NASPlay: A Serious Game for Air Traffic Control

Hayley J. Reynolds, Brian C. Soulliard, and Richard A. DeLaura

A serious game developed for training air traffic managers and for exploring new procedures in air traffic management enables participants to gain broad experience with traffic management decision making and the repercussions of the decisions. The game gives operators the opportunity to tackle in a day or two the decisions that they would normally encounter throughout a whole year or more.

»

In 1981, the National Airspace System (NAS) incurred a massive influx of new air traffic controllers brought on by the Reagan-era firings of more than 11,000 striking members of the Professional Air Traffic Controllers Association. As the new controllers gained experience, many advanced into positions of air traffic management, directing not only single aircraft but also large flows of many aircraft around bad weather. Because the controllers hired in 1981 are now retiring, the air traffic management domain is facing a void in experience, with most current traffic managers having five or less years of experience. If the current and new traffic managers are prepared with only "on-the-job" training, the necessary mental models of weather and traffic behavior they acquire will be based on the job experience of just five years or less. To combat this lack of experience, MIT Lincoln Laboratory researchers created an air traffic management serious game, called NASPlay, to enable a trainee to experience a year's worth of difficult days in only a day or two.

There are several reasons to train personnel through a gaming approach rather than through practice drills on a series of canned, realistic scenarios. By developing a model of the NAS (which could be scoped to various levels of complexity and fidelity depending on the purpose of the game), game designers necessarily develop hypotheses of the causes and effects of decisions within a complex environment. Because this model of the NAS can subsequently be revised on the basis of newly discovered data from the actual NAS, a more fully developed user mental model of the system can form so that cause-effect linkages between users' decisions and the resultant states of the emulated system can be clearly outlined and carried forward to the real system. Another benefit is that game players can gain decision-making experience more quickly than would be possible in the real world. A game also allows players to take risks and experiment in ways that would be unwise on the job. In addition, the gaming environment has the potential for time manipulation of the game. For example, if the user concludes that a decision was a poor one, then he or she could go "back in time" to modify or delete the action taken, resulting in a different outcome. This back-in-time capability—along with the ability to view quantitative, scored outcomes—would encourage the user to modify and optimize a strategy through iterative trial and error, and a robust scoring metric would challenge the user to replay scenarios in order to maximize scores.

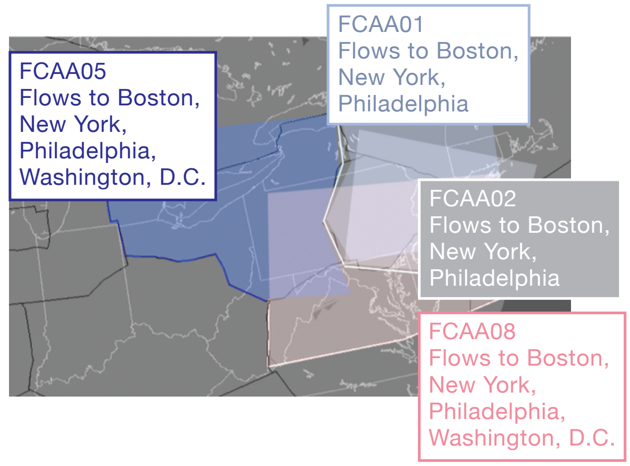


FIGURE 1. In the air traffic manager's computer display, the Flow Control Areas are demarked by lines "in the sky" that provide a means to control the rate of traffic traveling either north/south or east/west.

Traffic Flow Management

Traffic flow management (TFM) is performed at air traffic control facilities in the NAS to assess whether the traffic demand on the system exceeds the capacity of the system. If the demand exceeds capacity, traffic management coordinators (TMCs) must decide if and how to reduce that demand. This assessment of demand and capacity imbalances takes into account NAS resources, such as runways, routes, fixes (points along routes), and sectors. Options to reduce demand include delaying departure of flights (ground delay programs or GDPs), stopping the departure of flights to a particular airport altogether (ground stops), putting flights into holding patterns in the air, or changing the route that flights have requested. Airlines may also request that flights land at an airport different from the one planned (diversions) or cancel flights that they cannot complete because of capacity and resource constraints. At the national level, mass movements of traffic demand can occur if capacity is reduced across a large segment of the United States, as often happens during large thunderstorm fronts. Options that the national TMCs have include airspace flow programs (AFPs), which reduce the traffic demand over large airspace segments by delaying the departures of any flight flying through a Flow Control Area. Some examples of Flow Control Areas are shown in Figure 1.

National TMCs also have the option to reroute flights, using common strategic reroutes available in the National Severe Weather Playbook [1]. Examples of different scopes of reroutes are shown in Figure 2.

To assess the demand and capacity imbalances that are present across NAS resources, TMCs have several information systems available to them. Traffic demand for a resource can be assessed through either the Traffic Situation Display or the Flight Schedule Monitor, shown in Figure 3. The Traffic Situation Display allows the TMC to visually identify the points of possible congestion, both at the current time and the time projected into the future, through either a manual time slider or an automatic movie-like projection. The Flight Schedule Monitor's set time bins, which span into the future, provide aggregated counts of flights demanding a particular resource (e.g., airport or FCA). Fair-weather capacity for resources is shown in a horizontal line to enable TMCs to easily evaluate demand and capacity imbalance.

To determine if capacity has been impacted by weather, several weather tools are available. The most critical weather tool to assess convective weather (conditions that lead to thunderstorms) out to 2 hours into the future is the Corridor Integrated Weather System (CIWS) [2], which provides convective weather information and 0- to 2-hour forecasts covering the United States and southern Canada. The Consolidated Storm Prediction for Aviation (CoSPA) is CIWS's strategic counterpart, a

