Risk-hedged Multistage Stochastic Programming Model for Setting Flow Rates in Collaborative Trajectory Options Programs (CTOP)

Guodong Zhu* and Peng Wei[†]

Iowa State University, Ames, IA, 50011, U.S.A.

Robert Hoffman[‡] and Bert Hackney[§]

Metron Aviation, Herndon, VA, 20171, U.S.A.

As a new tool in the NextGen portfolio, the Collaborative Trajectory Options Programs (CTOP) combines multiple features from its forerunners including Ground Delay Program (GDP), Airspace Flow Program (AFP) and reroutes, and can manage multiple Flow Constrained Area (FCA) with a single program. A key research question in CTOP is how to traffic flow rates under airspace capacity uncertainties. In this paper, we first investigate existing CTOP optimization models and point out their roles in CTOP and in more general air traffic flow management research, and their advantages and disadvantages in terms of Collaborative Decision Making (CDM) compatibility, model flexibility, performance, and practicality. Having identified the missing piece in the family of CTOP rates optimization models, we proposed a multistage stochastic model to fill the niche, study its connections with other models, and test its performance on a realistic CTOP use case. We discussed in detail the variance and risk issues in air traffic flow management applications, which have thus far received little attention, and showed how to hedge system performance with variance or risk measures in the stochastic models. The models and discussions in this work are not only useful in implementing and analyzing CTOP programs, but are also valuable for the general multiple constrained airspace resources optimization problem.

Nomenclature

Notations in Enhanced Stochastic Optimization Model (ESOM)

R	Set of resources, including FCAs and PCAs
\mathcal{C}	Set of ordered pairs of resources. $(r, r') \in \mathcal{C}$ iff r is connected to r' in directed graph
$\Delta^{r,r'}$	Number of time periods to travel from PCA r to r' . Defined for all pairs $(r, r') \in \mathcal{C}$
T	Number of time periods, $t = 1, \dots, T$
Q	Number of scenarios, $q = 1, \dots, Q$
p_q	Probability that scenario q occurs
$f_t^{r,r'}$	Fraction of flights from resource r directed to resource r' in time period t
$M_{t,q}^r$	Real capacity of PCA r in time period t under scenario q
$L_{t,q}^r$	Number of flights that actually cross PCA r in time period t under scenario q
$L_{t,q}^r \\ A_{t,q}^r$	Number of flights taking air delay before crossing PCA r in time period t under scenario q
c_g	Cost per time period for holding one aircraft on the ground
c_a	Cost per time period for holding/delaying one aircraft in the air

Notations in Multistage Semi-Dynamic Stochastic Model

^{*}Graduate Research Assistant, Aerospace Engineering Department, Iowa State University, AIAA Student Member

[†]Assistant Professor, Aerospace Engineering Department, Iowa State University, AIAA Senior Member

[‡]Vice President, Advanced Research and Engineering Services, Metron Aviation, AIAA Senior Member

[§]Senior Analyst, Advanced Research and Engineering Services, Metron Aviation, AIAA Member

 $\begin{array}{ll} t_s & \text{Time period in which stage s begins} \\ S^r_{s,t} & \text{Number of flights originally scheduled to depart in stage s and arrive in FCA r (direct demand) in time period t \\ X^{r,q}_{s,t,t'} & \text{Number of flights originally scheduled to depart in stage s and arrive in FCA r in time period t, rescheduled to arrive in time period t' under scenario q \\ P^r_{t,q} & \text{Planned acceptance rate at FCA r in time period t under scenario q \\ B & \text{Total number of branches in the scenario tree} \\ N_b & \text{The number of scenarios corresponding to branch $b \in \{1, \dots, B\}$} \end{array}$

I. Introduction

Collaborative Trajectory Options Programs (CTOP) is a new Traffic Management Initiative (TMI) which is used by air traffic managers to balance demand with available airspace capacity. Compared with its predecessors Ground Delay Program (GDP), Airspace Flow Program (AFP) and reroutes, the introduction of CTOP brings two major benefits: the ability to handle multiple Flow Constrained Areas (FCAs) with a single program, and the flexibility it gives to airspace users to express their conditional preference over different route choices by allowing them to submit a set of desired reroute options (Trajectory Options Set or TOS)^[1].

An important research question in TMI optimization has been determining airport/FCA planned acceptance rates, which is the maximum number of aircraft to be accepted at each time period at a constrained resource. Because planned acceptance rates have to be set hours in advance so that flights can absorb delays on the ground or reroute to avoid the congested airspace, when the capacity reduction is caused by weather activity, decision making has to deal with the uncertain nature the weather forecast. Most of the literature on TMI optimization has focused on single airport GDP planning, and the dominant decision making under uncertainty approach has been stochastic programming, which minimizes the expected cost under different weather scenarios. Since the 1990s, the Federal Aviation Administration (FAA) has made significant changes in their air traffic flow management, moving from a centralized system to one called Collaborative Decision Making (CDM). Many decision support tools for air traffic managers and airline personnel are developed under this CDM paradigm^[2]. Representative work differs in two aspects: the degree to which traffic managers can modify or revise flights' controlled departure times and the compatibility with current CDM software. In [3, 4], the ground delay decisions are made at the beginning of the planning horizon and the two models are CDM-compatible [5]. In [6], the ground delay decisions are made at the flights' original scheduled departure stage to make use of the updated weather information. In [7], ground delay decisions consider the fact this flight may be further ground delayed later on (plan to replan). Both [6] and [7] are not compatible with current CDM software, since these two models assume that the control of individual flights is possible.

CTOP traffic flow rates planning is substantially more challenging than GDP or AFP planning due to two reasons. First, because there are now multiple FCAs, the locations of FCAs can be in parallel or in series and the rate of one FCA may affect the traffic volume to adjacent or downstream FCAs. Therefore the rates for CTOP FCAs need to be determined in an integrated way. Second, in general demand estimation and capacity information are both needed to determine the optimal FCA rates [8]. However, in CTOP the air traffic demand for the constrained regions is uncertain, since one flight now may have more than one route option.

