Alternatives for Scheduling Departures for Efficient Surface Metering in ATD-2: Exploration in a Human-in-the-Loop Simulation

Bonny K. Parke[1], Lindsay K. S. Stevens[2], William J. Coupe[2], Hanbong Lee[2], Yoon C. Jung[2], Deborah L. Bakowski[1], and Kimberly Jobe[1]

[1]San Jose State University Research Foundation
NASA Ames Research Center
P. O. Box 1, Moffett Field, CA 94035-0001
[2]NASA Ames Research Center
P. O. Box 1, Moffett Field, CA 94035-0001
Bonny.Parke@nasa.gov

Abstract. A Human-in-the-Loop (HITL) simulation was conducted to explore the impacts of various surface metering goals on operations and Ramp Controllers at Charlotte Douglas International Airport (CLT). Three conditions were compared: (1) Baseline, with no surface metering, (2) instructions to meet advisory times at the gate only, and (3) instructions to meet advisory times at the gate as well as the times at the scheduled taxiway spot, where aircraft are delivered to Air Traffic Control (ATC). Results showed increased compliance for taxiway spot times when compliance was first met for gate advisories. Instructing Ramp Controllers to meet advisory times at the gate improves spot time compliance and therefore surface scheduling predictability at CLT. Results also demonstrated there was increased compliance overall with gate and spot times in the second condition. This was likely due to higher Ramp Controller workload in the third condition.

Keywords: ATD-2 · Airport Surface Scheduling · Integrated Arrival, Departure, and Surface (IADS) Traffic Management · Human Factors Assessment in Field Operations · Surface Metering · HITL Simulation

1 Introduction

Airspace Technology Demonstration 2 (ATD-2) is an ambitious NASA project in coordination with the Federal Aviation Administration (FAA) and aviation industry partners that aims to integrate, for the first time, multiple concepts and technologies to reduce delays in the National Airspace System (NAS) [1]. The umbrella concept is Integrated Arrival, Departure, and Surface (IADS) operations. The FAA selected Charlotte Douglas International Airport (CLT) and surrounding air traffic control (ATC) facilities as a test bed. CLT is the sixth busiest airport in the US based on the number of operations in 2018 and the busiest on the East Coast [2]. It is known as having the runway capacity of a major airport with the ramp capacity of a mid-level airport. Among the concepts and technologies being integrated into new tools for Ramp and Air Traffic Controllers at CLT is surface metering, as outlined by the FAA's Surface Collaborative Decision Making (Surface CDM) Concept of Operations [3].

A goal of metering, or time-based scheduling, on the surface is to reduce aircraft wait times in the departure runway queue, with its attendant fuel burn and emissions. Benefits of using surface metering include increased predictability in timing of aircraft movement across the airport surface, as well as economic and environmental benefits. At CLT, initial assessment of quantitative measures have shown a marked improvement in fuel burn and emissions in banks of aircraft that are metered, with an average reduction of 1,000 pounds of fuel and 3,000 pounds of carbon dioxide emissions per bank (CLT typically has nine banks per day) [4].

In the Surface CDM concept, surface metering works by redistributing some of the time that would otherwise be spent in the departure runway queue back to the ramp area, typically at the gate. Each aircraft subject to surface metering is assigned a time to be delivered to the "spot," or the point at which Air Traffic Control (ATC) takes control of the aircraft. During surface metering, Ramp Controllers deliver aircraft through the ramp to the spot in compliance with each aircraft's assigned spot arrival time. ATD-2's surface scheduler works to achieve the goals of the Surface CDM concept by generating target event times for aircraft to push back off gates, be delivered to the spot, and arrive at the runway. Scheduler-generated gate advisories are used by Ramp Controllers to aid them in deciding when to release aircraft from the gate and begin moving them across the ramp.