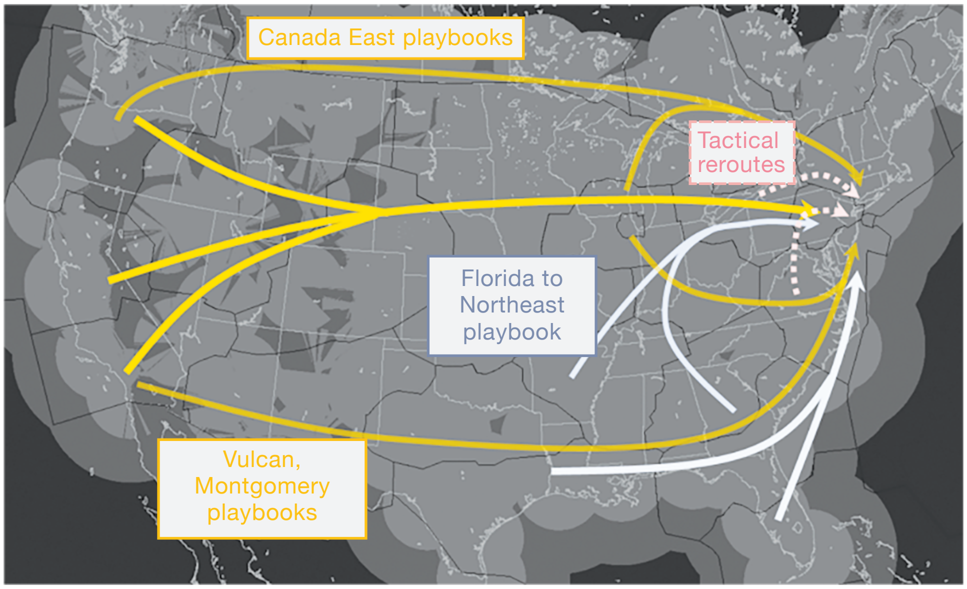
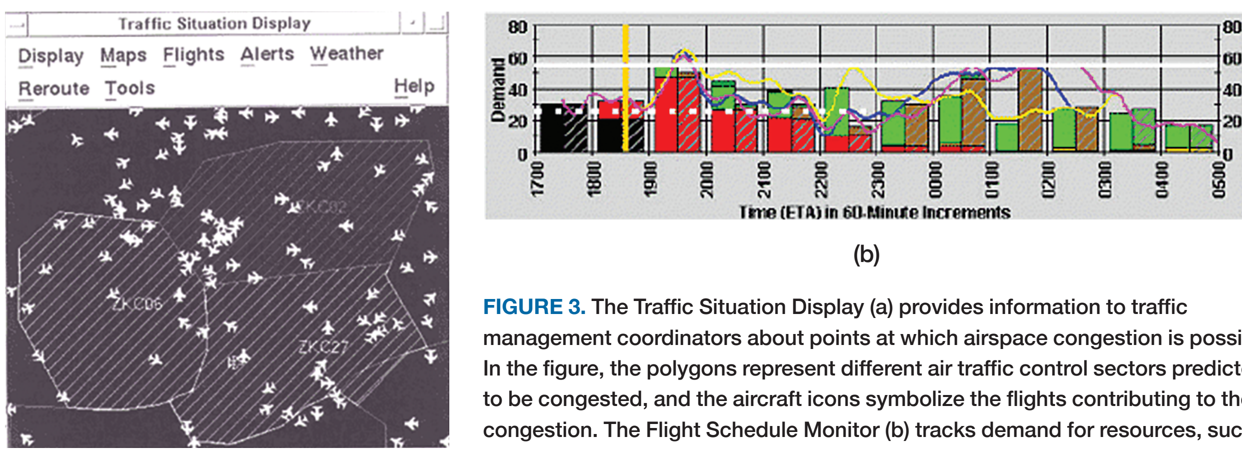


FIGURE 2. National severe weather reroute options are available to traffic management coordinators. Each colored line represents a different option in the National Severe Weather Playbook.



(b)

(a)

FIGURE 3. The Traffic Situation Display (a) provides information to traffic management coordinators about points at which airspace congestion is possible. In the figure, the polygons represent different air traffic control sectors predicted to be congested, and the aircraft icons symbolize the flights contributing to the congestion. The Flight Schedule Monitor (b) tracks demand for resources, such as airports or routes, in color-coded bins. Black represents flights that have already landed, red represents flights that are in the air currently, and green represents flights that have not yet departed.

prototype that provides deterministic weather projection 0 to 8 hours into the future [3]. The CIWS and CoSPA displays are shown in Figure 4.

The TMCs face several challenges in assessing and addressing demand and capacity imbalances. It can be difficult to evaluate the impact of adverse weather on route capacity. The varying severity of weather, pilots' unwillingness to fly through bad weather, and the inherent uncertainty in forecasts of weather and traffic demand all contribute to making the prediction of weather's impact on capacity an art. The Federal Aviation Administration (FAA) has a strong interest in keeping the NAS running at full capacity because any under-delivery of traffic costs the airlines, the FAA's "customers," money. In addition, multiple TMCs at different air traffic facilities can choose to address the demand in different, and often overlapping or conflicting, ways. Again, it is an art to determine which problems should be addressed at the national level (i.e., through AFPs and playbook reroutes) and which should be addressed tactically (i.e., through holding, ground stops, and tactical reroutes).

Applications of Gaming to Traffic Flow Management

Lincoln Laboratory's investigation into the benefits of a gaming approach to training revealed obvious applications of gaming to issues plaguing traffic flow management. Several issues with the NAS's TFM had been identified during the prototyping efforts on CIWS, the Route Availability Planning Tool (RAPT), Integrated Departure Route Planning, and CoSPA [4, 5]. Firstly, across the NAS, basic concepts of TFM are misunderstood by TMCs at the FAA facilities. To inform their decision making, TMCs are constantly striving for information that predicts evolving weather conditions as far in advance as possible; however, when more reliable information becomes available, they often fail to revisit their strategic decisions in tactical timeframes. In convective weather situations, when decision making has the most impact on traffic flow, TMCs often do not regularly reevaluate their strategic and tactical decisions to account for the fast-changing situation. Frequently, TMCs make decisions too early or too late, and they do not base decisions on the most diagnostic information available to them, preferring to rely on familiar information sources. In addition, traffic management experience is acquired slowly. Each day of convective weather provides only a single data point for the TMC to add to his or her experience. Moreover, traffic managers may not have the opportunity to learn from their decisions; for example, a TMC may make a decision at 8:00 a.m. but be off-shift before that decision shows (or does not show) results at 6:00 p.m. As researchers seek ways to improve TMC training, human-in-the-loop experiments to assess new training methods can be costly and TFM components can be difficult to replicate in the laboratory. Furthermore, because human-in-the-loop experiments are usually run in real time, they are time-constrained to address only a limited number of the wide range of scenarios that are likely to be encountered in real operations.

The gaming approach provides a means of addressing these TFM issues, particularly TMC training. Currently, training is performed in a classroom setting with TFM concepts conveyed in PowerPoint. The concepts that the lecturer is covering may or may not be directly connected with real traffic management scenarios or data and the information sources available to TMCs. A gaming environment could not only ensure that the TFM concepts are directly connected with NAS data and the information sources available to TMCs, but the TMCs could also experience dozens of situations that demonstrate the concepts. This immersive form of training could better encourage the development of strategic heuristics that, if executed correctly, could transfer positively to the real-world environment. Gaming is also a natural extension of the methods used to train air traffic controllers, who spend hours working a simulated traffic control environment.

