduction of binary variable $B^q_{i,\lambda,\mu}$.

$$\begin{aligned} &\sum_{\lambda,\mu} B^q_{i,\lambda,\mu} = 1 \\ &\sum_{\lambda} \lambda B^q_{i,\lambda,\mu} = \sum_{t=1}^{T} (t - \text{Dep}_i) Y_{qit} \\ &\sum_{\mu} \mu B^q_{i,\lambda,\mu} = \sum_{j=1}^{N_i} \sum_{k \in \Omega_{ij}} a^k_{q,ij} \end{aligned} \tag{36}$$

When making flight substitutions, we assume a flight can choose any route from its TOS. The other constraints in the

Fig. 2. Weather forecast for 2210z, taken at 1522z on July 15, 2016

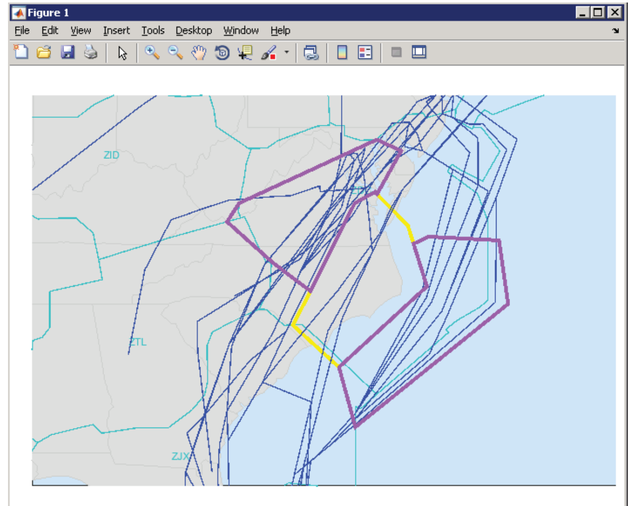

Fig. 3. Traffic routing around the original PCA

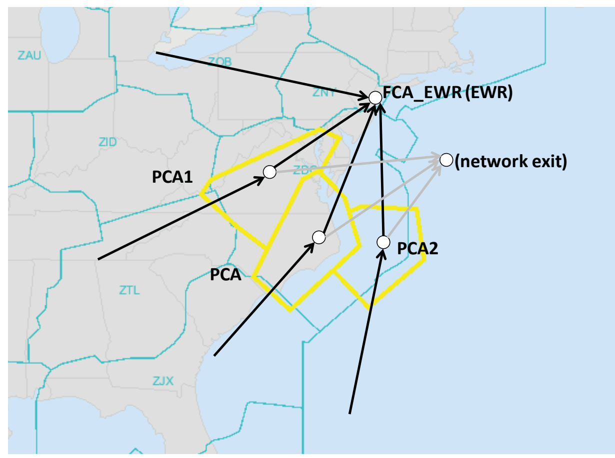

Fig. 4. Geographical display of a PCA network

dynamic substitution model are the same as (23) to (29). We only need to replace $M_{q,k}^t$ with $v_{a,q,k}^t$.

Each airline a will minimize its total ground delay and air delay cost function:

$$\min \quad \sum_{q=1}^{Q} p_q \sum_{i \in F_a} \{ \sum_{\lambda,\mu} c(i,\lambda,\mu) B_{i,\lambda,\mu}^q + \sum_{t=1}^{T} \sum_{j=1}^{N_i} c_{ij} \delta_{tij}^q \} \tag{37}$$

VII. Experimental Results

To demonstrate the performance of the proposed models, we created an operational use case based on actual events from July 15, 2016 [22].

A. Southern ZDC and EWR with Convective Activity

This use case primarily addresses convective weather activity in southern Washington Center (ZDC). Figure 2 shows the pattern of convective weather activity for that day. Southern ZDC is adversely impacted by the weather. We further assume there is demand-capacity imbalance at EWR airport. In principle, the EWR imbalance could be addressed by an isolated GDP. However, much of the traffic bound for EWR passes through southern ZDC; therefore, we show how the EWR arrival traffic can be folded into the same CTOP that addresses southern ZDC. Note that the traffic congestion at southern ZDC is comparable to an AFP with two 'wing' FCAs added, shown in Figure 3. The PCA network is shown in Figure 4. We assume there is a four-hour capacity reduction in ZDC/EWR from 2000z to 2359z.

