

Simulation Study of Technology for Predicted Flight Deck Alerting of Energy

Timothy J. Etherington
Crew Systems and Air
Operations Branch
NASA Langley Research Center
Hampton, Virginia
timothy.j.etherington@nasa.gov

Lynda J. Kramer
Crew Systems and Aviation
Operations Branch
NASA Langley Research Center
Hampton, Virginia
lynda.j.kramer@nasa.gov

Laura Smith-Velazquez
Advanced Avionics
Collins Aerospace
Columbia, Maryland
laura.smith-velazquez@collins.com

Maarten Uijt de Haag
Institute of Aeronautics and
Astronautics
Technical University of Berlin
Berlin, Germany
maarten.uijtdehaag@tu-berlin.de

Abstract— Automation surprise and automation mode confusion are contributory causal factors in accidents and incidents in commercial aviation world-wide. Complex algorithms within the flight management system computer are used to calculate and then execute optimized paths based on navigation and procedure constraints while operating the aircraft as efficiently as possible. However, little pilot training is provided for where pilots can gain an understanding of what the automation is doing and if the automation will make future path constraints.

Flight simulation studies were conducted at NASA Langley Research Center as part of the Automation and Information Management Experiment (AIME) series. Prediction of future events in response to clearance changes and off-nominal conditions were evaluated using the Predicted Alerting of Energy display technology in two of these AIME studies. Display changes based on pilot intent were used to evaluate pilot understanding of stabilized approach criteria in response to clearance changes late in the approach. Acceptability and usability analysis and pilot observations were used to make recommendations on future study. The data showed that prediction technologies can contribute to pilot confusion due to uncertainty of input information and the time horizon for the prediction. Training recommendations are provided based on crew observations for use in a simulation study.

Keywords—flight crew error, aviation safety, predicted alerting, energy state prediction

I. INTRODUCTION

The Commercial Aviation Safety Team (CAST) is a government-industry partnership with a goal of reducing the commercial aviation accident rate by 50% in successive decades. Loss-of-control inflight was the greatest cause of accidents in the decade starting in 2010 and CAST recently completed an extensive study [1] of these accidents and incidents. Twelve common themes were identified and in the majority of accidents, more than one of the twelve themes were identified as a causal event. Implementation strategies were developed [2] and detailed safety enhancements (SE) were recommended for implementation. SE's may recommend new procedures, new training material, or they may be new design features. Some recommendations required further study and six research SE's were created. NASA, in cooperation with the FAA, was tasked with the research. The research was conducted in the framework of the NASA's Airspace Operations and Safety Program.

The Automation and Information Management Experiment (AIME) series were several spiral development experiments conducted at NASA Langley Research Center to address two of the six research SE's. SE-207 [3] was a study of attitude and energy state awareness technologies and to find technologies that could be raised to a technical readiness level of five or greater. SE-208 [4] had the same objectives for airplane state awareness.

Energy state awareness by a pilot follows the traditional situation awareness (SA) paradigm of perception (Level 1 SA), comprehension (Level 2 SA), and projection into the future (Level 3 SA) [5]. Past technology development has centered on the perception of the current state (Level 1 SA) and troubleshooting of aircraft systems when an event is occurring, providing Level 2 SA in the form of improved alerting, system information depiction, process automation, and improved planning tools; however, pilots are still surprised due to automation complexity [6] and lack of transparency in even routine operations. Though current technology has improved Level 2 SA, the improvements are still largely static, current state, reactive improvements. In reality, the conditions encountered by an aircraft in flight are dynamic. State awareness technologies need to include the impact of how a data-driven process changes the goal-driven decision making and the projection of future state in a continuous dynamic loop to include proactive technology solutions as decision aids [7]. The intent here is to aid the pilot in Level 3 SA, providing a good understanding of the situation combined with an understanding of the automation and aircraft performance projected into the future state and predicting hazards proactively.

Predicted Alerting of Energy (PAE) technology was developed at Ohio University in conjunction with Collins Aerospace under the direction of NASA Langley Research Center. PAE aids the flight crew with the information to provide the Level 3 SA - allowing them to understand complex future behavior of the aircraft and its automation.

This paper details the testing of PAE in the so-called AIME-2.5 experiment. A pilot-in-the loop flight simulation study was conducted at NASA Langley Research Center to evaluate the effectiveness of PAE technologies while conducting area navigation (RNAV) approaches in visual conditions. Objective and subjective data was collected from aircraft parameters, questionnaires, audio/video recordings, head/eye tracking data, and pilot observations.

II. TECHNOLOGY CONCEPT

PAE technology has been previously tested and reported in the AIME [8] series of experiments. In AIME-1, PAE was used to evaluate pilot understanding of stall, overspeed, and unstable approaches. The learning algorithms were extensively tested, tuned and evaluated [9] and information was presented [10] to pilots on the vertical situation display (VSD) and the engine indicating and crew alerting system (EICAS). Fig. 1 shows the example rules [9] used to produce the alerts. PAE was tested again in AIME 2.5 to further evaluate the technology and to provide a framework for decision making while conducting last minute runway changes with a predicted pilot intent method. Alert criteria exercised in AIME 2.5 were primarily high-fast and low-slow. An aircraft was high-fast when path deviation was greater than 1000 feet high and indicated airspeed 20 knots greater than selected airspeed. Low-slow criteria were file selectable and set at path deviation greater than 250 feet low and indicated airspeed less than 5 knots slower than selected airspeed. Stable approach criteria were active when any parameter listed was predicted to be outside limits when the aircraft was less than 1000 feet above the ground. The detailed algorithm performance [11] will not be reported here but the method and display presentation are described.

