Ninth USA/Europe Air Traffic Management Research and Development Seminar (ATM2011)

Human-in-the-Loop Simulations of Surface Trajectory-Based Operations

An Evaluation of Taxi Routing and Surface Conformance Monitoring Decision Support Tool Capabilities

Emily K. Stelzer, Constance E. Morgan, Kathleen A. McGarry, Kathryn A. Klein, and Karol Kerns
Center for Advanced Aviation Systems Development
The MITRE Corporation
McLean, VA

Abstract— The Federal Aviation Administration's Next Generation Air Transportation System (NextGen) concept proposes a suite of decision support tools for use in the air traffic control tower to support safe and efficient operations. This paper describes the proposed set of automation capabilities to support ground and local controller activities in the NextGen mid-term (2018), including automated decision support tools to generate taxi routes and monitor pilot conformance to the assigned taxi route. To evaluate these capabilities, The MITRE Corporation's Center for Advanced Aviation System Development (CAASD) conducted two human-in-the-loop experiments to evaluate ground controller performance with these decision support tools, focusing on automation to support taxi route generation and conformance monitoring. These simulations specifically examined capabilities provided in surface automation designed by Mosaic ATM. This paper describes the results of those two experiments and discusses necessary future research to refine and validate the NextGen surface decision support tool concepts.

Keywords-surface automation; taxi routing; conformance monitoring; human performance

I. Surface Trajectory-Based Operations

The Next Generation Air Transportation System (NextGen) is the Federal Aviation Administration's (FAA's) concept for modernizing the National Airspace System through 2018. Through NextGen, the FAA is addressing the impact of air traffic growth by increasing national airspace (NAS) capacity and efficiency while simultaneously improving safety, reducing environmental impacts, and increasing user access to the NAS.

High-density airports in the NextGen mid-term timeframe (~2018) are expected to operate with a set of integrated automation decision support tools (DSTs), which include the following attributes: access to timely and dynamically updated information; a means for collaboration with stakeholders; maximization of airport capacity; and improved efficiency of airport and surface operations [1]. The DSTs described in the NextGen concept are the basis for an operational concept known as Surface Trajectory-Based Operations (STBO) that will manage traffic flows and resources on the airport surface and will enable surface trajectory-based operations in the far-term. This program is managed by the FAA's Aviation Research and Technology Development Office, and includes contributions from Mosaic ATM, MITRE, MCR LLC, and the John A. Volpe National Transportation Systems Center.

A. *Concept Overview*

The need for STBO is derived from a set of identified shortfalls in the current operations that do not meet NextGen needs [2]. In today's operations, periods of high demand often result in departure delays and long physical queues. While current surface operations tend to be tactical and reactive in nature, operations under NextGen will need to be strategic and predictive.

STBO is expected to increase airport surface efficiencies through shared situation awareness and local collaboration, allowing operators to work proactively to manage high surface demand and dynamic surface and airspace constraints. The shared situation awareness generated by STBO will support all stakeholders in surface operations. A common interface with flight operators, airport operators, and surface traffic management systems will enable collaboration.

In addition to supporting improved situation awareness, STBO is also expected to support better decision-making through local information sharing and automated DSTs [3]. These DSTs will share data with other NAS domains, which will contribute to an increase in operational efficiency. The capabilities of these DSTs will assist in the execution of Air Traffic Control Tower (ATCT) responsibilities by providing recommendations to improve throughput and reduce delays for surface operations. These recommendations will take into account flight plan data, flight operator preferences, constraints from other NAS domains (e.g., flow constraints), and local airport operational data [4]. The automated capabilities envisioned as part of the STBO concept are expected to assist tower controllers in performing their tasks and may reduce their workload [2].

STBO capabilities are expected to be developed and introduced in phases [2]. An initial implementation, to be deployed by the NextGen midterm (~2018), will include flight-specific data exchange between FAA facilities and stakeholders and will provide foundational DST capabilities to assist the tower in managing the airport configuration, assigning runways for departing flights, and managing runway queues. Moving through the NextGen mid-term timeframe, these initial DST capabilities will be enhanced and additional DSTs will be implemented to develop recommendations for optimizing airport configurations and flight-specific runway assignments; to provide for collaboration with flight operators for surface scheduling; and to introduce two-dimensional taxi routing and surface conformance monitoring capabilities. Into the far-term, DST capabilities will be enhanced to use surface trajectory modeling as the basis of automated decision making and conformance monitoring; surface operations will be integrated with airborne trajectory-based operations.

B. *STBO Capabilities in the NextGen Mid-term*

The STBO concept elements will be implemented as automated DSTs that support surface operations. These automated DSTs are envisioned to support (1) airport configuration management, (2) runway assignment, (3) scheduling and sequencing, and (4) taxi routing[1]. Descriptions of the mid-term capabilities for the STBO DSTs follow.

1) *Airport Configuration Management*

Airport configuration management capabilities will assist Air Traffic Control in planning and decision-making for airport configuration changes to improve the utilization of runways. These automated capabilities will be based on arrival and departure demand, with consideration of external factors, such as weather. The prescribed airport configuration will include runways for arrivals, departures, or mixed use; taxiway segments, whether open or closed; and standard operating procedures for the airport. The airport configuration capabilities support collaborative decision making among the NAS domains, coordination with the airport authority, and dissemination of the configuration change to the stakeholders. The concept for airport configuration management will apply to both surface arrival and departure operations and will enhance the efficiency of these operations.