B. Creating Capacity Profiles

For comparison purposes, we use the same capacity data as in [19]. The capacity changes can be modeled by a scenario tree, shown in Figure 5. Though this is a very simple tree, it has more than one branching point, which is more complex than a scenario bush. We expect the dynamic models will take advantage of the structure information and outperform the static model. The three scenarios correspond to good, average, and bad demand-capacity imbalances. The capacity information is listed in Table I.

Note in GDP optimization, we usually add one extra time period to make sure all flights will land at the end of the planning horizon. Because CTOP has multiple constrained resources, we need to add more than one time period depending on the topology of the PCA network. In this use case, we add four extra time periods, because the maximum average travel time between the three en route PCAs and EWR is around 1 hour (4 time periods). We assumed nominal capacity for the extra four time periods in Table I.

C. Traffic Demand

We used historical flight data from September 8, 2016, a representative clear-weather day for traffic demand. We avoided using the actual flight data from July 15, 2016, because flight plans and airline operational schedules were likely influenced by weather forecasts and related ATFM events on that day. We only kept flights which pass through one of the 3 PCAs created in ZDC plus all EWR arrivals. The resulting set contains 1098 flights, among them 890 flights that traverse the PCAs in their active periods. To form the base (preferred) route for each flight, we drew historical

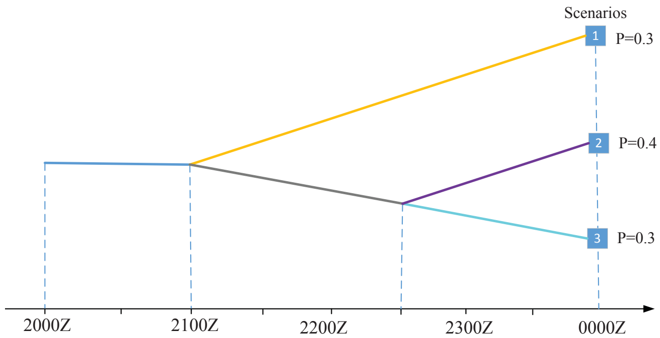

Fig. 5. Scenario tree used in the experiment

filed flight plans (from Sept. 8, 2016) from System Wide Information Management (SWIM) data.

A typical TOS package that might be submitted for this day would have one route for each PCA and one route that avoids all PCAs. To model the TOSs that airlines might submit in response to a CTOP, we drew from a combination of reroute TMIs from SWIM data and from Coded Departure Routes (CDR). There are in total 1368 TOS options for 890 flights, on average 1.54 options per flight.

D. Model Comparisons

From Figure 4 we know there are 7 possible paths: direct demand to EWR, cross one of the three PCAs then land at EWR, or cross one of the PCAs then exit the system. We require all the CTOP captured flights to land at EWR or to exit the PCA network at the end of the planning horizon.

To compare the results with aggregate models in [19], we first restrict flights to only take their most preferred route. The results are summarized in Table II. To show the minimum system cost we can potentially achieve in a CTOP, flights are allowed to choose a route from their TOSs before departure. The results are shown in Table III. The optimization models are solved using Gurobi 7.5.2 on a workstation with 3.6 GHz processors and 32 GB RAM [23].