Safety-related rules	
Safety	C_s
Stall	$V_{stall}(\hat{x}_t, \hat{u}_t) < \hat{V}_t$
Overspeed	$\hat{V}_t < V_{max}(\hat{x}_t, \hat{u}_t)$
Bank angle	$ \hat{\phi}_t < \phi_{max}(\hat{x}_t, \hat{u}_t)$
Pitch angle	$\hat{\theta}_t < \theta_{max}(\hat{x}_t, \hat{u}_t)$
Load factor	$n_{min}(\hat{x}_t, \hat{u}_t) < n < n_{max}(\hat{x}_t, \hat{u}_t)$
Angle of attack	$\hat{\alpha}_t < \alpha_{max}(\hat{x}_t, \hat{u}_t)$
High-fast	$(h - h_{path}) > 1000ft$ $(V - V_{target}) > 20kts$
Low-slow	$(h - h_{path}) < TBDft$ $(V - V_{target}) < TBDkts$

Procedure-related rules	
Procedure	C_p
Stable approach	$V_{ref} \leq V \leq V_{ref} + 20 kts$ $\dot{h} > -1000 fpm$ ILS: $ loc < 1 dot, gs < 1 dot$ on path, configured, etc.
RNP navigation	$ANP_t < RNP_t(\hat{x}_t, \hat{u}_t)$ etc.
Waypoint constraints	$h_{@wpt} \leq h_{below}, or$ $h_{@wpt} \geq h_{above}, or$ $h_{above} \leq h_{@wpt} \leq h_{below}$

Fig. 1. Safety and procedure rules.

Aural and textual alerts were presented on the EICAS. The navigation display (ND) along with the integrated VSD was used to indicate the spatial location of the alerts as they related to the current aircraft navigation plan. Internal automation state is provided from an ARINC 702A-4 standard flight management system (FMS) intent bus. PAE indications on the ND and VSD are displayed on top of the trajectory prediction systems (TPS) [12] technology. Using TPS, the spatial location of PAE symbology is more accurate than if placed on the magenta-colored planned path trajectory from the FMS on the ND and VSD alone. Table 1 shows the predicted alerts shown on the EICAS, ND and VSD. A small triangle symbol preceded each

EICAS message to attempt to follow the Boeing philosophy where no symbol indicated a status message with no checklist and a small square indicates a message that has an associated checklist. The triangle was used to indicate the message was a predicted message to differentiate that message from all other EICAS messages. The standard color scheme was utilized where memo or status messages were white and cautionary messages were amber. The urgency of the message was determined from the prediction time and can be derived from Table 1. A message would usually be displayed as a memo first and when the prediction time dropped below a threshold, the message would be elevated to a caution. Fig. 2 shows a stall prediction memo example with the cyan stall prediction shown on the VSD and the white stall EICAS message.

Table 1. EICAS Message and Prediction Times

	EICAS message	ND/VSD symbol	Predicted Time
Overspeed memo	'Δ OVERSPEED PRED' (white)	OVSPD (cyan)	45 sec
Overspeed caution	'Δ OVERSPEED PRED' (yellow)	OVSPD (yellow)	10 sec
Stall memo	'Δ STALL PRED' (white)	STALL (cyan)	45 sec
Stall caution	'Δ STALL PRED' (yellow)	STALL (yellow)	10 sec
Unstable approach memo	'Δ UNSTABLE APP PRED' (white)	UNSTAB (cyan)	45 sec
High and fast memo	'Δ HIGH FAST PRED' (white)	HIGH FAST	300 sec
Low and slow memo	'Δ LOW SLOW PRED' (white)	LOW SLOW	300 sec
Runway specific memo	'Δ HIGH FAST 13L' (white)	-	45 sec
Runway specific memo	'Δ LOW SLOW 13L' (white)	-	45 sec

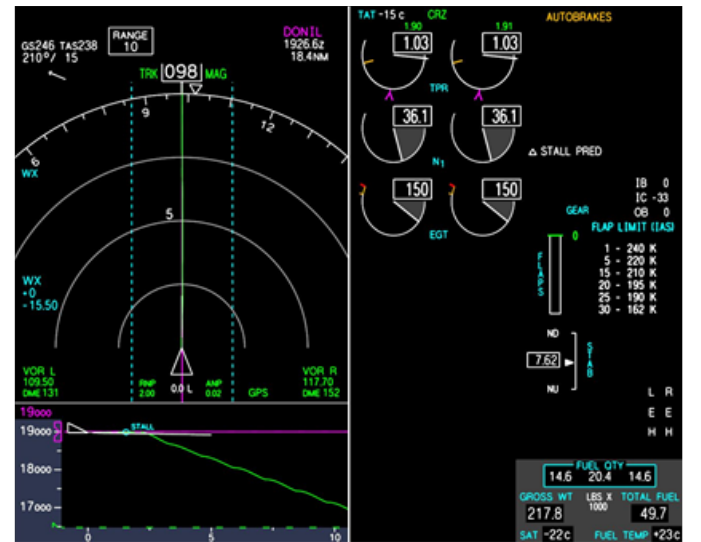


Fig. 2. Stall prediction example.

