CHAPTER XX

Extending Validated Human Performance Models to Evaluate NextGen Concepts

Brian F. Gore[1], Becky L. Hooey[1], Eric A. Mahlstedt[1], David C. Foyle[2]
[1]San Jose State University at NASA Ames Research Center

[2]Human-Systems Integration Division / MS 262-4

Moffett Field, CA 94035-0001 USA

E-mail: Brian.F.Gore@nasa.gov; Becky.L.Hooey@nasa.gov; Eric.Mahlstedt@nasa.gov; David.C.Foyle@nasa.gov

ABSTRACT

To meet the expected increases in air traffic demands, the National Aeronautics and Space Administration (NASA) and the Federal Aviation Administration (FAA) are researching and developing Next Generation Air Transportation System (NextGen) concepts. NextGen will require substantial increases to the data available to pilots on the flight deck (e.g., weather, wake, traffic trajectory projections, etc.) to support more precise and closely coordinated operations (e.g., self-separation, RNAV/RNP, and closely spaced parallel operations, CSPOs). These NextGen procedures and operations, along with the pilots' roles and responsibilities, must be designed with consideration of the pilots' capabilities and limitations. Failure to do so will leave the pilots, and thus the entire aviation system, vulnerable to error. A validated Man-machine Integration Design and Analysis System (MIDAS) v5 model was extended to evaluate anticipated changes to flight deck and controller roles and responsibilities in NextGen approach and land operations. Compared to conditions when the controllers are responsible for separation on decent to land phase of flight, the output from these model predictions suggest that the flight deck response time to detect the lead aircraft blunder will decrease, pilot scans to the navigation display will increase, and workload will increase.

Keywords: NextGen CSPO, MIDAS v5, Human Performance Model

1   INTRODUCTION

The National Airspace System (NAS) in the United States is currently being redesigned because it is anticipated that the current air traffic control (ATC) system will not be able to manage the predicted two to three times growth in air traffic in

the NAS (JPDO, 2011). To meet the expected increases in air traffic demands, the National Aeronautics and Space Administration (NASA) and the Federal Aviation Administration (FAA) are researching and developing Next Generation Air Transportation System (NextGen) concepts to alleviate bottlenecks caused by the anticipated growth.

One such bottleneck is anticipated to be in the decent, approach, and landing phases of flight. Closely Spaced Parallel Operations (CSPO) are expected to enable paired approaches to minimum runway spacing in instrument meteorological conditions (IMC) while maintaining an acceptable level of risk (Cox, 2010). The current requirement for landings in IMC is at least 4300 ft of lateral runway spacing (as close as 3000 ft for runways with a Precision Runway Monitor) whereas operations in visual meteorological conditions (VMC) require lateral runway spacing to be equal to or greater than 750 ft It is feasible for aircraft to perform both arrival and departure operations in IMC using VMC parallel separation standards as advanced navigation technology, sophisticated wake avoidance algorithms, 4-D flight management systems, and advanced flight deck displays become more widely available (Rutishauser, et al., 2003).

In the highly automated CSPO environment envisioned by NextGen, a paradigm shift might be required that would transfer the responsibility for separation from ATC, as is currently the case, to the flight deck. As flight decks are modified to accommodate the new suite of automation tools and displays required to support this, research must be conducted to ensure that they are designed and implemented in a safe manner without leaving pilots vulnerable to errors or excess workload. These NextGen procedures and operations, along with the pilots' roles and responsibilities must be designed with consideration of the pilots' capabilities. Failure to do so will leave the pilots, and thus the entire aviation system vulnerable to performance inefficiencies caused by error.

1.1   Using HPMs to Evaluate NextGen Concepts

There are large challenges associated with evaluating novel NextGen concepts such as CSPO and changes to pilot / ATC roles and responsibilities. Because NextGen concepts are still in the early stages of the design lifecycle, operator roles and tasks are often not well defined, and NextGen technologies have not necessarily reached a level of sufficient maturity to allow for physical prototypes. These factors limit the feasibility of full-mission human-in-the-loop (HITL) simulations. However, human performance models (HPMs) can be used to make meaningful contributions early in the design lifecycle, particularly for concepts that have high consequences associated with their failure. Models can be advantageous because they are cost effective, and eliminate concerns often associated with HITL testing of new concepts such as novelty and training effects. Furthermore, models are advantageous as compared to HITLs simulations where you have to build prototypes, details, or emulate details because HPMs represent information symbolically and therefore allow rapid prototypes of concepts to be generated and tested early in the design phase. One such HPM tool, the Man-machine Integration

Design and Analysis System (MIDAS), is discussed next.

