# Demonstrating Early-Adopter Benefits of Submitting Multiple Trajectory Options for Airlines

Gita S. Hodell[1]
*NASA Ames Research Center, Moffett Field, CA, 94035*

Connie L. Brasil[2], Nathan J. Buckley[3], Conrad Gabriel Jr.[4]
*San Jose State University/NASA Ames Research Center, Moffett Field, CA 94043*

Hyo-Sang Yoo[5]
*Universities Space Research Association/NASA Ames Research Center, Moffett Field, CA 94043*

Scott N. Kalush[6]
*Human Solutions Inc., Washington D.C., 20024*

and

Nancy M. Smith[7], Paul U. Lee[8]
*NASA Ames Research Center, Moffett Field, CA, 94035*

**A workshop at NASA Ames Research Center was held with airline industry stakeholders to demonstrate the impact of using Trajectory Options Sets (TOSs) during a Collaborative Trajectory Options Program (CTOP) for severe weather operations. The demonstration was conducted using a part-task Human-in-the-Loop (HITL) simulation of the Integrated Demand Management (IDM) concept, which is an air traffic management method that uses CTOP to deliver preconditioned traffic to the Time-Based Flow Management (TBFM) region. The demonstration addressed the following research objectives: first, determine who receives a greater benefit, TOS-participating or TOS-excluded airlines? Second, determine which method of trajectory selection yields a better solution, human/manual selection or automation? Finally, obtain feedback from stakeholders on their impression of the concept and recommendations for future work. The results showed that TOS-participating airlines received greater benefit in terms of total ground delay, ground delay savings, number of reroute options, and additional flight time, compared to TOS-excluded airlines. However, this result was dependent on the situational context, such as the number and location of flights. We found that all airlines benefitted when just a subset of airlines submitted TOS, but the greatest benefit went to the TOS-participating airlines. These benefits to the TOS-participating airlines were diminished as the number of TOS-participants in the system increased. Therefore, there was an "early-adopter" effect that suggested**

---

[1] Research Psychologist, Human Systems Integration Division, NASA ARC; *gita.s.hodell@nasa.gov*
[2] Senior Research Associate, Human Systems Integration Division, SJSU/NASA ARC
[3] Research Associate, Human Systems Integration Division, SJSU/NASA ARC
[4] Research Associate, Human Systems Integration Division, SJSU/NASA ARC
[5] Senior Engineer, Aerospace, Aviation Systems Division, USRA/NASA ARC
[6] Simulation Engineer, Human Systems Integration Division, HSI/NASA ARC
[7] Research Psychologist, Human Systems Integration Division, NASA ARC
[8] Research Engineer, Human Systems Integration Division, NASA ARC

**airlines could benefit by becoming the first to equip TOS without causing unfair disadvantages to those who do not equip TOS. In addition, we found that manual selection of trajectory options performed similarly to CTOP, but the CTOP solution was more efficient in terms of number of reroutes, additional flight time, and average ground delay for rerouted flights. Feedback from the stakeholders was solicited, and their overall impressions of the demonstrations were positive. They remarked that CTOP could do a better job than current day solutions, and they thought airlines could benefit from continuing to develop TOS capabilities.**

## I. Nomenclature

AAR = Airport Acceptance Rate
AC = Adjusted Cost
AFP = Airspace Flow Program
ANSP = Air Navigation Service Providers
ATCSCC = Air Traffic Control System Command Center
ARTCC = Air Route Traffic Control Center
CTOP = Collaborative Trajectory Option Program
EDCT = Expect Departure Clearance Time
ETA = Estimated Time of Arrival
EWR = Newark Liberty International Airport
FAA = Federal Aviation Administration
FCA = Flow Constrained Area
FET = Flow Evaluation Team
GDP = Ground Delay Program
HITL = Human-in-the-loop
IAT = Initial Arrival Time
IDM = Integrated Demand Management
MACS = Multi-Aircraft Control System (NASA ATC simulation platform)
MIT = Miles-in-Trail
NAS = National Airspace System
nCTOP = NASA CTOP emulation
NOAA = National Oceanic Atmospheric Administration
PGUI = Plainview Graphical User Interface (TBFM)
RTA = Required Time of Arrival
RTC = Relative Trajectory Cost
SME = Subject Matter Expert
STA = Scheduled Time of Arrival
TBFM = Time-Based Flow Management
TFM = Traffic Flow Management
TFMS = Traffic Flow Management System
TMI = Traffic Management Initiative
TOS = Trajectory Options Set
TRACON = Terminal Radar Approach Control
XM = Extended Metering (TBFM)

## II. Introduction

Integrated Demand Management (IDM), is a NASA developed Traffic Flow Management (TFM) concept that uses the Collaborative Trajectory Options Program (CTOP) to precondition traffic flows into the Time-Based Flow Management (TBFM) region, helping Air Navigation Service Providers (ANSPs) manage imbalances between demand and capacity in the National Airspace (NAS) [1]. CTOP is part of the Federal Aviation Administration's (FAA) Traffic Flow Management System (TFMS) automation. It can be used as a Traffic Management Initiative (TMI), like a Ground Delay Program (GDP) or an Airspace Flow Program (AFP) to regulate demand to a capacity-limited airport or Flow Constrained Area (FCA). Unlike other TMIs, CTOP is unique because it can also handle multiple parallel constraints within a single program by issuing Expect Departure Clearance Times (EDCTs), or predeparture reroutes via Trajectory Option Sets (TOSs). The TOS represents a list of alternate routes and the trigger conditions for selecting the routes that individual flights are willing to take if their currently filed route becomes inviable. Reasons for routes becoming inviable could be that convective weather has impacted a portion of the route, or there is a particular airspace resource that is scheduled to be over capacity. The amount of ground delay incurred by waiting until the constraint clears could be excessively high and costly to the airline, in which case a flight may be better off taking an alternate route. Airline operators coordinate TOSs in advance of departures with the Air Traffic Control System Command Center (ATCSCC) according to their business model.

Currently, procedures to generate sets of multiple trajectories do not exist for most airlines, and for airlines to submit TOS to CTOP they would need to be able to generate their own TOS. IDM works most effectively with TOS participation, but incorporating TOS capability into the airline business model is non-trivial. Although CTOP is an existing traffic management tool, it is not widely used by ANSPs so there is currently little impetus for airlines to invest in TOS capability. There are barriers to entry such as lack of software infrastructure, new changes to personnel roles and responsibilities, and cross-facility coordination procedures that all take time and financial resources to make TOS participation happen. Furthermore, during current day operations, dispatchers already utilize available reroutes in order to direct their fleet away from a weather constraint. However, this process is done manually by refiling multiple flight plans. It is unclear if using TOS would generate a better outcome than the strategies that are currently being used, so there needs to be more demonstrable evidence that TOS capability would benefit the airlines from a business perspective.

This paper will present the findings from a workshop held at NASA Ames Research Center in 2018 with members of the Flow Evaluation Team (FET), a sub-team of the joint government/industry Collaborative Decision Making working group, consisting of FAA and airline representatives. The purpose of this workshop was to explore the use of TOS during IDM operations from an airline perspective through a series of HITL simulations, and to collect feedback from the FET on the implications for their company operations through table-top discussions and question and answer sessions.

## III. Background

### A. Past IDM research

IDM was originally developed to address the issue of large, routine airport delays in the New York area. It leverages pre-existing NextGen technology to approach TFM in a novel way. Currently, ANSPs manage separation standards for aircraft with Miles-in-Trail (MIT), whereas IDM focuses on schedule-based (i.e. time-based) separation. Figure 1 presents a high-level overview of CTOP for IDM operations. The CTOP schedule is created during the strategic timeframe, hours in advance of departure for most flights. The CTOP schedules arrival times to the FCA boundary based on allotted capacity. After crossing into the TBFM region, which is 400 nautical miles from the New York TRACON, the TBFM schedule supersedes the CTOP schedule as flights are tactically managed all the way to the TRACON. IDM has been developed in multiple stages of concept validation, beginning with clear weather operations [1, 2] and following with severe weather operations [3]. The concept has demonstrated TFM system-wide benefits such as increased predictability and efficiency for flight operators, increased equity for short-haul flights by eliminating “double-delay” penalties and maximizing throughput at the airport.