Fig. 3 provides an example of a high and fast indication on the VSD. The magenta line represents the FMS-planned vertical path, the green line comes from TPS and represents the mode transition at 4000 feet to an altitude hold. The cyan circle and HIGHFAST text indicate the spatial location when the aircraft is predicted to be too high and/or fast to make a stabilized approach. Note this represents a trivial example of the flight crew forgetting to set the missed approach altitude on an RNAV approach but one that is often experienced daily in the national airspace system.

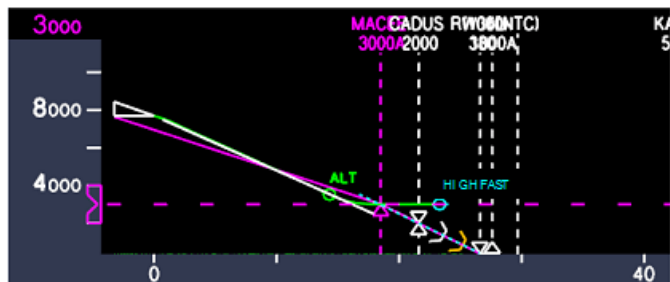


Fig. 3. High and fast example.

To further investigate, a predictor was developed in an attempt to determine pilot intent when changing runways late in the approach when there may not be time to update the original plan stored in the FMS. This change in plan could be due to emergencies or a late clearance. Using cross track deviation, extended runway centerline, bank angle, and the aircraft's spatial location, a learning algorithm was applied to determine when leaving the automaton plan, so it could be detected that the pilot was changing runways. If the lateral deviation is greater than a predefined amount, the aircraft is not correcting, and a parallel runway exists on that side, it is inferred that the pilot is changing runways. Alert criteria were then applied to the new runway using the prediction rules described above.

III. METHOD

AIME-2.5 was conducted at NASA Langley Research Center as part of the AIME series. Prediction of future events in response to clearance changes and off-nominal conditions were evaluated using the PAE technology concept. Display changes based on pilot intent were used to evaluate pilot understanding of stabilized approach criteria in response to clearance changes late in the approach and in response to off-nominal events. Acceptability and usability analyses and pilot observations were used to make recommendations on future study.

A. Scenarios

A baseline run was flown for crew familiarity and training using a flight from takeoff on KDCA Runway 19 to KJFK Runway 13L, designated Scenario 0a. Weather for the baseline run was ceiling 1000 feet, visibility 3 miles.

Experimental scenarios were designed to place crews into several plausible high/fast or low/slow events for exercising PAE and the pilot intent elements of the technology concept. These scenarios started with the aircraft initially on the CAMRN4 arrival into KJFK using the RNAV procedure. Note that the RNAV procedure to Runways 13L and 13R is initially the same approach path until near the airport boundary. During visual conditions the procedure is annotated for pilots to expect

clearance to either runway end. Runway 13L is 10,000 feet long and 150 feet wide. Runway 13R is 14,511 feet long and 200 feet wide. For the experimental scenarios weather was clear and unlimited visibility; the winds were 210 degrees at 15 knots resulting in a direct right cross wind.

For Scenario 9a, the crew was initially cleared for an RNAV approach to Runway 13L. A bird strike resulted in a left engine failure while descending on the RNAV 13L arrival near the time when the procedures to the two parallel runways are still coincident. Right after the failure, ATC contacted the crew and indicated smoke from the left engine and offered the closer Runway 13R, which was also longer and wider. (If the crew declined, the principal investigator instructed the crew to take 13R anyway.) Maneuvering to the new runway caused the crew to initially be high/fast for the approach.

For Scenario 9b, the crew was initially cleared on the RNAV approach to Runway 13R. Right before the turn inbound, the crew was cleared to the parallel Runway 13L which initially placed the aircraft in the low and/or slow condition. Radio transmissions on tower frequency indicated Runway 13R was closed due to a landing aircraft that blew a tire and emergency vehicles were being dispatched.

For Scenario 9c, the crew was cleared on the RNAV approach to Runway 13R. Just as the crew was descending near 3000 feet of altitude, ATC contacted the crew about a drone sighting and instructed the crew to fly right at a heading of 050 degrees and maintain 3000 feet. The heading toward the airport boundary and altitude clearance kept the aircraft from descending and placed the aircraft in a very high-altitude condition for landing. When the aircraft was just outside a right downwind to the runway, the crew was instructed by ATC that they were clear of the conflict and they were cleared to land on Runway 13R. No further instructions were given unless the crew asked for heading vectors.