1.2 The Man-Machine Integration Design and Analysis System (MIDAS)

NASA's MIDAS is a dynamic, integrated HPM that facilitates the design, visualization, and computational evaluation of complex man–machine system concepts in simulated operational environments (Gore, 2008). MIDAS symbolically represents many mechanisms that underlie and cause human behavior including the manner that the operator receives/detects information from an environment, comprehends and registers this information in a memory store, decides on a response, and responds to the information within the context of operational rules and human performance capacities. MIDAS combines these symbolic representations of cognition with graphical equipment prototyping, dynamic simulation, and procedures/tasks to support quantitative predictions of human-system effectiveness, and improve the design of crew stations and their associated operating procedures. MIDAS provides an easy to use and cost-effective means to conduct experiments that explore "what-if" questions about domains of interest.

2   A Process for Evaluating NextGen Concepts using HPMs

One challenge associated with developing valid models of NextGen concepts is the lack of HITL data with which to validate the models. In this paper we propose a candidate process for developing and validating HPMs for the evaluation of NextGen concepts. The process includes four steps: 1) Develop a baseline (current-day) model; 2) Validate the baseline (current-day) model; 3) Extend the baseline scenario to NextGen; and 4) Conduct iterative what-if scenarios to explore early design concepts. This process addresses the validation challenge by first developing models of known, current-day, operations, which are well-defined and proceduralized, and for which HITL data exists to enable validation. Then, the validated model platform is modified by integrating assumptions about likely NextGen changes that will be made to the flight deck equipage and pilot tasks. Confidence is attained that the validity of the model is preserved through the documentation of assumptions and through the small iterative model changes to the validated model. This process, as applied to the CSPO concept, is discussed next.

2.1   Develop the Baseline Model

A MIDAS v5 high-fidelity model of a two-pilot commercial crew flying current-day area navigation (RNAV) approach and landing operations was developed using a methodical, multi-dimensional approach (Gore, et al., 2011). The model represented a Boeing 777 flying from 10,000 ft to touchdown at Dallas Fort Worth (DFW) airport. The modeled scenario began with the aircraft at an altitude of 10,000 ft and 30nm from the runway threshold (see Figure 1). The cloud ceiling was 800 ft, with

a decision height (DH) of 650 ft at which point the modeled pilots disconnected the autopilot and manually flew the aircraft to touchdown. The model assumed that the "pilot flying" (PF) was in the cockpit's left seat and the "pilot monitoring" (PM) was in the right seat. The model scenario included communications with DFW Regional Approach Control, tower, and ground control, as well as intra-cockpit communications. In total, over 970 pilot tasks were included in the model.
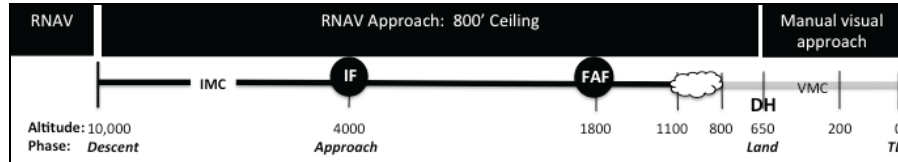

Figure 1. Baseline RNAV Model of Approach and Landing.

Notes: DH = Decision Height; FAF = Final Approach Fix; IF = Initial Fix; IFR = Instrument Flight Rules; RNAV = Area Navigation; TD = touchdown; VMC = Visual Meteorological Conditions.

