

# Evaluation of Multiple Flow Constrained Area Capacity Setting Methods for Collaborative Trajectory Options Program

Gita Hodell  
*San Jose State University at  
NASA Ames Research Center  
Moffett Field, CA, USA  
gita.s.hodell@nasa.gov*

Hyo-Sang Yoo  
*San Jose State University at  
NASA Ames Research Center  
Moffett Field, CA, USA  
hyo-sang.yoo@nasa.gov*

Connie L. Brasil  
*San Jose State University at  
NASA Ames Research Center  
Moffett Field, CA, USA  
connie.l.brasil@nasa.gov*

Nathan Buckley  
*San Jose State University at  
NASA Ames Research Center  
Moffett Field, CA, USA  
nathan.buckley@nasa.gov*

Conrad V. Gabriel  
*San Jose State University at  
NASA Ames Research Center  
Moffett Field, CA, USA  
conrad.v.gabriel@nasa.gov*

Scott Kalush  
*Human Solutions Incorporated  
Washington, DC, USA  
scott.kalush@humansolutionsinc.  
com*

Paul U. Lee  
*NASA Ames Research Center  
Moffett Field, CA, USA  
paul.u.lee@nasa.gov*

Nancy M. Smith  
*NASA Ames Research Center  
Moffett Field, CA, USA  
nancy.m.smith-1@nasa.gov*

**Abstract**— The purpose of this study was to compare flow constrained area (FCA) capacity setting methods for a Collaborative Trajectory Options Program (CTOP) as they pertain to the Integrated Demand Management (IDM) concept. IDM uses flow balancing to manage air traffic across multiple FCAs with a common downstream constraint, as well as constraints at the respective FCA locations. FCA capacity rates can be set manually, but generating capacities for multiple, interdependent FCAs could potentially over-burden a user. A new enhancement to CTOP called the FCA Balance Algorithm (FBA) was developed at NASA Ames Research Center to improve the process of capacity allocation. The FBA evaluates predicted demand and capacity across multiple FCAs and dynamically generates capacity settings for the FCAs that best meet capacity limits for all identified constraints. In a human-in-the-loop simulation study, subject matter experts were asked to use three different methods to allocate capacity to three FCAs, either (1) manually for every 60-minute time window, (2) manually for every 15-minute time window, or (3) by using the FBA capability to automatically generate capacity settings. Results showed no differences in terms of overall system performance, indicated by similar ground delay and airport throughput numbers between methods. However, differences in individual strategies afforded by the manual methods allowed some participants to achieve system-wide delay that was much lower than the average. The FBA was the fastest method of capacity setting, and it received the lowest subjective rating scores on physical task load, mental task load, task difficulty and task complexity out of the three methods. Finally, participants explained through qualitative feedback that there were many benefits to using the FBA, such as ease of use, precision, and low risk of human input error. These results suggest that the FBA automation enhancement to CTOP maintains system performance while improving human performance.

## I. INTRODUCTION

Integrated Demand Management (IDM) is a Trajectory Based Operations (TBO) concept developed by NASA to address demand/capacity imbalances in the National Airspace System (NAS). IDM coordinates Traffic Flow Management System (TFMS) initiatives with Time Based Flow Management (TBFM) by using the Collaborative Trajectory Options Program (CTOP) [1]. CTOP has two central features; first, it handles multiple flow constrained areas (FCAs) in a single program. Second, it issues ground delays and/or pre-departure reroutes selected from a user-submitted Trajectory Option Set (TOS) [2]. When CTOP is used in the context of IDM, it preconditions demand for TBFM. By coordinating these systems, IDM distributes arrival demand across multiple capacity-limited airspace resources that feed a capacity-limited airport. Detailed descriptions and evaluations of the end-to-end concept, from pre-departure planning through arrival at the airport, are available in Ref 1 and 3. This paper will focus on the CTOP setup operations associated with IDM.

In current day operations, excess demand can be curtailed by assigning airspace (upstream) and airport (downstream) capacity constraints in parallel, but independently from one another. CTOP allows a user to manage adjacent airspace constraints, but is not designed to meet upstream and downstream constraints simultaneously. IDM explores solutions to manage both airspace and airport capacity in an integrated manner. One way to manage both is through “flow balancing,” or balancing demand across multiple flows without exceeding airport capacity. The following section will describe the concept of flow balancing as it is used for IDM, and discuss some challenges that arise.