2) *Runway Assignment*

Runway assignment capabilities will assist controllers in early planning of flight-specific departure runway assignment based on factors such as the filed departure route, the airport configuration, aircraft type, projected runway loading, and flight operator preferences, requirements, and limitations [5]. Departure runway assignments for a flight and any updates will be shared with flight operators and available to NAS domains. The automation capabilities will display runway assignment recommendations to Ground Control, who will be able to enter or modify a runway assignment for a flight. Runway Assignment capabilities will enable planning for surface operations by flight operators and air traffic management as well as will improve airport throughput by balancing departure runways.

3) *Scheduling and Sequencing*

Scheduling and Sequencing capabilities will assist controllers in managing the surface schedule and runway sequences for both arrivals and departures and will support collaboration with flight operators. Sequencing support provides guidance on the suggested order of aircraft in a queue, while scheduling support involves time-based recommendations. The surface schedule and sequence will satisfy traffic management constraints and controlled departure times and will optimize the use of surface resources to meet the demand. The automated capabilities will provide sequencing recommendations based on the data known about the flight, such as the time the flight entered to the airport movement area and the location of the entrance to airport movement area relative to the departure runway. Scheduling and sequencing capabilities will improve airport throughput, provide solutions that comply with NAS constraints, improve flight operator efficiencies, and support collaboration with flight operators to meet their business needs.

4) *Taxi Routing*

Taxi Routing capabilities address both taxi route generation and surface conformance monitoring. Taxi route generation capabilities will consider current aircraft position, aircraft surface destination (e.g., the assigned departure runway), and the airport configuration (runway use and taxiway or runway closures), user preferences for gate and taxiway, and other relevant factors to propose to the controller an appropriate taxi route for a flight. In earlier operational timeframes, these are pre-defined taxi routes, while later timeframes may also include ad hoc routes, and farther into the future, full surface trajectories. Taxi route generation capabilities also support data link of the taxi route to the aircraft. Surface conformance monitoring capabilities will assist the controllers in monitoring the aircraft's conformance with its assigned taxi instructions. Taxi routing and surface conformance monitoring capabilities serve as the focal point for the set of simulations described in this paper. Thus, a more detailed description of the taxi routing decision support tools and a set of use cases for these capabilities are described below.

II. TAXI ROUTING DECISION SUPPORT TOOLS

A taxi route describes the path on the surface that an aircraft uses to traverse the airport movement area from one location to another. In current operations, ground control receives a pilot request for taxi and uses information he or she knows about the airport configuration, operational procedures, and flight information (e.g., departure route and aircraft size) to develop an operationally acceptable taxi route. These routes may be pre-defined standard routes that are commonly used at an airport facility or may be determined ad hoc to fit a specific situation. An operationally acceptable taxi route is one that is efficient and does not compromise safety. In current operations, the taxi route is voiced to the pilot over the radio frequency, which may result in misunderstandings or read-back errors. Today's tactical response to a pilot's call for taxi can

[1] Also supporting surface operations is an additional DST that supports departure routing. This capability is also planned to be supported on the TFDM platform as one of the integrated DSTs, but is outside the scope of current STBO concepts and will not be addressed further in this paper.

inhibit or limit efficiency in the current operation. In addition, the taxi route information is not stored electronically in any way so it is not available to be used by other capabilities or systems.

A. Taxi Routing under Nominal Conditions

Automated taxi route generation capabilities will consider a flight's surface origin and destination and take into account the airport configuration (i.e., both current and planned configurations); the runway assignment, which is provided by the runway assignment DST; flight operator preferences; and other relevant factors to propose an appropriate taxi route to the controller. The taxi route generation capabilities in the NextGen mid-term will also provide automation support for selection of a pre-defined taxi route. To enhance safety, a hold-short will be inserted automatically when a taxi route crosses a runway. To ensure that the controller always has final authority over an assigned route, the ground controller will be able to modify a suggested route, add any additional instructions (for example, a hold short at a merge point to sequence traffic), or develop a manual ad hoc route. Under the mid-term vision, a controller will then verbally issue the route to the pilot.

Taxi route generation will improve efficiencies in surface operations by optimizing taxi routes and providing a digital route instruction for use with other automation capabilities, including data link to the pilot which will reduce taxi route misunderstandings and read-back errors. To demonstrate the benefit of the taxi route generation automation, a scenario is presented below to highlight the use of this tool under nominal conditions.

In this scenario for a departing flight, the ramp area is controlled by a ramp tower. When a flight plan is filed and an expected departure runway assignment is generated, an expected taxi route will also be generated and these data and any updates will be shared with flight operators and available for use by other NAS domains. Both the flight operator and the surface automation will use the expected runway assignment and expected taxi route in surface operations planning. The pilot of a departing flight will call the ramp tower when the pilot is ready for pushback. The ramp tower manages and coordinates with ATC regarding the flight's pushback and provides taxi instructions for the flight to maneuver to a location where the pilot can enter the active movement area. As the pilot begins pushing back, the taxi routing capability in the STBO automation in the ATCT provides a suggested taxi route assignment to Ground Control, displayed as part of the electronic flight data. When ready to enter the airport movement area (AMA), the pilot calls Ground Control for taxi instructions and clearance onto the AMA. When the pilot calls for taxi, Ground Control reviews the suggested departure runway (surface destination of the taxi route), the suggested taxi route, and reviews the current situation (via a surface surveillance display, electronic flight data display, or the window view). In this scenario, the taxi route to runway 17R is to taxi via Kilo, turn right at K7, turn left at Lima hold short of Zulu, proceed on Lima, right on EH. See Figure 1.