There are several key observations from these two tables:

- The overall system cost could decrease by over 50% if flight operators submit TOS options for flights. This shows the benefit of allowing rerouting in the face of congestion.
- There are discrepancies between fully disaggregate model and partially disaggregate model results. The discrepancies are caused by rounding down the travel times between PCAs into 15-minute intervals.
- The second part of Table II has exactly the same results as the aggregate models paper [19]. This is anticipated and verifies the correctness of the results in both papers.
- The two-stage solution outperforms the deterministic policy (SCEN1-3), as it should, since it explicitly considers the uncertainty when making holding decisions.
- The semi-dynamic model solution is better than the two-stage model solution, and the dynamic model in turn performs better than the semi-dynamic model, which is also expected, because the dynamic model uses more weather evolution and flight schedule information than the two-stage static model.
- Partially disaggregate models are in general faster than fully disaggregate models. This is one of the motivations to develop partially disaggregate models.
- The computation times of deterministic, two-stage, and semi-dynamic models are all very short for a flight-by-flight level optimization problem. For fully dynamic models, the computation time is still acceptable. In all cases, the optimality gap between early stop solutions at 3 minutes and optimal values are less than 1%.

VIII. Conclusions

In this paper, we proposed six stochastic programming models for CTOP to test varying levels of flight aggregation and dynamics in the planning process. The performance of these six models is tested on a realistic CTOP use case, in which we have shown the overall system delays cost could decrease by over 50% if flight operators submit TOS options for flights, and a dynamic stochastic model could outperform a deterministic model by around 13%. The models are also promising in terms of computation time, and can be coded in a decision support tool to help air traffic managers understand, initiate, and perform post-analysis for CTOP programs.

The future work includes testing on more realistic capacity data and testing using a larger New York Metroplex use case, investigating the impact of cost ratio, air holding limit, lead time, etc. on the model solutions, investigating the value of weather forecast information, incorporating demand uncertainty, incorporating more realistic constraints on route and ground delay assignment, further improving the computational efficiency by employing optimization and computation techniques, and comparing the six models with FCA rate-planning models in [17][18].

IX. Acknowledgement

The authors would like to gratefully acknowledge Heather Arneson, Antony Evans, Banavar Sridhar, Deepak Kulkarni, Paul Lee, and Nancy Smith at NASA Ames Research Center and James Jones at MIT Lincoln Laboratory for their helpful discussions. This work was partly funded by NASA Research Announcement contract #NNA16BD96C.

X. Appendix

A. Acronyms

ATFM	Air Traffic Flow Management
ATC	Air Traffic Control
TMI	Traffic Management Initiative
GDP	Ground Delay Program
AFP	Airspace Flow Program
FCA	Flow Constrained Area
PCA	Potential Constrained Area
CTOP	Collaborative Trajectory Options Program
TOS	Trajectory Options Set
FAA	Federal Aviation Administration
RTC	Relative Trajectory Cost

	Resource/Time Bin	20:00	15	30	45	21:00	15	30	45	22:00	15	30	45	23:00	15	30	45	00:00	15	30	45
SCEN1	PCA	13	13	13	13	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25
	PCA1	44	44	44	44	50	50	50	50	50	50	50	50	50	50	50	50	50	50	50	50
	PCA2	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5
	EWR	8	8	8	8	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
SCEN2	PCA	13	13	13	13	13	13	13	13	13	13	25	25	25	25	25	25	25	25	25	25
	PCA1	44	44	44	44	44	44	44	44	44	44	50	50	50	50	50	50	50	50	50	50
	PCA2	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5
	EWR	8	8	8	8	8	8	8	8	8	8	10	10	10	10	10	10	10	10	10	10
SCEN3	PCA	13	13	13	13	13	13	13	13	13	13	13	13	13	13	13	13	25	25	25	25
	PCA1	44	44	44	44	44	44	44	44	44	44	44	44	44	44	44	44	50	50	50	50
	PCA2	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5
	EWR	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	10	10	10	10