B. Environment

The experiment was conducted at NASA Langley Research Center in the Research Flight Deck (RFD). The RFD is one of a set of modular cabs in the Cockpit Motion Facility (CMF). The simulator uses a high-fidelity Boeing B-757-200 aircraft model with extended post-stall behaviors. The overhead panel, mode control panel, and much of the center isle stand are classic B-757.

The RFD was modified, see Fig. 5, for AIME testing to represent a modern flight deck to ensure technologies being tested were at least as good or better than the best current design. This reference aircraft was chosen as the B-787 and the flight deck was modified with three large landscape displays in the multi-function display (MFD) to reflect the B-787 flight deck design. The lower display was replaced, and multi-function keypads were installed to replicate the B-787 virtual control display unit (CDU) functionality. Electronic Flight Bags (EFB) were installed outboard in the same location with the same functionality as the B-787. Datalink functionality was added and used during testing to represent future flight deck operations.

The RFD was placed on a hexapod motion base and the experiment was conducted under full motion. Realistic high-quality airport scenes were generated using photorealistic terrain

data. The scenes were presented on a collimated out-the-window, wide field-of-view display projected from a Collins Aerospace graphics system. The installed head-up displays were stowed for this evaluation.



Fig. 4. Research flight deck simulator.

C. Subjects

Twenty pilots (10 crews), representing three major airlines, participated in this experiment. 17 subjects were male and 3 were female. Each pilot held an Airline Transport Pilot rating and was current in the wide-body Boeing aircraft as either Captain or First Officer. Crews were paired from the same airline to minimize inter-crew conflicts in Standard Operating Procedures (SOPs) and Crew Resource Management (CRM) training. The Captains' average age was 60 years with average flight hours of 25,400. The First Officers' average age was 55 years with average flight hours of 14,831.

Pilot flying (PF) and pilot monitoring (PM) roles were assigned based on the experimental matrix.

D. Experiment Design

PAE was tested as part of a larger experiment within AIME-2.5 and some results [12, 13] were previously reported. Crews were briefed on the overall flight deck and specifically on all new technology. New technology and operation with the display control panels was covered followed by a training flight from takeoff to landing where each pilot could use the new technology. Scenario order and pilot roles were randomized within each crew.

IV. RESULTS

PAE was evaluated during visual approach conditions, with and without off-nominal events. The results presented represent subjective measurement of workload, acceptability and usability, situation awareness, and trajectory performance while conducting the approach. ANOVAs, using a significance value of 0.05, with pilot role and scenario as the main factors were conducted on the subjective measures. Additionally, subjective results are presented using an interval plot showing a 95% confidence interval for the mean. Median value symbols (circle with an x) are included on the interval plots. Objective measurements included successful landing and being able to establish stabilized approach criteria after a side-step maneuver during the approach to landing. The approach criteria included excessive airspeed deviation, excessive path/glidepath

deviation, and excessive descent rate. Since the scenarios used were off-nominal, including engine failure, side-step maneuvers and curved approaches with no straight segment until less than 500 feet above the runway, stable approach criteria were applied at 500 feet with a final evaluation at 300 feet. These objective results are presented using individual value plots.

A. Workload

Workload was measured using an industry standard multi-index measurement developed by NASA [14] and designated as task load index, or TLX. An ANOVA revealed no significant workload differences ($p > 0.05$) for pilot role or scenario. Workload is reported in Fig. 5 for pilot role (PF or PM) for each different scenario. The baseline run, Scenario 0a, was conducted as a training scenario and this may be reflected in the scores. Workload is generally rated low to moderate, with increased workload reflected in the more complex scenarios and in pilot role. Note that the highest workload is for PF on scenario 9a, the bird strike and loss of engine scenario. Even with the complex scenario, the side-step to the parallel runway was accomplished with only moderate workload reported (mean value of 50 for PF and 40 for PM). Slight differences in workload are reflected in the complexity of the scenarios flown. PF workload was greater than PM for all scenarios (mean value of 44 for PF and 38 for PM).

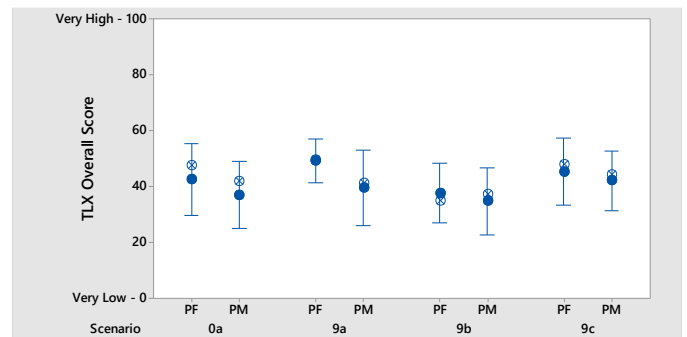


Fig. 5. Workload scores.