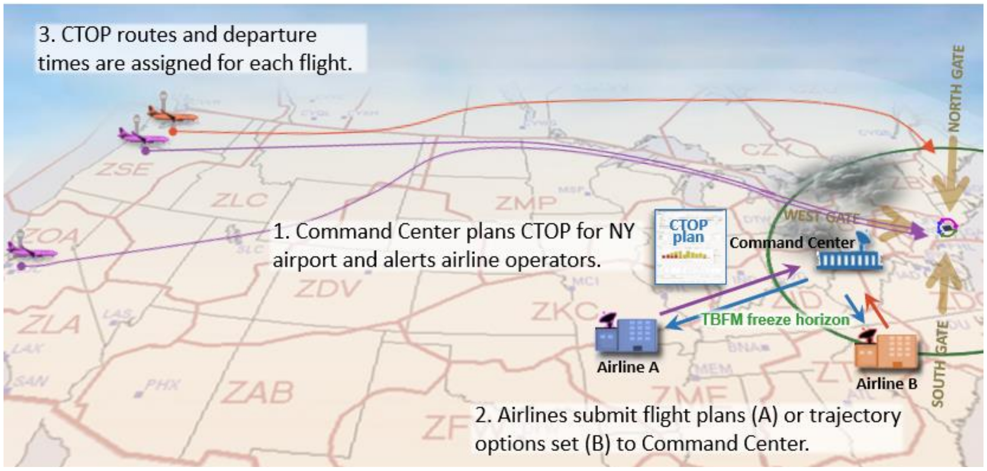

Fig 1. CTOP for IDM operation

To emulate CTOP operations, NASA developed nCTOP (NASA CTOP) software for simulation research purposes. During severe weather operations, nCTOP for IDM uses three FCAs co-located at the airport runway threshold (Figure 2) so the program can simultaneously manage both the upstream airspace and downstream airport constraints. To manage airspace capacity, the FCAs are subdivided into three flows, each one filtering for a single arrival flow (North, South, and West). The capacity on the West FCA is restricted due to weather, and flights scheduled on the West flow receive EDCTs or TOS reroutes to the North or South flows. To manage the runway capacity, the program includes all traffic flying to the airport, not just the flights that are affected by the weather. The combined capacity of the three FCAs cannot exceed the maximum capacity at the airport. Therefore, flights that are scheduled on the North or South flows can also receive EDCTs or TOS reroutes to ensure the target throughput rate is met.

**Fig 2. FCA placement for IDM**

Two nCTOP algorithms reduce traffic demand in situations of over-demand either through the en-route airspace or at the airport, the FCA Balancing Algorithm (FBA) and the TOS algorithm. The FBA, a NASA-developed capability, is a capacity setting automation enhancement for multiple FCAs which distributes capacity according to demand across multiple constraints. Currently, if a traffic planner wanted to set capacities on multiple FCAs in CTOP they would have to do so manually, and one at a time. The FBA for nCTOP allows the user to set a single maximum capacity for the combined FCAs. The FBA then distributes the slots automatically to the FCAs according to the proportional demand in each 15-minute time bin. Further details regarding the FBA can be found in [4].

The TOS algorithm incorporates the nCTOP FBA capacity settings to generate first-come-first-served slot-based assignments. Much like the current CTOP modeling capability, the TOS algorithm selects the sub-group of flights that would benefit from taking a TOS based on the user-submitted Relative Trajectory Cost (RTC) values. Actual RTC values are airline specific and calculated based on a number of factors. For IDM studies, RTC values are calculated as follows:

$$RTC = 1.5 \times \textit{flight time of alternate trajectory} - \textit{flight time of current trajectory}$$

Flights are only recommended for a TOS reroute if the adjusted cost (RTC + ground delay) for the original route is higher than the adjusted cost of a TOS reroute. Therefore, the flights that benefit from taking a TOS route are affected.

In an initial IDM concept validation study, the overall impact of TOS submission on the Traffic Flow system was examined [3]. Five simulation runs representing five levels of TOS participation by the airline carriers, a) 0%, b) 25%, c) 50%, d) 75%, and e) 100% were observed. The results suggested that any TOS submission above level 0% reduced total ground delay, and TOS submission rates of 50% or above resulted in the best overall system-wide outcome. In this study, the scenario was impartial to which specific airlines were TOS capable, but the reduction in ground delay was extended to all flights, including both those who took TOS reroutes and those who did not. However, flights who were TOS rerouted received less total ground delay than flights who were not, suggesting an "early-adopter" advantage which could be of interest to the airline industry. If an individual airline were the first in the field to equip TOS, there is evidence to suggest they would receive the greatest benefit in reduction of ground delay, which could give them an advantage over their competitors.

### B. The current study

One of the goals of the IDM project is to engage industry stakeholders early and throughout the development of the concept. The results from previous research suggested potential benefits for the system and the early adopters in a generic use case with controlled TOS participation levels, but when we presented the benefits to the airline industry stakeholders they had various concerns regarding the individual effects to the airlines rather than the global effects to the TFM system. First, if an airline were the sole TOS-participating carrier and others were not, would they be penalized by the system forcing them to give their slot to another airline? Second, if the reverse situation were true and an airline was the only TOS-excluded carrier, would they be penalized with excessive ground delay while other airlines got rerouted? Third, is using multiple trajectory options better than the current practice of refiling a flight plan from a constrained route to a preferred alternate route?

To address these concerns, a two day workshop with the FET was held at NASA Ames Research Center which consisted of part-task Human-in-the-Loop (HITL) simulations, and several structured discussions which addressed past research on the IDM concept, CTOP for IDM, and the implications for airlines incorporating TOS capability into their business model. During the discussions of the workshop, several simulation options using a LaGuardia (LGA) use case scenario were presented to the FET as potential ways to explore the issues of primary concern. They were interested in seeing what would happen if Delta, the largest carrier at LGA, was either the sole TOS-participating, or sole TOS-excluded carrier in the simulation. In addition, the FET wanted to simulate manually choosing their own reroutes and compare their solution with the CTOP generated solution. A series of runs were conducted in which either no airlines submitted TOS (Run 1: No TOS), all airlines submitted TOS (Run 2: All TOS), a subset of airlines submitted TOS (Run 3: Delta TOS and Run 4: American, JetBlue, Southwest, and United TOS), or FET participants assumed the roles of airline representatives and personally chose alternate flight plans from a list of pre-scripted routes (Run 5: FET picks route). Following each run, the FET participants were given airline specific data depicting how individual airlines were affected, and their reactions and feedback were collected. This paper will address the following research objectives:

1) Determine how the runs compare on performance metrics such as total ground delay, ground delay savings for those that took reroutes, additional flight time of those re-routes, and airport throughput.
2) Determine any "sole-carrier" effects, pertaining to the effects that could impact either the TFM system or other airlines due to a single carrier being the only one who is TOS-participating, or TOS-excluded.
3) Determine how this manual trajectory selection compares with the TOS algorithm performing the task through automation.
4) Elicit reactions from participants regarding the IDM concept, it's feasibility in today's world, and describe what could be improved or expanded upon to make the concept more robust.

## IV. Method

### A. Participants

Twelve members of the FET representing three airlines (American, Southwest, and United), three ANSP facilities (Cleveland Center, New York Center, and the ATCSCC), and one FAA Headquarters representative participated in this study. All participants were subject matter experts in the areas of airline industry, the FAA, or academic/technical expertise.

### B. Traffic scenario

An arrival scenario for LGA with traffic demand higher than the target airport capacity was used. This established the runway capacity as one of the constraints that needed to be satisfied. Furthermore, convective weather partially blocked the West flow and reduced capacity through Cleveland Center (ZOB), establishing airspace capacity in ZOB as a second constraint. Together, the two constraints of consistent over-demand at the runway and West gate provided a use case in which a CTOP TMI was appropriate.

The scenario consisted of 185 inbound flights, with the arrival demand at LGA between 40-43 flights per hour over four hours. The traffic distribution by airline included in the CTOP program, excluding exempt traffic, was Delta: 79 flights (56%), American: 34 flights (24%), JetBlue: 3 flights (2%), Southwest: 10 flights (7%), United: 4 flights (3%), Others: 12 flights (8%). The traffic formed three arrival flows. The South flow came through Washington Center (ZDC), the West flow came through Cleveland Center (ZOB), and the North flow came through Boston Center (ZBW). The scenario was generated from historical Center TRACON Automation System (CTAS) data from the NASA ATM warehouse [5].

A convective weather cell graphic was displayed over the ZOB airspace. The weather images were generated from radar data of a day in the NAS when weather affected the Northeast corridor (July 14th, 2015). The data was accessed via the National Centers for Environmental Information (NECI) database [6].

