NASA/CR–2020-220441

Honeywell Enhancing Airplane State Awareness (EASA) Project: Final Report on Refinement and Evaluation of Candidate Solutions for Airplane System State Awareness

Stephen D. Whitlow, and Michael B. Dillard
Honeywell Aerospace Advanced Technology, Plymouth, Minnesota

February 2020

NASA STI Program . . . in Profile

Since its founding, NASA has been dedicated to the advancement of aeronautics and space science. The NASA scientific and technical information (STI) program plays a key part in helping NASA maintain this important role.

The NASA STI program operates under the auspices of the Agency Chief Information Officer. It collects, organizes, provides for archiving, and disseminates NASA's STI. The NASA STI program provides access to the NTRS Registered and its public interface, the NASA Technical Reports Server, thus providing one of the largest collections of aeronautical and space science STI in the world. Results are published in both non-NASA channels and by NASA in the NASA STI Report Series, which includes the following report types:

- TECHNICAL PUBLICATION. Reports of completed research or a major significant phase of research that present the results of NASA Programs and include extensive data or theoretical analysis. Includes compilations of significant scientific and technical data and information deemed to be of continuing reference value. NASA counter-part of peer-reviewed formal professional papers but has less stringent limitations on manuscript length and extent of graphic presentations.

- TECHNICAL MEMORANDUM. Scientific and technical findings that are preliminary or of specialized interest, e.g., quick release reports, working papers, and bibliographies that contain minimal annotation. Does not contain extensive analysis.

- CONTRACTOR REPORT. Scientific and technical findings by NASA-sponsored contractors and grantees.

- CONFERENCE PUBLICATION. Collected papers from scientific and technical conferences, symposia, seminars, or other meetings sponsored or co-sponsored by NASA.

- SPECIAL PUBLICATION. Scientific, technical, or historical information from NASA programs, projects, and missions, often concerned with subjects having substantial public interest.

- TECHNICAL TRANSLATION. English-language translations of foreign scientific and technical material pertinent to NASA's mission.

Specialized services also include organizing and publishing research results, distributing specialized research announcements and feeds, providing information desk and personal search support, and enabling data exchange services.

For more information about the NASA STI program, see the following:

- Access the NASA STI program home page at http://www.sti.nasa.gov

- E-mail your question to help@sti.nasa.gov

- Phone the NASA STI Information Desk at 757-864-9658

- Write to:
  NASA STI Information Desk
  Mail Stop 148
  NASA Langley Research Center
  Hampton, VA 23681-2199

NASA/CR–2020-220441

Honeywell Enhancing Airplane State Awareness (EASA) Project: Final Report on Refinement and Evaluation of Candidate Solutions for Airplane System State Awareness

Stephen D. Whitlow, and Michael B. Dillard
Honeywell Aerospace Advanced Technology, Plymouth, Minnesota

National Aeronautics and
Space Administration

Langley Research Center
Hampton, Virginia 23681-2199

Prepared for Langley Research Center under
Contract NNL16AA06B/NNL17AA31T

February 2020

Acknowledgments

The authors would like to thank the NASA PM, Dr. Steven Young, for his support and guidance during the execution of this program. The authors would acknowledge the valuable contribution of the Honeywell Team including Dr. Umut Orhan, Aaron Leander, Dr. Barbara Holder, Dr. Aaron Gannon, and John Suddreth. The technology evaluation documented in this report, AIME3, was successful due to a close collaboration with the Langley Research team lead by Lynda Kramer, James Barnes, and Tim Etherington.