A task network model architecture such as the one embedded in MIDAS contains top level tasks (e.g., land) that are subsequently decomposed into finer-grained tasks, generally to the button-press level for physical control input tasks, scan fixation points for the visual system, and verbal strings for the vocal communication output. The tasks in the network are then tied to a set of behavioral primitives. These tasks then wait for release conditions to be satisfied by the environment, the operators, the controls, or the displays. The behavioral tasks in MIDAS are termed the operator primitives. The task network model illustrated in Figure 2 illustrates a subset of the entire task network model, a snapshot of the flight deck's flap-setting procedure required when landing the aircraft in the simulation.
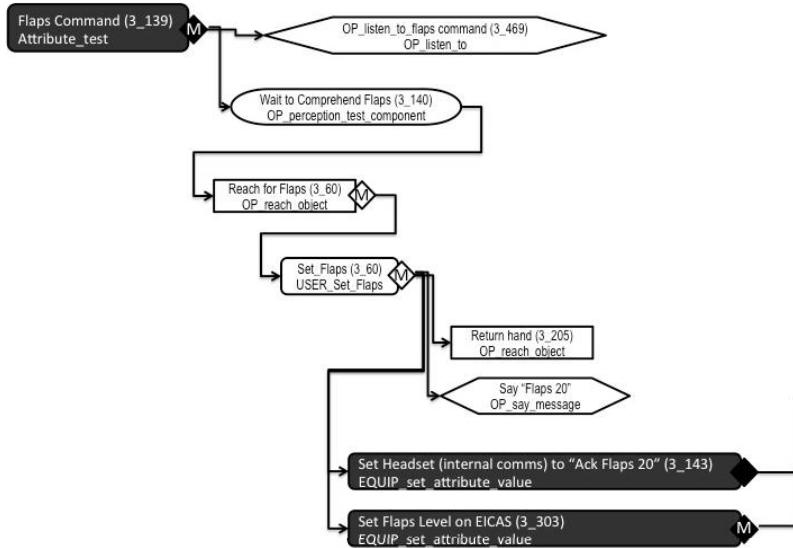

Figure 2. Task Network Model Implementation of a Set Flaps Sequence.

2.2 Validate the Baseline Model

The model inputs, including the task trace and input parameters, were validated using focus group sessions comprised of a total of 8 commercial pilots with glass-cockpit aircraft and RNAV flying experience. The pilot-centric scenario-based cognitive walkthrough approach captured the context of operations from 10,000' to Touchdown and enabled pilots to assess the modeled tasks and identify tasks that were missing, or in the wrong sequence. Out of 74 tasks in the MIDAS RNAV application presented, the focus group pilots identified 12 tasks that should be removed, reordered, or added. The pilots also completed quantitative rating scales, which were used to validate the model input parameters for workload and visual attention. Thirty-nine tasks were rated on the visual, auditory, cognitive, and motor dimensions (as relevant for the task) resulting in 75 ratings. Four pilot tasks were modified and three new primitives were created. The model was refined based on the results of this input validation process. Next the model outputs, workload and visual attention, of the refined model were statistically compared to existing HITL data (Anders, 2001; Mumaw, Sarter and Wickens, 2001; Hüttig, Anders, and Tautz 1999). The workload model output correlated (r^2=.54) with a comparable HITL study with for overall workload. The individual workload dimensions also correlated positively with the HITL study (r^2=.55 to .94). Visual scan time correlated with three HITL studies (r^2 = .99). These results provide confidence that the model validly represents pilot performance.

2.3. Extend Validated Model to CSPO Scenarios

The RNAV scenario was modified to reflect the CSPO concept based on assumptions about changes to: 1) flight deck equipage (e.g., the addition of data communications, augmented wake and traffic information on the Primary Flight Display (PFD) and Navigation Display (Nav), and visual and auditory wake threat alerts; and, 2) flight crew tasks (e.g., identifying and tracking paired traffic, receiving and accepting datalink, monitoring wake displays). This assumed an operational environment consistent with NextGen goals of reduced landing minima, specifically, a cloud ceiling of 200 ft and a DH of 100 ft (see Figure 3). The assumptions were made based on interviews with NextGen concept developers and scenario-based focus groups with pilots experienced with current-day Simultaneous Offset Instrument Approaches (SOIAs).
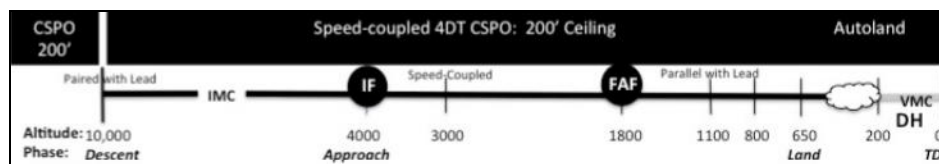