## II. IDM FOR NEWARK LIBERTY AIRPORT (EWR)

The IDM concept was tested using EWR arrival scenarios with convective weather limiting arrivals from the west gate and requiring demand to be offloaded to the north and south gates. To manage the capacity of each flow, a CTOP setup with three collocated FCAs at the EWR runway threshold, each corresponding to one of three arrival gates (west, north, or south) was used. The number of filed flight plans through each FCA was not to exceed the maximum flow rates, which were 12 flights per hour through the west, 24 through the north, and 28 through the south. Although the upstream arrival flows had a combined capacity of 64 aircraft per hour, capacity allocation was limited to meet the EWR airport acceptance rate of 44 aircraft per hour. Therefore, any demand in excess of either the airspace or airport constraints had to take a CTOP assigned ground delay, or a pre-departure reroute selected from a TOS. The goals of CTOP preconditioning for IDM were to provide demand that (1) did not over or under deliver to the arrival gates, (2) insured that the aggregate demand across all gates did not over or under deliver to the arrival airport, and (3) minimized the adverse impact on flight operations, such as excessive departure delays or unnecessary reroutes.

The CTOP interface within TFMS includes bar chart displays of expected demand using 60, 30, or 15-minute time bins. When setting CTOP capacity values manually, the user must establish capacities for every FCA, ensure that the sum across all FCAs does not exceed the airport capacity, then manually enter the values for each time bin into CTOP for the duration of the program. This is a fairly cumbersome manual task, particularly when the 15-minute bin size is used.

As an alternative to manual capacity setting, methods for algorithmically generated capacity settings are currently being explored [4,5]. During a workshop held at NASA Ames Research Center in 2016 with a CTOP subject matter expert (SME), an automated decision support capability was proposed to help generate capacity settings and balance the load across multiple FCAs. As a result, NASA developed prototype automation called the FCA Balancing Algorithm (FBA), which allocates capacity that is proportional to demand while simultaneously matching an airport acceptance rate.

A key component of the IDM concept is the ability to manage airspace and airport constraints in the same program. To achieve this, the capacity settings for individual FCAs must be user generated, either through manual methods or automated decision support. One of the primary questions regarding different methods of FCA capacity setting is; whether any of the methods, either manual or FBA, are operationally feasible within the context of IDM? In addition, how do they compare in terms of the quality of the solutions

they provide? In order to address these questions, a better understanding of how each method affects overall system performance, and human performance (i.e. workload and strategies) is needed. This paper will discuss the following research questions:

- Are either or both of these methods (i.e. manual vs. FBA) feasible?
- Are the methods different in terms of system performance? If so, which is better?
- Are the methods different in terms of human performance? If so, which is better?
- Are there individual differences in user strategies associated with different methods?
- What are the benefits and limitations of using different methods?

A part-task human-in-the-loop (HITL) simulation was conducted at NASA Ames Research Center in 2017 to explore the previous questions. We evaluated the system and human performance of two manual (60 or 15-minute time bins) and one automated (FBA) methods. We also collected qualitative feedback on strategies, benefits, and limitations of the methods.

## III. METHOD

### A. Participants

Five retired FAA personnel participated in this study. One participant had experience working at the Air Traffic Control System Command Center (ATCSCC). The other four participants worked in Air Route Traffic Control Centers (ARTCCs) as traffic management coordinators (TMCs) in the Oakland, Washington D.C., New York, and Cleveland Centers.

### B. Traffic scenario

The traffic scenario consisted of 195 aircraft inbound to EWR. The arrival traffic formed three inbound flows competing for one landing runway (EWR22L). The heaviest flows came from the south through Washington Center (ZDC), and the west through Cleveland Center (ZOB). The northern flow included both domestic and transatlantic arrivals through Boston Center (ZBW). For realistic representation of air traffic, historical Center TRACON Automation System (CTAS) data from the NASA ATM data warehouse [6] was used. The scheduled demand on the runway was approximately 48.8 aircraft per hour. Three ring shaped FCAs were co-located at the EWR runway threshold. Each FCA filtered for aircraft from a single flow (i.e. one from the north, south, and west). The combined capacity for the three FCAs was set to accept 44 aircraft per hour.

A semi-permeable weather cell was placed over the ZOB airspace for the purpose of restricting airspace capacity on the

western flow (see Fig. 1). The weather was displayed on the traffic planner station to provide visual information about the location and intensity of precipitation (green=low, yellow=moderate, red=high). The weather images were generated from actual weather radar data collected on July 14<sup>th</sup>, 2015 and stored in the National Centers for Environmental Information (NECI) database [7].