B. Usability and Acceptability

The System Usability Scale (SUS) [15] was used to gauge how pilots perceived usability of the PAE technology. SUS is not a percentile scale but as described in [16], an SUS score between 68 and 80.3 is considered a “good” design and a score above 80.3 is considered an “excellent” design. Acceptability was measured using a standard Likert-type scale from 1 to 7, where 1 was very unacceptable, 4 was average, and 7 was very acceptable. Usability and acceptability are reported in Fig. 6 and Fig. 7 according to pilot role for each different scenario flown.

An ANOVA revealed no significant ($p > 0.05$) differences in SUS scores for pilot role or scenario flown. PAE was rated as a good to excellent design across all scenarios, see Fig. 6. Median values ranged between 75 and 94 for the PF and 74 to 90 for the PM. Acceptability scores were high with median scores ranging between 5.5 and 6 for all scenarios and crew position, see Fig. 7. Acceptability nearly mirrors usability across all conditions.

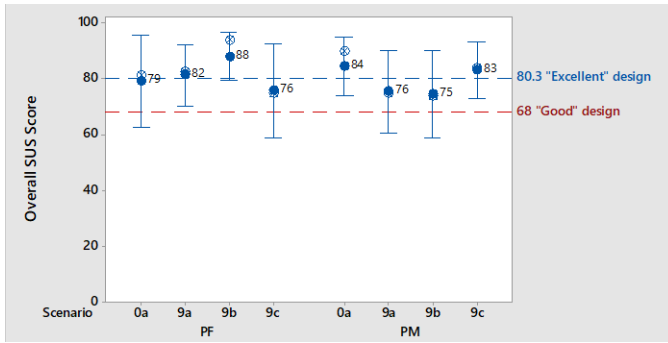


Fig. 6. Usability scores.

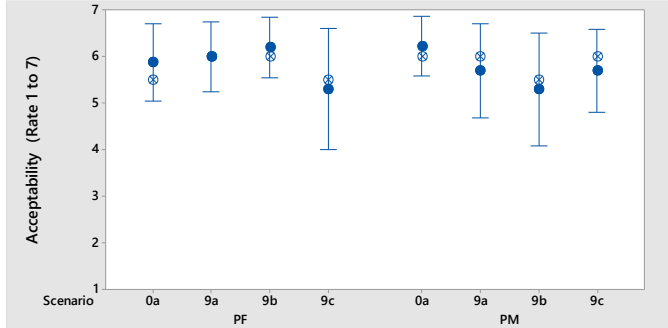


Fig. 7. Acceptability ratings.

C. Situation Awareness

Situation awareness (SA) was measured by using the 3-D Situational Awareness Rating Technique (SART) [17]. It calculates SA based on three dimensions: (1) Demand on pilot attentional resources, (2) Supply of pilot attentional resources, and (3) Understanding by the pilot of the situation. Pilots rated these three components after each flight on a scale of 0 (very low) to 100 (very high). From these ratings, SA is calculated as:

$$SA = \text{Understanding} - (\text{Demand} - \text{Supply})$$

The SART technique provides relative SA pilot ratings among the conditions evaluated.

An ANOVA revealed no significant ($p > 0.05$) differences in SART ratings due to pilot role or scenario. However, as detailed in Fig. 8, perceived SA appears dependent on the scenario flown. The PF reported SA gains with PAE in Scenarios 9b and 9c compared to the baseline scenario (without a runway change) but reported reduced SA in Scenario 9a compared to the baseline. This result was most likely influenced by the complex scenario design of 9a and high workload environment experienced by the PF. PAE only supported part of the PF task for reacquiring a new runway path.

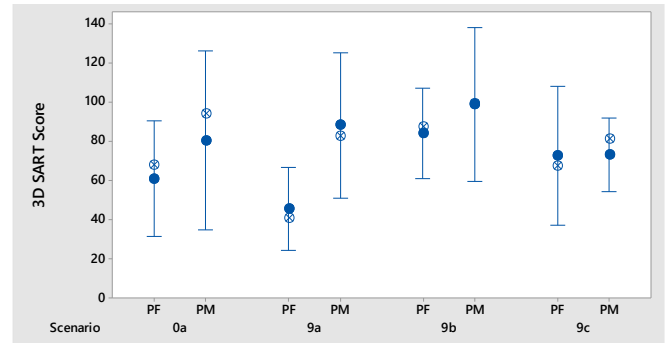


Fig. 8. Situation awareness scores.

D. Trajectory Performance

Fig. 9 provides a sample of actual trajectories (color-coded by crew) for the two side-step scenarios. For visual clarity, only a subset of the 10 runs flown in each scenario were used in this figure. Scenario 9a was the bird-strike and loss of the left engine. A notice was provided by ATC of smoke and an offer for the longer parallel runway allowed for evaluation of the PAE technology for the side-step. The ATC call was timed to coincide with approaching the right extended centerline, so a glance out-the-window provided a tantalizing view of the closer and longer parallel runway. Note the right cross-wind and a left engine failure with associated right yaw from the operating good engine both made a side-step to the right runway more difficult. It is not surprising that there was some overshoot from the extended runway centerline. As soon as the crew maneuvered inside the path from RNAV 13L, each crew received a high/fast advisory for Runway 13R indicating PAE sensed the pilot intent to proceed to the parallel runway, even though the plan in the FMS had not been updated. Although primarily a visual maneuver, PAE provided indications when the aircraft was high on the new approach path and when the aircraft was stable on the new path. If the alerts were timely, crews would have the opportunity to correct the path and achieve a stable approach. Workload precluded crews from changing V_{ref} after the engine failure, but it is good crew practice to add speed as the single engine approach speed would be about 12 knots faster. The increased speed will necessitate increased vertical speed. Maneuvering close to the runway and trying to capture path from above, many crews experienced an additional alert for low and slow Runway 13R while still high and before crews corrected the descent rate for a stable approach. Although a technically correct predictor behavior, these indications caused crew confusion.