Figure 3. CSPO 200 Scenario Timeline.
Notes: IF = Initial Fix, FAF = Final Approach Fix, DH = Decision Height, IMC = Instrument Meteorological Conditions, VMC = Visual Meteorological Conditions, TD = Touchdown

In order to ensure the verifiability and validity of the CSPO model, the specific CSPO task changes and input parameters were validated using the same pilot focus group sessions described previously. In the focus group sessions, after the pilots completed the task trace and input parameter worksheets for the RNAV model, the CSPO concept was introduced. The pilots were briefed on the goals of NextGen, expected changes to flight deck equipage, and pilot procedures. Examples of the wake displays on both the PFD and Nav and the visual and auditory wake warnings and alerts were presented. A video of two pilots completing CSPO procedures from a HITL simulation (Verma et al., 2008) was also presented.

2.4. Conduct "What-if" Scenarios

Using the validated CSPO model as a research platform, the model was then exercised to explore a number of CSPO design concepts including varying the flight crew task allocation, pilot-ATC roles and responsibilities, and the format and location of wake and spacing information. In total, 26 model-based scenario manipulations were completed. Analyses of pilot performance measures, including time required to complete tasks, pilot workload, pilot scan patterns and response times to off-nominal events were used to draw conclusions regarding the information requirements necessary to support NextGen CSPO concepts (see Hooey, Gore, Mahlstedt, & Foyle, 2012).

3.0 CSPO CONCEPT EVALUATION RESULTS

In this section, results of one of the CSPO 'what-if' evaluations are presented. Specifically, the model was modified to evaluate the effect of delegating separation to the flight deck for CSPO operations. Two scenarios were modeled: 1) **Pilot-responsible separation**, a scenario to represent NextGen CSPO conditions in which the pilots are responsible for separation delegation and for detecting and initiating emergency escape maneuvers; and, 2) **ATC-responsible separation**, a scenario to represent current-day conditions in which the ATC are responsible for separation and for detecting and initiating emergency escape maneuvers. As the aircraft was on final approach, a wake threat occurred in which the wake of the lead aircraft extended into the ownships' trajectory requiring a missed approach. In the pilot-responsible for separation scenario, pilots were provided with a dynamic wake display and a two-stage alert system that first issued a visual and auditory warning as the wake threat developed and a final alert commanding an immediate take-off / go-around (TOGA) procedure. The dynamic breakout maneuver was shown on the Nav. In the ATC-responsible separation scenario, ATC received the two-stage wake warning and alert, who then issued a verbal 'Go-Around' command to the aircraft, with a missed approach path that accounted for metroplex traffic, terrain, and wind conditions.

Ten Monte Carlo simulation runs were generated to evaluate the pilots' time to initiate the emergency escape maneuver in response to the wake threat alert, the

pilot's scan (percent dwell time, PDT) performance, and the pilots' workload were assessed.

3.1 Response Time to Wake Threat Alert

The response time to the wake threat alert as a function of current-day ATC-responsible or NextGen Pilot-responsible for separation can be found in Figure 4. This figure illustrates that pilots were only negligibly faster to initiate the emergency escape maneuver in the pilot responsible condition (2.9s) than in the ATC-responsible condition (3.2s); (t(9)=.07, p>.05).
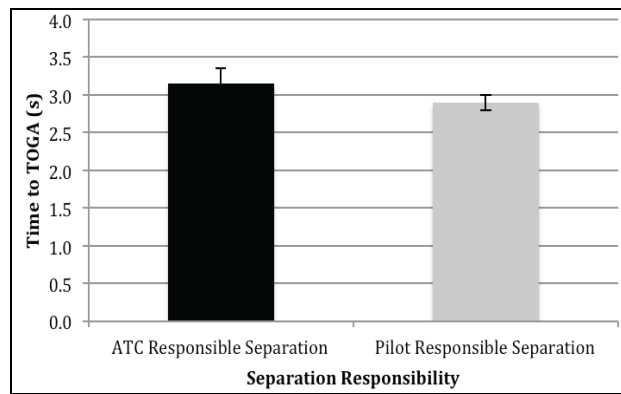

Figure 4. Time to TOGA as a Function of Crew Responsibility for Separation.