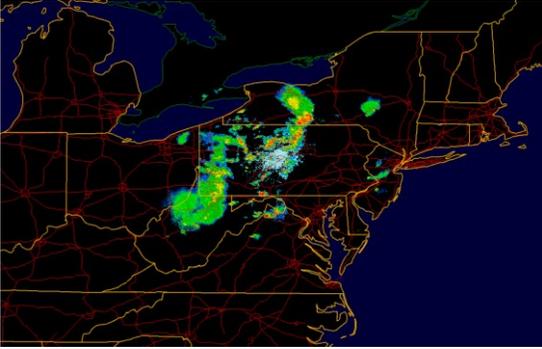


Fig. 1. Convective weather cell restricting EWR arrivals from the west

Trajectory option sets (TOSs) were included for 50% of the eligible flights in the scenario. TOS capability is a key function of CTOP which assigns alternate routes to flights that are impacted by CTOP assigned ground delay. The TOS algorithm calculated the cost to the flight operator of each trajectory option as the sum of the ground delay plus a relative trajectory cost (RTC), then assigned the trajectory with the lowest cost. For this study, the  $RTC = 1.5 \times$  flight time of alternate trajectory – flight time of current trajectory. An in-depth description of the IDM concept of TOS usage and how the TOSs were generated can be found in Ref 8.

### C. Tools

a) *MACS*: The Multi-Aircraft Control System (MACS) is simulation software designed to support human-in-the-loop airspace operations research [9]. In this study, MACS was used to emulate a TFMS planner workstation. The station emulated an En Route Automation Modernization (ERAM) display, scaled to present a full national airspace system (NAS), views of traffic demand with graphics of the convective weather, and a sector load table.

b) *nCTOP*: CTOP emulation software developed by NASA called nCTOP was used alongside MACS on the traffic flow management system (TFMS) planner station. For the purpose of comparing current day CTOP capabilities with the FBA enhancements, two different configurations of nCTOP were used, (1) fielded CTOP, and (2) enhanced nCTOP.

The fielded CTOP configuration was designed to be a realistic emulation of current features and capabilities that are available with the FAA’s CTOP. The user interface displayed traffic demand on FCAs individually. A bar chart represented projected demand over the course of several hours divided

into discrete time bins. The bin size could be set according to user preferences in either 60, 30, or 15 minute increments (see Fig. 2). To view more than one FCA included in the program, the user toggled between FCAs via drop down menu. To input capacity for FCAs, the user selected either 15 or 60-minute time bins (CTOP does not currently allow capacity to be set in 30-minute bins), and manually input numerical values into each bin with a keyboard.

The enhanced nCTOP user interface maintained all of the functionality of fielded CTOP with a few added features. First, the bar chart of projected demand could be viewed either by individual FCA, or by combined view which displayed total demand color coded by flow. Additionally, the FCA bar charts could be viewed in a stacked format (shown in Fig. 3), in which the bar charts for all FCAs were visible on one page. This ability to view all the FCAs at once eliminated the need to toggle back and forth between each one. Second, the enhanced nCTOP version contained the FBA.

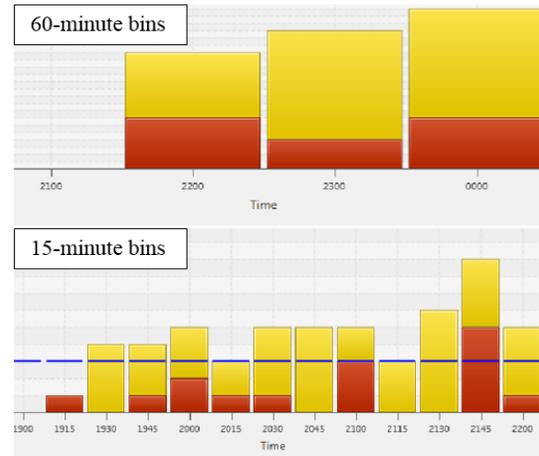


Fig. 2. nCTOP “fielded CTOP” version of the bar chart showing scheduled demand on single FCAs in 60 or 15-minute bins

The FBA required the user to specify a target combined rate per hour across all FCAs. For this study, individual caps were also placed on the three FCAs to limit capacity to 12, 24, and 28 flights per hour on the west, north, and south FCAs. The FBA determined what the corresponding 15-minute bin rate would be and assigned capacity to each FCA that was proportional to the demand at that time.