Another TFM application of gaming is in the exploration of new procedures in consideration for their implementation into the NAS. Often, a new procedure is proposed by a TMC and then is modified and honed during actual operations, to the potential detriment of the flights exposed during the early, less effective iterations. If a game environment was available, new procedures could be explored to identify when to enact or retract a procedure, which traffic flows to target, and what amount of action to implement (e.g., what number of miles should separate aircraft) without adversely affecting any flights in the NAS until the procedure is mature. A third application is concept development for TFM products. New tools developed for TFM purposes could be tested and iterated in TFM games before being deployed into the field.

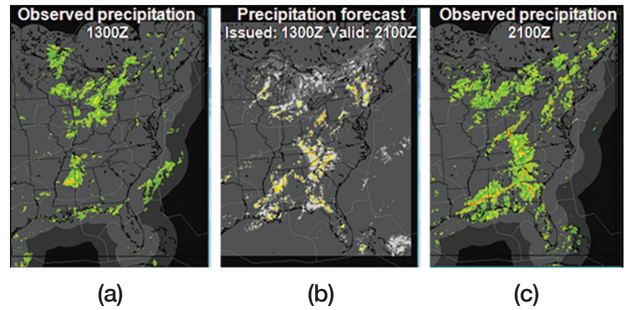


FIGURE 4. The Corridor Integrated Weather System (CIWS) shows the observed precipitation at 1300 zulu[1] (a). The CoSPA precipitation forecast issued at 1300 zulu for 2100 zulu is shown in (b). The actual precipitation at 2100 zulu is shown in (c). In the (a) and (c), green indicates light precipitation; yellow indicates medium precipitation, and orange/red is heavy precipitation. In (b), white represents light precipitation; yellow is medium precipitation, and red is heavy precipitation. While the forecast did not capture much of the light precipitation, it did capture the heavy precipitation and characterized well the type of storm (patchy, heavy in some areas, clear in others as opposed to a stationary impenetrable front). This forecast enables the air traffic manager to have sufficient information to make reroutes around the heavier precipitation.

[1]Zulu is a term used to indicate that the time referenced is the Coordinated Universal Time (in essence the same as Greenwich Mean Time). Referencing time as zulu (presented on a 24-hour scale marked with Z) assures there are no misunderstandings as to actual time meant.

Requirements for a Traffic Flow Management Serious Game

For a serious game to be successful in aiding traffic flow management, it must meet multiple requirements. Many of these requisites involve creating a game environment that faithfully simulates the actual world in which TMCs work. The game needs to accurately represent the resources in the NAS, including airports, routes, fixes, sectors, and facilities. The dynamics of the different aircraft types must also be accurately represented, and flight behavior must reflect real flights (e.g., filed flight levels, standard speeds, and structured routings). The game has to provide realistic depictions of current and forecasted weather, similar to the CIWS displays used in operation. The assessments of capacity impacts have to be accurate; aircraft holding and other signs of demand and capacity imbalances should correctly reflect the simulated weather conditions.

The game environment must also present information in ways that replicate the systems and displays that TMCs use. The game needs to be able to emulate the Traffic Situation Display and Flight Schedule Monitor to allow TMCs to assess demand, and it needs to emulate CIWS and CoSPA to allow TMCs to assess capacity. Useful graphical user interface behaviors, such as filtering flights, zooming in and out, and specifying airport resources, must also be replicated. When the game is being used to test the impacts of new information on decision making and operational outcomes, it must support the incorporation of new components.

An effective TFM game will provide trainees with an experience that is true to what they will encounter on the job. The scenarios in the game must replicate the complexities of an actual operational day, including information uncertainty and multiple overlapping decisions. To support the rapid incorporation of lessons learned into operations and to enable the creation of a large and varied library of experiences for trainees, the game must utilize automation for reducing the time and effort needed to create new scenarios. The decision choices offered in the game should be representative of decisions that TMCs would actually make and may make in the future. The means to address demand and capacity imbalance should be based on the same choices that TMCs have now—airspace flow programs (AFPs), ground delay programs (GDPs), ground stops, and reroutes. These decisions should be offered at different times throughout the scenario day to replicate the conditions under which controllers currently operate. If the game is to be used for evaluating new procedures, it should provide information on the procedure's potential impact, reflect the decision process, and model the resulting outcomes.

A useful game will meet several functional requirements. The game should allow a scenario day to be simulated in only a few minutes. The game player needs to be able to replay and make different decisions for the same "day" so that he or she can view and assess the varying outcomes of decisions. The game should also provide some objective feedback about the system's performance, given the decisions made during the game. These metrics or scores should reflect operational metrics (e.g., delays) that are used currently to diagnose NAS issues or experimental metrics (e.g., number of times that a traffic manager views information to make a particular decision) that may provide new insights into operations. Ideally, the game is web-based to allow access for players who may have limited alternative access to the final game product.

Development of the NASPlay Serious Game Prototype

While a serious gaming architecture could support the multiple applications described in the previous section, Lincoln Laboratory pursued the following development goal for NASPlay: develop a serious game architecture that supports ingesting data from actual operational days and provides the game player with a choice of alternative traffic management initiatives that result in an operationally relevant score for each alternative.

Figure 5 illustrates the architecture and assignment of functional capabilities in the current NASPlay system. There are three major components, each with specific and well-defined functional capabilities and interfaces: the NAS simulation engine (NSE), the NASPlay game server, and the game interface. Trainees would interact directly with the game interface.

The computational performance of the simulation is insufficient to run in real time while the game is being played. To accommodate the realities of current simulation performance limitations, the NASPlay developers formulated a constrained-choice concept of operations