Work was funded by the Federal Aviation Administration

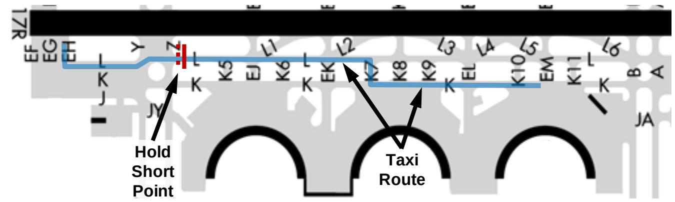

Figure 1. Taxi route.

This is a standard taxi route and the initial segment to the hold short of Zulu is voiced to the pilot as: "…runway 17R, taxi via Kilo, Kilo Seven, Lima, hold short of Zulu." The pilot acknowledges the route and begins to taxi while the Ground Controller pushes a single button to acknowledge the taxi route and indicate in the automation that the flight is taxiing. To release the flight from the hold short, the Ground Controller contacts the pilot and voices: "…continue taxi via Lima and Echo Hotel, monitor tower." The pilot acknowledges the remainder of the route while the Ground Controller pushes a button to remove the hold short from the automation and to forward the electronic flight data to the Local Controller who manages arrivals and departures on the runway.

B. Surface Conformance Monitoring

In surface conformance monitoring, an aircraft's location in the airport movement area is compared with its assigned taxi route to determine whether the aircraft is on its assigned route or is out of conformance with its assigned route. In today's operation, conformance monitoring is one of Ground Control's responsibilities and depends on human observation out the window. Today's surface surveillance, Airport Movement Area Safety System/Airport Surface Detection Equipment Model X (AMASS/ASDE-X), provide alerts for imminent safety issues: when taxiing aircraft in the airport movement area enter a runway that has moving aircraft on or approaching the runway. Currently, there is no automation support for controllers in monitoring conformance to taxi routing. When Ground Control observes an aircraft deviate from its assigned taxi route or when there is an alert by surface surveillance of an imminent safety issue, the controller focuses on the situation until it is resolved. This focus diminishes Ground Control's attention to other aircraft or tasks, potentially causing other tasks to be delayed until this issue is resolved.

In the NextGen timeframe, surface conformance monitoring will be reliant on surface surveillance capabilities. The automation will compare the aircraft's current location with its assigned taxi route and will determine if the aircraft is in conformance or out of conformance. When the aircraft is out of conformance, alerting will be based on severity of the impact of the nonconformance. The concept envisions a visual alert to the controller and an additional aural alert for safety-related surface conformance issues. There will also be support for out-of-conformance alerts to the flight crew via data link. Out-of-conformance conditions include, for example, an aircraft failure to turn when there is a turn in the assigned taxi route and an aircraft turning off the assigned taxi route when the turn is not in the assigned taxi route.

In addition to these location-based conditions, out-of-conformance conditions that relate to aircraft velocity will also be defined. One example includes an aircraft failing to move when instructed to move, such as an aircraft that fails to begin the take off roll when cleared for take-off. Another example of a velocity-based nonconformance condition includes an aircraft moving without being instructed to move, such as an aircraft passing a hold-short point or beginning a take-off roll without clearance. Automated surface conformance monitoring will support increased efficiency of surface operations by ensuring taxiing aircraft compliance with assigned taxi routes and airport safety by detecting out of conformance aircraft movement that may contribute to runway incursions, potentially reducing the opportunities leading to runway incursions.

Using the scenario described above (see Figure 1), the identified flight can be out of conformance with its assigned taxi route in several ways. First, the flight may make a turn at Kilo Eight, turning off of the prescribed route. In this case, Ground Control would be alerted of this out-of-conformance condition with visual alerts on the map and flight display. An aural alert would not be provided, as this event was not currently safety critical as the flight is not threatening a runway. In responding to the visual alert, Ground Control would assess the situation and take corrective action to address the nonconformance condition. The corrective action may involve the controller modifying the taxi route in the automation to include the wrong turn or communicating with the pilot to clarify the remainder of the route. When the taxiing flight is back in conformance with its assigned route, the alert would clear.

Second, the flight may fail to make the turn on Lima, continuing straight on Kilo Eight and approaching the runway. If the aircraft failed to hold short of the runway, a safety critical alert would be issued to the controller in both the visual and auditory modality.

Assuming the aircraft has made the appropriate turn at Lima and is heading along the route displayed in Figure 1, and additional nonconformance event could occur if the aircraft fails to hold short of Zulu. This non-safety critical event would be alerted visually to Ground Control. In responding to the alert, Ground Control would assess the situation and take necessary action to address the error by instructing the pilot to stop or hold short of a subsequent intersection if needed.

III. RESEARCH QUESTIONS

Extensive work has been conducted defining the concept of use for these new technologies [4, 5]. The scenarios described above provide a concrete representation of how these technologies are envisioned to work and be used under NextGen. Despite these accomplishments, a series of research questions need to be addressed before these technologies can be implemented in the operational environment. These questions address both the measurable benefit associated with the taxi routing and conformance monitoring decision support tool and the acceptance of these capabilities by the operational community:

- Do controllers assign automation-generated taxi routes to departure aircraft?
- Do surface conformance monitoring capabilities improve controllers' response to nonconformance events?
- Do controllers trust the automation?
- Do changes in traffic load impact these findings?

IV. HUMAN-IN-THE-LOOP SIMULATIONS

A roadmap has been developed by The MITRE Corporation's Center for Advanced Aviation System Development (CAASD) to outline a set of human-in-the-loop simulations to address the research questions stated above. This roadmap describes a systematic approach for conducting a total of six evaluations, which will empirically evaluate the five STBO decision support tools that are envisioned under NextGen. These simulations are projected to occur biannually through 2012.