Available from:

NASA STI Program/Mail Stop 148
NASA Langley Research Center
Hampton, Virginia 23681-2199
Fax: 757-864-6500

ABSTRACT

The loss of pilot airplane state awareness (ASA) has been implicated as a factor in several aviation accidents identified by the Commercial Aviation Safety Team (CAST). These accidents were investigated to identify precursors to the loss of ASA and develop technologies to address the loss of ASA. Based on a gap analysis, two technologies were prototyped and assessed with a formative pilot-in-the-loop evaluation in NASA Langley's full-motion Research Flight Deck. The technologies address: 1) data source anomaly detection in real-time, and 2) intelligent monitoring aids to provide nominal and predictive awareness of situations to be monitored and a mission timeline to visualize events of interest. The evaluation results indicated favorable impressions of both technologies for mitigating the loss of ASA in terms of operational utility, workload, acceptability, complexity, and usability. The team concludes that there is a feasible retrofit solution for improving ASA that would minimize certification risk, integration costs, and training impact.

Table of Contents

List of Figures

List of Tables

Acronyms

Acronym	Description
AOA	Angle of Attack
AOSP	Airspace Operations and Safety Program
ASA	Airplane State Awareness
CAST	Commercial Aviation Safety Team
COA	Course of Action
EICAS	Engine-Indicating and Crew-Alerting System
FMC	Flight Management Computer
FO	First Officer
IDAD	Input Data Anomaly Detection
IRB	Institutional Review Board
MEL	Minimum Equipment List
PF	Pilot Flying
PFD	Primary Flight Display
PITL	Pilot-in-the-Loop
PM	Pilot Monitoring
RA	Radar Altimeter
SART	Situation Awareness Rating Technique
SOP	Standard Operating Procedures
SW/SSOT	Status Widget & Support Suite of Tools
SUS	System Usability Scale

Introduction

Under NASA's Airspace Operations and Safety Program (AOSP), Honeywell is developing advanced technologies to address loss of airplane state awareness (ASA) by pilot. First, the team reviewed past accidents and incidents to identify situations and precursors to situations where pilots constructed inaccurate or incomplete ASA. These included events identified by Commercial Aviation Safety Team (CAST) (CAST 2014a; CAST, 2014b), as well as 24 loss of awareness events between January 2007 and January 2016 identified by the Honeywell team. Next, the team cataloged information support for these situations as provided by modern commercial flight decks. Based on these analyses, the team then generated conceptual designs for technologies that may help to improve airplane system state awareness. These were refined through feedback session with Honeywell pilots and system engineers (Whitlow et al., 2017).

We have employed a design philosophy that ASA should be seamlessly integrated into legacy displays that support aviate and manage systems tasks. These displays command the lion-share of flight crew attention. Continuously and explicitly conveying ASA with these displays is the most cognitively efficient means and should not only enhance ASA in normal situations, but also will facilitate safe operation during non-normal situations by co-locating ASA elements with primary flight displays.

Conceptual designs have been developed that we believe will address the following research aims related to the broader goals of this class of technologies:

- Eliminate the opaque wall between crew and automation
- Improve the crew's ability to recognize and respond to non-normal situations that accommodate inherent attention and monitoring limitations

In Year 2 of the effort, we conducted a pilot-in-the-loop (PITL) formative evaluation in Honeywell's Redmond B737-NG Simulator to demonstrate prototypes in a simulated flight environment and to elicit initial pilot feedback on the usability and acceptability of the designs. This was to ensure we are pursuing a fruitful path and assess whether the design would have lower certifiability risk and could be implemented in a cost-effective manner.

We found that pilots rated both technology interventions favorably in terms of subjective workload, acceptability, complexity, and usability. In addition, pilots provided many suggestions to improve the designs and expand the functionality. Most importantly, pilots were supportive of an integrated solution that would include a simpler indication on the Primary Flight Display (PFD) while maintaining the valuable information content on the Engine-Indicating and Crew-Alerting System (EICAS) display. This would better conform to display conventions while reducing certification risk.

Accordingly, we matured the conceptual designs for the present evaluation in the NASA Langley Cockpit Motion Facility's Research Flight Deck. The work reported here is related to previously reported evaluations conducted at NASA (Young et al., 2016a; Young et al., 2016b).