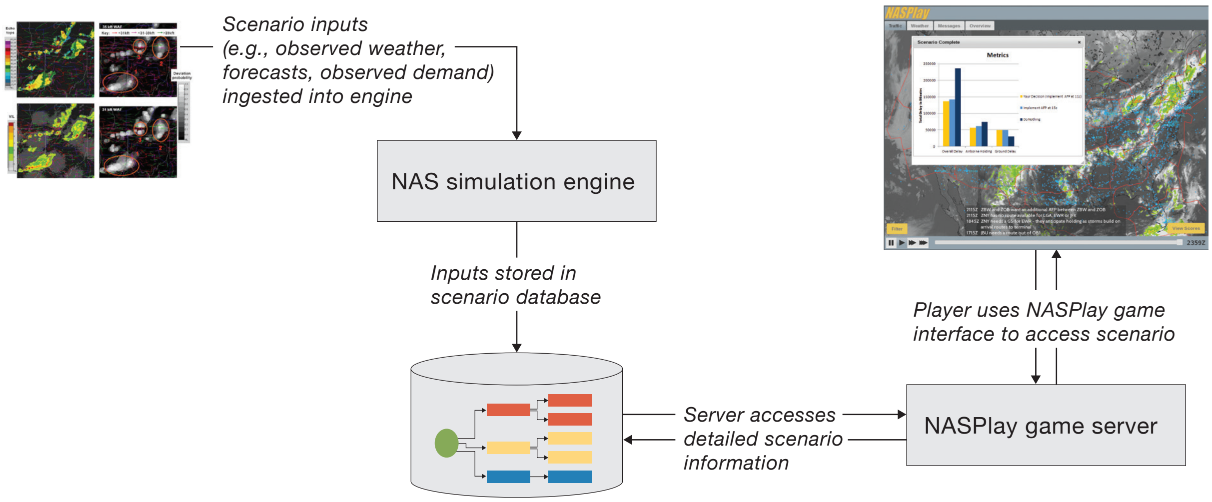


FIGURE 5. The NASPlay system architecture contains scenario inputs, the NAS simulation engine, NASPlay game server, a scenario database, and the NASPlay game interface. The scenario inputs are ingested into the NAS simulation engine and stored in a scenario database. When the scenario is accessed by the NASPlay game server (through the NASPlay game interface by a player), the database is accessed for the detailed scenario information, including the action choices indicated by the colors in each branch above.

in which a limited set of traffic management decisions is made available to the game player at a sequence of discrete decision points during gameplay. As a result, the NAS simulation is decoupled from the actual gameplay, and changes in NAS states that result from players' decisions are precalculated, stored on disk, and provided to the player by the server.

The use of the constrained-choice concept has several implications. Clearly, fewer decision options are available to the game player, and the extent of option exploration is limited to those that were considered by the author of the scenario. However, it is still possible to include a huge number of choices that reflect all the decisions that a TMC could realistically make. The specification of an explicit set of decision options also makes it possible to make clear comparisons between different traffic management strategies and choices. The decoupling of the game from the simulation also makes it possible to provide a large library of game scenarios that may be accessed as part of a progressive training regimen or a concept-engineering and validation exercise. This approach puts much less demand on the network and server, allowing virtually any number of players to access the game at once. With some forethought in scenario development, the constrained-choice concept can provide a rich and varied environment to address many of the challenges in evaluating and training traffic management planning and decision making.

NAS Simulation Engine

The NSE implements the rules for NAS behavior (e.g., its response to external inputs such as scheduled demand and weather impacts) that define specific gameplay scenarios. The NSE schedules flight departures, models flight trajectories, and implements default behaviors of the NAS in response to external events or conditions that arise as the simulation proceeds (e.g., what happens to flights entering an air traffic control [ATC] sector when that sector is at capacity). The NSE also provides the capability to model commonly used traffic management initiatives, such as ground delay programs, on the basis of forecast traffic demand and constraints. The NSE provides the capability to harvest data about the evolving state of the NAS during the simulation (e.g., the location of flights and their delay status) and calculates NAS-wide and local performance metrics. Finally, the NSE provides fast-time simulation capabilities to facilitate the generation of NAS outputs corresponding to each branch of the constrained-choice decision tree.

To create the level of realism required for meaningful game development, the NSE must provide fine-grained control of flight trajectories and air traffic control actions, particularly in response to external events such as thunderstorms. The NSE must provide a way to extend or replace default behaviors (e.g., pilot decisions to accept or reject routes through convective weather-impacted airspace or the response of ATC to weather impacts) with models such as the Convective Weather Avoidance Model for pilot decision making in convective weather–impacted airspace, or the Controller Workload Model for sector capacity.

After a technical evaluation of an array of existing simulation products, AirTOpsoft's simulator, AirTOp [6], was chosen because of its agent-based foundation and flexible development environment. AirTOp's implementation enables fine-grained control over several key elements of NAS operation and simulation:

- Dynamic capacity constraints. Simulations may be initialized with time-varying capacity constraints on any airspace resource that is defined in the NAS adaptation.
- Options for tactical weather avoidance. AirTOp provides mechanisms to implement tactical weather avoidance options, such as no-notice holding and trajectory vectoring to avoid weather. In addition, thresholds (often referred to as hooks) can be set to trigger diversions, ground stops, and other tactical responses to airspace constraints.
- Hooks for calculation of default and custom performance and scoring metrics. AirTOp supports the specification of software watch points that can trigger data analysis and the output of user-specified simulation state data for incorporating the generation of performance and scoring metrics into the simulation.
- Data are stored in easily modified text files.

After the development staff spent several weeks familiarizing themselves with the AirTOp environment, they input the baseline structure for the NAS (e.g., Air Route Traffic Control Center boundaries, ATC sector boundaries, navigation fixes, jet routes, aircraft types, and airports) into the simulation. Data for a full day's flight plan schedule were assembled and input to AirTOp. As is common with navigation data, a significant amount of "cleaning" of the data was required to return reasonable output:

- Correction (where possible) or removal of ambiguously or incorrectly specified navigation fixes from flight plans
- Assignment of aircraft performance statistics when the aircraft type is unknown to AirTOp
- Proper sequencing of departure times to ensure temporal continuity of flight plans that have multiple stops and continuation legs
- Filtering of flight plans that are outside the scope of the game scenario to reduce simulation run time
- Specification of realistic cruise altitudes and air speeds for flight plans that are missing this information
- Conversion of scheduled flight plans into AirTOp's input format
- Determination of which entry is most accurate if the same flight plan appears multiple times in the data; removal of any "loops" from routings
- Conversion of units of measurement, especially for speed (e.g., Mach, indicated airspeed, true airspeed)

Five critical weather-impact capabilities were also implemented in NASPlay: time-varying winds for flight trajectory modeling; time-varying air traffic control sector capacity constraints that include considerations for convective weather impacts; time-varying air traffic flow capacity constraints that account for convective weather impacts; time-varying airport capacity; and time-varying fix capacity.

In addition, several initial performance metrics were implemented: individual flight delay, separated into ground and airborne portions, as well as planned and unplanned portions; hourly measurement of aggregate delay and holding; time of flight; fuel burned; and cancellations and diversions.