CLT has unique constraints that may impact the Ramp Controllers' ability to focus on delivering aircraft to the spot in compliance with an assigned spot time. Not only does CLT have the ramp capacity of a mid-level airport with limited ramp real estate, it also possesses extended areas within the ramp where traffic flow is reduced to a single-lane, so traffic can only flow one direction at any given time. Once an aircraft is inserted into the flow of traffic in the ramp, few, if any, options exist to adjust the aircraft's rate of travel to the spot.

A question exists as to whether Ramp Controllers can utilize alternative strategies to achieve the Surface CDM goal of spot time compliance during surface metering at airports like CLT with ramp area constraints, and the impact of those methods on surface operations and Ramp Controllers. This question was explored in a Human-in-the-Loop (HITL) simulation conducted at NASA Ames Research Center in June, 2018.

2 Method

2.1 Human-in-the-Loop Simulation

Conditions. To explore the impact of various methods of ramp operations for complying with times at the spots, the HITL compared three conditions at a simulated CLT:

1) **Baseline:** No metering/no advisory condition; normal operations,
2) **TOBT-Only:** Metering with Ramp Controller instructions to meet gate advisories for departures (Target Off-Block Times or TOBT) within a ±2 minute window, and
3) **TOBT+TMAT:** Metering with Ramp Controller instructions to meet gate advisories (TOBT) within a ±2 minute window for departures and, in addition, arrival

times at the spots (Target Movement Area entry Times or TMAT) within a ±5 minute window.

HITL Overview. Each condition (i.e., Baseline, TOBT-Only, and TOBT+TMAT) was tested for three runs for a total of nine experimental runs. Three 70-minute traffic scenarios were used during the study; each scenario was used once in each test condition. Each of these traffic scenarios was duplicated from CLT live traffic recordings on different days during Bank 2 (the heaviest traffic bank at CLT) and thus had similar traffic loads. The three conditions were counterbalanced over time to counteract any order effects of training and experience. On each run, four Ramp Controllers rotated through each of the four CLT ramp positions: North, East, South, and West Sector. Ramp Controller training took place on the first morning and included familiarization with the ATD-2 IADS Ramp Traffic Console (RTC) interface and the three conditions.

Participants. Participants in the HITL included four experienced Ramp Controllers, one Ramp Manager, four ATC Tower Controllers, and one ATC Traffic Management Coordinator (TMC), along with eight Pseudo-Pilots and two Air Route Traffic Control Center (ARTCC) confederates.

Instruments. In addition to quantitative metrics such as target time compliance, human factor metrics were collected. Workload Assessment Keypads (WAKs) were adapted to chime every five minutes during the HITL runs, at which time Ramp Controllers rated their workload on a 1 to 5 scale, with 1 being "Very low workload," and 5 being "Very high workload." Ramp Controllers also rated their workload on post-run surveys using five of the NASA Task Load Index (TLX) rating items [5]. Situational awareness was derived from an adapted version of the 3D Situational Awareness Rating Technique[1] (SART) [6]. Ramp Controllers were also asked to rate the acceptability of various aspects of ramp operations and to describe how they handled aircraft in the TOBT+TMAT condition. A post-study survey and debrief followed.

Displays. Fig. 1 shows the new ATD-2 IADS Ramp Traffic Console (RTC) which was available to Ramp Controllers in all conditions; detailed descriptions can be found elsewhere [7].

[1] The 3D SART consists of a score obtained by adding an item rating "your Understanding of the traffic situation" to an item rating "the Supply of your attentional resources" and subtracting an item rating "Demand on your attention" (i.e., SA = U + (S-D)).