3.2 Flight Crew Scan Performance

The output of the flight crew scan performance during the alert phase was analyzed to determine the impact that the change from ATC-responsible to Pilot-responsible had on the pilots' scanning during the wake threat event. Scan performance was measured by the percent dwell time (PDT) that the pilot spend on three main areas of interest (AOIs): Primary Flight Display (PFD), Navigation Display (Nav), and Out the Window (OTW). As shown in Figure 5, when pilots were responsible for separation they spent more time monitoring the Nav containing wake information than when ATC was responsible for separation (t(19)=27.59, p<.01). As a result, *both* pilots spent less time monitoring OTW and critical flight performance data on the PFD in the pilot-responsible paradigm.

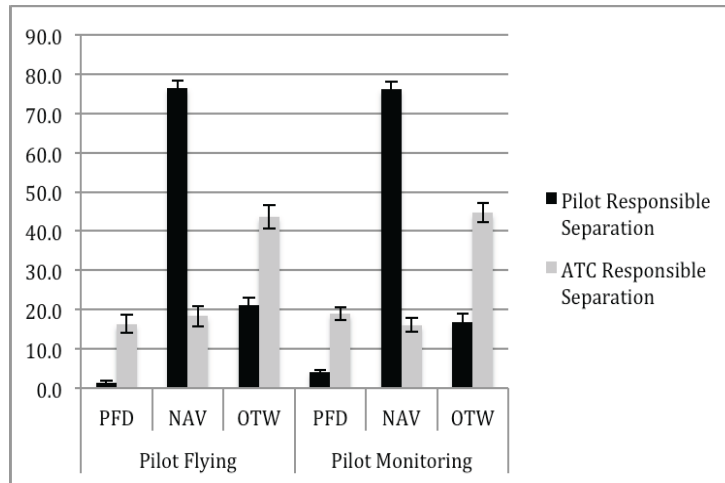**Figure 5. PF and PM Scan over the Total Alert Phase**

3.3 Mean Workload

Figure 6 presents mean workload over the total alert phase when ATC was responsible for separation and when pilots were responsible for separation. It is evident that the workload is predicted to be higher for both members of the flight crew when flying CSPO approaches when the responsibility for separation is shifted to the cockpit ($F(1,18)=31.48$, $p<.01$). This increased workload occurs because the PF and PM spend more time on workload-inducing tasks like visually tracking the lead aircraft, making mental comparisons to displayed information (on the Nav), determining the missed approach response, and making control inputs to maintain the correct spacing, tasks that are completed by the ATC when ATC is responsible for separation.
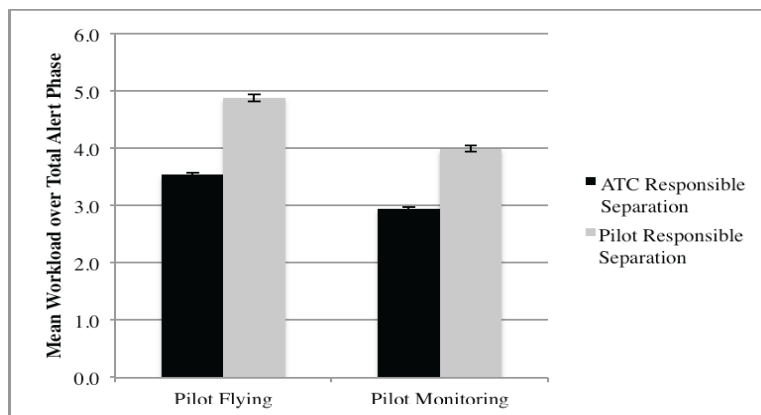

Figure 6. PF and PM Mean Workload Ratings Across the Total Alert Phase.

DISCUSSION AND IMPLICATIONS

An application model of commercial airline pilots conducting approach-and-landing procedures was created using the MIDAS software following a methodical development and validation approach. The premise that guided the current work was that model validity is a process, not solely a single value at the conclusion of a model development effort. Valid inputs lead to valid outputs. It is therefore necessary to follow an iterative input validation process as well as an iterative output validation process. Conducting only one of these validation processes may lead to invalid models. This is especially true as the complexity of the operational environment and tasks increase.