This paper details the methods and results from the first two simulations, which took place during 2010. These simulations were designed to examine the human performance implications associated with the use of initial prototype taxi route generation capabilities, including the surface conformance monitoring tools that are projected to be implemented in the mid-term plan for NextGen. These simulations examined the use of these technologies by trained air traffic controllers, employed by the Federal Aviation Administration (FAA), in a high-fidelity tower simulation. The following sections outline the methods used to evaluate the stated research questions, as well as results produced from the simulations.

A. Simulation 1

1) Methods

a) Participants: Thirteen FAA employees, who reported an average of 22.5 years of experience controlling traffic, participated in the study. At the time of the simulation, seven individuals were serving in roles at the FAA Headquarters, while the remaining six participants came from facilities including Minneapolis – St. Paul International (MSP), Great Falls International (GTF), Phoenix Sky Harbor International Airport (PHX), Long Beach (LGB), Baltimore Washington International (BWI), Albuquerque International Sunport (ABQ). The participants ranged in age from 32 to 56 years (M = 46.7 years). Three of the participants were female, while the remaining ten particpiants were male. None of the participants had worked at Dallas Fort Worth International Airport (DFW), which served as the simulated site for the study.

b) Simulation Environment: The simulation was conducted at the MITRE Corporation's Center for Advanced Aviation System Development's (CAASD) Aviation Integration Demonstration and Experimentation for Aeronautics (IDEA) laboratory. CAASD's air traffic control tower simulator was used to generate a high fidelity representation of DFW during a standard south flow operation (see Figure 2).

Figure 2. Air traffic control tower simulator.

Configurable workstations were located directly below the out-the-window scene and included displays to emulate the Digital Bright Radar Indicator Tower Equipment (DBRITE) system. In addition, workstations were outfitted with an electronic flight strip display and a surface surface map display that were developed by Mosaic ATM. These displays are shown in Figure 3.

Figure 3. Electronic flight strip display and surface map, developed by Mosaic ATM.

c) Procedure: Each controller was provided with a half day of training, during which they were introduced to DFW procedures and layout and familiarized with the displays. Participants were then asked to serve as the ground controller in the tower, directing departure traffic from the ramp area to the runway edge. Participants were responsible for verbally assigning a taxi route to pilots, monitoring pilots' conformance to that route, and revising the route when necessary. Participants were given one 45-minute practice trial and four 45-minute experimental trials. After each trial, participants completed questionnaires to capture subjective ratings of their trust in the provided systems and acceptance of the automated tools.

To examine participants' ability to detect and respond to pilot deviations, four nonconformance events were randomly distributed within each of the scenarios. These events were comprised of lateral deviations that occurred from erroneous turns or from pilots' failure to turn at a correct intersection. When participants detected a pilot deviation, they were asked to press a button on the electronic flight strip display and verbally respond to the pilot to correct the error. None of the nonconformance events in this simulation resulted in a safety or sequencing concern.

d) Experimental Design: In order to address the proposed research questions, two independent variables were manipulated in a within-subjects design: (1) the presence of automation and (2) traffic load. First, automation was manipulated to produce two levels. In the manual condition, participants were provided with the displays shown in Figure 3; however, no taxi routing or conformance monitoring capabilities were provided. In the automation condition, decision support tools were used to generate taxi routes based on the assigned departure fix. In addition, automation was also provided to alert the controller to a pilot deviation. When a nonconformance event was identified, the electronic flight strip for the aircraft was highlightged in red (see Figure 3) and an aural alert was presented. While the aural alert was terminated after five seconds, the visual alert persisted until the aircraft had rejoined the route or until a new route had been manually entered on the strip. In addition to manipulating the presence of automation, traffic load was also systematically varied. In the low traffic load scenarios, traffic flows included 41 departures and 23 arrivals per hour. In the high traffic load scenarios, traffic flows were increased by nearly 15% to 48 departures and 25 arrivals per hour.

2) Simulation 1 Results

The following sections outline participants use and subjective assessment of the taxi route generation and surface conformance monitoring capabilities. A more detailed review of these findings is also available [6, 7]

a) Taxi Route Generation: In the automation condition, participants were provided with capabilites that automatically generated taxi routes and presented these routes on the strip. In all cases, these capabilities performed without error.

An analysis was conducted to examine the frequency with which participants manually modified routes that were generated by the automation. This analysis indicated that, in all cases, controllers used the automation-generated taxi routes and never modified these routes. This finding is not necessarily surprising, as the automated taxi route assignments always matched the rules provided to participants during training. To further explore this issue, an additional analysis was performed to examine how frequently participants modified routes in response to a pilot's deviation from the assigned route. This analysis also indicated that controllers never modified taxi routes when a nonconformance event occurred. Rather, controllers always provided corrective taxi instructions to return deviating aircraft to their original route, even if the original route was no longer the most efficient.

Additional discussions with each participant indicated that these individuals were often reluctant to make changes when nonconformance events occurred because characteristics of the route editing interface made these changes difficult. Participants noted that the interface, which provided an alphabetical keyboard for input, would require significant modifications before being suitable for implementation in an operational environment.

In addition to examining participants' use of the taxi routing capabilities, participants were also asked to provide subjective ratings of trust in the tool. On average, participants rated their trust of this decision support tool as 8.58 (out of 10.0), which is remarkably high for a prototype system that has not been extensively used, but not necessarily surprising given that the performance of the tool was perfect.