The CTOP included three FCAs co-located at the airport runway threshold. The airport capacity was set at 36 flights per hour. The maximum capacity on the West FCA was restricted to four flights per hour, and the North and South FCAs shared the remaining 32 slots. The summed capacity of the three FCAs could not exceed the airport maximum.

Of the 185 flights in the scenario, 138 were "TOS eligible," meaning they had more than one trajectory option for their approach to the airport. Pre-scripted TOS were used for the subset of flights which included up to three trajectory options (see Table 1). For example, a flight departing from Denver (KDEN) and scheduled to arrive through the West gate into the New York airspace could be rerouted before departure through the North or the South gate. Options were ranked by RTC values. The remainder of flights in the scenario either had no TOS options, were exempt from reroutes, or were excluded from the CTOP. Flights with no TOS options were required to stay on their original route and only accept EDCT if they were included in the CTOP. Exempt flights were either too close to their departure time, or already airborne. Excluded flights were scheduled outside of the time window when CTOP was active.

**Table 1. Pre-scripted Trajectory Options Sets**

| Callsign | FCA | TOS Option | Flight Plan |
|---|---|---|---|
| UAL556 | WEST | 1 | KDEN./.ZIRKL..MCK..LNK.J60.DJB..YNG..ETG.MIP4.KLGA |
| UAL556 | SOUTH | 2 | KDEN./.PER..RZC..ARG.J46.BNA.J42.BKW.J42.GVE.KORRY4.KLGA |
| UAL556 | NORTH | 3 | KDEN./.BRYCC..TAYOT..DAYYY..RUBKI..SIKBO..TULEG..RKA.HAARP3.KLGA |
| UAR4314 | WEST | 1 | KCLE./.FAILS..JFN..ETG.MIP4.KLGA |
| UAR4314 | NORTH | 2 | KCLE./.FAILS..ERI..JHW..MEMMS..WILET..RKA.HAARP3.KLGA |
| UAR5706 | WEST | 1 | KORD./.MOBLE..ADIME..GERBS.J146.ETG.MIP4.KLGA |
| UAR5706 | NORTH | 2 | KORD./.HANKK..EXTOL..RKA.HAARP3.KLGA |
| UAR5706 | SOUTH | 3 | KORD./.EARND..ELANR..EMMLY..ERECO..IIU.J526.BKW.J42.GVE.KORRY4.KLGA |
| UAR6256 | SOUTH | 1 | KIAD./.AGARD.KORRY4.KLGA |

### C. Tools

#### *1. MACS*

The Multi-Aircraft Control System (MACS) is software designed at NASA Ames Research Center to support high fidelity airspace operations research [7]. MACS provides a simulation environment in which HITL tests can be conducted. In this study, MACS was used to simulate a TFMS planner station with custom graphics displaying traffic, weather, and sector loadings.

#### *2. nCTOP*

'NASA' Collaborative Trajectory Options Program (nCTOP) is a NASA developed emulation of CTOP that has similar capabilities to the fielded version. In addition, nCTOP includes user interface enhancements that were added to aid researchers with information visualization, user-controlled automation support, modeling, and monitoring the TMI.

## D. Measures

Metrics were collected to illustrate the impact that different TOS submission levels had on the overall TFM system as well as the impact to individual airlines. To evaluate system and airline performance, summary statistics were collected to measure total ground delay, TOS associated flight time increase, number of TOS-rerouted flights, throughput at the runway, and ground delay savings. Ground delay savings is the metric calculated by nCTOP which equals the ground delay (minutes) of a TOS rerouted flight minus the ground delay that the flight would have incurred if it stayed on its original route. To evaluate the subjective opinions of participants on the overall concept of IDM operations, all live demonstrations and table-top discussions were recorded using digital audio recorders.

## E. Procedure

Prior to the demonstrations, participants were assigned role-playing ownership of an individual airline. The airlines represented in the simulation were Delta, American, JetBlue, Southwest, and United. If a participant was already professionally affiliated with an airline, they were assigned to that airline during the simulation, as was the case for American, Southwest, and United airlines. Otherwise ownership was randomly assigned.

No TOS (Run 1) and All TOS (Run 2) conditions were completed in advance of the live demonstrations. The No TOS condition represented a baseline condition in which no TOS options were included in the program, and ground delay assigned through EDCT was the only method CTOP used to precondition traffic. The system-wide impact on traffic flows was intended to illustrate a hypothetical worst-case scenario, in which ground delay would be excessively high, and all flights would be forced to stay on their original route. In the All TOS (Run 2) condition, every flight that had available route options in the scenario submitted them to CTOP, and the program used both EDCT and TOS reroutes to precondition traffic. This condition represented a hypothetical best-case scenario, in which all airlines were equally equipped with TOS so the benefits could be shared by everyone. Both Run 1 and Run 2 were modeled after previous research [3] that used 0% and 100% TOS conditions to determine the benefits of different TOS submission levels. The ground delay, ground delay savings, additional flight time, and number of TOS rerouted flights for each airline were printed on paper and given to the participants to review and discuss with researchers in a group.

### *1. Demonstration 1: Sole-carrier analysis*

Following the review and discussion of Runs 1 and 2, a lead researcher briefed the participants on options for a live demonstration of scenarios, and they chose to compare Delta TOS (Run 3), and American, JetBlue, Southwest, and United TOS (Run 4). The Delta TOS (Run 3) condition was designed to show what would happen if a single carrier with the majority of flights (Delta: 56%) were the only TOS-participating airline. In this case, only Delta flights were given multiple route options, while all other airlines included in the CTOP received EDCTs. The American, JetBlue, Southwest, and United TOS (Run 4) condition was designed to show what would happen if Delta were the only TOS-excluded carrier. Therefore, everyone but Delta was given multiple route options, and Delta was assigned EDCTs.  Both Run 3 and Run 4 were designed to characterize the variation of impact on different airlines in a mixed-equipage scenario. After each run, participants were given their airline specific ground delay, ground delay savings, additional flight time, and TOS reroute information on printed paper. Participants reviewed and discussed the results with researchers in a group.

### *2. Demonstration 2: Automation vs. human selection of trajectory assignment*

After the conclusion of Runs 3 and 4, a lead researcher briefed participants on Run 5, the final live demonstration in which FET airline participants picked the routes for their own respective airlines. This run was intended to approximate current day operations. Given the severe weather problem that was presented, participants were asked to role-play as airline dispatchers by choosing which flights from their fleet they would like to reroute, and the route they should take around the weather. Researchers provided the participants with a list that contained the same set of routes that were used for TOS in the previous runs. Participants chose one flight plan option from the set of routes, given the prior knowledge of the weather problem and results from the previous runs. The simulation scenario was updated to reflect the participants reroute choices, and a GDP-like CTOP was used to issue EDCTs with no TOS options. Participants were given a print-out of the impact to their airline to review, and a group discussion of the results followed.

## V. Results

The purpose of this study was threefold. First, to demonstrate the impact to both the system and individual airlines by varying levels of TOS participation in CTOP. Second, to compare airline performance between automatic selection of predeparture reroutes, or human/manual selection by the airline representative. Third, to elicit feedback from the FET participants on the demonstration of the IDM concept. In Runs 1 – 4, the independent variable was the proportion of traffic that had TOS available to be used in CTOP. Four levels of the independent variable were used, Run 1: No TOS, 0%; Run 2: All TOS, 100%; Run 3: Delta TOS, 56%; Run 4: American, Southwest, JetBlue and United, 36%. In Run 5: FET picks route, the independent variable was the number of flights that were manually picked to receive a predeparture reroute from a pre-scripted set of routes (non-TOS). The dependent variables that were measured were airport throughput, CTOP assigned ground delay, ground delay savings from TOS reroute, additional flight time from that reroute, and the number of TOS reroutes awarded.

This section will discuss the part-task HITL results in the following four parts. First, the high-level system performance results from Runs 1 – 5 are summarized. Second, the analysis of a single, major airline as the sole carrier to either a) participate, or b) not participate in TOS reroutes is presented (Runs 1 – 4). Third, the analysis of predeparture reroute selection by either a) CTOP automation, or b) human participant is presented (Runs 2 and 5). The final section will present the recorded comments made by participants giving their opinions on the results.