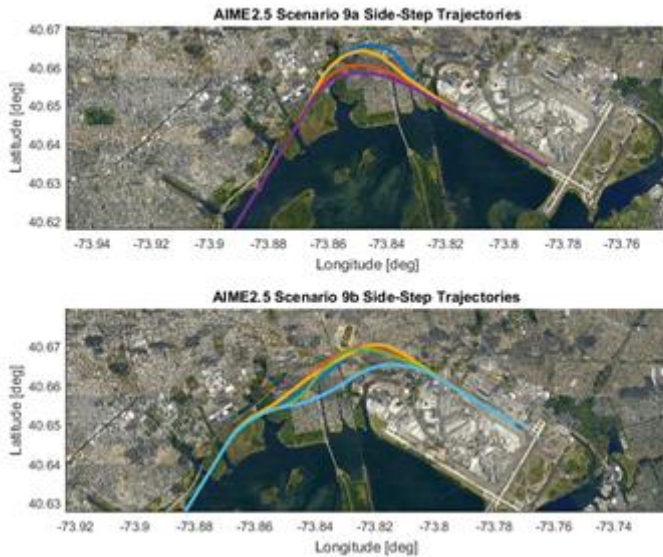


Fig. 9. Side-Step trajectories.

Objective data was evaluated using the Flight Safety Foundation's new stabilized approach and go-around guidelines [18]. The recommendations are for the aircraft to be configured for landing by 1000 feet, evaluate if stable by 500 feet but allow for correcting and if not stable by 300 feet, the crew should go-around. The criteria are indicated airspeed between V_{ref} and $V_{ref}+10$ knots, localizer and glideslope deviation less than 1 dot or path deviations less than half scale, and vertical speed is 1000 fpm or less. Fig. 10 shows the airspeed difference slightly higher than the V_{ref} entered before the engine failure but well within range of the speed reflected from single engine operations at the 500-foot evaluation. Airspeed difference is within 10 knots at the 300-foot go-around decision. Fig. 11 shows the descent rate higher than desired but reflective of the higher single engine airspeed and the difficulty of capturing the path after being forced high on the approach. All crews continued and all but one met the stable approach criteria by the 300-foot go-around decision. One crew was slightly over the 1000 fpm threshold as they held a higher airspeed throughout the approach. For the final straight segments of the approach to the runway for all the scenarios, all crews were configured for landing and within $\frac{1}{2}$ dot on both localizer and glideslope. All crews executed a successful landing within the lateral confines of the runway and touchdown zone.

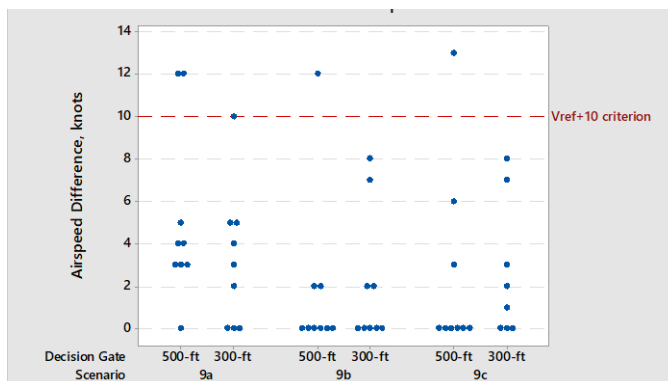


Fig. 10. Airspeed Stabilization at 500 and 300 ft decision points.

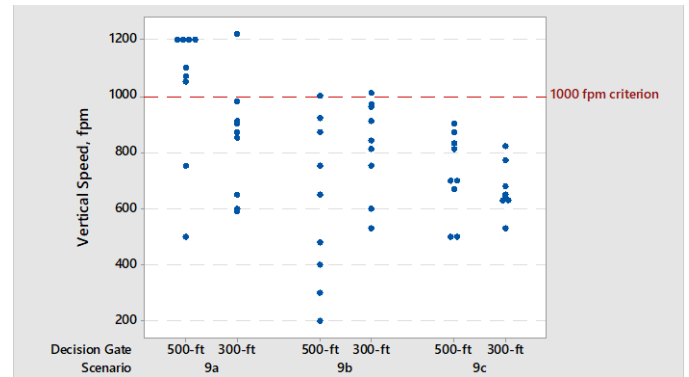


Fig. 11. Vertical speed stabilization at 500 ft and 300 ft decision points.