This work is an endeavor to optimize CTOP traffic flow rates under airspace capacity uncertainties in a coordinated and dynamic way while considering variance and risk issues. The main contributions of this paper are as follows:

- 1. We conduct a literature review and categorized the existing CTOP traffic flow rates stochastic optimization models based on five criteria:
 - (a) How dynamic or flexible the model is in terms of assigning ground delays (and reroutes)
 - (b) Does the model assume the routes have been assigned
 - (c) Whether we explicitly differentiate FCA, which is a flow control mechanism, with Potentially Constrained Areas (PCA), the actual constrained airspace region [9]
 - (d) The aggregate level of decision variables
 - (e) CDM-CTOP software compatibility

Like the periodic table of elements that can guide the finding of new chemical elements, Table 1 can guide researchers to develop new models for the classical and future TMIs, which is of great value for air traffic flow management research.

	Aggregate Model	Disaggregate Models					
	FCA-PCA Rate Planning	PCA Rate Planning	Flight Level Planning for PCA Network				
	Multi-commodity Flow Approximation	Multi-commodity Flow	Lagrangian	Lagrangian-Eulerian			
Static	ESOM ^[9]	[10] §IV	[11] §IIIA	[11]§IIIB			
Semi-dynamic	Semi-dynamic ESOM	[10] §V	[11] §IVA	[11]§IVB			
Dynamic	$ ext{D-ESOM}^{[12]}$	[10] §VI	[11] §VA	[11]§VB			

Table 1: Comparison of Stochastic Models for CTOP

- 2. We point out that among all CTOP rates optimization models, the Enhanced Stochastic Optimization Model (ESOM)^[9] is most consistent with current CTOP and CDM software. Due to its important role, we further investigate the properties of this model, reveal its shortcomings and propose some potential remedies.
- 3. Having identified the missing piece in the family of CTOP rates optimization models from Table 1, we propose a multistage semi-dynamic stochastic model to fill the niche, study its connections with other models, and tested its performance on a realistic CTOP use case. This model performs better than ESOM in terms of delay costs, and we can see more characteristics of FCA-PCA type of model as models become more dynamic.
- 4. Minimizing expected cost under different weather scenarios makes more sense if the stochastic event repeats over and over again. This is not true for the air traffic flow management problem, especially for the en route case. Air transportation is a risk-averse industry. Apart from expected performance of the solution, air traffic managers also care about predictability (the variance issue) and want to control the impact of extreme events (tail behavior or risk issue). In this paper, we will discuss in details various variance and risk measures.

This paper is organized as follows: in section II, we discuss some important concepts which are used throughout this paper. From section III to section V, we talk in detail about items 1 to 4 listed above. In section VI, we test the models on a realistic use case. In section VII, we summarize the major findings and point out future work.

II. Preliminary Concepts

In this section, we talk about five key components of various models described in the following sections.

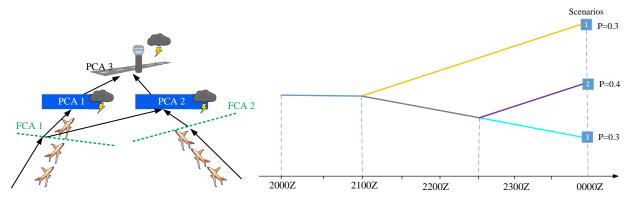


Figure 1: FCAs and PCAs

Figure 2: Scenario Tree

A. Potential Constrained Area and Capacity Scenarios

Potential Constrained Areas (PCAs) are the airspace regions in which air traffic demand may exceed capacity. Though an imbalance resulting from pure demand increase is possible, we usually focus on the case in which adverse weather causes capacity reduction. In the stochastic programming framework, future capacity realization is represented by a finite set of scenarios arranged in a scenario tree, and each scenario is associated with a probability. A branch point in a scenario tree corresponds to the time when we acquire new information and similar scenarios evolve into distinct scenarios. For example, there are three scenarios in Figure 2. Scenario 1 to 3 correspond to optimistic, average and pessimistic weather forecasts, respectively.

B. Flow Constrained Area and Planned Acceptance Rate

Flow Constrained Area (FCA) was first introduced in AFP to model constrained airspace resources. In the CTOP setting, different from a PCA which coincides with a physically constrained area and whose future capacity is stochastic, a FCA is an artificial line or region in the airspace and serves like a valve to control traffic flows into a region. Figure 1 depicts both FCAs and PCAs in a traffic management setting. The key research question we want to answer in this paper is determining the planned acceptance rate for each FCA. FCA rates are selected by air traffic manager and can be different from the real physical capacity or PCA capacity.

In CTOP, we explicitly distinguish the control mechanism (FCA with one set of acceptance rates) with the source of the problem (PCA with multiple sets of possible capacity values), which facilitates the multiple constrained resources mathematical modeling. In AFP, this distinction was not necessary, since we do not need to model the flow of the traffic from one constrained resource to another one.

C. FCA-PCA network

A related concept is the FCA-PCA network, which is a directed graph that links the FCAs and PCAs and models the potential movement of traffic between them. This concept gives structure to our problem and allows us to use various network flow optimization techniques.

III. Review and Remarks on Existing CTOP Models

The key input to all stochastic models is a scenario tree, which models the weather evolution and encodes the capacity uncertainty. The other model inputs include CTOP planning horizon, FCA/PCA network topology, flight TOS information and unimpeded entry times to the FCAs and PCAs for each TOS route. Table 1 lists twelve CTOP traffic flow rates stochastic programming models. Our new model is in the shaded cell. In this section, we discuss in detail about these models from five perspectives.

A. Differences and Connections of CTOP Models

1. Multistage Model Vs. Two-stage Model

The first criterion that differentiates these models is how flexibly the flights' controlled departure times (and reroutes) can be modified or revised. Similar to single airport GDP planning, if the ground delay (and reroute) are determined only using scenario information but not the branch point information, and usually assigned at the beginning of the planning horizon, the model is classified as a static model; if the ground delay (and reroute) decisions are made at a flight's original scheduled departure stage or other predetermined time, and uses the latest weather information, the model is semi-dynamic; if the ground delay (and reroutes) can be revised multiple times and each decision takes these future revisions into account, then the model is dynamic. The essence of the difference is how much information about the structure of the scenario tree and flight schedule is exploited by a model. Both the semi-dynamic models and dynamic models are multistage stochastic models, and static models are two-stage stochastic models. The dynamic models are more flexible than semi-dynamic models, which in turn are more flexible than static models. The more flexible a model is, the better system delay performance it can achieve, but the less predictable the flight schedule will be.