TABLE I

CAPACITY SCENARIOS

Disaggregate Model Fully VS. Partially	Ground Delay Periods If This Scenario Occurs:			Air Holding Periods If This Scenario Occurs:			Expected Cost	Running Time	Early Stop	Early Stop
	SCEN1	SCEN2	SCEN3	SCEN1	SCEN2	SCEN3		Mins	at 1 Min	at 3 Mins
SCEN1	106	106	106	0	194	383	491.0	0.05		
SCEN2	300	300	300	0	0	197	418.2	0.08		
SCEN3	489	489	489	0	0	0	489.0	0.30		
Two-Stage Model	300	300	300	0	0	189	413.4	0.52		
Semi-Dynamic Model	182	300	424	0	0	67	342.0	1.12	342.0	
Dynamic Model	145	302	478	0	0	12	314.9	9.41	324.1	317.1
Perfect Information	106	300	489	0	0	0	298.5			
SCEN1	88	88	88	0	198	392	481.6	0.03		
SCEN2	280	280	280	0	0	209	405.4	0.03		
SCEN3	470	470	470	0	0	0	470.0	0.03		
Two-Stage Model	280	280	280	0	0	197	394.0	0.02		
Semi-Dynamic Model	156	280	403	0	0	67	319.9	0.32		
Dynamic Model	116	280	463	0	0	7	289.9	1.38	292.6	
Perfect Information	88	280	470	0	0	0	279.4			

TABLE II

DETERMINISTIC VS. STOCHASTIC SOLUTIONS COMPARISION (DELAY COST RATIO $\beta/\alpha = 2$) **WITHOUT** TOS

Disaggregate Model Fully VS. Partially	RTC Costs in Mins If This Scenario Occurs:			Ground Delay Periods If This Scenario Occurs:			Air Holding Periods If This Scenario Occurs:			Expected Cost	Running Time	Early Stop	Early Stop
	SCEN1	SCEN2	SCEN3	SCEN1	SCEN2	SCEN3	SCEN1	SCEN2	SCEN3		Mins	at 1 Min	at 3 Mins
SCEN1	76	76	76	75	75	75	0	127	232	325.93	0.10		
SCEN2	222	222	222	125	125	125	0	0	49	184.00	0.07		
SCEN3	262	262	262	148	148	148	0	0	0	182.93	0.35		
Two-Stage Model	250	250	250	125	125	125	0	0	25	173.33	0.33		
Semi-Dynamic Model	216	254	266	107	124	133	0	0	16	164.03	0.32		
Dynamic Model	216	254	266	95	124	136	0	0	15	160.73	4.93	163.49	161.93
Perfect Information	76	222	262	75	125	148	0	0	0	142.24			
SCEN1	76	76	76	53	53	53	0	129	249	315.73	< 0.01		
SCEN2	222	222	222	100	100	100	0	0	50	159.6	< 0.01		
SCEN3	288	288	288	112	112	112	0	0	0	150.40	< 0.01		
Two-Stage Model	270	270	270	100	100	100	0	0	16	145.60	0.02		
Semi-Dynamic Model	216	254	266	73	99	109	0	0	11	133.63	0.37		
Dynamic Model	216	254	266	61	99	113	0	0	7	128.83	3.43	134.83	128.83
Perfect Information	76	222	288	53	100	112	0	0	0	115.9			

TABLE III

DETERMINISTIC VS. STOCHASTIC SOLUTIONS COMPARISION (DELAY COST RATIO $\beta/\alpha = 2$) **WITH** TOS

RCL	Rate Computation Loop
RBS	Ration by Schedule
SAGHP	Single Airport Ground Holding Problem
DMU	Decision Making Under Uncertainty
ETA	Estimated Time of Arrival

REFERENCES

[1] Federal Aviation Administration, "CTOP, Collaborative Trajectory Options Program, system operations programs," Washington D.C., 2010, https://www.nbaa.org/ops/airspace/tfm/tools/faa-ctop-overview-091610.pdf.

[2] K. Chang, K. Howard, R. Oiesen, L. Shisler, M. Tanino, and M. C. Wambsganss, "Enhancements to the FAA ground-delay program under collaborative decision making," *Interfaces*, vol. 31, no. 1, pp. 57–76, 2001.

[3] CSC/TFMM-13/1744, "TFMS Functional Description, Appendix C: Traffic Management Initiative (TMI) Algorithms," Tech. Rep., 2014.