The second side-step shown in Fig. 9 for Scenario 9b is when the crew was given a new runway clearance to Runway 13L due to emergency vehicles on the original cleared runway, 13R. As soon as the crew flew past the extended runway centerline for 13R and maneuvered to Runway 13L, a low/slow runway 13L alert was generated. Although guidance was not provided for the new path and the crews maneuvered visually, a low approach to a runway is relatively easy to correct by shallowing the descent for a short period of time. No false negatives or reversals of alert were generated by the predictor. Except for the one crew that was 12 knots fast, Fig. 10 shows all crews stable at 500 feet. By 300 feet all crews were stable on airspeed. Fig. 11 reflects a stable descent rate at both 500 foot and 300 foot decision points (one crew's descent rate of 1010 fpm was considered close enough to meet 1000-fpm criterion). For the final straight segments of the approach to the runway for all the scenarios, all crews were configured for landing and within $\frac{1}{2}$ dot on both localizer and glideslope. All crews executed a successful landing within the lateral confines of the runway and touchdown zone.

Fig. 12 shows the trajectory for Scenario 9c, the drone encounter. A normal downwind pattern altitude would be around 1500 feet and spaced laterally near the location that the crew is cleared to land. During the encounter, the crew was instructed to maintain 3000 feet, so the aircraft was very high compared to a stabilized approach to the runway, even if allowed an extended downwind. Note the one crew that flew a diagonal intercept to the path. This crew had even less distance to lose the required altitude and this crew conducted a go-around.

Many crews commented on lack of guidance or even lack of indication to the severity of the off-path condition. The most common comment was "why didn't the system just change the plan in the FMS". Flown correctly, this scenario had the longest straight final approach segment after the drone encounter and that is reflected in all criteria with small tolerances. Except for one crew that was 13 knots fast at 500 ft, Fig. 10 shows all crews stable on airspeed at both 500 foot and 300 feet decision points. Fig. 11 show all crews stable at 500 feet and 300 feet in descent rate. For the final straight segments of the approach to the runway for all the scenarios, all crews were configured for landing and within $\frac{1}{2}$ dot on both localizer and glideslope. All crews, except for the one that conducted a go-around, executed a successful landing within the lateral confines of the runway and touchdown zone.



Fig. 12. Trajectories for drone encounter scenario.

V. CONCLUSIONS

Workload was consistent across roles and scenarios so no increase in workload is associated with using the PAE technology and similarly, no improvement in workload either. Since the baseline workload measurements were collected from a nominal training flight with no approach changes, it is not surprising that workload didn't improve. PF workload was reported as moderate with scenario dependent increases in workload as expected. PM workload was low across all scenarios with slight increase in workload reported for the bird-strike and drone scenarios.

Using the SUS scores, PAE was rated as a good to excellent design by the PF across all experimental scenarios. For the PM, PAE was rated slightly lower for the two side-step maneuvers, where the SUS score was just under an excellent design. For the PM task, this slight reduction is likely due to missing guidance or deviation to the parallel runway needed to perform the monitoring task. This is opposite for the PF, where the SUS scores were higher for the side-step tasks when the PAE technology provided pertinent alerts but lower for the drone scenario where more maneuvering was required which the technology as-tested did not support as well.

SA was dependent on the scenario, and for the PM appears to contradict the usability and acceptability ratings, as the PM reported an increase in SA for the two side-step maneuvers compared to the baseline and a decrease in SA for the drone scenario. The PF reported a decrease in SA compared to the baseline for scenario 9a, the bird-strike scenario. This is likely attributed to the technology providing no help for the engine failure or the loss of thrust.

Algorithm performance [9] was previously reported and indicates appropriate tracking for pilot intent and for prediction of unstable approaches. The learning algorithms performed well in the study. Typically, within 20 seconds of starting a side-step maneuver, the algorithm detected the change and provided alerts to the parallel runway for the rest of the event. Apart from the drone event, the two side-step maneuvers did not require excessive descent rates or complex maneuvers. A simple vertical deviation indicator could have provided some indication for altitude deviation from a 3-degree path to the runway. The

simple indicator works well when the path can be easily reached without lateral maneuvering, but if significant maneuvering is required, as in the pattern flown for the drone encounter, some form of guidance may be required. Many transport category aircraft have extreme difficulty in intercepting any stabilized approach path from well above the path. Finding the optimum guidance would require further research. Another potential method would be to provide a highway-in-the-sky when a runway change is detected. Having an indication that shows current deviation integrated with current descent path could be a powerful tool if it shows that closure rate is sufficient to regain the stabilized approach path before reaching 500 feet AGL. These concepts will require further study. Except for the crew that conducted a go-around, all crews but one were considered to have met the stable approach criteria by 300 feet and all conducted a successful landing.