If the total amount of demand across all of the FCAs exceeded the available capacity in a given time bin, the amount of remaining demand was added to the next time bin and the process was repeated. After capacity values were assigned to each 15-minute bin, the algorithm adjusted the values based on flight exemptions (i.e. flights that were already airborne or within 30 minutes of departure time). Exempt aircraft were given priority for slots in the schedule, always given a controlled time of arrival (CTA) equal to their

estimated time of arrival (ETA), and were not considered for TOS reroutes. For any given 15-minute bin, for each FCA, if the number of exempt aircraft exceeded the capacity it was allocated, the algorithm attempted to reallocate spare (non-exempt) capacity from other FCAs until the capacity accounted for the exempt demand.

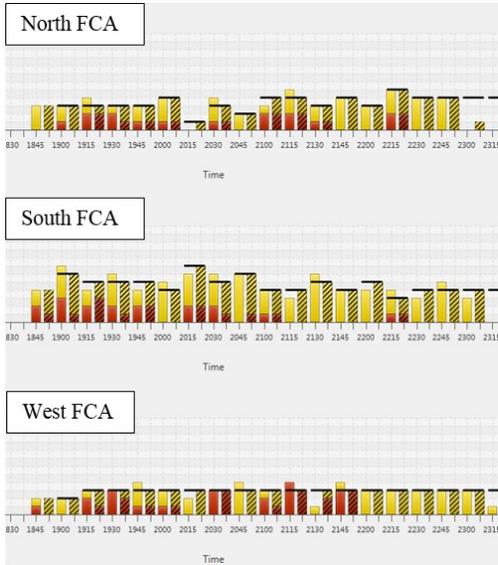


Fig. 3. Enhanced nCTOP optional view of stacked demand bar chart, FBA allocated capacity lines, and model bars (greyed out)

#### D. Measures

To evaluate the subjective workload of the participants, a computerized post-run questionnaire with a 1-5 Likert type scale response format was administered. Four workload related items were developed to measure Task Complexity (1 = Very simple, 5 = Very complex), Task Load—physical and Task Load – mental (1 = Very low, 5 = Very high), and Task Difficulty (1 = Very easy, 5 = Very difficult).

To gain a better understanding about the strategies, benefits, and drawbacks associated with each setting method, subjective feedback was collected through a combination of post-run and post-test questionnaires. The questionnaires used an open-ended response format and asked participants questions regarding (1) the strategy they used to solve the flow balancing problem, (2) the information they used to conclude they had a viable solution, (3) the pros and cons of each method, and (4) what was feasible and unfeasible about each method.

#### E. Procedure

Participants were given a 30-minute training session prior to each test condition. The part-task study consisted of three test conditions, (1) 60-minute/manual, (2) 15-minute/manual, and (3) FBA. Each participant completed a test condition once, for a total of three runs per participant.

Participants were instructed to use tools available in the CTOP GUI, such as the demand bar chart, to analyze the demand in the scenario. They were told that the weather cell in ZOB had a maximum permeability of 12 flights per hour, and the rest of the demand had to be redistributed between the two alternate flows in the north and south. Furthermore, the north could not exceed 24 aircraft per hour, and the south could not exceed 28 per hour. Finally, the combined demand on the EWR airport could not exceed 44 aircraft per hour. Under these conditions, participants were asked to find the optimal capacity rates for the three FCAs.

First, in the 60-minute/manual condition, participants calculated and manually entered hourly capacity rates. Second, in the 15-minute/manual condition participants calculated and manually entered quarter-hourly capacity rates. Participants were offered pencil, paper, and a calculator (tools actually used in the field) to help complete their calculations. Finally, the FBA condition used automation to calculate the rates for 15-minute time bins and automatically input the rates into the nCTOP GUI.

In all conditions, participants had the opportunity to review the settings by modeling the system-wide outcome of their solution before executing the CTOP program. By modeling their solution, participants could analyze the distribution of delays, and the re-distribution of demand that the system would incur if the participant accepted the capacity settings. If the participant thought the capacity rates could be improved after looking at the modeled outcome, they could change the settings and re-model. If the participant thought that the modeled outcome was acceptable, they initiated the sending of pre-departure reroutes and ground delay to the flights.

After each participant completed a test condition they filled out a computerized post-run questionnaire. After all conditions had been completed, participants were given a final post-test questionnaire.