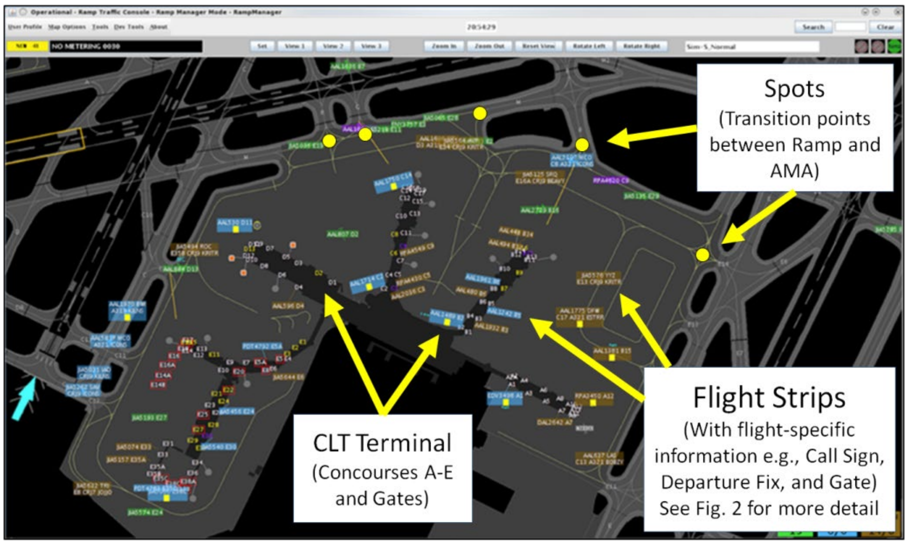

Fig. 1. Display of the Ramp Traffic Console (RTC) tool showing the CLT Ramp area with five Concourses (A-E), ramp area (made up of four ramp sectors), and surrounding taxiways and runways. Example "spots," the transition points between the ramp area and the ATC-controlled Airport Movement Area (AMA), are depicted. Each flight is represented by a digital flight strip, which indicates flight-specific information.

Close-ups of the flight strips are shown in Figs. 2-3.

Fig. 2. Digital flight strips on the RTC display in the Baseline, or no metering condition, (left) at the gate prior to pushback and (right) following pushback, while taxiing in the ramp area.

Fig. 3. Flight strips for aircraft in both metering conditions: (left) at gate showing gate hold advisory countdown ("4 min") to the TOBT (Target Off-Block Time); (middle) at which time "PUSH" is indicated; (right) after pushback with TMAT (Target Movement Area entry Time) showing "1941" for arrival at the spot.

3 Results

3.1 Quantitative Results

TMAT Compliance Increases for Aircraft Compliant with TOBT. During the runs compliance with the TOBT times at the gate and TMAT times at the spot were measured for all aircraft subject to surface metering. In both the TOBT-Only and TOBT + TMAT condition we observed that compliance with the TOBT advisory increased the chance of complying with the TMAT advisory. Similar findings appear in analysis of operational data collected at CLT during the Phase 1 field evaluation [8]. This indicates that the TOBT advisory is helpful in achieving TMAT compliance. This is an important finding as future systems will be built around controlling the flow of surface traffic via the TMAT at the spot, in contrast to ATD-2 which controls the flow of traffic via the TOBT at the gate.

Table 1. TOBT and TMAT compliance metrics for flights subject to metering.

Compliance	TOBT-Only condition	TOBT+TMAT condition	Sig. Level Chi Square
TOBT (±2 min)	61.7% (21/34)	57.1% (24/42)	p = .68, 0.2 (*df* 1)
TMAT (±5 min)	85.3% (29/34)	69.0% (29/42)	p = .10, 2.7 (*df* 1)
TMAT given TOBT compliance	95.2% (20/21)	75.0% (18/24)	p = .07, 3.3 (*df* 1)

Additionally, we observed that TMAT compliance in the TOBT-Only condition was higher than TMAT compliance in the TOBT + TMAT condition. Given the relatively small sample size of HITL data the significance of this finding is unclear. There exist various factors affecting the traffic situation in each simulation run, but one possible reason for lower TMAT compliance in the TOBT + TMAT condition is the increased workload of having to keep track of aircraft scheduling times at both the gate and the taxiway spot. This increased workload could have lowered the ramp controllers' situation awareness and interfered with meeting the TMAT times. Evidence for this possibility will be explored in the next sections.