If the capacity information is known precisely, i.e. there is only one possible future scenario, the models in each column of Table 1 will produce exactly the same results. For example, the CTOP TOS allocation

model [13] can be seen as a special case of the models in [11] where deterministic capacity information is available.

2. Ground Delay Model Vs. Ground Delay and Route Assignment Model

The second criterion is whether we assume the route assignment has been made before running the model. For models in columns 1 and 2 of Table 1, we assume the routes have been allocated and the following question is answered: given route assignment, how can we best manage this demand, in terms of system delay costs, through the congested regions by dynamically assigning delays to flights. For models in columns 3 and 4, we have the freedom to choose routes for flights and we aim to show the best total system performance we can potentially achieve in terms of total reroute and delay costs.

A feature of CTOP is that it will assign not only the ground delay but also the reroute. Therefore, models in columns 1 and 2 only solve part of the CTOP planning problem, and have to be paired with TOS allocation algorithm. A TOS allocation algorithm takes planned acceptance rates as input, creates and assigns slots to flights in an equitable way [14;15]. Models in columns 3 and 4, on the other hand, solve the ground delay and reroute assignment problem at the same time in a centralized way.

3. FCA-PCA Model Vs. PCA Model

The third criterion is whether we use both FCA and PCA concepts in the model. In the current CDM-CTOP software implementation, only the concept of FCA is considered. When implementing a CTOP, air traffic managers will input the FCA planned acceptance rates and then run the TOS allocation algorithm. Therefore only if a model can generate FCA planned acceptance rates, can it be used by the current software. Hence FCA-PCA models have profound practical values.

Without the artificial FCA concept, the mathematical formulations of PCA models are neater and the "physical pictures" are clearer. Because the current TOS allocation algorithm only supports FCA planned acceptance rates as input, designing a TOS allocation algorithm that supports PCA rates input is a promising idea.

FCA-PCA models and PCA models are related. In [10], we show that from PCA planned acceptance rates we can explore the locations of FCAs and also obtain FCA planned acceptance rates.

To clarify different "rates" in FCA-PCA Models (column 1), we have FCA (conditional) planned acceptance rates or FCA rates for short; in PCA models (column 2), we have PCA (conditional) planned acceptance rates or PCA rates; when we talk about traffic flow control in CTOP in general, we vaguely call the (FCA or PCA) rates CTOP traffic flow rates, or CTOP rates.

4. Aggregate Model Vs. Disaggregate Model

The fourth criterion is the aggregate level of decision variables. The CTOP rate planning problem is naturally a multi-commodity problem, since flights will traverse different congested airspace regions and reach different destinations. We can choose to use the pre-calculated flow split ratio to approximate the traffic flows between resources (column 1), or choose to explicitly deal with a multi-commodity flow problem (column 2). In the latter case, we need to group flights by the PCAs they traverse. The former one is more aggregate than the latter one. In multistage semi-dynamic and dynamic models, we need to further group flights by departure stages or en route times, and models will become even less aggregate. We can only talk about (conditional) planned acceptance rates for aggregate models.

Models in columns 3 and 4 are essentially at a flight-by-flight level and therefore called disaggregate models. They have to be at flight level since these models also determine the reroute assignment and the composition of TOSs is different for different flights.

5. CDM-CTOP Compatibility

The fifth criterion is the compatibility with current CDM-CTOP software. It is important to first define what is CDM-compatibility. In the literature, one definition of CDM-compatibility is that the model should be able to accommodate FAA and airline operations including slots compression, intra-airline cancellation and substitution [7]. In this loose sense, all twelve models can be made to be consistent with CDM philosophy. Here we adopt a stricter definition: the model should be compatible with the current CDM-CTOP software implementation.

All semi-dynamic and dynamic models are not compatible with the current CDM-CTOP software. One of the reasons is that the current software doesn't support conditional delay decisions. On a side note, software compatibility is just one reason for their lesser practicality compared with static models. Another reason is a weather forecast scenario tree with relatively accurate branch points is not easily obtainable yet for multiple constrained en route resources. But this is an important line of research in the aviation weather community [16–19]. Static PCA models and disaggregate models, even though they do not directly optimize FCA rates, can provide guidelines on setting FCA rates.

The models that do not conform to the current CDM-CTOP implementation are still valuable in theory, and can be used as benchmarks for the compatible models and as references for future software development.

B. Enhanced Stochastic Optimization Model (ESOM)

In this section, we discuss one of the twelve models which is of particular interest to us due to its practical value. ESOM extends the classical single resource GDP model to the multiple constrained resources case ^[9]. The main advantage of ESOM is that it is most consistent with current CDM-CTOP software. From ESOM we can compute planned acceptance rates that can be directly implemented in decision support tools. Another advantage of ESOM is its computation speed, since it is the most aggregate model in Table 1. The third reason why we want to revisit ESOM is that some of its characteristics are inherited by Semi-dynamic ESOM and D-ESOM in column 1. Improvements to ESOM will also be valuable for the development of those models.

For ease of reference, ESOM is listed below. In summary, the objective function (1) minimizes ground delay and expected air delay costs; constraint (2) plans the ground delay at each FCA; constraint (3) gives the number of air delayed flights at each PCA; constraint (4) reflects the physical capacity constraint at each PCA; constraint (5) models the traffic flow split between resources; constraint 7 is the boundary condition; constraints (8, 9) model the travel time between resources. The reader is referred to [9] for the detailed derivation of this model.

$$\min \sum_{r \in FCA} \sum_{t=1}^{T} c_g G_t^r + \sum_{r \in PCA} \sum_{t=1}^{T} \sum_{q=1}^{Q} c_a p_q A_{t,q}^r$$
(1)

s.t.
$$P_t^r = \text{UpFCA}_t^r + D_t^r - (G_t^r - G_{t-1}^r) \quad \forall r \in \text{FCA}, \forall t$$
 (2)

$$L_{t,q}^r = \text{UpFCA}_t^r + \text{UpPCA}_{t,q}^r - (A_{t,q}^r - A_{t-1,q}^r) \quad \forall r \in \text{PCA}, q, t$$
(3)

$$M_{t,q}^r \ge L_{t,q}^r \quad \forall r \in \text{PCA}, \forall q, \forall t \tag{4}$$

$$\sum_{(r,r')\in\mathcal{C}} f_t^r = 1 \tag{5}$$

$$G_t^r, P_t^r, L_{t,q}^r, A_{t,q}^r \ge 0 \quad \forall r, \forall q, \forall t$$
 (6)

$$G_0^r = G_{T+1}^r = A_{0,q}^r = A_{(T+1),q}^r = 0 (7)$$

$$UpFCA_t^r = \sum_{(r',r)\in\mathcal{C}} f_{t-\Delta^{r'},r}^{r',r} \cdot P_{t-\Delta^{r'},r}^{r'}$$
(8)

$$UpPCA_{t,q}^{r} = \sum_{(r',r)\in\mathcal{C}} f_{t-\Delta^{r'},r}^{r',r} \cdot L_{t-\Delta^{r'},r,q}^{r'}$$

$$\tag{9}$$

ESOM has some shortcomings. These result from the difficulty in dealing with CTOP's inherent demand uncertainty and the approximation to the multi-commodity flow. As a result, in general we cannot obtain the theoretical optimal FCA planned acceptance rates.