Technology Description

During this phase, we integrated two related, synergistic technology concepts to better support pilots maintaining awareness of airplane states. Both technologies are based on intelligent modules that analyze and reason on the state of aircraft and mission as it relates to monitoring requirements. While these technologies were evaluated separately in Year 2, based on the Honeywell evaluation feedback, we integrated the two technologies to operate synergistically.

The Input Data Anomaly Detection (IDAD) technology identifies when data critical to pilot awareness and/or autoflight function is likely anomalous. The Status Widget & Support Suite of Tools (SW/SSOT) that provides a timeline indication on EICAS to enable the pilot monitoring (PM) to quickly assess the state of monitoring requirements. The SW provides high-level symbology, while the SSOT provides lower-level details that pilots may optionally 'pull-up' to explain SW indications. This includes relevant information to explain events depicted in the SW and support pilot decision making.

IDAD

The intended function of the IDAD technology is to identify when data critical to pilot awareness and autoflight function is likely anomalous, for example in the case of a frozen pitot tube, and then notify the crew per the needs of the situation. This can serve to augment existing voting methods with robust statistical and analytic methods with respect to how flight critical input data sources are validated. This involves developing a statistical model of sensor input to evaluate whether current values deviate or are trending toward a deviation. If a deviation is detected, then the system executes decision logic to determine which actions or notifications the system should make for the observed anomaly.

In addition to detection, the module is intended to be diagnostic in that it can identify the most likely anomalous parameter to help pilots understand which information source is reliable and which is not. Consideration will include several factors (e.g., whether the parameter is in use; whether there is a flight path impact), to determine the appropriate response. Based on incident analysis and pilot reviews, an important function of response to anomalies is to prevent a “first bad move” by pilots, possibly exacerbated by startle/surprise response to erroneously presented information.

Predicting Anomalies

In addition to detecting data anomalies that are not currently monitored, the IDAD technology also addresses startle/surprise by predicting when data anomalies may occur. Based on the data analytic approach used to determine when a data point is anomalous, under certain situations it is possible to observe data points over time and predict when an anomaly may occur based on a trend.

For a given sensor or metric, a threshold is established to distinguish typical expected values from atypical unexpected values (i.e., anomalous values). If this threshold is crossed, an anomaly is said to be happening at present (“current anomaly”). Before this happens, if the observed values are trending toward an anomaly, prediction is performed to estimate when the threshold will be crossed in the future. For the version tested, the timing of the predicted anomaly is calculated via a linear extrapolation. The figure below provides an example of how anomalies are predicted. As the observed values (represented as blue dots) approach the anomaly threshold for this metric (standard deviation of airspeed), a prediction is calculated (represented as yellow line). By calculating where the prediction line will cross the anomaly threshold, we estimate how far in the future that is from the present.