## IV. RESULTS

The purpose of this study was to evaluate the impact of three different capacity setting methods on system and human performance, and to identify strategies, benefits and drawbacks, associated with each method. Three independent variables were used, the 60-minute/manual, the 15-minute/manual, and the FBA capacity setting conditions. To assess operational feasibility, data on a number of dependent variables were collected. System performance metrics included total ground delay and airport throughput assigned by each method. Human performance measures included task performance time, and subjective workload ratings on Task Difficulty, Task Load – mental, Task Load – physical, and Task Complexity. In addition, we gathered qualitative feedback regarding strategies, benefits, and drawbacks, associated with each method.

## A. System performance

a) *Total assigned ground delay*: The total cumulative ground delay assigned by each capacity setting method is shown as a box-plot (see Fig. 4). The 60-min/manual condition had the highest median ground delay, with a moderate amount of variability, *Mean* ( $M$ ) = 46.4 hours, *Median* ( $Med$ ) = 47.8 hours, *Standard Deviation* ( $SD$ ) = 3.42 hours), the 15-minute/manual had the second highest median ground delay and the highest variability ( $M$  = 43.32 hours,  $Med$  = 46.97 hours,  $SD$  = 7.84 hours), and the FBA had the lowest median ground delay with the least variability ( $M$  = 46.56 hours,  $Med$  = 46.78 hours,  $SD$  = 0.88 hours).

Overall, the FBA performed slightly better than the 60 and 15-minute/manual conditions in terms of median ground delay. However, performance in the manual conditions was much better than the FBA on specific runs, denoted by the lower averages and minima of the two manual methods. We attribute these variations in performance to differences in individual strategy, which will be discussed in section C. Finally, the two manual conditions experienced moderate to high variability of outcomes while the FBA solutions were all tightly grouped around the mean. This suggests that the automated solution was frequently used without modification because the automation will reliably produce similar results when given the same set of inputs.

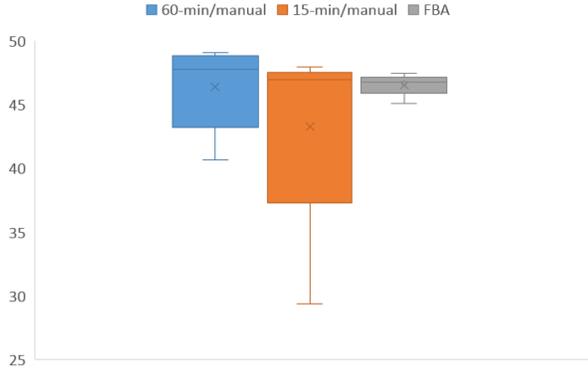


Fig. 4. Average total ground delay (hours) assigned to non-exempt flights by each method

To determine whether or not there was at least one difference between conditions in terms of the amount of assigned ground delay, a one-way ANOVA was conducted. There were no statistically significant differences found in the amount of ground delay assigned between the three conditions,  $F(2,12) = 0.77$ ,  $p = 0.5$ . Therefore, the evidence indicates that the three methods performed similarly in terms of ground delay.

b) *Airport throughput*: To compare each condition in terms of cumulative airport throughput, the linear relationship between the total number of planes landed and elapsed time was plotted. The scenario used in the simulation was designed

to have fewer airplanes landing at the beginning and the end of the scenario. Therefore, a piecewise regression was used to analyze a period of steady throughput ranging from 1.5 to 3.5 hours into the simulation.

For each condition, the slopes of the fitted regression lines were averaged across five participants. The 60-minute/manual ( $M=44.5$ ,  $SD=0.31$ ), 15-minute/manual ( $M=44.25$ ,  $SD=0.16$ ), and FBA ( $M=44.13$ ,  $SD = 0.04$ ) conditions performed similarly, delivering to the target rate (44 aircraft per hour).

## B. Human Performance

a) *Task performance time*: Table 1 presents summary statistics for task performance time. In Table 1 we observe that participants spent more minutes on task in the 15-minute/manual condition ( $M=12.6$  minutes,  $SD=2.1$  minutes) than in any other condition. The condition with the second highest task performance time was the 60-minute/manual ( $M=11.45$  minutes,  $SD=6.51$  minutes), and the FBA condition had the lowest task performance time ( $M=3.19$  minutes,  $SD=0.69$  minutes). Overall, we can see the FBA was the fastest of the three methods.