Pilot confusion was observed for some of the highly dynamic conditions encountered in the off-nominal events. With further training and explanation about how predictors react, this confusion can be eliminated. This training was not provided to keep pilots from guessing what problems they were going to encounter in the evaluation, but further testing could use more realistic training. To understand potential confusion, one example is if an aircraft is high on the approach path and correcting with a high vertical speed. In this dynamic condition with a vertical speed greater than that required to maintain the stable path, the predictor initially provides an indication that the aircraft is high and fast. When the aircraft nears the appropriate trajectory but before maneuvering is initiated to capture the stabilized path, the predictor, looking into the future, determines that if the aircraft continues along that path it will now be low and slow. This second alert is confusing without understanding explicitly how the predictor works. Because the predictor works over a long-time scale, it needs to be carefully tuned to reduce those effects if possible. Other methods will be explored to determine if the second alert may be suppressed or the time scale altered until passing through the stable path.

ACKNOWLEDGMENT

This work was sponsored by NASA's Airspace Operations and Safety Program, System-Wide Safety project, led by John Koelling, Project Manager. This work is also covered in part under an Inter-Agency Agreement with the Federal Aviation Administration. Special thanks to Sabreena Azam, Ashley Awward, and John Stuber of the FAA for their support and guidance along the way.

REFERENCES

- [1] Commercial Aviation Safety Team; "Airplane State Awareness Joint Safety Analysis Team Final Report," Skybrary Bookshelf, Jun 2014.
- [2] Commercial Aviation Safety Team; "Airplane State Awareness Joint Safety Implementation Team Final Report," Skybrary Bookshelf, Jun 2014.
- [3] Commercial Aviation Safety Team; "Safety Enhancement SE 207 ASA Research – Attitude and Energy State Awareness Technologies," Skybrary Bookshelf, Jun 2014.
- [4] Commercial Aviation Safety Team; "Safety Enhancement SE 208 ASA Research – Airplane Systems Awareness," Skybrary Bookshelf, Jun 2014.

- [5] Endsley, M. R. (2000). Theoretical underpinnings of situation awareness: A critical review. In M. R. Endsley & D. J. Garland (Eds), *Situation awareness analysis and measurement* (pp. 3-32). Mahwah, NJ: LEA.
- [6] Mumaw, R.J., Sarter, N., Wickens, C., Kimball, S., Nikolic, M., Marsh, R., Xu, W., & Xu, X. (2000). Analysis of pilots' monitoring and performance on highly automated flight decks. (Final Project Report: NASA Ames Contract NAS2-99074). Seattle, WA: Boeing Commercial Aviation.
- [7] Fan, T.P., Hyams, D.S., Kuchar, J.K., (1998). Study of In-Flight Replanning Decision Aids. AIAA Guidance, Navigation, and Control Conf., Boston, MA, August 10-12, 1998.
- [8] S. Young, T. Daniels, E. Evans, E. Dill, M. Uijt de Haag and T. Etherington, "Flight Simulation Study of Airplane State Awareness and Prediction Technologies," in 35th AIAA/IEEE Digital Avionics Systems Conference (DASC), Sacramento, CA, September 25-29, 2016.
- [9] P. Duan, "Predictive Alerting for Improved Aircraft State Awareness," Electronic Thesis or Dissertation, Ohio University, 2018.
- [10] P. Duan, M. Miltner, M. Uijt de Haag, M. Yocius, J. Engler and T. Schnell, "Human-In-The-Loop Evaluation of an Information Management and Notification System to Improve Aircraft State Awareness," in AIAA Scitech Forum, Kissimmee, Florida, January 5-9, 2015.
- [11] M. Uijt de Haag, J. Engelmann and C. Mourning, "Performance Evaluation of Predictive Alerting of Energy Methods using Flight Simulator Data," in AIAA SciTech Forum, San Diego, California, January 7-11, 2019.
- [12] T. Etherington, L. Kramer, K. Shish, S. Young, E. Evans and T. Daniels, "Increasing Pilot's Understanding of Future Automation State - An Evaluation of an Automation State and Prediction System," in 35th AIAA/IEEE Digital Avionics Systems Conference (DASC), San Diego, CA, September 9-12, 2019.
- [13] E. Evans, L. Kramer, T. Daniels, S. Young, Y. Santiago-Espada, J. Barnes and T. Etherington, "Usability Evaluation of Indicators of Energy-Related Problems in Commercial Airline Flight Decks," in AIAA/IEEE Digital Avionics Systems Conference (DASC), San Diego, CA, September 9-12, 2019.
- [14] Hart, S.G. and Staveland, L.E., "Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research." In P.A. Hancock & N. Meshkati (Eds.), *Human Mental Workload*. Amsterdam: North-Holland, 1988, pp. 139-183.
- [15] "System Usability Scale (SUS)," <https://www.usability.gov/how-to-and-tools/methods/system-usability-scale.html>.
- [16] J. Sauro, *A Practical Guide to the System Usability Scale*, Denver: Measuring Usability, 2011.
- [17] FAA, "Guide to Human Performance Measurement," American Institute of Aeronautics and Astronautics, Washington D.C., 1993.
- [18] Flight Safety Foundation, "Go-Around Decision-Making and Execution Project Final Report", March 2017. Retrieved on June 8, 2020 at https://flightsafety.org/wp-content/uploads/2017/03/Go-around-study_final.pdf.