To examine the impact of traffic load on trust, a paired t-test was conducted on these data. This analysis revealed that traffic load significantly impacted ratings of trust ($t_{(12)} = 2.16$, $p = 0.05$). Specifically, participants reported higher levels of trust in the automatically generated taxi routes when traffic load was high ($M = 8.70$) as compared to when it was low ($M = 8.45$). This finding likely resulted from participants' need for greater reliance on this capability when workload was elevated.

b) Surface Conformance Monitoring: Statistical analysis indicated that the presence of surface conformance monitoring automation strongly influenced participants' performance ($F_{(1, 11)} = 3.41$, $p = 0.001$). While participants detected less than half of events when no automation was provided ($M = 46.7\%$), *all* events were detected with the aid of the nonconformance alerts. Average response time was found to be 23.1 s, though automation was not found to improve the speed with which participants responded ($p > 0.10$). Variations in traffic load were not found to affect these measures.

When participants were asked to rate their trust in the nonconformance alerting capabilities, average trust was reported to be 7.48 (out of 10.0). In addition, ratings of trust were found to be higher when traffic load was high ($M = 7.46$) relative to low ($M = 7.19$; $t_{(12)} = 2.16$, $p = 0.05$), mirroring the results reported for the taxi routing capability.

While the reported value of trust was relatively high, additional analysis indicated that trust was largely impacted by the performance of the automated alerts. The simulation was designed to examine perfect alerting performance, though there were documented cases where the logic of the algorithms, when paired with certain actions or states in the simulation environment, produced false alerts. Because a significant body of research exists to suggest that frequent false alerts can impact trust [8], a correlation was conducted to examine the relationship between the number of false alerts to which a given participant was exposed with subjective ratings of trust on that trial. This analysis indicated a significant negative correlation between false alarm frequency and subjective ratings of trust ($r = -0.35$, $p = 0.10$). This relationship is shown in Figure 4.

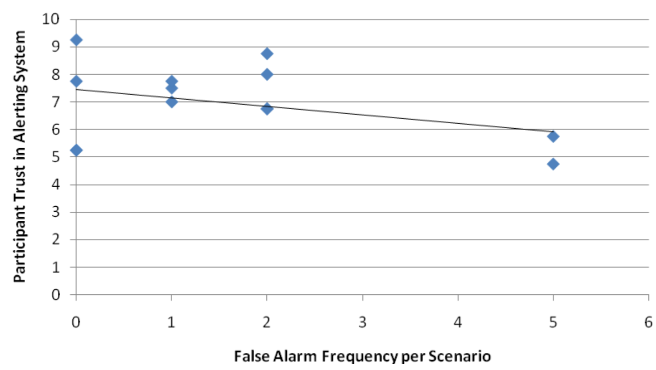

Figure 4. Correlation of false alarm frequency and trust.

c) Summary: The described results from the study demonstrate the potential performance benefits associated with the surface conformance monitoring capability, included in the STBO concept. In addition, participants indicate high trust in the alerting and taxi route generation capabilities needed to support these future operations. However, the nonconformance events presented in Simulation 1 were limited in scope and did not include an examination of the effectiveness of these systems in supporting the detection of safety critical deviations. Simulation 2 explored controllers' use of the described technologies for a broader range of events.

B. Simulation 2

1) Methods

a) Participants: Twelve participants took part in the simulation and all were current employees of the FAA at the time of the study. Nine of the participants were male, while the remaining three participants were female. The participants ranged in age from 27 to 51 years ($M = 43.3$ years). On average, participants reported 17.2 years of experience controlling traffic. Participants came from FAA Headquarters, as well as facilities including Washington Dulles International (IAD), Los Angeles International (LAX), Ronald Reagan Washington National (DCA), St. Paul Downtown Airport/Holman Field (STP), John F. Kennedy (JFK) and Raleigh Durham (RDU). Note that no participants currently or previously worked at Dallas-Fort Worth International Airport (DFW), which served as the simulated location for the study.

b) Simulation Environment: The second simulation used the same environment as the first simulation, including the electronic flight strip display and surface map shown in Figure 3.

c) Procedure: As in the first simulation, each controller was provided with a half day of training, and was asked to serve as the ground controller in the tower during data collection. Participants had the same responsibilities as in the first simulation, and were given the same number of trials. After each trial, participants completed questionnaires to capture their trust in the provided systems and acceptance of the automated tools.

To examine participants' ability to detect and respond to pilot deviations, between six and seven nonconformance events were randomly distributed within each of the scenarios. Some of these events were comprised of lateral deviations that occurred from erroneous turns or from pilots' failure to turn at a correct intersection. In addition to lateral path errors, hold short deviations occurred in each scenario, during which a pilot failed to hold short at either a taxiway or a runway crossing. The path deviations and the hold short deviation at the taxiway crossing were non-safety critical events. The hold short deviation at the runway crossing was a safety critical event that resulted in a runway incursion but not one that would have generated an ASDE-X alert. The simulation confederate local controller held departing traffic around the time of the safety critical nonconformance event to insure there was no potential for a collision. When participants detected a pilot deviation, they were asked to press a button on the electronic flight strip display and verbally respond to the pilot to correct the error.

d) Experimental Design: The second simulation utilized the same experimental design as the first simulation, with two levels of automation (automation and no sutomation) and two level of traffic load (low and high) crossed in a within-subjects design. In the low traffic load scenarios, traffic flows included 46 departures and 21 arrivals per 45 minutes. The high traffic load scenarios contained 53 departures and 21 arrivals per 45 minutes. The low traffic load scenarios contained seven nonconformance events, while there were six events in the high traffic load scenarios. The order of scenario presentation was randomly selected across the thirteen participants from the full set of permutations associated with a completely counterbalanced design.

2) Simulation 2 Results

Key results from the simulation are summarized in the section below. A full review of these findings is available in a published technical report [9].

a) Taxi Route Generation: After each of the automation scenarios, participants responded to six questionnaire items that addressed whether and when they used system-generated taxi routing. They also responded to questions about taxi route modifications.