TABLE I. SUMMARY STATISTICS OF TASK PERFORMANCE TIME (MINUTES)

Method	Average	SD	Median
60-min	11.45	6.51	9.83
15-min	12.6	2.10	12.47
FBA	3.19	0.69	3.25

A one-way ANOVA was conducted to determine whether or not there were differences between the three conditions on task performance time. There was at least one statistically significant difference in the amount of time spent on task between the three conditions,  $F(2,12) = 8.35$ ,  $p < 0.01$ . Therefore, the three methods are different in terms of task performance time, and visual comparison of the means suggest that FBA is the fastest method.

b) *Subjective workload ratings*: Subjective workload ratings were collected in the form of four questionnaire items related to workload, Task Difficulty, Task Load – mental, Task Load – physical, and Task Complexity. The workload items were measured using a 5-point Likert type scale (1 = low levels of difficulty, load, or complexity, 5 = high levels of difficulty, load or complexity). In a table of summary statistics (see Table 2) we observed for all four subjective workload items, the FBA was frequently given the lowest possible score (Task Difficulty,  $M = 1$ ,  $SD = 0$ , Task Load – mental,  $M = 1$ ,  $SD = 0$ , Task Load – physical,  $M = 1.2$ ,  $SD = 0.4$ , Task Complexity,  $M = 1$ ,  $SD = 0$ ). It is notable that on three out of the four items (Task Difficulty, Task Load – mental, and Task Complexity) that the FBA ratings had zero variability, indicating unanimous ratings. Scores for both the

15-minute/manual and the 60-minute/manual conditions were always at or below 3.4, indicating low to moderate workload was associated with the manual methods.

A one-way ANOVA was conducted to determine whether or not there were differences between the three conditions on the four subjective workload item ratings. Results indicated that there was at least one statistically significant difference between the conditions on all four of the questions, Task Difficulty ( $F(2,12) = 9.73, p < 0.01$ ), Task Load - mental ( $F(2,12) = 21.71, p < 0.01$ ), Task Load - physical ( $F(2,12) = 13.5, p < 0.01$ ), and Task Complexity ( $F(2,12) = 23.4, p < 0.01$ ). Therefore, the evidence indicates that the three methods perform differently on all items of subjective workload. Visual inspection of the mean and median summary statistics suggests that the FBA has the lowest ratings on all four subjective workload items of all the conditions. The 60 and 15-minute/manual conditions appear to be very similar to each other in terms of mean and median.

TABLE II. SUMMARY STATISTICS OF SUBJECTIVE WORKLOAD QUESTIONNAIRE

Subjective workload items	Average	SD	Med	F	n
Task Difficulty					
60	2.6	1.1	3		5
15	2.8	0.4	3		5
FBA	1	0	1		5
				9.73**	
Task Load - mental					
60	1.8	0.4	2		5
15	3	0.7	3		5
FBA	1	0	1		5
				21.71**	
Task Load - physical					
60	3	0.7	3		5
15	3	0.7	3		5
FBA	1.2	0.4	1		5
				13.5**	
Task complexity					
60	2.8	0.8	3		5
15	3.4	0.5	3		5
FBA	1	0	1		5
				23.4**	

Scores with \* were significant at the .05 level, \*\* were significant at the .01 level

c. *Subjective ratings by performance time correlation:* To determine whether or not there was a relationship between the amount of time spent on task and answers on the subjective rating questionnaire, a Pearson product-moment correlation coefficient was computed. There was a moderate to strong positive correlation between task performance time and Task Difficulty ( $r = 0.74, n = 5, p < 0.01$ ), Task Load - mental ( $r = 0.60, n = 5, p < 0.05$ ), Task Load - physical ( $r = 0.65, n = 5, p < 0.01$ ), and Task Complexity ( $r = 0.56, n = 5, p < 0.05$ ). Overall, the relationship between task

performance time and subjective workload scores was such that the longer a participant spent performing the capacity setting task, the higher the scores were on all four items.

### C. Strategy

To investigate whether or not participants formed individual strategies to solve the flow balancing problem, an open-ended response questionnaire was used. Participants were asked to provide information about how they solved the traffic flow problem, and what information they used to validate their solution. From their responses, we extracted two major strategic themes. First, demand-based settings, and second, regional knowledge of the airspace and TOS settings.

a) *Strategy 1: Demand-based settings:* Four out of five participants chose to allocate capacity that was proportional to the amount of projected demand over each arrival flow. It is notable that this is the same method that is employed by the FBA. Due to the characteristics of the traffic scenario, demand in the south flow was higher than the north flow, therefore participants who chose demand-based settings allocated more capacity to the south. The availability of TOS reroutes was not considered in the decision making process. To validate their solution, participants modeled their solution and used information about the median, average, and total ground delay over each flow to decide if they were evenly distributed between the north and the south. Participants who used this strategy reported using it during both the 60 and the 15-minute manual conditions.