1. As we discussed in [8], if we do not use saturation technique, we will likely end up with bad planned acceptance rates policy because of the demand shift issue. On the other hand, we proved that even in the simpler GDP case, saturation technique in general cannot give the optimal solution and can only be used as a heuristic to find a better suboptimal solution.

One possible way to avoid the demand shift problem is to use ESOM as a simulation model or use other fast-time simulation models, treat P_t as input parameters and use simulation-based optimization. This is still ongoing research and the model no longer falls into the category of stochastic programming, so it is not the focus of this paper.

2. In ESOM, we cannot follow flights' route schedule precisely. For example, if 10 flights are scheduled to pass PCA1 and then land at the airport in Figure 1, in ESOM we cannot guarantee these 10 flights will eventually travel the scheduled planned route and land by the end of planning horizon.

This is because split ratios $f_t^{r',r}$ are pre-calculated parameters. If we delay some flights to the next time period, we have to use the split ratios in the next period. As an extreme example, in Figure 1 suppose at time period t the split ratio from FCA 1 to PCA1 is 1, and from FCA 1 to PCA 2 is 0;

Time period	FCA1-PCA1	FCA1-PCA2
t	1	0
t+1	0	1

Table 2: Example of Pre-calculated Ratios

at time period t + 1, split ratio from FCA 1 to PCA1 is 0, and from FCA 1 to PCA 2 is 1. If we delay some flights from t to t + 1, then these flights will be rerouted to PCA 2 instead of PCA 1 by ESOM. The flow approximation results in the violation of the original flight route schedules. In some examples, a flight could even be rerouted to a different destination airport.

- 3. Since $f_t^{r',r}$ are continuous numbers, when we solve ESOM, we cannot impose integrality constraints for decision variables. It is well known that objective value of a linear programming relaxation can be very different from the objective value of integer solutions [20]. After we solve ESOM we need to round the planned acceptance rates P_t^r to integers. It is not yet clear what a good rounding strategy is.
- 4. In ESOM, we cannot strictly enforce boundary conditions. In TMI rates optimization, we usually require all the affected flights to land at airport or exit the constrained airspace by the end of the planning horizon. Imposing $G_{T+1}^r = A_{T+1,q}^r = 0$ is sufficient for single resource GDP planning, but not sufficient in CTOP planning since in CTOP there can be en route flights between constrained resources. There are two potential ways to alleviate the problem:
 - As suggested in [9], we can choose a sufficiently long planning horizon. Since ground delay and air delay are penalized in the objective function, all CTOP captured flights will eventually exit the FCA-PCA network. There are two disadvantages with this approach: first, this is not an explicit constraint, therefore we do not know if the boundary condition is indeed met; second, it is not obvious how long the planning horizon should be
 - We can explicitly enforce at each PCA that the total number of landed/exited flights during the entire planning horizon $\sum_{t}^{t} L_{t,q}^{r}$ should equal to the total number of flights that are scheduled to land/exit at that PCA [10]. The latter value can be obtained by route pre-processing. The main problem with this approach is that due to the approximation of flow splits, this constraint may not be exactly met and the problem can become infeasible.

Since traffic flow schedules cannot be strictly followed, boundary conditions may not be exactly met and integrality constraints are not enforced, the objective function is only an approximation to the true objective value.

The root cause of items 2 to 4 is that we approximate the essentially multi-commodity flow as a single-commodity traffic flow. To our knowledge, this approximation has to be made in order to directly optimize FCA rates. In other words, to better satisfy the CDM-CTOP compatibility requirement, we could only do a compromised traffic flow control and optimization.

5. In constraint (2) of the original version of ESOM, for FCA r the incremental number of flights taking ground delay can be as high as $UpFCA_t^r + D_t^r$. But flights in $UpFCA_t^r$ have already taken off and may have incurred ground delay before. A better way to model ground delay could be introducing intermediate variables $P_t^{\prime r}$ and these constraints:

$$P_{t}^{\prime r} = D_{t}^{r} - (G_{t}^{r} - G_{t-1}^{r})$$

$$P_{t}^{r} = P_{t}^{\prime r} + \text{UpFCA}_{t}^{r}$$

$$P_{t}^{\prime r} \ge 0$$
(10)

IV. Multistage Semi-Dynamic ESOM Model for FCA Rate Planning

In this section, we introduce the multistage semi-dynamic version of FCA-PCA model. The first set of constraints is the conservation of flow constraints

$$\sum_{t'=t}^{T} X_{s,t,t'}^{r,q} = S_{s,t}^{r} \quad \forall s, \forall q, \forall r \in \text{FCA}$$

The scenario dependent demand for each FCA r in time period t is

$$\sum_{s} \sum_{t' < t} X_{s,t',t}^{r,q} \quad \forall q \tag{11}$$

At each FCA, we have the planned acceptance rates equal to the summation of the upstream demand and direct demand from the airport

$$P_{t,q}^{r} = \text{UpFCA}_{t,q}^{r} + \sum_{s} \sum_{t' < t} X_{s,t',t}^{r,q}$$
(12)

Constraint (12) is similar to (10) in that we only ground-delay flights that directly fly from the departing airports. A key difference here is that planned acceptance rates $P_{t,q}^r$ now become scenario dependent. As a result, UpFCA_{t,q} is also scenario dependent.