NASPlay Game Server

The primary role of the NASPlay game server is to provide to the game client the current state of the NAS resulting from the player's decision choices up to that point. The NAS state includes the flight plans and locations of all flights, outputs from operational models for the current weather and weather forecasts, and stakeholder comments or tactical responses (e.g., a request for a ground stop or diversion) derived from external sources (e.g., the National Traffic Management Log) or automatically generated by the NSE during scenario preparation. The server also records player decisions and interactions that will be used for postgame analysis.

Game Client

The game client is the player's window on the NAS world. It renders the game display that provides (1) the current state of the NAS, such as flight locations and plans, current NAS performance statistics, and emulation of commonly used tools such as the Flight Status Monitor; (2) feedback

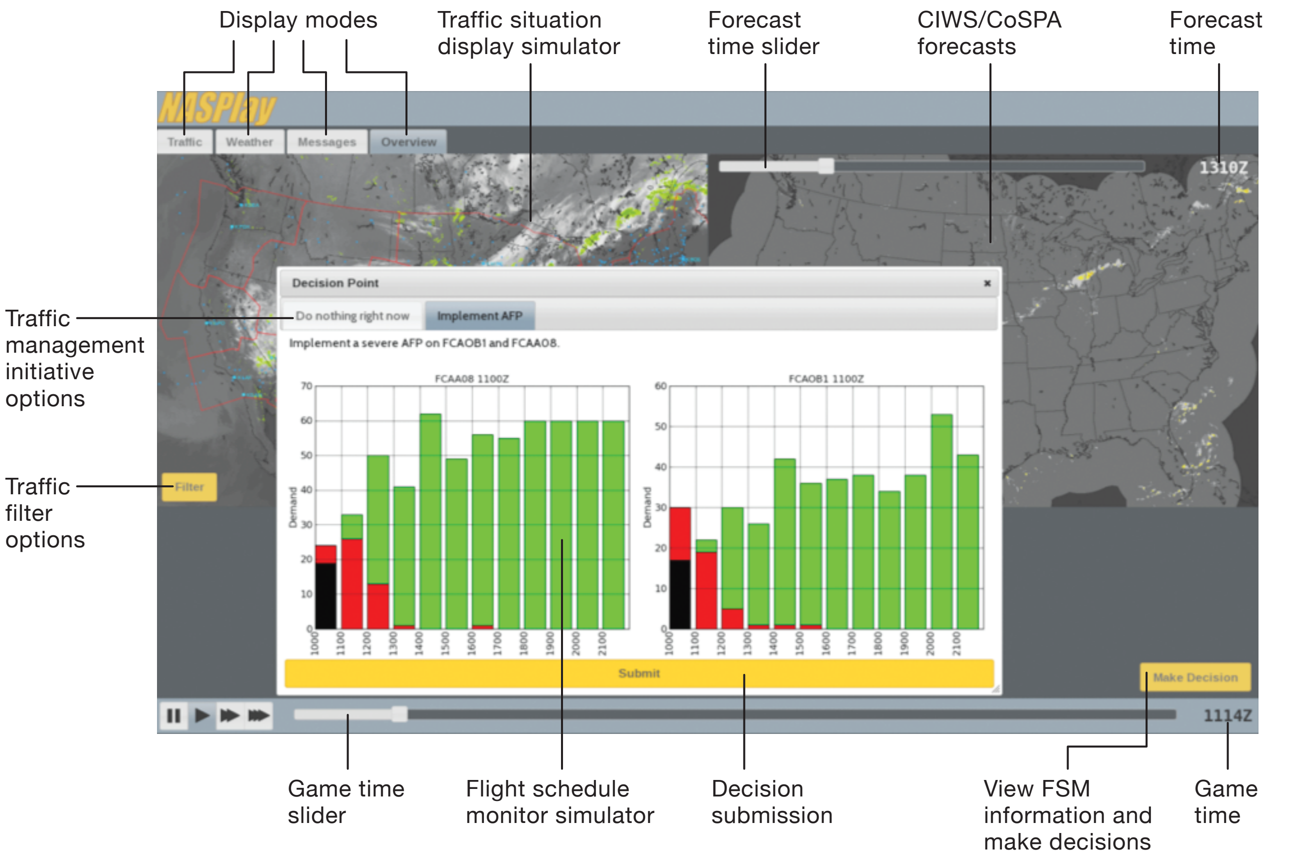


FIGURE 6. The NASPlay user interface includes display modes, time sliders, filtering options, game time, and other data critical to the decisions required, such as the Flight Schedule Monitor.

comments from other NAS stakeholders; and (3) external factors, such as the weather. It provides the game clock control, allowing the player to start and pause the action or rewind to review the previous state of the world. The client prompts the player for decisions, providing NAS modeling information relevant to the decision options. Finally, the client passes selected player interactions to the server for logging and postgame analysis. The client in the current NASPlay prototype is shown in Figure 6.

Game Scenario

The development of the game scenario is key to the success of the game. Gaining TMCs' acceptance of a traffic management serious game would be impossible if the scenario did not capture the complexities of traffic management and the subtleties of the operational environment. To this end, the NASPlay developers chose a particularly impactful day that had questionable traffic management decisions made during the operation. A similarly impactful weather day was further explored with respect to forecast uncertainty and decision making [7].

On 11 September 2013, a group of severe thunderstorms developed between Maine and Tennessee around 1600Z (4 p.m.), impacting eastbound arrivals starting around 1700Z. Traffic managers at the Command Center in Virginia opted to address the capacity constraint imposed by these storms by rerouting New York–bound flights from Fort Worth, Houston, and Memphis centers south through the Vulcan Playbook[2] (VUZ) reroute and AZEZU Playbook reroutes, and by tactically managing traffic through ground stops. Managers also implemented Airspace Flow Programs at 1650Z for two flow-control areas (shown in Figure 7) from 1915Z and 1945Z. There was significant NAS disruption, including 69 diversions,

[2] A playbook contains a set of standard routes that ATC can utilize to fit a particular set of circumstances when the preferred routes are not available. These routes were created to allow for rapid implementation of rerouting as needed.