[4] P. Liu and M. Hansen, "Scenario-free sequential decision model for the single airport ground holding problem," in *Proc. 7th USA/Eur. Air*

Traffic Manage. R&D Seminar, 2007.

[5] J. Cox and M. J. Kochenderfer, "Ground delay program planning using Markov Decision Processes," *Journal of Aerospace Information Systems*, pp. 134–142, 2016.

[6] G. Clare and A. Richards, "Air traffic flow management under uncertainty: application of chance constraints." Proceedings of the 2nd international conference on application and theory of automation in command and control systems, 20–26, IRIT Press, 2012.

[7] J. C. Jones, R. DeLaura, M. Pawlak, S. Troxel, and N. Underhill, "Predicting & quantifying risk in airport capacity profile selection for air traffic management," in *14th USA/Europe Air Traffic Management Research and Development Seminar (ATM2017), Seattle, USA*, 2017.

[8] J. Chen and D. Sun, "Stochastic ground-delay-program planning in a metroplex," *Journal of Guidance, Control, and Dynamics*, vol. 41, no. 1, pp. 231–239, 2017.

[9] S. Gupta and D. J. Bertsimas, "Multistage air traffic flow management under capacity uncertainty: a robust and adaptive optimization approach." AGIFORS-International Federation of Operational Research Societies, 2011.

[10] L. S. Cook and B. Wood, "A model for determining ground delay program parameters using a probabilistic forecast of stratus clearing," *Air traffic control quarterly*, vol. 18, no. 1, p. 85, 2010.

[11] O. Richetta and A. R. Odoni, "Solving optimally the static ground-holding policy problem in air traffic control," *Transportation Science*, vol. 27, no. 3, pp. 228–238, 1993.

[12] ——, "Dynamic solution to the ground-holding problem in air traffic control," *Transportation Research Part A: Policy and Practice*, vol. 28, no. 3, pp. 167–185, 1994.

[13] M. O. Ball, R. Hoffman, A. R. Odoni, and R. Rifkin, "A stochastic integer program with dual network structure and its application to the ground-holding problem," *Operations Research*, vol. 51, no. 1, pp. 167–171, 2003.

[14] B. Kotnyek and O. Richetta, "Equitable models for the stochastic ground-holding problem under collaborative decision making," *Transportation Science*, vol. 40, no. 2, pp. 133–146, 2006.

[15] A. Mukherjee and M. Hansen, "A dynamic stochastic model for the single airport ground holding problem," *Transportation Science*, vol. 41, no. 4, pp. 444–456, 2007.

[16] G. Zhu and P. Wei, "An Interval-based TOS Allocation Model for Collaborative Trajectory Options Program." AIAA Aviation, Atlanta, GA, June 2018.

[17] R. Hoffman, B. Hackney, G. Zhu, and P. Wei, "Enhanced Stochastic Optimization Model (ESOM) for setting flow rates in a Collaborative Trajectory Options Programs (CTOP)." AIAA Aviation, Atlanta, GA, June 2018.

[18] G. Zhu, P. Wei, R. Hoffman, and B. Hackney, "Risk-hedged multistage stochastic programming model for setting flow rates in Collaborative Trajectory Options Program (CTOP)." Submitted to AIAA Scitech 2019.

[19] ——, "Aggregate multi-commodity stochastic models for Collaborative Trajectory Options Program (CTOP)." International Conference on Research in Air Transportation (ICRAT), Barcelona, Spain, June 2018.

[20] ——, "Saturation technique for optimizing planned acceptance rates in traffic management initiatives." The 21st IEEE International Conference on Intelligent Transportation Systems, Hawaii, USA, 2018.

[21] J. R. Birge and F. Louveaux, *Introduction to stochastic programming*. Springer Science & Business Media, 2011.

[22] Metron Aviation, "Collaborative Trajectory Options Program (CTOP) use cases to assess in the SMART-NAS test bed," Tech. Rep., 2017.

[23] Gurobi Optimization, Inc., "Gurobi optimizer reference manual," 2016. [Online]. Available: http://www.gurobi.com