The flow at each PCA node is:

$$L_{t,q}^r = \text{UpFCA}_{t,q}^r + \text{UpPCA}_{t,q}^r - (A_{t,q}^r - A_{t-1,q}^r) \quad \forall r \in \text{PCA}, \forall q, \forall t$$
(13)

The maximum capacity at each PCA is:

$$M_{t,q}^r \ge L_{t,q}^r \quad \forall r \in \text{PCA}, \forall q, \forall t$$
 (14)

Fractional flows sum to 1.0:

$$\sum_{(r,r')\in\mathcal{C}} f_t^{r,r'} = 1 \quad \forall r \in \text{FCA, PCA}, \forall t$$
(15)

Decision variables must be nonnegative:

$$G_{t,q}^r, P_{t,q}^r, L_{t,q}^r, A_{t,q}^r \ge 0 \quad \forall r, \forall q, \forall t$$
 (16)

We assume the planning horizon is sufficient long:

$$G_0^r = G_{T+1}^r = A_{0,q}^r = A_{(T+1),q}^r = 0 (17)$$

In multistage stochastic programming we need nonanticipativity constraints, which ensure that decisions made at time t are solely based on the information available at that time. These can be written as follows:

$$X_{s,t,t'}^{r,q_b^b} = \dots = X_{s,t,t'}^{r,q_{N_b}^b} \tag{18}$$

Flow from upstream resources is captured in these auxiliary variables:

$$\operatorname{UpFCA}_{t,q}^{r} = \sum_{(r',r)\in\mathcal{C}} f_{t-\Delta^{r',r}}^{r',r} \cdot P_{t-\Delta^{r',r},q}^{r'}$$

$$\operatorname{UpPCA}_{t,q}^{r} = \sum_{(r',r)\in\mathcal{C}} f_{t-\Delta^{r',r}}^{r',r} \cdot L_{t-\Delta^{r',r},q}^{r'}$$
(19)

The objective function minimizes expected ground delay and air delay costs:

$$\min \sum_{q=1}^{Q} p_q \{ \sum_{r \in FCA} \sum_{s} \sum_{t'=t}^{T} (t'-t) c_g X_{s,t,t'}^{r,q} + \sum_{r \in PCA} \sum_{t=1}^{T} c_a A_{t,q}^r \}$$
 (20)

A. Remarks

As we have mentioned, the main advantage of the semi-dynamic model is that in theory it performs at least as well as the static ESOM model in terms of system delay costs. We also stressed that the semi-dynamic FCA-PCA model has all the imperfections ESOM has. Moreover,

- since $P_{t,q}^r$ now depends on which scenario happens, $P_{t,q}^r$ policy is no longer directly implementable in current CDM-CTOP software.
- across a range of scenarios, the flow of the air traffic can be quite different. But for all scenarios, $P_{t,q}^r$ and $L_{t,q}^r$ both use the same split ratio $f_t^{r,r'}$. It is easy to imagine the split ratios will not be good approximations for at least some scenarios.

The static model can be considered as a special case of semi-dynamic model. By imposing

$$X_{s,t,t'}^{r,1} = X_{s,t,t'}^{r,2} = \dots = X_{s,t,t'}^{r,Q} \quad \forall s, t, t', \forall r \in \text{FCA}$$
 (21)

the semi-dynamic model can be reduced to static model.

B. Additional Modeling Considerations

In the original formulation, we assume we obtain the scenario tree (Figure 3b) from a weather forecast, and we only know the new operational conditions if they have actually changed. In other words, we have to wait until the branch point to know which scenario actually materializes. In reality, since the short-term weather forecast is fairly accurate, we can assume we know which scenario will happen a few time periods in advance. We assume a flight is ready to take off anytime at or after the scheduled takeoff time. It is more reasonable to require the model to make the ground delay decision a certain time before the scheduled departure time so that airlines can decide when to let passengers board.

1. Short-term Weather Forecast

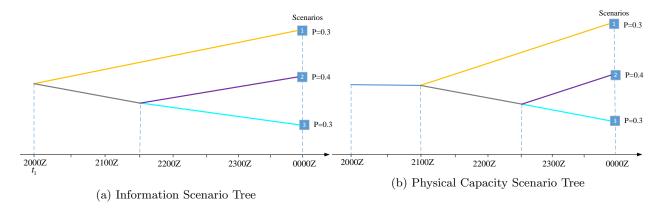


Figure 3: Effect of Short-term Weather Forecast

Assuming that the one-hour weather forecast is exactly accurate, then the scenario tree about the weather information (Figure 3a) is the left translation of the scenario tree for actual physical capacity (Figure 3b). For example, at 2100Z or t_0 , which is the beginning of stage 1, we only need to impose

$$X_{1\,t,\,t'}^{r,2} = X_{1\,t,\,t'}^{r,3} \quad \forall r, t' \tag{22}$$

Even though physical capacities are still $M_{t_1,1}^r = M_{t_1,2}^r = M_{t_1,3}^r$. Compared with the original implementation, because of the new short-term weather forecast, we have less restrictive nonanticipativity constraints and will have lower system delay costs.

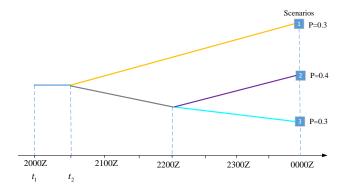


Figure 4: Effect of Minimum Notification Time and Short-term Weather Forecast

2. Minimum Notification Time

Assume we need to let each flight know the ground delay decision 30 minutes in advance before scheduled departure time. The effect of this minimum notification time requirement is that it will translate the scenario tree to the right (Figure 4). For flights scheduled to depart at 2000Z (t_1) , we can only use the weather information at 1930Z, and at that time we do not yet know whether scenario 1 will happen or not. For flights departing at 2030Z (t_2) , we can utilize the weather information at 2000Z. Therefore at t_1 and t_2 we have

$$X_{1,t_1,t'}^{r,1} = X_{1,t_1,t'}^{r,2} = X_{1,t_1,t'}^{r,3} \quad \forall r, t'$$

$$X_{2,t_2,t'}^{r,2} = X_{2,t_2,t'}^{r,3} \quad \forall r, t'$$
(23)

It can be seen that the longer the minimum notification time, the less information we can use and more restrictive nonanticipativity constraints will have to be enforced.

The results of this section can be similarly adopted for all semi-dynamic and dynamic models in Table 1.