b) *Strategy 2: Regional knowledge of the airspace and TOS:* One participant with past professional experience controlling traffic through the North East airspace used his prior knowledge to inform his decision making process. His strategy was based on the regional characteristics of the traffic scenario. For instance, a flight operator coming from the west would be more likely to prefer a TOS reroute to the north rather than the south because the northern route is shorter. Furthermore, he could see via modeling that ground delays in the north were higher when more capacity was assigned to the south. Hence, this participant assigned extra capacity to the north FCA to encourage flights to take a TOS to the north, even though overall demand was higher in the south. The participant validated his decision by monitoring the average ground delay assigned to each flow and determining they were roughly equal. Also, he used his best judgment to determine that the number of TOS reroutes assigned was reasonable.

c) *Comparison of Strategy 1 vs. Strategy 2:* To investigate differences in the overall impact on system performance that resulted from different strategies, tables of capacity settings, and the results of settings in terms of ground delay, flight time difference, and number of reroutes were generated. We selected two participants from the 15-minute/manual condition (Participant A and Participant B) to demonstrate the differences between strategies in system

performance. Participant A used demand-based settings, by strictly assigning capacity which was proportional to the amount of predicted demand on each flow. Participant B used regional knowledge of airspace configuration and TOS availability, to encourage flights to reroute to the north instead of south. The reason a single test condition was selected was because Participant B only used this strategy once.

We can see in Table 3 that Participant B assigned seven more slots to the north than Participant A. As a result, Participant B assigned less total ground delay without increasing airborne delay (i.e. the difference between original route flight time and TOS reroute flight time), or violating the maximum rate through the north gate, with the same number of TOS reroutes as Participant A (see Table 4).

TABLE III. TOTAL NUMBER OF SLOTS ASSIGNED TO FLOWS BY EACH PARTICIPANT IN THE 15-MINUTE/MANUAL CONDITION

FCA	Participant A	Participant B
North	46	53
South	74	68
West	45	45

TABLE IV. GROUND DELAY, FLIGHT TIME DIFFERENCE, AND TOS ASSIGNMENTS BY EACH PARTICIPANT

System performance metrics	Participant A	Participant B
Total ground delay (minutes)	2818	1765.2
Flight time difference (minutes)	219.9	175.6
Number of reroutes	21	21
West to North	9	15
West to South	12	6

The results of the strategy comparison indicate that in some situations, capacity allocation based on regional airspace knowledge may positively impact system performance.

#### D. Benefits, limitations, and feasibility of capacity setting methods

To gain a better understanding of the initial, subjective opinions of participants on each of the capacity setting methods, an open-ended response questionnaire was administered. The questionnaire was used to collect qualitative responses regarding the benefits, limitations, and feasibility of each method.

According to the participants, the 60-minute/manual method was more feasible than the 15-minute/manual method in terms of workload, but potentially less accurate. First, the 60-minute/manual method took less time to mentally compute and manually input the capacity settings since there were fewer time bins for each FCA. Second, participants could see

a general view of the demand, which in many situations could be sufficient. As one participant stated:

*“I loved the problems with the 60 minute bins... as in it takes a lot less time due to less math. I think 60 minutes would work in most circumstances but not all...the problem is that it is difficult to spot which times it would not work.”*

Although the settings for 60-minute time bins were less cumbersome to compute, there were limitations to the method regarding precision. First, because participants only saw overall demand, there was potential for a few unused slots to go unnoticed, which could have made it more difficult to feed the target delivery rate. Second, it was difficult to monitor the distribution of demand within the 60-minute period. If the traffic happened to be heavily loaded in the first 30-minutes of the problem, this could have negatively impacted downstream sectors by having too many planes in the air that were not properly spaced out. Participants stated that the ability to use 30-minute time bins might be a better solution.

The 15-minute/manual method was less feasible in terms of workload than the 60-minute/manual method for participants, however it was more precise. First, the 15-minute/manual method allowed participants to observe and control demand with finer granularity which encouraged optimal capacity utilization. Second, it was easier to monitor the distribution of demand which mitigated concerns of overloading downstream sectors. The primary limitations associated with the 15-minute/manual method were workload and human error. As one participant stated:

*“The possibility of making errors while inputting the capacity numbers is very large.”*

Not surprisingly, many participants stated that inputting the larger number of capacity settings while toggling back and forth between FCAs was time consuming and cumbersome. Furthermore, participants had to spend time scrolling through multiple time bins and correcting errors. Having the ability to see all the active FCAs in a combined format on a single screen was identified as a potential feature that could help alleviate some of the workload.