Over half of the participants (54%) indicated that they used the automated route generation function in one or both of the low and high traffic scenarios. In half of those responses, participants responded to a follow-on question and estimated how frequently they used it. These results indicated that the predominant pattern (66% of responses) was to use the automated route generator most of the time, as shown in Figure 5. Note that this finding indicates less reliance on the automated taxi routing capabilities than were found in Simulation 1. This difference may be attributed to changes that were made in the route editing interface for Simulation 2.

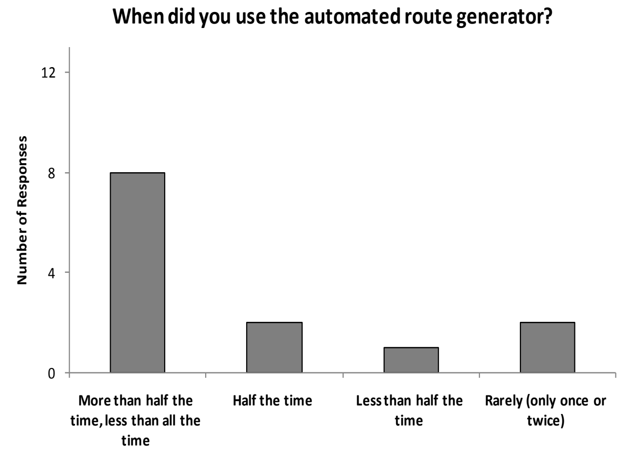

Figure 5. Relative frequency of use of the automated route generation function.

In this simulation, the keyboard used in Simulation 1 was augmented with a graphical editor interface that allowed editing on a map-based display by clicking on key taxiway intersections. These two route editing capabilities were available regardless of the automation condition. The ten questionnaire items that addressed the basic editing capabilities were analyzed as a group. Specifically, participants were asked to describe and assess their use of taxi route amendments to change the initial taxi route or to reestablish route conformance following a nonconformance alert. According to the responses, less than a quarter of the participants modified a taxi route in any scenario, for any reason, regardless of the automation condition or traffic load. Responses further indicated that only 19% of the participants used the keyboard while only 12% used the graphical interface to amend taxi routes. Thus, while we observed additional cases in which edits were made to the taxi routes, these cases were still rare, regardless of the editing methods used.

In fact, participant assessments of the graphical editor were moderately negative, indicating that they did not find the new capability to be particularly helpful. In fact, participants indicated a moderate preference for using the keyboard over the graphical editor for modifying routes. Commentary about the graphical editor capability expressed concern over the extensive head-down effort involved. These results indicate that the graphical editor interface should be modified to make it easier for the controllers to input both initial and modified taxi routes. One possible method is to only show the area of interest (taxiways and runways) specific to the selected aircraft. As the taxi route is selected, the algorithm can expand the area if needed.

b) Surface Conformance Monitoring: An analysis of variance (ANOVA) was conducted to examine the impact of both automation and traffic load on nonconformance event detection accuracy. This analysis revealed a significant main effect for automation ($F_{(1, 8)} = 6.65$, $p = 0.033$), indicating that participants detected less than half of events when no automation was provided for support ($M = 44.2\%$) and a majority of events when alerting was provided ($M = 86.44\%$).

This effect is depicted graphically in Figure 6. The main effect of traffic load was not significant, nor was the interaction between traffic load and automation presence. This finding was comparable to the results described for Simulation 1, where the detection rate in the no automation condition was 46.7%. In Simulation 1, however, the automation resulted in an increase in the detection rate to 100%, while here the increase was to 86.44%. This difference in Simulation 2 can be attributed to participants who failed to indicate that they had detected the nonconformance event by pressing an indicated key on the display interface. However, the data in both simulations indicate a clear benefit to detection accuracy.

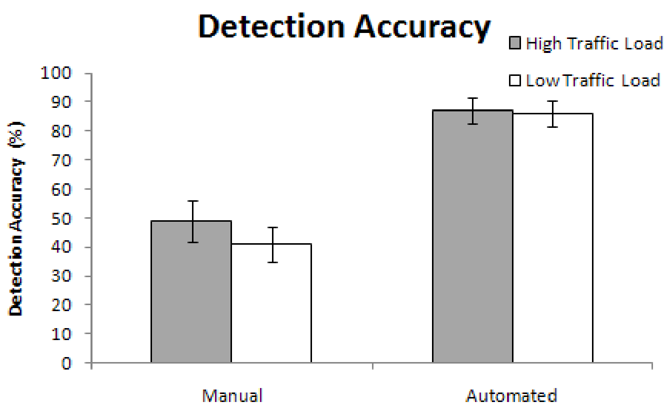

Figure 6. Main effect of automation on nonconformance detection accuracy.

To further examine participants' detection performance, analyses were conducted on non-conformance detection response times. Paired t-tests were conducted to evaluate the impact of automation and traffic load separately as main effects. As with detection accuracy, automation had a significant impact on detection response time ($t_5 = 2.93$, $p = 0.03$), indicating participants detected nonconformance events faster under the automated condition ($M = 8.1$ s) than under the manual condition ($M = 22.9$ s), as shown in Figure 7. The effect of traffic load on detection response time was not significant.