V. Variance and Risk Issues in Air Traffic Flow Management

Variance and risk are important topics in stochastic programming theory and application, especially in finance [21;22]. Compared with other fields, variance and risk issues receives less attention in air traffic flow management for several reasons: stochastic programming models are very often used to compare with deterministic models in a proof of concept to demonstrate their advantages in terms of reducing delay costs; by increasing the air cost coefficients, we could indirectly get a more conservative ground delay policy; we usually get scenarios from simulations, then do clustering and pick only a few representative scenarios so that it would be easy for the analysts to check. Due to the difficulty to get realistic scenario data, it is also common for researchers to make up the scenario data. With only a small number of scenarios, it is not very necessary to study the tail behavior of the cost distribution.

In the case of CTOP, which has multiple constrained resources, there can be many more weather scenarios. With the development of ensemble weather forecast techniques and decision support tools, more scenarios are desirable to capture the edge cases. We believe now it is meaningful to directly use the variance and risk management techniques in air traffic flow management.

Air traffic flow management models have similar objective function structures: the summation of ground delay and air delay (and reroute delay). Here we will use ESOM model as the example. The goal of this section is to show how we can apply some of the classical variance and risk measures to air traffic flow management problems.

A. Variance Measures

1. Markowitz Mean-variance Approach

In ESOM, the objective is to minimize ground delay and air delay cost. Once the planned acceptance rates are set, the G_t can be determined. A_t are random variables depending on the realization of weather scenarios.

Denote r_q as the recourse cost for scenario q, that is $r_q = c_a \sum_{t=1}^T A_{t,q}$. Mean recourse cost $\bar{r} = \sum_{q=1}^Q p_q r_q$. The variance-sensitive cost is

$$\min \sum_{t=1}^{T} c_g G_t + \mathbf{E}\{r_q\} + \gamma \mathbf{var}\{r_q\} = \sum_{t=1}^{T} c_g G_t + \sum_{q=1}^{Q} p_q r_q + \gamma \sum_{q=1}^{Q} p_q (r_q - \sum_{q=1}^{Q} p_q r_q)^2$$
(24)

Where γ is the risk-aversion parameter. Compared with the original objective function, we now have an additional quadratic term. In ESOM, we do not impose integrality constraints on the decision variables, therefore we can easily solve this Quadratic Programming problem. However, for general air traffic flow management problems, we do require the decision variables to be integers. Therefore we want to use variance and risk measures that can keep the mathematical formulation in mixed integer linear form.

2. Maximum Absolute Deviation

The deviation of the recourse cost of the q-th scenario from the mean recourse cost is $d_q = r_q - \bar{r}$. Let $D = \max_q |d_q|$ denote the maximum deviation. We can add a penalty term to the objective function

min
$$\sum_{t=1}^{T} c_g G_t + \sum_{q=1}^{Q} p_q r_q + \gamma D$$

$$|r_q - \bar{r}| \le D \quad q = 1, \dots, Q$$
other constraints
$$(25)$$

By adjusting γ , we can make tradeoffs between maximum deviation and mean delay cost.

In [23], instead of using maximum absolute deviation as a soft constraint as shown above, the authors impose it as a hard constraint:

min
$$\sum_{t=1}^{T} c_g G_t + \sum_{q=1}^{Q} p_q r_q$$
$$|r_q - \bar{r}| \le \alpha D$$
other constraints

where D is now a parameter and is the maximum absolute deviation of risk-neutral solution. By adjusting $\alpha, 0 \le \alpha \le 1$, a Pareto-optimal curve can be drawn where the x and y axes are maximum deviation and mean delay cost, respectively.

3. Mean Absolute Deviation

A similar dispersion metric is the mean absolute deviation, which is defined as $\mathbb{E}\{|r_q - \bar{r}|\} = \sum_{q=1}^{Q} p_q(|r_q - \bar{r}|)$.

The variance-sensitive formulation is

min
$$\sum_{t=1}^{T} c_g G_t + \sum_{q=1}^{Q} p_q r_q + \gamma \sum_{q=1}^{Q} p_q d_q$$
$$|r_q - \bar{r}| = d_q \quad q = 1, \dots, Q$$
other constraints (26)

We can choose to only penalize the derivation that is above the expectation by replacing constraints $|r_q - \bar{r}| = d_q$ with

$$d_q \ge r_q - \bar{r} \tag{27}$$

B. Risk Measures

1. Minimize Worst Realization

A direct idea to control the risk is to minimize the cost of the worst realization. We also want the expected cost below a certain level μ_0 . The Minimax model can be easily formulated as

min
$$y$$

$$r_q \leq y \quad q = 1, \dots, Q$$

$$\sum_{t=1}^{T} c_g G_t + \sum_{q=1}^{Q} p_q r_q \leq \mu_0$$
 Other constraints

The problem with such a such model is that sometimes focusing only on the worst case scenario may result in poor expected value.

2. Value at Risk (VaR)

An extension to worst case analysis is quantile analysis. Value-at-Risk (VaR) is the greatest loss that can occur with a given probability α . Usually α takes values like 0.95 or 0.99.

$$VaR_{\alpha}(\sum_{t=1}^{T} c_{g}G_{t} + r_{q}) = \inf\{l | \mathbb{P}(\sum_{t=1}^{T} c_{g}G_{t} + r_{q} > l) \le 1 - \alpha\} = \inf\{l | \mathbb{P}(\sum_{t=1}^{T} c_{g}G_{t} + r_{q} \le l) \ge \alpha\}$$
 (28)

We could formulate a MILP model to get VaR_{\alpha}

min
$$y$$

$$\sum_{t=1}^{T} c_g G_t + r_q \le y + M z_q \quad q = 1, \dots, Q$$

$$\sum_{q=1}^{Q} p_q z_q \le 1 - \alpha - \epsilon, \quad z_t \in \{0, 1\}$$

$$\sum_{t=1}^{T} c_g G_t + \sum_{i=1}^{Q} p_q r_q \le \mu_0$$
Other constraints

where M is a sufficiently large constant and ϵ is an arbitrarily small positive constant. Binary variable z_q takes value 1 whenever $r_q > y$. The second constraint ensures that the total probability $r_q > y$ is less than $1 - \alpha$. The use of ϵ is just a convention when probabilities are discrete.