The current research has evaluated proposed changes to flight deck technologies, pilot procedures, operations, and roles and responsibilities that are likely with NextGen CSPOs through the use HPM outputs of time required to complete tasks, pilot scan performance, and pilot workload. The three measures of pilot performance (blunder response time, scan time, and workload) provide insights into potential costs of transitioning to a CSPO paradigm in which pilots are responsible for separation. CSPOs are predicted to influence the flight deck by directing a greater scan percentage to the Nav. Blunders in CSPOs are predicted to be detected faster (negligibly) when the flight deck is given responsibility for separation but this comes at a cost of increased workload for the PF and the PM and a reduced spread of visual attention across cockpit and out-the window. The flight-crew scan performance output reveals that in CSPOs, both pilots attend to the Nav during the alert phase the majority of their time. This is important because neither of the pilots is looking at the PFD, the primary display used for flight. This has important implications for NextGen CSPOs because it may be better for one of the pilots to focus on the Nav display, while the other pilot focuses his/her attention on the PFD. The appropriate allocation of attention is something that needs to be defined.

Future research, using HITL methods with both pilots and ATC, is required to better quantify the actual response-time benefit offered by the pilot-separation concept. In addition, system studies are required to determine if the response-time benefit provides a practical advantage in operational environments (i.e., is a .3s benefit practically meaningful?). Furthermore, HITL research is required to fully assess the workload and visual attention decrements associated with the pilot-separation scenario.

REFERENCES

Abbott, T.S. and D.M. Elliott. 2001. *Simulator Evaluation of Airborne Information for Lateral Spacing (AILS) Concept.* NASA/TP-2001-210665, March 2001.

Anders, G. 2001. *Pilot's Attention Allocation during Approach and Landing: Eye- and Head-Tracking Research in an A330 Full Flight Simulator.* Paper presented at the 11th International Symposium on Aviation Psychology, Columbus, OH.

Cox, G. 2010. Closely Spaced Parallel Runway Operations (CSPO). Presented at the CSPO Working Group Meeting, June 3, 2010. Available at: http://www.faa.gov/about/office_org/headquarters_offices/avs/offices/afs/afs400/afs440/cspo/meetings/media/cspo_working_group.pdf

Gore, B.F. 2008. Chapter 32: Human performance: Evaluating the cognitive aspects. In. V. Duffy ed., *Handbook of digital human modeling.* Boca Raton: CRC Press Inc. pp. 32:31-32:18.

Gore, B.F., B.L. Hooey, C. Socash, N. Haan, E. Mahlstedt, D.L. Bakowski, A.M. Gacy, C.D. Wickens, M. Gosakan, and D.C. Foyle. 2011. Evaluating NextGen Closely Spaced Parallel Operations Concepts with Human Performance Models. *HCSL Technical Report (HCSL-11-01).* Moffett Field, CA: NASA Ames Research Center.

Hooey, B.L., B.F. Gore, E. Mahlstedt, and D.C. Foyle. 2012. *Evaluating NextGen Closely Space Parallel Operations Concepts with Validated Human Performance Models: Flight Deck Guidelines (Part 1 of 2).* HCSL Technical Report (HCSL-12-01). Human Centered Systems Laboratory (HCSL), Moffett Field, CA: NASA Ames Research Center.

Hüttig, G., G. Anders, and A. Tautz. 1999. *Mode Awareness in a modern Glass Cockpit – Attention Allocation to Mode Information.* Paper presented at the 10th International Symposium on Aviation Psychology, Columbus, OH.

Mumaw, R.J, N. Sarter and C.D. Wickens. 2001. *Analysis of pilots' monitoring and performance on an automated flight deck.* Paper presented at the 11th International Symposium on Aviation Psychology, Columbus, OH.

Joint Planning Development Office (JPDO), 2011. Targeted NextGen Capabilities for 2025. JPDO: December.

Rutishauser, D., G. Lohr, D. Hamilton, R. Powers, B. McKissick, C. Adams, and E. Norris. 2003. Wake Vortex Advisory System (WakeVAS) Concept of Operations *NASA/TM-2003-212176, April.*

Verma, S., S. Lozito, T. Kozon, D. Ballinger, and H. Resnick. 2008. Procedures for off-nominal cases: Very closely spaced parallel runways approaches, Digital Avionics System Conference, St. Paul, MN.