3.2 Subjective Results

Perceived Efficiency in Ramp Operations. After each run, Ramp Controllers were asked "During the busiest time in this run, how acceptable were the following in terms of operational efficiency?" As can be seen from Fig. 4, the general trend was that operations were less acceptable in the TOBT+TMAT condition (red bar). Hold times at the hardstands (holding areas for aircraft) were rated as significantly less acceptable in the TOBT+TMAT condition than Baseline.

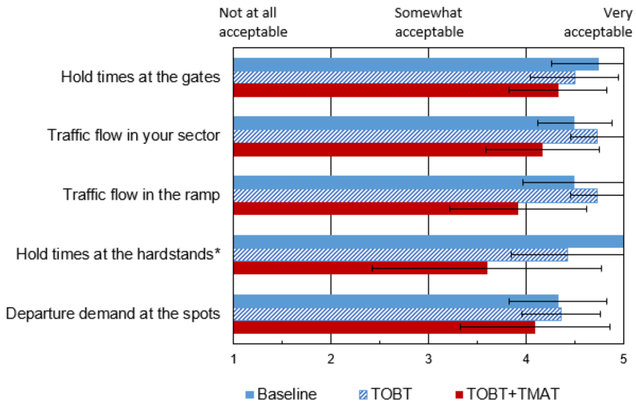

Fig. 4. Ramp Controllers' perceptions of the acceptability of ramp conditions in terms of operational efficiency. N= 36, (12 ratings in each of 3 conditions, 4 controllers in each of 3 runs). Error bars = 95% Confidence Intervals. Significant ANOVAs are marked with customary asterisks (e.g. "Hold times at the hardstands" was significant at $p \leq .05$).

Working with TMATs. In the TOBT+TMAT condition, Ramp Controllers were asked, "How frequently in this run did you use TMATs to make decisions about sequencing aircraft?" The average response was "About half the time."

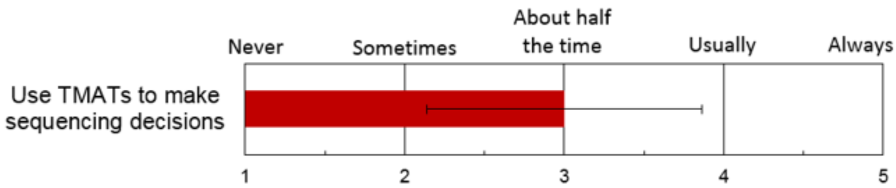

Fig. 5. Average responses of Ramp Controllers in the TOBT+TMAT condition on frequency of TMAT usage, N = 12 ratings (4 controllers in 3 TOBT+TMAT runs). Error bar is 95% Confidence Interval. Of the 12, 2 responses were "Never."

To explain this finding, a typical comment on the post-run survey was, "Things were flowing a bit fast. I didn't have enough time to really sequence the TMAT times." Hence high workload and time pressure appear to be contributing factors as to why more Ramp Controllers didn't use TMATs to sequence aircraft.

Ramp Controllers were asked to "Please rate how appropriate the times of the TMATs were in this run for aircraft coming from the gates in your sector and from other sectors." For aircraft originating in their own sector, TMAT times were thought to be "About right," as shown in Fig. 6.