### A. Summary of results by run

To observe overall system performance in terms of number of reroutes, total ground delay, flight time increase, ground delay savings, and airport throughput, a summary table was created (see Table 2). Run 1: No TOS had the largest ground delay (118.5 hours) and the lowest throughput (33) of all conditions. Run 2: All TOS had the second lowest total ground delay (44.6 hours), the second highest flight time increase (5 hours), and the lowest ground delay savings (11.9 hours). Run 3: Delta TOS had the lowest number of reroutes (16) and the least flight time increase (3.3 hours) of all conditions but had the second highest ground delay (52.9 hours). Run 4: American, JetBlue, Southwest, United TOS had a moderate amount of flight time increase (4 hours) and the highest ground delay savings (28.7 hours). Run 5: FET picks route had the largest number of reroutes (31), the least total ground delay (44 hours), and the largest flight time increase (6.8 hours). Run 1 underdelivered to the target airport throughput rate and Runs 2 – 5 met the target rate. Ground delay savings were not reported for Runs 1 and 5 because all flights in those runs had a single option flight plan, therefore no planes were TOS rerouted.

**Table 2. Summary of system performance results by run**

| Run | Ground delay (hours) | Flight time increase (hours) | Ground delay savings (hours) | TOS-rerouted flights | Throughput |
|---|---|---|---|---|---|
| 1. No TOS | 118.5 | -- | -- | 0 | 33 |
| 2. All TOS | 44.6 | 5 | 12.8 | 23 | 36 |
| 3. Delta TOS | 52.9 | 3.3 | 13.7 | 16 | 36 |
| 4. American, JetBlue, Southwest, United TOS | 48.5 | 4 | 28.7 | 19 | 36 |
| 5. FET picks route | 44 | 6.8 | -- | 31* | 36 |

*Note.* *Rerouted manually from TOS list, not from CTOP automation

As expected, Run 1 performed the worst in terms of ground delay and airport throughput. Any condition associated with either TOS or manual predeparture reroutes (Runs 2 – 5) resulted in better performance in terms of ground delay and throughput. However, there was a tradeoff between ground delay and flight time increase. For example, in the conditions with the fewest reroutes (Run 3 and 4), flight time increase was low but ground delay was higher when compared to conditions with more reroutes (Run 2 and 5). Conversely, Runs 2 and 5 had more reroutes than Runs 3 and 4, and hence larger flight time increases, but also less ground delay.

Run 4 stood out by having more than double the ground delay savings than the other runs. This is because of the proportion of traffic controlled by Delta airlines. In the traffic scenario, Delta accounted for 56% percent of the traffic, with 23% of Delta flights filed through the West gate. If those Delta flights arrived earlier in the schedule, the amount of potential ground delay that could accumulate on the West flow would be huge due to Delta being the only TOS-excluded airline in Run 4. Other airlines could leave the West flow for less saturated gates, and therefore saved more ground delay because the ground delay savings metric represents actual ground delay minus potential ground delay.

The difference between Run 2 and Run 5 was notable because they were the most similar in terms of the number of available reroutes in the scenario. Participants in Run 5 chose the reroutes based on the same set of alternate routes that were used for TOSs in Run 2. The ground delay in Run 2 (44.6 hours) and Run 5 (44 hours) was similar, however Run 5 had more reroutes (31) and more flight time increase (6.8 hours) than Run 2 (Reroutes: 23; flight time increase: 5 hours). This means that overall, the solution from Run 5 would have been more costly to airlines because airborne delay is more expensive than ground delay.

Target throughput at the airport was reached in all conditions with TOS or manual reroutes. This finding is consistent with previous research [3] in-which TOS-participation levels of 0% and 25% resulted in under-delivery to the airport rate, although the rate gradually increased as the participation volume increased, with maximum throughput being reached at 50% or above. In Run 3 there was 56% TOS-participation, and in Runs 2 and 5 there was 100% TOS or reroute participation of eligible flights. However, it is interesting that in Run 4 only 36% of eligible flights were TOS-participating, yet the target airport rate was reached, which means the number of TOS-participants in Run 4 was sufficient enough to make use of all available capacity. Therefore, there were no differences in throughput between any of the TOS or reroute associated conditions.

## B. Demonstration 1: Sole carrier analysis

### *1. Summary of Results by Airline*

To compare the performance of individual airlines when one dominant airline is either the sole carrier participating in TOS, or the sole carrier excluded from TOS, airline performance was measured in terms of ground delay, ground delay savings, number of reroutes, and additional flight time. Results from Run 1: No TOS, Run 2: All TOS, Run 3: Delta TOS, and Run 4: American, JetBlue, Southwest, and United TOS are compared in Table 3. All airlines in Run 1: No TOS experienced the highest levels of ground delay except for JetBlue. In Run 3: Delta TOS, Delta received the least ground delay (17.1 hours), the most reroutes (16), most additional flight time (3.3 hours), and most ground delay savings (13.7 hours) of all conditions. Likewise, in Run 4: American, JetBlue, Southwest, and United TOS, all TOS participating airlines except for JetBlue received the least amount of ground delay (American: 6.4 hours; Southwest: 1.9 hours, United: .9 hours), the most reroutes (American: 10; Southwest: 6, United: 3), most additional flight time (American: 1.9 hours; Southwest: 1.3 hours, United: .8 hours), and the most ground delay savings (American: 14.6 hours; Southwest: 9.1 hours, United: 14.6 hours), of all conditions. JetBlue did not have any available TOS routes in Run 4.

**Table 3. Summary of airline results for Runs 1 - 4**

| Run | Delta | American | JetBlue | Southwest | United |
|---|---|---|---|---|---|
| 1. No TOS | | | | | |
| Ground Delay (hours) | 48.3 | 33.6 | .5 | 21.1 | 14.4 |
| Number of reroutes | -- | -- | -- | -- | -- |
| Additional flight time (hours) | -- | -- | -- | -- | -- |
| Ground delay savings (hours) | -- | -- | -- | -- | -- |
| 2. All TOS | | | | | |
| Ground Delay (hours) | 23.7 | 10.7 | 1.7 | 3.2 | 1.2 |
| Number of reroutes | 10 | 6 | 0 | 4 | 2 |
| Additional flight time (hours) | 2 | 1.2 | 0 | 1 | .6 |
| Ground delay savings (hours) | 4.5 | 3.3 | 0 | 2.7 | 3.3 |
| 3. Delta TOS | | | | | |
| Ground Delay (hours) | 17.1 | 16.9 | 1.1 | 8.4 | 5 |
| Number of reroutes | 16 | -- | -- | -- | -- |
| Additional flight time (hours) | 3.3 | -- | -- | -- | -- |
| Ground delay savings (hours) | 13.7 | -- | -- | -- | -- |
| 4. American, JetBlue, Southwest, United TOS | | | | | |
| Ground Delay (hours) | 34.4 | 6.4 | 1.3 | 1.9 | .9 |
| Number of reroutes | -- | 10 | 0 | 6 | 3 |

|  |  |  |  |  |  |
|---|---|---|---|---|---|
| Additional flight time (hours) | -- | 1.9 | 0 | 1.3 | .8 |
| Ground delay savings (hours) |  | 14.6 | 0 | 9.1 | 4.9 |

A general trend was that airlines with or without TOS rerouting capabilities in Runs 2 - 4 performed better than the baseline Run 1: No TOS condition. Furthermore, airlines with TOS capabilities, i.e., Delta in Run 3, and all airlines except Delta in Run 4, performed better than they did in Run 2: All TOS. This is consistent with the expectation that the TOS algorithm selects the best option for the flight and will reroute the flight if the ground delay on the original route becomes too high. Fewer TOS participants in the system create more opportunities for TOS capable flights to get their preferred reroute. The other airlines that would normally compete for the preferred slots are held back by ground delay, as shown by non-TOS participating airlines American, Southwest, and United in Run 3, and Delta in Run 4, performing worse than Run 2. The exception to this trend was JetBlue, who experienced higher ground delay in Runs 2 - 4 because all their flights were scheduled through the South flow with no options to reroute into either the West or North flow. As flights rerouted out of the West into the South flow, delays for flights scheduled on the South flow increased. Therefore, the results for JetBlue were scenario specific for this simulation.

We observed the sole TOS-participating carrier (Delta, Run 3) received the greatest benefit in terms of the lowest overall ground delay and highest ground delay savings of all runs. This finding confirms that airlines were not penalized for being TOS equipped if others did not. Finally, we observed that the sole TOS-excluded carrier (Delta, Run 4) received higher ground delay than they would have if they were TOS-participating (Run 3), but still lower than the baseline of Run 1. Therefore, TOS-participating airlines benefit the most, but TOS-excluded airlines can also benefit from the overall system delay reduction.