3. Conditional Value at Risk (CVaR)

Conditional Value at Risk (CVaR) quantifies the expected losses that occur beyond the VaR and was introduced to address the shortcomings of VaR^[24]. In the finite number of scenarios case, the CVaR measure is well defined by the following optimization problem:

$$CVaR_{\alpha} = \max_{u_t} \left\{ \frac{1}{1-\alpha} \sum_{q=1}^{Q} \left(\sum_{t=1}^{T} c_g G_t + r_q \right) u_q : \sum_{q=1}^{Q} u_q = 1 - \alpha, 0 \le u_q \le p_q \ \forall q \right\}$$
(30)

It can be easily seen that when $\alpha \to 1_-$, CVaR_{α} gives the worst case value, and when $\alpha = 0$, $\text{CVaR}_{\alpha} = \bar{r}_q$. When r_q have already been computed, the above formula can be used to obtain CVaR_{α} . When r_q are also

decision variables, we could use the dual form of (30)

$$\text{CVaR}_{\alpha} = \min_{\eta, e_q} \quad \eta + \frac{1}{1 - \alpha} \sum_{q=1}^{Q} p_q e_q$$

$$e_q + \eta \ge \sum_{t=1}^{T} c_g G_t + r_q \quad \forall q$$

$$e_q \ge 0$$

$$\sum_{t=1}^{T} c_g G_t + r_q \le \mu_0$$

$$\text{Other constraints}$$

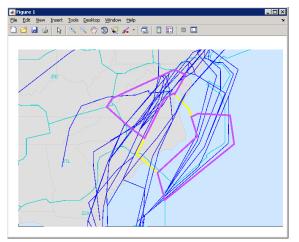
$$(31)$$

where η is the dual variable to the equality constraint and e_q is the dual variable to $u_q \leq p_q$.

The study of variance and risk is a relatively mature research area in stochastic programming. The measures we listed in this section are just a few representative ones. We wanted to demonstrate that it is straightforward to incorporate a linear form of variance and risk terms in air traffic management problem formulation.

VI. Experimental Results

To compare the performance of the proposed semi-dynamic model with ESOM and PCA models, we continue using the test case with convective weather activity in southern Washington Center (ZDC) and EWR airport. We assume there is a four-hour capacity reduction in ZDC/EWR from 2000Z to 2359Z. By analyzing the traffic trajectory and weather data, we can build the FCA-PCA network. In this use case, each FCA directly lies atop of the corresponding PCA, so the four FCAs are omitted in Figure 6.



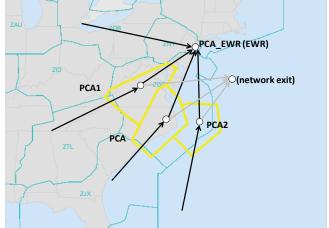


Figure 5: Traffic Trajectory Visualization

Figure 6: Geographical display of a FCA-PCA network

A. Capacity Profile and Traffic Demand

In this work, we directly manipulate the capacity profiles from the base forecast to create alternative capacity profiles. This gives us full control over the capacity data for experimental purposes. The capacity information is shown in Table 3. We can see that in scenario 1 at 2100Z PCA1's 15-minute capacity changes from 44 to 50, the EWR's capacity changes from 8 to 10; in scenario 2 at 2230Z, the capacities of PCA1 and EWR return to the nominal values. These two changes correspond to the two branch points in the scenario tree shown in Figure 2.

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

	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
SCENZ	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 3: Capacity Scenarios

		nd Delay F Scenario			Holding Pe Scenario			Total Cost Scenario		Expected Cost	Running Time
	SCEN1	SCEN2	SCEN3	SCEN1	SCEN2	SCEN3	SCEN1	SCEN2	SCEN3		Seconds
ESOM Semi-Dynamic ESOM	281.07 87.57	281.07 281.07	281.07 478.67	0	0	197.60 0	281.07 87.57	281.07 281.07	395.19 478.67	399.63 282.30	0.025 0.130
Two-Stage PCA Model ^[10] Semi-Dynamic PCA Model ^[10] Dynamic PCA Model ^[10]	280 156 116	280 280 280	280 403 463	0 0 0	0 0 0	190 67 7	280.0 156.0 116.0	280.0 280.0 280.0	660.0 537.0 477.0	394.0 319.9 289.9	0.109 0.612 1.480

Table 4: FCA-PCA vs. PCA Models Stochastic Solutions Comparision $(c_a/c_g=2)$

time period depending on the topology of the FCA-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). For any time periods outside the CTOP start-end time, e.g. the extra four time periods in Table 3, we assumed nominal capacity.

We use historical flight data for traffic demand modeling. We only keep 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 traverse the PCAs in their active periods and 130 flights land at EWR airport.

B. Model Comparisons in terms of System Delay Costs

The optimization models were solved using Gurobi 7.5.2 on a laptop with 2.6 GHz processors and 16 GB RAM. The main results are listed in Table 4. There are some key observations from this table:

- 1. Even though we did not enforce integrality constraints on decision variables in ESOM, still the objective value of ESOM is larger than the corresponding two-stage PCA model, in which all the variables are required to be integer. This unexpected result is exactly due to items 2 and 4 discussed in section III B. In this use case, EWR is a rather constrained resource. If the fraction of traffic from enroute PCAs to WER in t+1 is larger than in time period t, and some flights are delayed from t to t+1. The effect is that we create some flights to EWR out of nothing. Actually after we solve ESOM we know the number of flights that land at EWR is 132.08, which is larger than the scheduled number 130!
- 2. The objective value of semi-dynamic ESOM is smaller than semi-dynamic PCA model. Even though the result is the other way around, the underlying explanations are similar. If we delay some flights to a time period in which the fraction of traffic to the constrained resources is small, some of the demands essentially vanish. That is the one of the reasons we see a lot of ground delays in scenario 3 for semi-dynamic ESOM.
- 3. In this example, the more flexible the model is, the larger the difference between FCA-PCA model and PCA model. We think this is because a more flexible model will more heavily exploit the imperfections of the FCA-PCA model.

4. The computation times of ESOM and semi-dynamic ESOM models are very short, and semi-dynamic model performs better than ESOM in terms of total delay costs, which are all expected.

VII. Conclusions and Future Work

In this paper, we reviewed several CTOP rate planning models that had been previously presented. We classified them based on various properties and identified one missing piece from this "family". We then formulated and presented that model in this paper and conducted numerical tests on it. We pointed out a number of differences and connections among the various models, and highlighted some limitations inherent in all FCA-PCA models resulting from CDM-compatibility requirements.