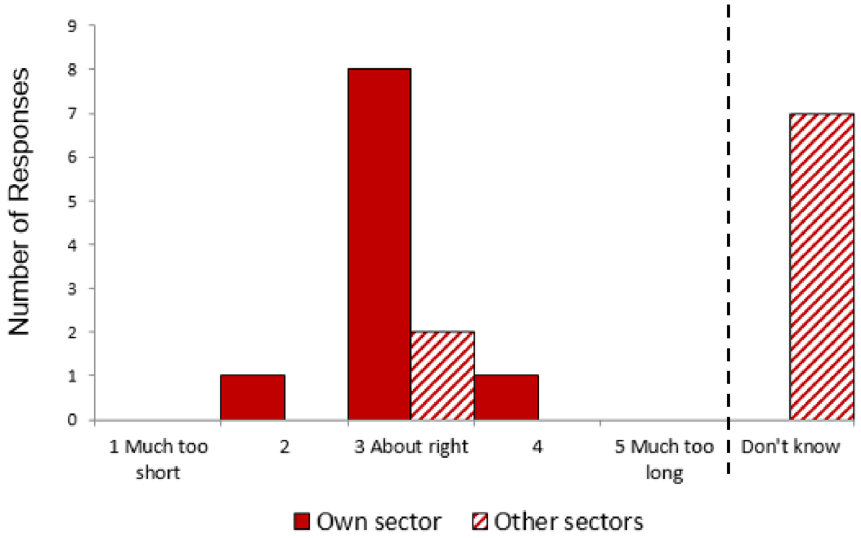

Fig. 6. TMAT times were judged "About right," unless aircraft came from another sector, in which case the Ramp Controllers selected "Don't know." $N = 10$ ratings for "own sector" and 9 for "other sectors" from Ramp Controllers who used TMATs to make sequencing decisions.

As can be seen, Ramp Controllers frequently marked "Don't know" on this question for aircraft that had originated in another sector. This shows a lack of situation awareness in this condition and indicates it is likely they were not making sequencing decisions based on TMAT times for those aircraft.

Situation Awareness. The final SART scores in the TOBT+TMAT condition were significantly lower than in the Baseline condition, as shown in Fig. 7. The scores were not significantly lower in the TOBT-Only condition.

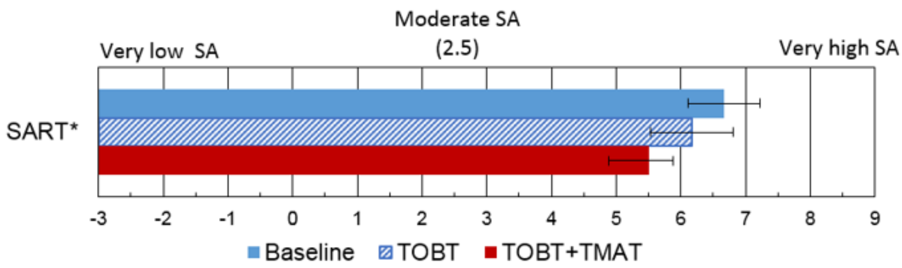

Fig. 7. Final SART scores in the three conditions. $N = 36$ (12 ratings in each condition). ANOVA is significant at $p<.05$; error bars = 95% Confidence Intervals.

Workload Assessment Keyboard (WAK) Results. Ramp Controller workload was significantly higher in the TOBT+TMAT condition than Baseline as indicated by the WAK.

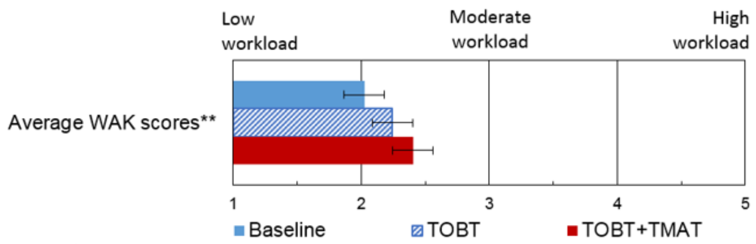

Fig. 8. WAK results in the three conditions. *N*s = 154,155,154 per condition. ANOVA is significant at *p*<.01; error bars = 95% Confidence Intervals.

Results of NASA Task Load Index (TLX) Items. Workload items "Time Pressure" and "Effort" were significantly higher in the TOBT+TMAT condition than in the other two conditions, as shown in Fig. 10. Other comparisons were not significant.