The FBA method was viewed as both feasible and accurate by participants. The FBA method eliminated the cumbersome task of mental computation and manual inputs, and also displayed demand within granular 15-minute bins. Participants did not enumerate any limitations to the FBA method. However, even with the ease of use the FBA provided, participants unanimously agreed that the option for the user to make manual adjustments to the FBA solution was imperative, i.e. the FBA solution should be used for decision support, but the user should make the final decision. As one participant stated:

“Just because the algorithm balances the demand it does not mean our goals are accomplished. Manually making adjustments allows us to encourage certain behaviors. Demand is only one factor in capacity. It could be impacted by weather, staffing, or other TMIs [traffic management initiatives]. All sorts of issues are not known to a dispatcher, dispatching the flight.”

## V. DISCUSSION

This study evaluated different methods for setting flow rates across multiple, interacting FCAs. The three methods were similar in terms of overall system performance, yet there was more individual variation in the 60 and the 15-minute/manual methods than the FBA method. Differences in individual strategies used to solve the traffic flow problem allowed some participants to achieve lower amounts of system-wide delay than the average. We attribute these differences to varying levels of knowledge of the regional airspace that was being monitored. For example, if a participant had expert knowledge of the airspace including preferred routes, then he could mentally approximate the demand after accounting for post-TOS demand reallocation. The FBA does not account for these specialized TOS characteristics when it computes the flow rates. Therefore, it is possible for an expert user to out-perform the FBA. In contrast, other participants used total demand as the primary heuristic for capacity distribution. Therefore, the majority of participants in the 60 and 15-minute/manual conditions performed similarly to the FBA because the demand-based strategy performed similarly throughout all conditions. By examining contrasting strategies, we found that some scenarios can benefit from a more nuanced approach than strictly demand-based capacity allocation. These findings suggest that there is potential to develop improvements to the FBA keeping these considerations in mind.

Out of the three methods, the FBA was the fastest, and easiest to use according to task performance time metrics, subjective workload ratings and qualitative feedback. We found that the capacity setting methods that took longer to complete were also reported as being more difficult, complex, and having higher task loads, both mental and physical. Participants also commented that the FBA involved less manual work. This is not surprising considering the method of rate computation is different between manual and automated (pencil and paper vs. a single button press). While not mathematically complex, the computations cannot easily be done in one's head, especially as the number of FCAs involved increases. Another contributing factor to subjective workload was the time it took to recover from human input

error, which occasionally increased task performance times. Hence, an automated tool to reduce calculation and input error makes sense from a workload perspective.

In conclusion, the initial evaluation of the FBA capacity setting method appears promising. There is potential to incorporate the FBA into CTOP as a means to mitigate human operator workload while maintaining throughput accuracy.

## ACKNOWLEDGMENT

The authors would like to thank all the subject matter experts who frequently offer their insight and contribute immensely to the development of the IDM concept.

## REFERENCES

- [1] N. M. Smith, C. Brasil, P. U. Lee, N. Buckley, C. Gabriel, C. P. Mohlenbrink, F. Omar, B. Parke, C. Speridakos, H. Yoo. "Integrated demand management: Coordinating strategic and tactical flow scheduling operations." *16th AIAA Aviation Technology, Integration, and Operations Conference*, 2016, pp. 4221.
- [2] National Business Aviation Association, "Collaborative Trajectory Options Program (CTOP)," <https://www.nbaa.org/ops/airspace/tfm/tools/ctop.php>, accessed: 2018-07-11.
- [3] 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.
- [4] C. Kaler, W. Hall, C. Brinton, B. Capozzi, G. Hunter, "Collaborative trajectory options program within the NAS flow advisory manager." In *18th AIAA Aviation Technology, Integration, and Operations (ATIO) Conference*, 2018.
- [5] R. Hoffman, B. Hackney, P. Wei, G. Zhu, "Enhanced stochastic optimization model (ESOM) for setting flow rates in collaborative trajectory options programs (CTOP). In *18th AIAA Aviation Technology, Integration, and Operations (ATIO) Conference*, 2018.
- [6] 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.
- [7] National Oceanic and Atmospheric Administration's National Centers for Environmental Information, "NEXRAD data archive, inventory and access," <https://www.ncdc.noaa.gov/nexraddinv>, accessed: 2017-11-29
- [8] H. Yoo, C. Brasil, N. 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 (ATIO) Conference*, 2018.
- [9] 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