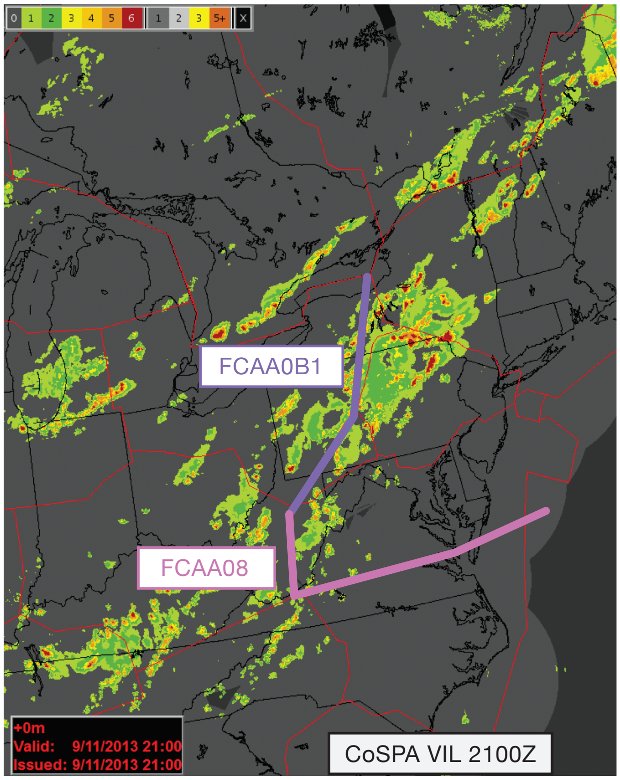


FIGURE 7. The map shows the flow control areas (FCAs)—FCAAOB1 and FCAA08—that were restricted and the convective weather existing on 11 September 2013 at 21:00 zulu. Data are from the CoSPA system, which uses vertically integrated liquid measurements to predict convective weather activity.

55 holding events (totaling 21 hours), 13 ground stops (totaling 7 hours), and 72 taxi-backs to the airport gates.

The critical aspects of a scenario for game reconstruction include the timeframe and area of interest, the decisions available, the information available, and the metrics by which the decisions are evaluated.

The area of interest for this scenario is the New York (ZNY)–District of Columbia (ZDC) area. Thus, it is important for the player to be able to view and filter flights arriving and departing this area of the NAS. Also important is the player's ability to zoom into this area to discriminate the local weather and traffic. The timeframe of interest is from 1700Z to 2100Z, a busy period during which the weather significantly affects the high-demand traffic areas. For decision-making purposes, it is important to be able to view not only the unfolding weather and traffic but also the forecasts of weather and traffic demand for this timeframe.

It is critical to identify the key decisions available to the TMC to address the demand and capacity imbalance issues for a particular scenario. The strategic traffic management decisions available to the TMCs for this scenario include reroutes, AFPs, GDPs, and ground stops. For convective weather impacts in ZNY and ZDC, the appropriate AFPs include FCAOB1 (the eastern boundary of Cleveland center) and FCAA08 (west/east line through Washington center). Some of the decisions must be made no later than four hours before the expected impact in order to have the desired effect on the traffic. Thus, forecasts for the 1800Z timeframe and beyond need to be available to the game player by no later than 1400Z. GDPs for the New York airports were made available as potential decisions as well.

An example constrained decision tree was created for the game scenario by using these key decisions. Figure 8 illustrates a traffic management initiative decision tree that enumerates the set of possible decisions for this scenario. At 1315Z and 1715Z, the player is able to choose whether to implement an AFP or a reroute and whether the AFP should be "mild" or "severe." If a reroute is chosen, then the player is also offered the choice at 1915Z to implement a GDP or not. Seven outcomes for this scenario are possible, and the possible decisions in this example are constrained. A full scenario would allow for the 10 to 100 choices that an actual national air traffic manager would experience in a convective weather day.

To adequately represent the scenario in a context familiar to the game player, the information presented must be consistent with that used by a TMC. Two information sources are used to assess demand over time: the Traffic Situation Display and the Flight Schedule Monitor, both shown in Figure 3. To assess capacity, TMCs must have adequate knowledge of the current and forecasted location and severity of the weather, such as is depicted in the CIWS and the CoSPA tools shown in Figure 4. Strategic TMCs also receive input from local air traffic control facilities (ARTCCs[3], TRACONs[4], and towers), as well as from their airline customers, about what decisions to implement

[3]Air Route Traffic Control Centers handle primarily en route aircraft on instrumented flight plans; 21 centers cover the regions over the United States.

[4]Terminal Radar Approach Control facilities handle ATC operations near major airports, primarily aircraft arrivals and departures.

via phone calls. To simulate this communication in the game environment, a chat window was implemented to allow facility and airline agents to provide their opinions on the decision options. Chat messages were derived from the National Traffic Management Log and created in response to simulated events.

To understand whether a game player's decision was "good" or "bad," operational performance metrics must be established. A common metric used by the ATC community is the amount of delay accrued during an event for the NAS. Additional metrics—airborne holding time, ground delay time, uncontrolled delay, fuel burn, and number of diversions and cancellations—were identified to indicate the quality of a decision. Filters identifying where and when the delays occurred also provide an indication of how the traffic was affected by decisions. The defining and weighting of performance metrics in scoring are areas of active research, and these are expected to evolve significantly as NASPlay development continues.

To acquire the data to ensure the scenario fidelity for the game, the following required data from 11 September 2013 were assembled:

- NAS definition data
- Scheduled traffic data
- Wind data
- Lincoln Laboratory's CIWS and CoSPA weather data archives for the NAS
- Lincoln Laboratory's Route Availability Planning Tool data for fix capacities
- Lincoln Laboratory's Traffic Flow Impact data for sector and flow capacities
- Command center teleconference and National Traffic Management Log data (what decisions were considered when, inputs by ATC facilities and airlines)

Emulations of the Traffic Situation Display, Flight Schedule Monitor, CIWS, and CoSPA were created to ensure realistic representation of the information consistent with the traffic management context.

Validation and Evaluation of NASPlay

Both validation and evaluation are required to ensure that NASPlay meets the needs of the NAS users. A detailed report of the validation for NASPlay is provided in Davison Reynolds, DeLaura, and Soulliard [8]. It