### *2. Flights by Assigned FCA*

To detect any differences between conditions in the traffic load on each arrival flow, a summary table of FCA assignments was created (see Table 4). Of all the runs, Run 1: No TOS had the largest number of flights on the West flow (37), and the fewest on the North (28) and South (77) flows. Run 3: Delta TOS and Run 4: American, JetBlue, Southwest, and United TOS had very similar traffic distributions to each other. Compared to Run 1, Runs 3 and 4 had a lower number of flights on the West flow (Run 3: 21, Run 4: 18). Run 2: All TOS had the fewest number of flights on the West flow (14), and the highest on the North (40) and South (88) flows.

**Table 4. Number of flights assigned to each FCA**

| Run | Total North | North reroute | Total South | South reroute | Total West |
|---|---|---|---|---|---|
| 1. No TOS | 28 |  | 77 |  | 37 |
| 2. All TOS | 40 | (12) | 88 | (11) | 14 |
| 3. Delta TOS | 37 | (10) | 82 | (6) | 21 |
| 4. American, JetBlue, Southwest, United TOS | 38 | (10) | 85 | (9) | 18 |

*Notes.* ( ) TOS rerouted from the West flow

In Run 1, flights impacted by the weather on the West flow were not offloaded to an alternate gate, so the West flow was more susceptible to delays. Runs 3 and 4 utilized TOS options to offload traffic from the West flow, therefore traffic increased in the North and South. Runs 3 and 4 also shared similar patterns of reroutes, with both runs allocating more reroutes to the North flow than the South. Run 2 demonstrated the largest reduction in traffic on the West flow due to the high number of available TOS options.

Overall, Runs 3 and 4 had similar distributions of flights assigned to each flow, and there were no important differences between the two runs. This demonstrated that a single major carrier either being TOS-participating or TOS-excluded did not have a strong effect on the way CTOP distributed traffic. The more TOS options that were available in the system, the more traffic on the West decreased, which is why Run 2 showed the biggest impact of offloaded traffic.

### *3. Delay averages by FCA*

#### *3.1 Average ground delay and flight time increase by FCA*

To determine the differences between conditions in average ground delay and average additional flight time per flow across all flights in the scenario, a summary table was created (see Table 5). In terms of ground delay, Run 1: No TOS resulted in the highest average delay on the West flow (Mean=179 minutes) and the most unequal distribution of delay (Range=178 minutes). Run 3: Delta TOS only and Run 4: American, JetBlue, Southwest, and United TOS only showed a reduction in delay on the West when compared to Run 1 (Run 3: 70 minutes; Run 4: 79 minutes, and an increase in equal distribution of delays (Run 3: Range = 54 minutes, Run 4: Range = 61 minutes). Run 2 had the lowest average delay on the West flow (Mean=30 minutes) and the most equal distribution of delay (Range=7 minutes). In terms of average flight time increase, all TOS rerouted flights had a flight time increase of 15 minutes or less for all runs, and no meaningful differences were detected between Runs 2, 3, or 4.

**Table 5. Average ground delay and flight time increase by FCA**

| Run | Average delay (minutes) | | | | Average flight time increase (minutes) | |
|---|---|---|---|---|---|---|
| | South | West | North | Delay Range | South from West | North from West |
| 1. No TOS | 6 | 179 | 1 | 178 | -- | -- |
| 2. All TOS | 25 | 30 | 32 | 7 | 14 | 12 |
| 3. Delta TOS | 16 | 70 | 33 | 54 | 15 | 11 |
| 4. American, JetBlue, Southwest, United TOS | 18 | 79 | 27 | 61 | 12 | 13 |

Runs 2 – 4 demonstrated that as the number of available TOS options increased the amount of ground delay on the West flow decreased, and delay was distributed more evenly between flows. There were marginal increases in delay in the South and North as more flights rerouted to those flows. Runs 3 and 4 performed similarly to each other, but not as well as Run 2 in terms of average ground delay distribution per flow. Overall, the number of available TOS routes was the biggest factor in determining ground delay distribution.

The different conditions did not seem to influence average flight time increases, and all additional flight time was considered acceptable by current industry standards. This is because of the way that each trajectory option was weighted by the TOS algorithm. The automation did not assign a route if the cost of the reroute was higher than the cost of the ground delay.

### *3.2 Average ground delay for flights that remain on original route vs. TOS rerouted flights by FCA*

To distinguish any differences between the average amount of ground delay assigned to flights that stayed on their original route and flights that took a TOS reroute, a summary table for Runs 2 – 4 only was computed (see Table 6). Run 4: American, JetBlue, Southwest, and United TOS only had the most similar delays between original route and TOS reroute. Between the original South route (9 minutes) and the South TOS route (6 minutes), there was a difference of three minutes, and between the original North route (13 minutes) and the North TOS route (11 minutes), there was a difference of only two minutes. By comparison, Run 2: All TOS had larger differences in delays. There was an eight minute difference between the original South route (15 minutes) and the South TOS route (8 minutes), and an eleven minute difference between the original North route (19) and the North TOS route (8 minutes).

**Table 6. Average ground delay for flights that remain on original route vs. TOS rerouted flights by FCA**

| Run | Original route (minutes) | | | TOS rerouted (minutes) | |
|---|---|---|---|---|---|
| | South | West | North | South from West | North from West |
| 2. All TOS | 15 | 13 | 19 | 7 | 8 |
| 3. Delta TOS | 10 | 40 | 17 | 4 | 10 |
| 4. American, JetBlue, Southwest, United TOS | 9 | 38 | 13 | 6 | 11 |

*Note.* There were no flights that rerouted into the West flow from another flow

In all runs, flights that stayed on their original route took more ground delay than flights that were TOS rerouted, but the differences were modest. Of the flights who took TOS reroutes, there was low variability across Runs 2 – 4 (Minimum=4 minutes, Maximum=11 minutes, Range= 7 minutes), indicating the benefit of lower ground delay in TOS rerouted flights was stable across different conditions.

### *4. Ground delay levels of acceptability*

To determine the differences between each condition in the amount of CTOP assigned ground delay, the length of ground delay was categorized into decreasing levels of acceptability, ranging from “on-time” (most acceptable) to “two or more hours” (unacceptable) (see Table 7). The total number of flights that fell into each range were calculated. We observed that Run 1: No TOS performed the best in terms of total number of flights that were scheduled to be on-time (108), but also performed the worst in terms of number of flights that received two or more hours of ground delay (25). Run 2: All TOS had the fewest on-time flights (79), and the lowest amount of unacceptable delay (2+ hours: 0). Run 3: Delta TOS (on-time: 92; 20-60 minutes: 34; 1-2 hours: 10; 2+ hours: 3) and Run 4: American, JetBlue, Southwest, and United TOS (on-time: 94; 20-60 minutes: 32; 1-2 hours: 15; 2+ hours: 0), showed similar downward trends as each other where the amount of ground delay decreased by each category.

**Table 7. Ground delay by levels of acceptability**

| Run | on-time | 20-60 minutes | 1-2 hours | 2+ hours |
|---|---|---|---|---|
| | Most acceptable | Acceptable | Less acceptable | Unacceptable |
| 1. No TOS | 108 | 3 | 6 | 25 |
| 2. All TOS | 79 | 62 | 1 | 0 |
| 3. Delta TOS | 92 | 34 | 10 | 3 |
| 4. American, JetBlue, Southwest, United TOS | 94 | 32 | 15 | 0 |

In Run 1, the restriction of traffic on the West flow resulted in the schedule under-delivering to the airport which is why so many flights from the North or South could fly on-time into the essentially unrestricted airport. However, the West flow was strongly burdened by unacceptably large delays (2+ hours), making the delay distribution inequitable between flows. Since delays were more evenly distributed between flows in Runs 3 and 4, the overall ground delay was reduced making it possible for a larger number of flights to receive between 0 – 60 minutes of delay. However, a moderate number of flights incurred unacceptable delays of one hour or more. Runs 3 and 4 performed similarly to each other in the amount of ground delay by category, and both performed better than Run 1. Run 2 had the fewest number of on-time flights, but there were more flights with 20-60 minutes of delay than any other run. Therefore, the majority of flights in Run 2 received an acceptable amount of ground delay except for the single flight that received 1 – 2 hours. It may be preferable to an airline to have the fewest number of flights possible with one hour or more of delay, in which case Run 2 resulted in the best outcome. However, it may be preferable to maximize the number of on-time flights while sacrificing a few flights to take larger delays, in which case Runs 3 and 4 had better outcomes.