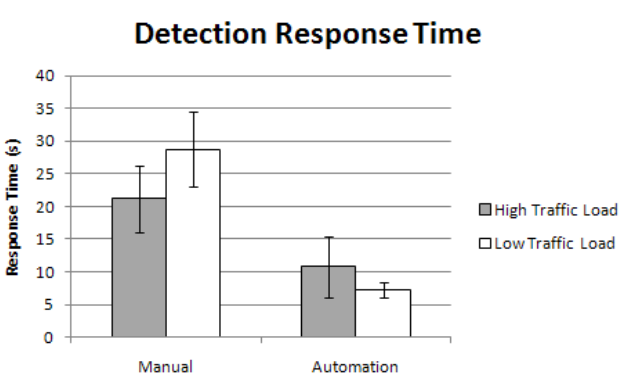

Figure 7. Effect of traffic load and condition on nonconformance detection response time.

To explore any possible differences in response time between those events that were considered safety critical (crossing a runway without clearance) and those that were not as safety critical (path deviation or hold short deviation) paired t-tests were run on the data. The analysis found a significant difference in response times between the two types of nonconformance events, with response time to the safety critical events shorter ($M = 9.3$s) than response times to the other events ($M = 27.4$s; $t_3 = -3.53$, $p = 0.04$), though this effect was only present in the manual condition. This relationship is shown graphically in Figure 8.

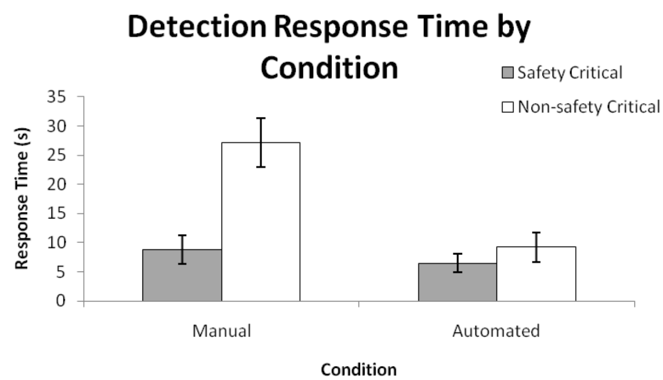

Figure 8. Detection response time to safety critical and non-safety critical nonconformance events.

c) Summary: Results from Simulation 2 provide additional evidence of the viability of the taxi routing decision support tool and surface conformance alerting, and the basic effectiveness of these capabilities in supporting efficient surface operations. The following section outlines the broad conclusions drawn from the two simulations, notes the potential limitations that must be considered, and highlights future research that must be done before STBO is used in operations.

V. Summary and Conclusions

The human-in-the-loop simulations described in this paper were designed to provide an empirical evaluation of the feasibility and potential benefits associated with the use of the taxi routing and surface conformance monitoring automation proposed in STBO. In the initial sections of this paper, we reviewed the concept of use for these decision support tool capabilities and provided scenarios that illustrate its two primary capabilities: taxi route generation and surface conformance monitoring. We also proposed a set of research questions that must be answered before the operational use of these technologies will be possible. The findings are summarized below.

A. Taxi Route Generation

Under STBO, decision support tool capabilities will provide controllers with recommended taxi routes to move departure aircraft from the ramp area to the runway. The two human-in-the-loop simulations conducted at CAASD in 2010 indicated that participants used the automation-generated taxi routes in nearly all cases. While participants did indicate that it was difficult to edit taxi routes, which may have attributed to their reluctance to readily make changes to these routes, these participants also indicated that the automation generating these routes was reliable and trustworthy. Collectively, these findings indicate the viability of the route generation concept, while simultaneously pointing to the need to develop intuitive and low workload methods for revising taxi routes on an electronic flight strip interface.

While these simulations provide a basic evaluation of the taxi route generation tool concept, additional research is still needed to closely examine taxi route assignment under more dynamic and complex scenarios. For example, extensive research is still needed to examine the impact of airport configuration changes or full airport turn-arounds, as well as the taxi routing changes that result from these changes, on the workload of local and ground controllers.

B. Surface Conformance Monitoring

Surface conformance monitoring for pilot deviations is a key capability in the STBO concept. In the simulations described here, the extent to which these capabilities improve controllers' responses to these deviations was evaluated. Results from the present study indicate that providing surface conformance monitoring automation results in nearly perfect detection of these events, doubling the detection rates found when automation was not present. Results also point to the finding that controllers detect these events more quickly with automated support, at least for non-safety critical deviations, which are typically detected very slowly. While this latter finding may not be associated with a direct safety benefit, it can be tentatively concluded that more rapid detection of non-safety critical events is likely to improve efficiency and may also reduce the likelihood that non-safety critical deviations become safety critical errors (e.g., a lateral deviation results in an aircraft approaching a runway entrance).

Participants also indicated that the surface conformance monitoring automation was trustworthy in both simulations. These individuals were found to trust the automation more when traffic load was amplified. This finding likely reflects participants' need to rely on this automation to a greater degree when surface traffic density increased. Finally, research from Simulation 1 provided some initial evidence that false alerts, which will exist in the operational world when surveillance data is imperfect, impacts trust and may be likely to impact controllers' use of these tools [6].

The STBO simulations provided a baseline evaluation of non-safety critical and safety critical nonconformance events. To fully exercise these technologies, additional research will be needed to examine a broader range of nonconformance events (e.g., taking off without clearance) and an assessment of multi-level alerting. Research is also necessary to evaluate the use of conformance monitoring for taxi routes related to performing scheduling and sequencing tasks.

C. Summary and Future Simulations

While the two presented simulations provide a foundation for understanding the use of basic taxi route generation and surface conformance monitoring technologies, significant research is needed to understand and quantify the impact of these technologies on surface operations. This research need will continue to be addressed through a series of human-in-the-loop simulations expected to evaluate the use of the STBO decision support tools in a simulated airport environment with heightened complexity and fidelity. Specifically, these evaluations will examine the use of these technologies by the local and ground controllers; the effectiveness of multi-level alerting for a broad range of nonconformance events (e.g., failure to takeoff with clearance, takeoff without clearance, failure to hold short at runway and taxiway intersections, time-based conformance during taxi); the impact of surveillance noise; and taxi routing for scheduling and sequencing.