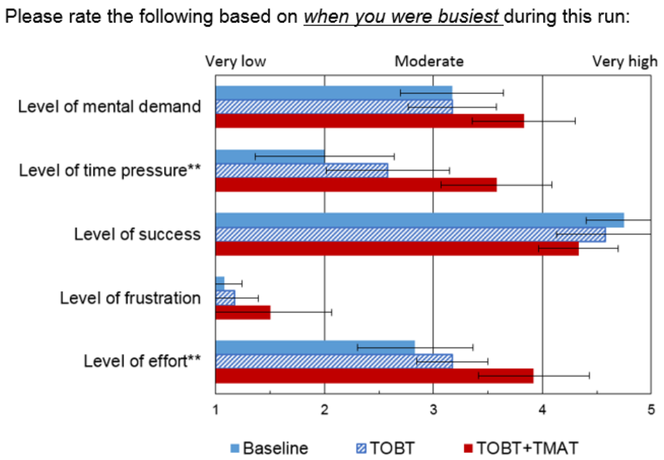

Fig. 9. Results on NASA TLX items in the three conditions. *N* = 36 ratings, (12 in each of 3 conditions). Error bars = 95% Confidence Intervals.

Post-Study Survey Results. Self-assessed workload on the post-study survey was significantly higher in the TOBT+TMAT condition than in the other two conditions and near "Very high." There was no significant difference between Baseline and the TOBT-Only condition.

Please describe your workload at the busiest times in each of the conditions in this simulation.

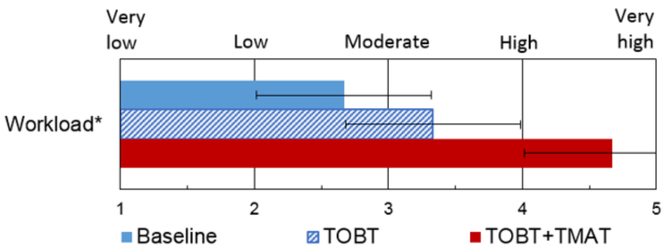

Fig. 10. Self-assessed workload in the three conditions on the post-study survey. $N = 4$. Repeated measures ANOVA, sphericity not assumed, $F(2,2) = 28$, $p = .03$. Error bars = 95% Confidence Intervals.

4 Summary and Discussion

In anticipation of the implementation of FAA's Surface CDM concept of surface metering, in this HITL different possibilities were explored for achieving TMAT compliance at the "spot," the point where ATC takes control, within a ±5 minute window. Three conditions were tested to understand the impacts of this goal on operations and on Ramp Controllers at a simulated CLT, where there is limited real estate for managing ramp traffic. Results from the HITL indicated that there was more compliance with the TMATs in the TOBT-Only condition, where Ramp Controllers were only asked to push aircraft back from their gates within a ±2 minute window, than in the TOBT+TMAT condition, where Ramp Controllers were given explicit instructions to also try to get aircraft to the spots within a ±5 minute window.

Reasons for this finding were explored in this paper. Increased workload and lowered situation awareness may have contributed to reduced TMAT compliance in the TOBT+TMAT condition. First, Ramp Controllers reported using TMATs to make decisions about sequencing only half the time, with comments explaining this in terms of workload and time pressure: "Things were flowing a bit fast. I didn't have enough time to really sequence the TMAT times." Second, on specific workload measures, the workload was significantly higher in the TOBT+TMAT condition than the other two conditions. This was the case for two items on the NASA TLX: level of time pressure and level of effort, and for the self-assessed workload on the post-simulation survey. Workload on this survey was rated on a 1 to 5 scale as 4.7 in the TOBT+TMAT condition compared to 3.3 in the TOBT-Only condition, and 2.7 in the Baseline condition.