The comparison of Runs 3 and 4 showed that a sole carrier either TOS-participating or TOS-excluded did not make a meaningful impact on ground delay acceptability. This was demonstrated by Runs 3 and 4 performing similarly to each other. Since both Runs 3 and 4 performed better than Run 1, and Run 2 performed the best overall, as the number of TOS-participating airlines increased the level of ground delay acceptability also increased. The evidence showed that the most important factor in determining the level of ground delay acceptability was the number of TOS participating airlines in the scenario.

## C. Demonstration 2: Automation vs. human selection of trajectory assignment

To determine whether there were differences in system performance or airline performance when automation vs. a human determined which aircraft were rerouted and which routes were assigned, the results from Run 2: All

TOS and Run 5: FET picks route were compared. The following section presents results that were analyzed in terms of ground delay, number of reroutes, additional flight times, and FCA assignments.

*1. Summary of Results by Airline*

The airline specific results are summarized in Table 8. In Run 5: FET picks route, Delta had lower ground delay (19.7 hours), higher reroutes (15), and higher flight time (2.7 hours), than in Run 2: All TOS (ground delay=23.7 hours; reroutes=10; flight time=2 hours).

American had higher ground delay (13.5 hours), higher reroutes (10), and higher additional flight time (2.5 hours) in Run 5 than in Run 2 (ground delay=10.7 hours; reroutes=6; flight time=1.2 hours). In Run 5 both Southwest and United had higher ground delays (Southwest=4.6 hours; United=2.1 hours) than in Run 2 (Southwest=3.2 hours; United=1.2 hours) despite having similar amounts of reroutes and additional flight time between both runs. JetBlue did not have any reroute options in either run.

**Table 8. Summary of airline results for all conditions**

| Run | | Delta | American | JetBlue | Southwest | United |
|---|---|---|---|---|---|---|
| 2. All TOS | | | | | | |
| | Ground Delay (hours) | 23.7 | 10.7 | 1.7 | 3.2 | 1.2 |
| | Number of reroutes | 10 | 6 | 0 | 4 | 2 |
| | Additional flight time (hours) | 2 | 1.2 | 0 | 1 | .6 |
| 5. FET picks route | | | | | | |
| | Ground Delay (hours) | 19.7 | 13.5 | 1.5 | 4.6 | 2.1 |
| | Number of reroutes | 15 | 10 | 0 | 4 | 2 |
| | Additional flight time (hours) | 2.7 | 2.5 | 0 | .9 | .6 |

In terms of ground delay, Run 5 performed slightly better than Run 2, but the difference was minimal. One of the important differences between the two runs is the number of reroutes. In Run 5, participants in charge of Delta and American chose to manually reroute more flights than the automation chose in Run 2. This was a contributing factor to the increased flight time that those airlines also saw in Run 5. Increased flight time is a necessary part of rerouting flights, but because airborne delay is more costly than ground delay, many industry professionals would prefer to take delay on the ground when possible. The TOS algorithm automation in Run 2 was expected to reroute the fewest possible flights, and only those flights that would benefit by prioritizing ground delay over additional flight time. Therefore, Run 2 generally had fewer flights rerouted and lower additional flight time than Run 5.

*2. Flights by Assigned FCA*

To assess the number of flights assigned to each flow a summary table was created (Table 9). In Run 2: All TOS the number of flights redirected from the West were split nearly evenly to the North (12) and the South (11) which caused this condition to have the highest number of flights assigned to the South flow (88). By comparison, in Run 5: FET picks route, the majority of reroutes from West flow were directed to the North flow (West to North: 22; South to West: 5). In addition Run 5 was the only condition in which flights were routed out of the South flow to the North (South to North: 3). This resulted in Run 5 having the highest total number of reroutes and the most flights assigned to the North route.

**Table 9. Number of flights assigned to each FCA**

| Run | Total North | North reroute | Total South | South reroute | Total West |
|---|---|---|---|---|---|
| 2. All TOS | 40 | (12) | 88 | (11) | 14 |
| 5. FET picks route | 53 | (22), 3* | 78 | (5) | 10 |

*Notes.* ( ) TOS rerouted from the West flow, *TOS rerouted from the South flow.

Compared to the CTOP automation in Run 2, Run 5 had a higher number of reroutes. The FET participants who were asked to pick the routes assumed that the South flow was saturated because it was the heaviest flow, so many of the reroutes went to the North flow. Since participants did not have access to actual demand statistics, they explained that predicting what other airlines would do was a factor in their decision making. In contrast, the automation used actual demand to calculate the fastest route with the least ground delay associated with it and ended up moving far fewer planes.

## *3. Delay averages by FCA*

### *3.1 Average ground delay and flight time increase by FCA*

The average ground delay and flight time increase separated by assigned FCA is summarized in Table 10. Run 5: FET picks route had higher average ground delay on the West (Mean=50 minutes) than Run 2: All TOS (Mean=30 minutes). Run 5: FET picks route had lower average ground delay on the North (Mean=12) than Run 2: All TOS (Mean=32). Run 2: All TOS had a more even distribution of ground delay between the three flows (Range=7 minutes) than Run 5: FET picks route (Range=38 minutes). There were no meaningful differences between the conditions in terms of average flight time increase.

**Table 10. Average ground delay and flight time increase by FCA**

| | Average delay (minutes) | | | | Average flight time increase (minutes) | | | |
|---|---|---|---|---|---|---|---|---|
| | South | West | North | Delay range | North | South | South from West | North from West |
| 2. All TOS | 25 | 30 | 32 | 7 | -- | -- | 14 | 12 |
| 5. FET picks route | 19 | 50 | 12 | 38 | 14* | 11* | -- | -- |

Note: * Denotes flights that were manually rerouted

It is noteworthy that in terms of total ground delay, Run 2 (44.6 hours) and Run 5 (44 hours) were similar, yet Run 5 had less equitable average delay between flows. Furthermore, despite their similarities in terms of average flight time increase by flow, Run 5 ultimately had more total flight time increase (6.8 hours) than Run 2 (5 hours). This means that the routes given to flights in Run 5 were more expensive to fly, and the West flow was more heavily impacted by delays than Run 2. Despite both conditions having the same number of predeparture reroute options available to them, results indicate the automation in Run 2 found a more optimal solution to the problem.

### *3.2 Average ground delay for flights that remain on original route vs. TOS rerouted flights by FCA*

The average ground delay assigned to flights who remained on their original route and flights that took either a TOS reroute (in Run 2) or a manual reroute (in Run 5) was computed in Table 11. Run 2: All TOS had higher average delay for flights that stayed on their original South route (Mean=15 minutes) and North route (Mean=19 minutes) than flights who took a South TOS (Mean=7 minutes) or a North TOS (Mean=8 minutes). Conversely, Run 5: FET picks route had higher average delays for flights that took TOS routes to the South (Mean=19) and North (Mean=16) compared to flights that stayed on their original South route (Mean=17 minutes) or North route (Mean=7 minutes).

**Table 11. Average ground delay for flights that remain on original route vs. TOS rerouted flights by FCA**

| Run | Original route (minutes) | | | TOS/Manual rerouted (minutes) | |
|---|---|---|---|---|---|
| | South | West | North | South | North |
| 2. All TOS | 15 | 13 | 19 | 7 | 8 |
| 5. FET picks route | 17 | 19 | 7 | 19* | 16* |

*Note.* * Denotes flights that were manually rerouted.

The differences between the two conditions in the amount of ground delay that was assigned to original routes vs. TOS reroutes were minimal. However, it is noteworthy that in previous runs (Runs 3: Delta TOS and Run 4:

American, JetBlue, Southwest, and United TOS) the TOS assigned route always received less ground delay than the original route, matching the pattern that was seen in Run 2. This makes Run 5 the only condition that did not demonstrate that pattern, as Run 5 experienced higher delay on the manually assigned TOS route than the original route. Therefore, contrary to previous runs, there was no additional benefit to manually rerouting flights in terms of reduced average ground delay. Possibly indicating that the wrong flights were moved at the wrong time. CTOP automation is more adept at picking the right route at the right time for the most equitable savings for the system.

### *4. Ground Delay Distribution*

Flights that were categorized into acceptable or unacceptable amounts of delay are reported in Table 12. Run 2: All TOS had a higher number of on-time flights (79) than Run 5: FET picks route (69). There were more flights with unacceptable amounts of delay in Run 5: FET picks route (1-2 hours=3; 2+ hours=2) than in Run 2 (1-2 hours=1; 2+ hours=0).