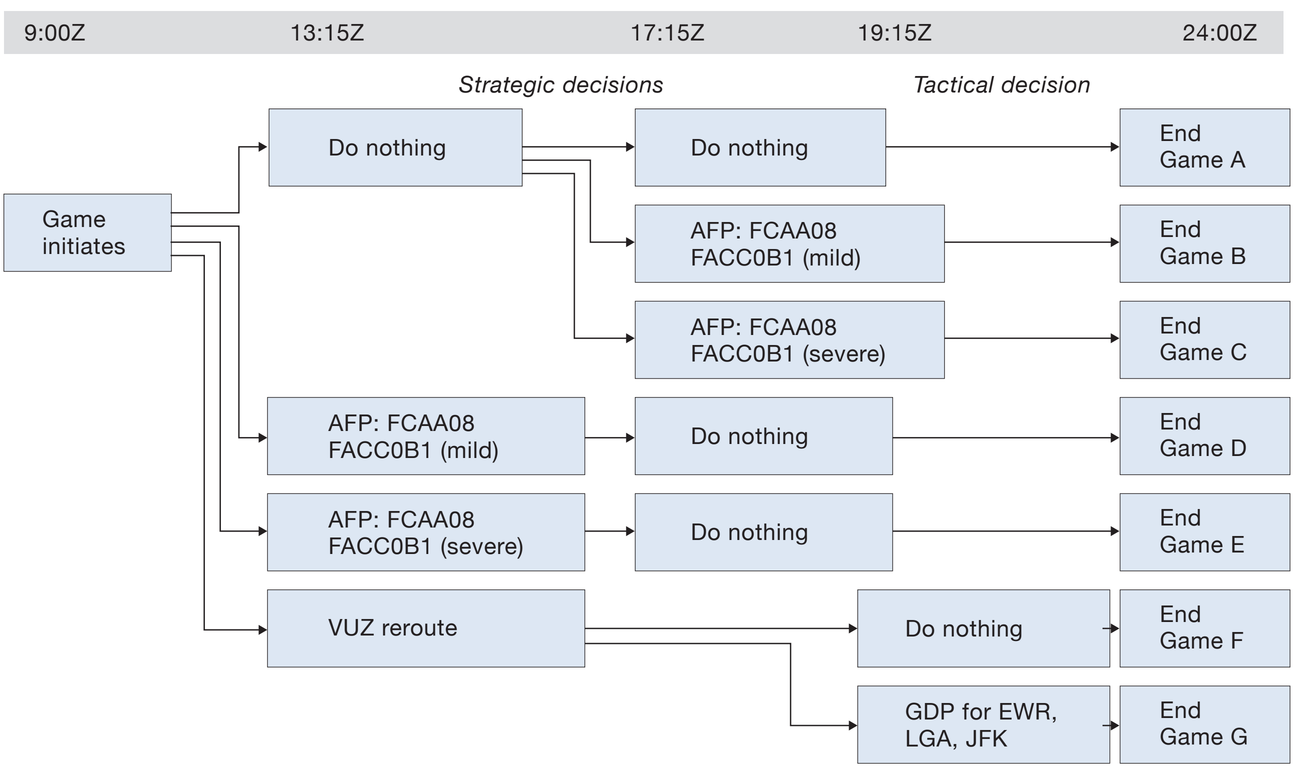


FIGURE 8. This simplified decision tree for a constrained choice game illustrates the choices available to the player at different times during the scenario.

is important that the simulation dynamics, the traffic demand, and the weather capacity algorithms all work in concert to provide a valid representation of the NAS because conclusions drawn from an invalid model would not transfer to the real NAS. Likewise, if the NAS users evaluate NASPlay and find that it does not represent the NAS in some critical way, NASPlay will not be accepted within the community. Thus, initial validation and evaluation have been attempted for NASPlay.

An initial validation was performed for a nominal, unconstrained (fair weather) operational scenario, taken from operations on 21 October 2012. The primary focus of the validation was to ensure that the schedule cleansing and wind data ingest resulted in reasonable flight plans, trajectories, and overall number of operations.

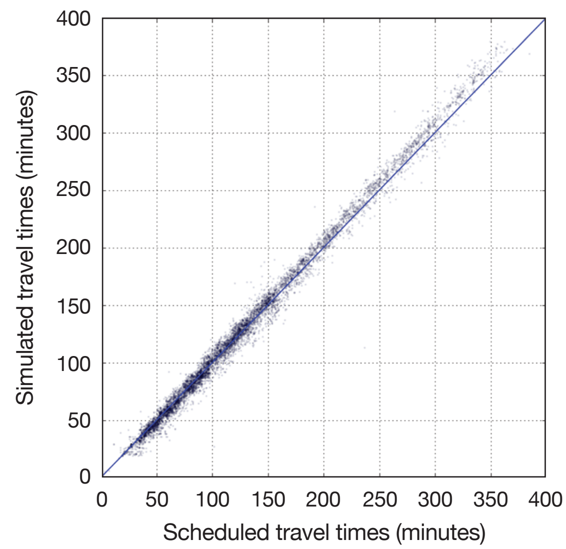


FIGURE 9. A comparison of simulated and observed flight times for major airport origin-destination pairs are shown.

Individual Flight-Level Validation—Flight Simulations

The total number of flights flown by the simulation was 34,928, a number roughly in line with the FAA's OPSNET[5] ASPM[6] 77 terminals' count of approximately 56,000 operations for the day. Note that ASPM operations include both arrivals and departures for domestic airports, so their count is roughly double the number of flights for 21 October 2012. However, the ASPM count does not include general-aviation flights.

AirTOp time of flight was compared to observed time of flight for each scheduled flight with a corresponding observed departure. The results of the comparisons for flights between the 34 largest airports in the continental United States are presented in Figure 9 and show good agreement between simulated and observed flight times. Top-down map views of several flight plans for different origin-destination pairs were inspected to ensure that "doctored" simulation flight plans were reasonable. The distribution of flight altitudes as a function of flight distance was also examined to confirm that cruise altitudes were sensible in NASPlay. Finally, an initial performance measurement analysis capability that will form the basis for the game scores was developed. The capability currently assesses ground delays (planned and unplanned), airborne delays, cancellations, and diversions, all adjusted for the types of flights impacted (e.g., passenger, cargo, or general aviation).

[5]Operations Network is the official source of data on NAS traffic operations and delays.

[6]Aviation System Performance Metrics is an online database of information on flights to and from the 77 U.S. airports.

Flow-Level Validation—Capacity Modeling

A time-variant capacity constraint was developed for flows and sectors. Each flow captures flights going through its area in a certain direction. The algorithm was previously developed and verified at Lincoln Laboratory. Testing indicates that the simulation is performing as the model predicts.