We also reviewed concepts of variance and risk, and showed how these can be easily incorporated into air traffic flow management models using additional terms and constraints. This could be done with any of the CTOP rate planning models we presented.

Future work includes testing the idea of simulation-based optimization, comparing ESOM with other models, and showing the effect of incorporating variance and risk terms into CTOP models using realistic capacity data.

VIII. Acknowledgement

This research was partly funded under NASA Research Announcement contract #NNA16BD96C.

References

- ¹ Federal Aviation Administration, "Introduction to Collaborative Trajectory Options Program (CTOP)," https://cdm.fly.faa.gov/?page_id=983, Accessed: 2018-11-19.
- ² Vossen, T. W., Hoffman, R., and Mukherjee, A., "Air traffic flow management," *Quantitative problem solving methods in the airline industry*, Springer, 2012, pp. 385–453.
- ³ Ball, M. O., Hoffman, R., Odoni, A. R., and Rifkin, R., "A stochastic integer program with dual network structure and its application to the ground-holding problem," *Operations Research*, Vol. 51, No. 1, 2003, pp. 167–171.
- ⁴ Richetta, O. and Odoni, A. R., "Solving optimally the static ground-holding policy problem in air traffic control," *Transportation Science*, Vol. 27, No. 3, 1993, pp. 228–238.
- ⁵ Kotnyek, B. and Richetta, O., "Equitable models for the stochastic ground-holding problem under collaborative decision making," *Transportation Science*, Vol. 40, No. 2, 2006, pp. 133–146.
- ⁶ Richetta, O. and Odoni, A. R., "Dynamic solution to the ground-holding problem in air traffic control," Transportation Research Part A: Policy and Practice, Vol. 28, No. 3, 1994, pp. 167–185.
- ⁷ Mukherjee, A. and Hansen, M., "A dynamic stochastic model for the single airport ground holding problem," *Transportation Science*, Vol. 41, No. 4, 2007, pp. 444–456.
- ⁸ Zhu, G., Wei, P., Hoffman, R., and Hackney, B., "Saturation Technique for Optimizing Planned Acceptance Rates in Traffic Management Initiatives," The 21st IEEE International Conference on Intelligent Transportation Systems, Hawaii, USA, 2018.
- ⁹ Hoffman, R., Hackney, B., Zhu, G., and Wei, P., "Enhanced Stochastic Optimization Model (ESOM) for Setting Flow Rates in a Collaborative Trajectory Options Programs (CTOP)," AIAA Aviation, Atlanta, GA, June 2018.
- ¹⁰ Zhu, G., Wei, P., Hoffman, R., and Hackney, B., "Aggregate Multi-commodity Stochastic Models for Collaborative Trajectory Options Program (CTOP)," International Conference on Research in Air Transportation (ICRAT), Barcelona, Spain, June 2018.

- ¹¹ Zhu, G., Wei, P., Hoffman, R., and Hackney, B., "Centralized Disaggregate Stochastic Allocation Models for Collaborative Trajectory Options Program (CTOP)," 37th AIAA/IEEE Digital Avionics Systems Conference (DASC), London, UK, 2018.
- ¹² Metron Aviation, "CTOP Rate Selection Techniques Extended, Interim Results," Tech. rep., 2018.
- ¹³ Zhu, G. and Wei, P., "An Interval-based TOS Allocation Model for Collaborative Trajectory Options Program," AIAA Aviation, Atlanta, GA, June 2018.
- ¹⁴ CSC/TFMM-13/1744, "TFMS Functional Description, Appendix C: Traffic Management Initiative (TMI) Algorithms," Tech. rep., 2014.
- Hoffman, R., Burke, J., Lewis, T., Futer, A., and Ball, M., "Resource allocation principles for airspace flow control," AIAA Guidance, Navigation, and Control Conference and Exhibit, p. 6470.
- Liu, P.-c. B., Hansen, M., and Mukherjee, A., "Scenario-based air traffic flow management: From theory to practice," Transportation Research Part B: Methodological, Vol. 42, No. 7, 2008, pp. 685–702.
- Steiner, M., Bateman, R., Megenhardt, D., Liu, Y., Xu, M., Pocernich, M., and Krozel, J., "Translation of ensemble weather forecasts into probabilistic air traffic capacity impact," Air Traffic Control Quarterly, Vol. 18, No. 3, 2010, pp. 229–254.
- ¹⁸ Clarke, J.-P. B., Solak, S., Ren, L., and Vela, A. E., "Determining stochastic airspace capacity for air traffic flow management," *Transportation Science*, Vol. 47, No. 4, 2012, pp. 542–559.
- Matthews, M., DeLaura, R., Veillette, M., Venuti, J., and Kuchar, J., "Airspace Flow Rate Forecast Algorithms, Validation, and Implementation," Project report atc-428, MIT Lincoln Laboratory, Lexington, MA, 2015.
- Wei, P., Cao, Y., and Sun, D., "Total unimodularity and decomposition method for large-scale air traffic cell transmission model," *Transportation Research Part B: Methodological*, Vol. 53, 2013, pp. 1–16.
- ²¹ King, A. J. and Wallace, S. W., *Modeling with stochastic programming*, chap. Modeling the Objective Function, Springer Science & Business Media, 2012.
- ²² Mansini, R., odzimierz Ogryczak, W., and Speranza, M. G., Linear and mixed integer programming for portfolio optimization, Springer, 2015.
- ²³ Sengupta, P., Tandale, M. D., and Menon, P. K., "Risk-hedged traffic flow management under airspace capacity uncertainties," *Journal of Guidance, Control, and Dynamics*, 2014.
- Rockafellar, R. T., Uryasev, S., et al., "Optimization of conditional value-at-risk," Journal of risk, Vol. 2, 2000, pp. 21–42.

IX. Appendix for Acronyms and Abbreviations

TMI	Traffic Management Initiative
GDP	Ground Delay Program
AFP	Airspace Flow Program
FCA	Flow Constrained Area
PCA	Potential Constrained Area
PAR	Planned Acceptance Rate
CTOP	Collaborative Trajectory Options Program
TOS	Trajectory Options Set
FAA	Federal Aviation Administration
ESOM	Enhanced Stochastic Optimization Model