**Table 12. Ground delay by levels of acceptability**

| Run | on-time | 20-60 minutes | 1-2 hours | 2+ hours |
|---|---|---|---|---|
|  | Most acceptable | Acceptable | Less acceptable | Unacceptable |
| 2. All TOS | 79 | 62 | 1 | 0 |
| 5. FET picks route | 69 | 69 | 3 | 2 |

Overall, the differences between the two runs were not extreme. However, considering both runs had a similar number of flights that received ground delay (Run 2: 142; Run 5: 143), and a similar amount of total ground delay (Run 2: 44.6 hours; Run 5: 44 hours), it is interesting that Run 5 had fewer on-time flights and more unacceptable ground delay than Run 2. This indicates that the flights selected for manual reroutes may not have been the optimal choices. The specific flights that are assigned reroutes matter, because their initial position on the schedule affects the positions of all subsequent flights. By comparison, the TOS automation calculates the cost for every flight in the program, then reroutes flights earlier in the program which reduces delay for flights later in the program.

## D. Feedback from the FET

After the completion of each in-person run (Run 1: No TOS and Run 2: All TOS were excluded from the live demonstration), verbal feedback was recorded to collect reactions, comments, and questions from the participants. The following section will present direct quotes from in-person discussions that were held during the demonstrations.

### *1. Feedback on Demonstration 1: Sole-carrier analysis*

The system-wide benefits of having 100% of carriers submit TOS are easily demonstrable, but what about the more likely scenario that only a subset of carriers will develop TOS capability in the near future? One of the main concerns raised by airline industry members is what affect would one dominant carrier, such as Delta in LGA, have on the rest of the system if they were either TOS-participating or TOS-excluded? Prior to the demonstration, participants questioned if accepting a longer route and flying extra miles would unfairly punish the TOS-participating carrier, while simultaneously reducing traffic on the constricted flow and shifting the benefit to TOS-excluded carriers. The following comments came from the participants after reviewing the results from Run 4: American, JetBlue, Southwest, and United TOS only, in which Delta was the only TOS-excluded carrier:

> Delta: *Some delays didn't change, and some got bigger when compared to when we were the only ones filing TOS.*
> American: *With everybody submitting TOS except Delta the delay reduction was huge.*
> Southwest: *We're almost all on time, all of them got rerouted.*
> United: *It was eye opening. I wasn't expecting it to turn out this way.*

All of the bigger carriers were satisfied with the results from Run 4. Only one carrier, JetBlue, did not directly benefit from other airlines filing TOS:

> JetBlue: *The only TOS that I could have filed would have been so much flight time that it wouldn't be worth it, and none of my delays are over a half hour…[I] would have gotten hit [with ground delay] in current day [operations] too, either way they are just stuck.*

Under current day operations with this specific traffic scenario, JetBlue received a realistic amount of delay. However, if the severe weather constraint had blocked the South flow, or if the destination airport was different, it is logical to conclude that the outcome for JetBlue would be different.

The final consensus from the sole-carrier runs (Runs 3: Delta TOS, and Run 4: American, JetBlue, Southwest, and United TOS) was that TOS-participating airlines benefitted from submitting TOS. As one participant summarized:

> *I'm happy that it turned out the way we hoped it would as far as who gets the benefit. The fact that just because you are the only carrier that is using TOS you're not simply going to be moving out of the way so that everybody else can get the benefit. You're still the one that's getting the most benefit out of it. So that's good for the airlines to see.*

From a larger TFM system perspective, it is important for the airlines to buy-in to TOS-participation if CTOP is going to be used as a TMI in the future. CTOP is a way for airlines to avoid costly mandatory reroutes by submitting user-preferred reroute options. However, developing TOS capability is an investment, so providing evidence of direct benefits to the airlines is critical. There should be an advantage if a company is going to take risks and modify their operations. Collecting positive feedback from the FET on the results of the sole-carrier demonstration was an important step.

### *2. Feedback on Demonstration 2: Automation vs. human selection of trajectory assignment*

In this study, Run 1: No TOS was used as a type of baseline condition, in which all flights received a single trajectory option based on their historical flight plan. When CTOP ran under those conditions it produced large ground delays, and those results were used to illustrate a "worst-case" scenario for participants. However, the ground delays from Run 1 were not realistic under current-day operations because the airlines would know significantly ahead of time if a severe weather situation was going to develop. This would spur them to refile flight plans for several flights over better routes. Run 5: FET picks route was developed to create a more realistic baseline where all the flights were given routes chosen by participants role-playing as the airline dispatchers.

Participants compared the CTOP with TOS solution to their self-generated solution to see which one worked better for the overall system as well as their own airline. The results from Run 2: All TOS and Run 5: FET picks route were generally similar in their performance, however there were some key differences that indicated the automation in CTOP generated a better overall solution. We believe this is because the CTOP algorithm can calculate the schedule flight by flight, on a first-come first-served basis for all flights in the program, which is why it can identify the best available route option for any given flight. It can determine which slots are available within the context of other traffic. Airline operators only have knowledge of the schedule and flight plan for their own fleet, and do not know the intent of other traffic in advance. This is why one participant commented:

> *A CTOP, I believe, is a more surgical way to manage.*

One conclusion widely supported by the demonstrations was that the use of any TOS in the TFM system is beneficial to the airlines because more planes can take their preferred route rather than taking a large ground delay or an unnecessarily long mandatory reroute. The take-home message for one participant was:

> *What the data seems to be showing at this point is no matter what the case is, running a CTOP is doing better than what we're doing today. That gets better if we precondition by doing single option TOS…if you don't have the multi[ple] option TOS possibility, that's better than the baseline. It's even better if 25% participate and gets better and better and better as you get more participation. That's exactly what we want the data to show, right? Because it basically says no matter what we're doing better.*

### *3. Feedback on next steps*

#### *3.1 Modeling capabilities would be useful to airlines*

One participant commented that the ability to model their own CTOP would be a major factor for industry buy-in. Currently, the users have only limited information about CTOP TMIs, and no ability to model CTOP impact on their own fleet. Members of the airline industry would like the ability to model CTOP programs for themselves to help inform their strategies and decision making.

#### *3.2 Testing TOS in other locations*

Every day, Chicago O'Hare airport (ORD) experiences routine delays because of high traffic density arriving from East coast airports. The Northeast and Southeast arrival gates are the most constrained resources, while the Western gates are open. Using TOS with the concept of fix balancing to offload the over-demand on the Eastern side of the airport to the Western side seemed like a good use case. The participants agreed that if the concept could be tested and validated in ORD, then it would be another step towards airlines becoming more invested in TOS capabilities as a part of their business model.

## VI. Discussion

### A. System performance

By comparing Run 1: No TOS (baseline) and Run 2: All TOS, we observed that system performance was better than baseline when TOS were included in the CTOP. In Run 1, total ground delay was higher and disproportionately distributed to the West flow, and the system under delivered to the airport. Conversely, ground delay was lower and more evenly distributed between the three flows, and the target throughput rate was met in Run 2. As more flights took TOS reroutes away from the West, the other flows experienced increased delays, but not an excessive amount.

This was the expected result as similar conditions had been explored in previous research [3]. These comparison modes were selected to show a hypothetical upper and lower bound illustrating how well the system could function with and without predeparture reroutes available. Run 3: Delta TOS and Run 4: American, JetBlue, Southwest, and United TOS were selected to represent middle-ground scenarios, in which the NAS was populated with mixed-equipage airlines. Therefore, from a system performance perspective, Runs 3 and 4 were expected to perform somewhere in between Runs 1 and 2, which was mostly true according to the results of the study. The biggest factor to influence system performance was the number of TOS available in the CTOP that include options to route out of the constrained West flow, such that as the number of TOSs increased the system performance got better. Therefore, we concluded that TFM operations were improved by airlines participating in TOS submission.

### B. Sole-carrier effects

In many cases, there were no adverse effects for TOS excluded airlines if there was one major carrier that was TOS-participating. The reverse was also true, in that there were no adverse effects for the sole TOS-participating airline if the rest of the carriers were TOS-excluded. We observed in Run 3: Delta TOS and Run 4: American, JetBlue, Southwest and United TOS, that both the TOS-participating and TOS-excluded airlines received lower ground delay that was more equitable between flows than they did in Run 1: No TOS. Therefore, there were benefits for all airlines associated with just a subset of airlines participating in TOS.