Although the SART situation awareness scale used in this study indicated that situation awareness was lowest in the TOBT+TMAT condition, this reached significance only in comparison with Baseline. It should be noted, however, that the frequent "Don't know" responses on the post-run survey also indicated a low situation awareness in the TOBT+TMAT condition. For example, Ramp Controllers frequently did not know if the TMAT times of aircraft originating from a different sector were appropriate. As one Ramp Controller put it, "Trying to think about the TMAT times and keeping them in order . . . can sometimes be demanding. Trying to keep order and recognize what other team members may have going on is demanding enough. Once I push and send an instruction to taxi, I usually don't have enough time to go back and see if the TMAT time is within limits. I think the system should monitor and adjust these numbers."

4.1 Conclusion and Next Steps

Therefore, for airports like CLT with unique limitations in the ramp area, instructing Ramp Controllers to meet scheduling times at the spots may not be feasible. Benefits of TMAT compliance can still be achieved by Ramp Controllers working toward TOBT compliance, as evidenced by improved TMAT compliance and decreased workload in the TOBT-Only condition. Additional anecdotal evidence was captured during a TOBT+TMAT run when one Ramp Controller commented, "I think once the ramp got congested, it basically made the TMAT times insufficient. I would like to have had another way to get a few of them out thru the traffic." Other ramp controllers echoed this in the post-simulation debrief stating that at CLT, once an aircraft is merged into the ramp traffic, there is very little a Ramp Controller can do to meet TMAT times. However, by following the TOBT gate advisories based on an algorithm that makes a best estimate of the time it takes for an aircraft to move from the gate to the spot, Ramp Controllers can, with little increased effort compared to non-metering operations, still deliver aircraft to the spots within the TMAT window.

Next steps are to explore scheduling alternatives in a HITL with more runs, to use an airport with a larger ramp area relative to runway capacity (Dallas-Fort Worth International Airport) and to include a TMAT-Only condition. With a larger ramp area, the Ramp Controllers may be able to meet a TMAT-Only condition.

Acknowledgments. Thanks are due to the many Ramp Controllers, ATC Controllers, and pseudo-pilots who participated in this simulation, along with the ATD-2 Team and the Future Flight Central Simulation Team that made the simulation possible.

References

1. Jung, Y. et al., "Airspace Technology Demonstration 2 (ATD-2) Phase 1 Concept of Use (ConUse)," NASA/TM-2018-219770, February 2018.
2. FAA Air Traffic Activity System (ATADS), Airport Rankings for 2018 by Operations (2018) https://aspm.faa.gov/opsnet/sys/opsnet-server-x.asp
3. FAA Air Traffic Organization Surface Operations Directorate: U.S. Airport Surface Collaborative Decision Making Concept of Operations (ConOps) in the Near-Term: Application of the Surface Concept at United States Airports. June (2012) https://faaco.faa.gov/index.cfm/attachment/download/33926
4. Aviation Systems Division, NASA Ames Research Center, Airspace Technology Demonstration 2 (ATD-2) Overview https://www.aviationsystems.arc.nasa.gov/research/atd2/index.shtml
5. Hart, S. G.,& Staveland, L. E. Development of a NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research. In: Hancock, P.S., Meshkati, N., Human Mental Workload, pp. 139--183, Elsevier Science Publishers B. V., Amsterdam (1988)

6. Taylor, R. M.: Situational Awareness Rating Technique (SART): The Development of a Tool for Aircrew Systems Design. In: Proceedings of the AGARD AMP Symposium on Situational Awareness in Aerospace Operations, CP478. Seuilly-sur Seine: NATO AGARD (1989)
7. Parke, B., Jobe, K., Chevalley, E., Verma, S., Dulchinos, V., & Stevens, L.: Impact of Data Exchange Provided by ATD-2 Tools at Charlotte-Douglas International Airport. 9th International Conference on Applied Human Factors and Ergonomics, Orlando, FL, 21-25 July (2018)
8. Coupe, W. J., Lee, H., Jung, Y., Chen, L., & Robeson, I.: Scheduling Improvements Following the Phase 1 Field Evaluation of the ATD-2 Integrated Arrival, Departure, and Surface Concept. 13th USA/Europe Air Traffic Management Research and Development Seminar, Manuscript in progress (2019)