NASPlay Evaluation

The NASPlay prototype was initially evaluated by several NAS user groups, including trainers and traffic management specialists from the Air Traffic Control System Command Center (ATCSCC), the manager of tactical operations in the Northeast United States, former en route/TRACON/tower controllers, and representatives from two airlines. An initial introduction to the prototype, which included the potential concept of operations and use for the tool, was provided to the evaluators. Users then played through the demonstration scenario, seeking out diagnostic information and making their own choices. Once the users completed the demonstration, they were asked what, if any, value the prototype concept would have in their jobs and what information or functionality is missing to achieve that value. Table 1 itemizes the

Table 1. NAS Users' Estimation of the Value of NASPlay to Their Operations

USER GROUP	VALUE OF NASPLAY TO OPERATION
ATCSCC trainers	Integration of NASPlay with their laboratory training environment to produce fast-time "what if" decisions to a set of specified scenarios
ATCSCC traffic management specialists	Ability to conduct over-the-shoulder, on-the-job training with new traffic management specialists; ability to better understand the interaction of traffic management initiatives with one another in a controlled environment
Manager of tactical operations in the Northeast United States	Capability to support continual offline demand and capacity imbalance identification; evaluation of traffic management decisions with objective metrics
Former air traffic controllers	Ability to try out and evaluate the effects of new procedures offline; training in severe weather decision making
Airline representatives	Ability to model and better understand the effect of traffic management initiatives on their businesses; this understanding could lead to effective lobbying for particular initiatives on strategic planning

Table 2. NAS Users' Suggested Improvements for NASPlay

USER GROUP	SUGGESTED IMPROVEMENT
ATCSCC trainers	Generate many severe weather scenarios; explore connecting NASPlay to training Flight Schedule Monitor
ATCSCC traffic management specialists	Incorporate airline cancellations and pilot diversions into functionality; make NASPlay multiplayer and web-based for a single scenario
Manager of tactical operations in the Northeast United States	Generate tactical scenarios for NASPlay focusing on a single en route center and/or TRACON; provide ability to continuously monitor airport surface status; make NASPlay scoring consistent with the FAA's internal AERO operational evaluation statistics page
Former air traffic controllers	Incorporate airline cancellations, pilot diversions, and tactical rerouting into functionality
Airline representatives	Make the prototype available for airline use

specific value to their jobs that the user groups saw for the prototype.

The users also offered suggestions for additional information and functionality to improve the ability of NASPlay to meet their identified needs; Table 2 itemizes these suggestions.

Future Development

The NASPlay prototype was developed to address serious shortfalls in current FAA capabilities for training air traffic managers, evaluating current and proposed NAS operational procedures, and developing and validating new operational concepts. Its platform integrates a commercial simulation capability with both the Laboratory's novel algorithms for severe weather capacity and its gaming interface. The prototype's output was validated in both fair and severe weather.

All the NAS users who evaluated NASPlay's operational value and functionality prioritization saw useful applications of NASPlay for their respective jobs. Many of the suggested improvements in information and functionality are possible and desirable to accomplish within the next year. A more detailed user evaluation is planned for the end of next year. The goal of that evaluation will be to gather input about both the available decision choices in the assembled scenarios and the usability of the current NASPlay prototype.

Over the next year, NASPlay will be expanded into the tactical traffic management realm, with the development of tactical scenarios (regionally focused rather than nationally focused). Additional possibilities for NASPlay's expansion include

- Multimodal performance scoring to evaluate decision making according to alternative performance criteria, such as environmental impact or passenger experience
- Multiplayer gaming for advanced training and evaluation of future concepts, such as dynamically configured airspace
- Agent-based Monte Carlo simulations to realistically assess the potential benefits of new forecast tools and procedures, accounting for limitations, such as forecast accuracy and uncertainty in the response of pilots and controllers to events
- Real-time simultaneous simulations and scoring of potential alternate outcomes to guide planners in operational decision making ■

References

1. Federal Aviation Administration, National Severe Weather Playbook, 2016, https://www.fly.faa.gov/PLAYBOOK/pbindex.html.
2. D.L. Klingle-Wilson and J. Evans, "Description of the Corridor Integrated Weather System (CIWS) Weather Products," MIT Lincoln Laboratory, Lexington, Mass., Project Report ATC-317, 1 August 2005.
3. M.M. Wolfson, W.J. Dupree, R. Rasmussen, M. Steiner, S. Benjamin, and S. Weygandt, "Consolidated Storm Prediction for Aviation (CoSPA)," 13th Conference on Aviation, Range, and Aerospace Meteorology, New Orleans, La., 2008, http://www.ll.mit.edu/mission/aviation/publications/publication-files/ms-papers/Wolfson_2008_ARAM_MS-30236_WW-14159.pdf.
4. D.L. Klingle-Wilson, "Integrated Terminal Weather System (ITWS) Demonstration and Validation Operational Test and Evaluation," MIT Lincoln Laboratory, Lexington, Mass., Project Report ATC-234, 13 April 1995.
5. M. Robinson, R. DeLaura, and N. Underhill, "The Route Availability Planning Tool (RAPT): Evaluation of Departure Management Decision Support in New York During the 2008 Convective Weather Season," Eighth USA/Europe Air Traffic Management Research and Development Seminar, Napa, Calif., 2009, http://www.ll.mit.edu/mission/aviation/publications/publication-files/ms-papers/Robinson_2009_ATM_WW-16318.pdf.
6. AirTOpsoft, AirTOp fast time simulator, 2016, http://www.airtopsoft.com/.
7. H.J. Davison Reynolds, R. DeLaura, J. Venuti, and M.M Wolfson, "Uncertainty and Decision Making in Air Traffic Management," AIAA Aviation, Technology, Integration, and Operations Conference, Los Angeles, 12–14 August 2013.
8. H.J. Davison Reynolds, R. DeLaura, and B. Soulliard, "How Good Is Good Enough? Exploring Validation for an Air Traffic Control Serious Game" chapter in *Advances in Human Factors, Business Management, Training and Education: Proceedings of the Applied Human Factors and Ergonomics 2016 Conference on Human Factors, Business Management, and Society*, J.I. Kantola, T.Barath, N. Salman, and T. Andree, eds. Switzerland: Springer International Publishing, 2017.

About the Authors

Hayley J. Reynolds is an assistant leader of the Informatics and Decision Support Group. Her expertise is in human-systems integration and systems engineering. Since joining Lincoln Laboratory in 2009, she has worked extensively on the design and development of aviation decision support systems, biosurveillance systems and emergency management systems. She also leads several programs centered on counter-human trafficking. She holds a bachelor's degree in psychology from the University of Illinois at

Urbana-Champaign, and a master's degree in aeronautics and astronautics and a doctoral degree in aeronautical systems and applied psychology, both from MIT.

Brian C. Soulliard is currently a staff member at Google. While at Lincoln Laboratory, he was the lead software developer for NASPlay. He graduated with a bachelor's degree in software engineering from the Rochester Institute of Technology with minors in computer science and in game design and development.

Richard A. DeLaura is a technical staff member in the Air Traffic Control Systems Group. Prior to joining Lincoln Laboratory in 2000, he was a research scientist at the University of Massachusetts Dartmouth, developing software and curriculum to introduce advanced mathematical concepts to middle and high school students. He has authored or coauthored several conference proceedings and Federal Aviation Administration research reports. He holds a bachelor's degree in chemistry and physics from Harvard University.