However, the TOS-participating airlines received greater benefits than the TOS-excluded airlines, which we have deemed the "early-adopter" effect. For example, in Run 3, Delta had the lowest assigned ground delay and highest ground delay savings, and in Run 4, American, Southwest, and United had the lowest assigned ground delay and highest ground delay savings of all runs. By comparison, in Run 2 Delta, American, Southwest, and United collectively received less ground delay, but individually the reduction was smaller. Therefore, with fewer TOS participants in the system, the airlines who participated received greater individual benefits, but these benefits were reduced as more airlines became TOS-participating. Therefore, we concluded that it is more advantageous for airlines to be TOS participants rather than not.

In general, more TOS in the system resulted in better outcomes for most airlines. The exception to this trend was JetBlue, who did marginally worse under the TOS conditions and better in Run 1: No TOS. Because JetBlue in this scenario had all flights coming from the South, there was little chance to reroute into another flow. As delay in the South increased, JetBlue was subjected to higher delays. As one participant observed, the amount of delay assigned to JetBlue in Runs 2 – 5 was not significantly higher than what might have occurred under standard operations. An initial concern was how TOS participating carriers would affect the overall TFM system. The LGA use case scenario highlighted that other factors to consider include where the flights originate from, whether they had any TOS options available to them, and their position in the schedule.

### C. Automation vs. human selection of trajectory assignment

In Run 2: All TOS, an algorithm selected the reroutes, whereas in Run 5: FET picks route, human operators were asked to hand-pick the reroutes from a scripted TOS list. Runs 2 and 5 represented two very similar conditions with the same number of reroutes available. The only difference was the method of selecting which planes received reroutes and where. The system-wide outcome for Run 5 was similar to that of Run 2: All TOS in terms of total ground delay and airport throughput.

Despite the general similarities in results between the two runs, Run 5 had less equitable delay among flows, higher average delay for the flights that took TOS reroutes, and a higher number of reroutes than Run 2. This increased the total additional flight time by adding almost two hours of unnecessary and costly airborne delay and resulted in Run 5 having the lowest number of on-time flights of all runs.

Under realistic circumstances, dispatchers would have access to more decision-support tools (DST) with flight specific information such as arrival time, connecting flights, fuel reserves, and crew scheduling. All this information could certainly help operators make educated guesses, however dispatchers' knowledge of the best routes to pick given the available slots would be limited because they would not have insight into the schedule for all flights at the shared resource. Indeed, when participants chose their best available option for manual rerouting in Run 5, the flights that took reroutes ended up with higher ground delay on average than flights that stayed on their original route, indicating worse performance for rerouted flights. By comparison, CTOP builds the schedule knowing the intent of all flights in the program. It locates available capacity at the FCAs and determines the lowest cost option to the flights. By ensuring that the RTC threshold values are met, CTOP has the ability to select candidates for TOS only when it can improve the performance of the flights. Therefore, we concluded that the CTOP algorithm was more efficient than manual selection at choosing which flights needed to be rerouted to achieve adequate system performance.

Before the demonstration, the FET participants voiced their concern over the possibility that accepting a longer route could clear the way for another competing carrier to gain a slot on the preferred route. If airline operators had past experience with degraded performance from manually rerouted flights, there would be an understandable aversion to allowing CTOP to select reroutes for them. We showed participants that the TOS algorithm selected the best option for the flight only when the RTC criterion was met. In practice, RTC is a user-submitted parameter, which gives the operators more control to ensure that their business objectives are met. An important outcome of the demonstration was convincing the participants to trust the automation to deliver an advantageous result. Going forward, there may be a need for more extensive training for operators on the underlying mechanisms that CTOP uses for trajectory selection. This could help their understanding that CTOP works differently than manual rerouting, and ease their fears of being at a disadvantage by taking a reroute.

Something else to consider is the workload, both cognitive and physical, that humans experience while analyzing reroutes for several flights at a time. Airline dispatchers will issue reroutes to any flights that conflict with a major airspace constraint, however the work is heavily manual. Alternatively, the process of maintaining a database of multiple trajectory options for every flight in the fleet lends itself easily to automation. Although, the larger implications for an airline's organizational structure to support that type of change is something left to be explored.

### D. FET feedback

Because the HITL used real traffic, representatives from United, Southwest, and American found the role-play to be particularly meaningful because they could identify with the traffic and routes. The specific ownership of the traffic allowed the representatives to immerse themselves in the problem. Their responses were overwhelmingly positive, and they were surprised and pleased at the outcome of the runs. All agreed that the concept of using CTOP with user-preferred routes to precondition traffic is a better way to serve the NAS customers as well as the overall system than current TMIs. In addition, the FET supported the idea of expanding TOS capabilities among the airline carriers and proceeding with testing in other locations around the country.

## VII. Conclusion

The FET workshop held at NASA consisted of a two part HITL simulation and a series of in-person discussions. The purpose of the workshop was to demonstrate how individual airlines can be impacted by using TOS during a CTOP, and to collect feedback on IDM operations from the perspective of the airlines. A primary concern that was specifically addressed by the part-task HITL was who receives the greater benefit, TOS-participating or TOS-excluded airlines? The results show that TOS-participating airlines receive greater benefit in terms of ground delay, number of reroute options, and additional flight time, than TOS-excluded airlines. However, this result is dependent on the number and location of flights. Therefore, airlines can experience improved performance by equipping TOS. In addition, we found that TOS-excluded airlines also benefit from other airlines participating in TOS, because the total system-wide ground delay is reduced as more TOS are introduced into the system. The evidence suggests that the benefits are distributed fairly, and there is no unfair disadvantage for airlines who cannot or choose not to equip TOS. Furthermore, participants concluded that using CTOP with the TOS reroute automation was better than current day TMIs which use mandatory reroutes and require manual re-filing of flight plans.

Ultimately, airlines would like better DST for CTOP modeling and TOS generation to support movement toward a multiple trajectory option paradigm. Field tests and high-fidelity simulations with location-specific adaptations can also help vet new traffic management concepts for the NextGen airspace.

## Acknowledgments

The authors would like to thank the Flow Evaluation Team for participating in this demonstration and offering their expert opinions and insights to the project. We would also like to thank all the subject matter experts who have contributed immensely to the IDM concept.

## References

[1] N. M. Smith, C. L. Brasil, P. U. Lee, N. J. Buckley, C. V. Gabriel, C. P. Mohlenbrink, F. G. Omar, B. K. Parke, C. G. Speridakos, & H. Yoo. **“**Integrated demand management: Coordinating strategic and tactical flow scheduling operations,” *In 16th AIAA Aviation Technology, Integration, and Operations Conference*, 2016, p. 4221. Washington, D.C.
doi: 10.2514/6.2016-4221

[2] H. Yoo, C. Brasil, N. M. Smith, P. U. Lee, C. P. Mohlenbrink, N. Buckley, A. Globus, and G. S. Hodell. "Integrated demand management (IDM)-Minimizing unanticipated excessive departure delay while ensuring fairness from a traffic management initiative," *In 17th AIAA Aviation Technology, Integration, and Operations Conference*, 2017, pp. 4100.
doi: 10.2514/6.2017-4100

[3] H. Yoo, C. L. Brasil, N. J. Buckley, G. S. Hodell, S. N. Kalush, P. U. Lee, & N. M. Smith. **“**Impact of Different Trajectory Option Set Participation Levels within an Air Traffic Management Collaborative Trajectory Option Program**,”** *In 18th AIAA Aviation Technology, Integration, and Operations Conference*, 2018, Denver, CO.
doi: 10.2514/6.2018-3040.c1

[4] Hodell, G. Yoo, H., Brasil, C., Buckley, N., Kalush, S., Lee, P.U., Smith, N.M. “Evaluation of Multiple Flow Constrained Area Capacity Setting Methods for Collaborative Trajectory Options Program,” *In 37th IEEE/AIAA Digital Avionics Systems Conference (DASC)*, 2018. St. Petersburg, FL.
doi: 10.1109/DASC.2018.8569872

[5] M. M. Eshow, M. Lui, S. Ranjan. “Architecture and capabilities of a data warehouse for ATM research,” *In 33rd IEEE/AIAA Digital Avionics Systems Conference (DASC),* 2014, pp. 1E3-1.

[6] “National Oceanic and Atmospheric Administration’s National Centers for Environmental Information, *NEXRAD data archive* [online database], URL: https://www.ncdc.noaa.gov/nexraddinv [retrieved: 29 November 2017].

[7] T. Prevot, P.U. Lee, T. Callantine, J. Mercer, J. Homola, N. M. Smith, E. Palmer. “Human-in-the-loop evaluation of NextGen concepts in the Airspace Operations Laboratory,” *AIAA Modeling and Simulation Technologies Conference,* 2010, pp.7