ACKNOWLEDGMENT

The authors would like to thank all of the participants who provided invaluable feedback to the experimental team. The authors would also like to thank Paul Diffenderfer for providing subject matter expertise that was critical in building a realistic operational environment for testing; Dave Tuomey for developing MITRE's simulation environment; and Mosaic ATM for software development efforts on the surface automation prototype.

REFERENCES

[1] Air Traffic Organization, NextGen and Operations Planning, Research & Technology Development, Air Traffic Systems Concept Development,, NextGen Mid-Term Concept of Operations for the National Airspace System, Version 2, April 2010, Federal Aviation Administration, Washington, DC.

[2] V. Nene, C. E. Morgan, P. A. Diffenderfer, and A. P. Colavito, , A Mid-Term Concept of Operations for a Tower Flight Data Manager (TFDM), MP090169R1, September 2010, The MITRE Corporation, McLean, VA.

[3] L. F. Audenaerd, C. S. Burr, P. A. Diffenderfer, and C. E. Morgan, Surface Trajectory-Based Operations (STBO) Mid-Term Concept of Operations Overview and Scenarios, MP090230R1, September 2010, The MITRE Corporation, McLean, VA.

[4] C. E. Morgan and P. A. Diffenderfer, Mid-Term Surface Trajectory-Based Operations (STBO) Concepts of Use: Summary, MTR100269V1, September 2010, The MITRE Corporation, McLean, VA.

[5] C. E. Morgan, Mid-Term Surface Trajectory-Based Operations (STBO) Concepts of Use: Departure Runway Assignment, MTR100269V4, September 2010, The MITRE Corporation, McLean, VA.

[6] K. A. Klein, E. K. Stelzer, E. T. Nelson, C. Brinton, S. Lent, Improving Efficiency with Surface Trajectory-Based Operations and Conformance Monitoring, Proceedings from the Digital Avionics Systems Conference, October 2010, Salt Lake City, UT.

[7] E. K. Stelzer, Human-in-the-Loop Concept Evaluation of Surface Conformance Monitoring Automation, MTR100188, June 2010, The MITRE Corporation, McLean, VA.

[8] R. Parasuraman, R. and V. Riley, "Humans and automation: Use, misuse, disuse, abuse," Human Factors, 39, 1997, pp. 230-253.

[9] K. A. McGarry, K. Kerns, Results of a Second (Controller) Human-in-the-Loop Simulation Study of Automated Capabilities Supporting Surface Trajectory-Based Operations, MTR100483, December 2010, The MITRE Corporation, McLean, VA.

AUTHOR BIOGRAPHIES

Emily K. Stelzer received a BA in psychology from the University of Cincinnati, Cincinnati, OH, USA in 2000; MA in engineering psychology from the University of Illinois at Urbana-Champaign, Champaign, IL, 2002; and PhD in engineering psychology from the University of Illinois at Urbana-Champaign, Champaign, IL, 2005.

She served as a Senior Scientist and Team Lead of the Human Automation Interaction group at Aptima, Inc. in Washington, DC from 2005 until 2009. She is currently working as a Multi-Discipline Systems Engineer at MITRE's Center for Advanced Aviation Systems Development (CAASD) in McLean, VA.

Dr. Stelzer is a member of Human Factors and Ergonomics Society (HFES).

Constance E. Morgan received an AB in liberal arts from Brown University, Providence RI, USA in 1973; and an M.S. in administration from George Washington University, Washington DC, USA in 1980.

She is currently working as a Lead Multi-Discipline Systems Engineer at MITRE CAASD in McLean, VA.

Kathleen A. McGarry received a BA in psychology from the College of New Rochelle, New Rochelle, NY, USA, 1993; a MA in applied experimental psychology from The Catholic University of America, Washington, DC, 2004; and a PhD in applied experimental psychology from The Catholic University of America, Washington, DC, 2007.

She was a Research Assistant in the Cognitive Science Laboratory, a Human Factors Engineer with Titan/L3 Communications, and is currently a Lead Human Factors Engineer with MITRE CAASD.

Dr. McGarry is a member of HFES.

Kathryn A. Klein was born in Leesburg, Virginia, USA in 1980 and received a BA in economics from the University of Virginia, Charlottesville, VA, USA in 2003 and an MS in systems engineering with a concentration in human factors from the University of Virginia, Charlottesville, VA, USA in 2005.

She is currently working as a Lead Simulation Modeling Engineer and Project Team Manager for future air traffic control tower concepts at MITRE CAASD.

Ms. Klein is a member of HFES.

Karol Kerns received her B.A. degree in Psychology from LaSalle University, Philadelphia, PA in 1974 and earned her M.S. and Ph.D. degrees in experimental psychology from Saint Louis University, St. Louis, MO in 1976 and 1980.

She currently serves as a Principal Research Scientist with MITRE CAASD in McLean, VA and has worked for MITRE CAASD since 1983. She has authored and coauthored numerous technical papers and book chapters in the areas of ATC/Flight Deck integration, data link and human factors in air traffic management. While at MITRE CAASD, she has also completed two, two-year interagency assignments stationed at Federal Aviation Administration (FAA) Headquarters. On each of these, she led human factors projects to plan and coordinate human factors research, development, and implementation activities in support of the Aeronautical Data Link and the Next Generation Air/Ground Communications System programs.