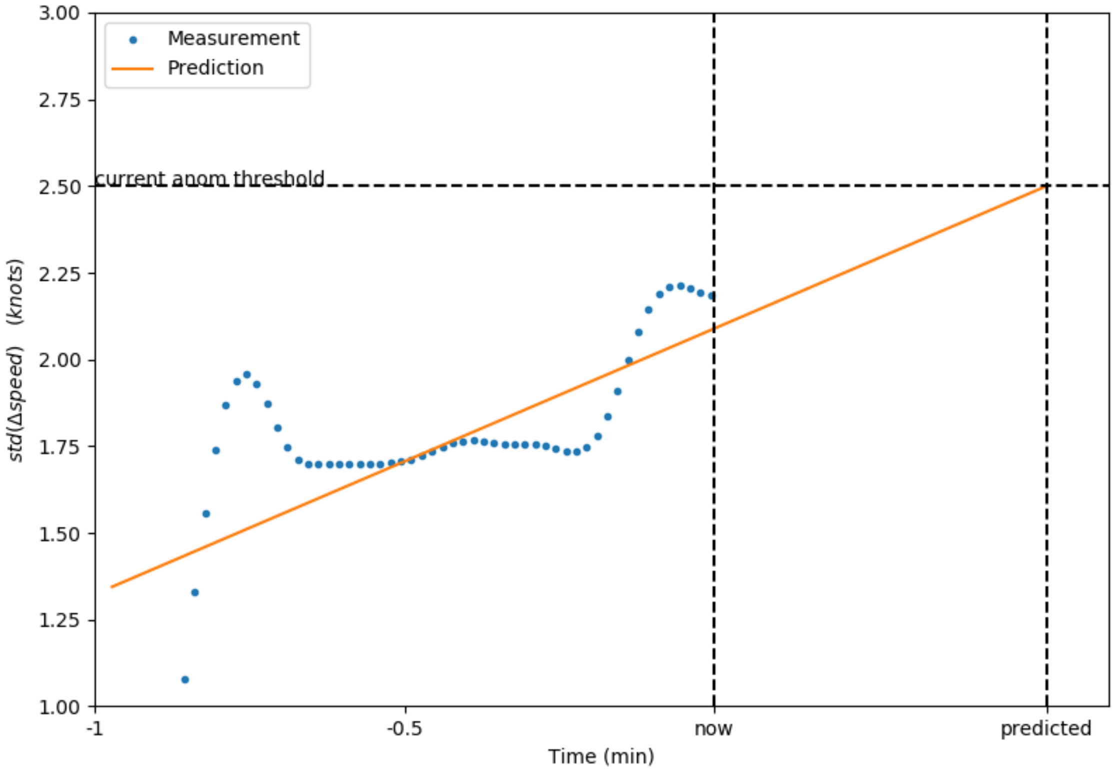


Figure 1. Example of anomaly trend detection of airspeed.

This calculation is performed repeatedly to update the prediction as changes occur in the observed values.

SW/SSOT

The Status Widget (SW) provides a timeline display to enable the crew to quickly assess the current and future state of monitoring requirements. The SW provides a visual “at a glance” indication conveying both urgency and criticality of monitoring tasks. The SW depicts both *expected* events, like upcoming checklists and the Top of Descent (T/D), as well as *unexpected* events such as current and predicted input data anomalies. The SW is interactive, allowing the crew to open the SSOT window for more detailed information about active events, then close the SSOT at their discretion. Events are graphically depicted along a fixed 15-minute timeline and move toward the "now" line as time passes. The timeline includes a "now" time gutter that catches active events to make them persistent to enhance crew awareness once they pass the "now" timeline, to prevent them from being missed. Likewise, there is "> 15 minute" time gutter to represent events beyond the fixed 15-minute time horizon.

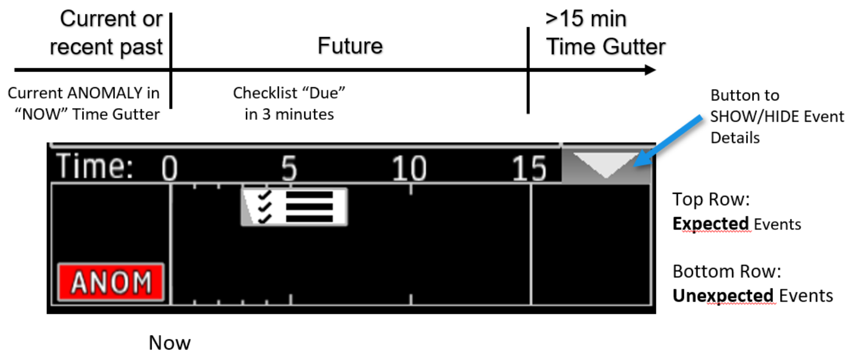


Figure 2. Annotated example of Status Widget with an expected event (Checklist) and unexpected event (anomaly).

The Support Suite of Tools (SSOT) presents relevant information to help the crew understand the active events and support their decision making. Information content for expected events varies by the type of event, but all include an estimate when the expected event is "due". For example, a pending Gate Report event could include the appropriate telephone number and a pending checklist could include checklist items. The color coding of the both the SW and SSOT should conform to flight deck conventions.

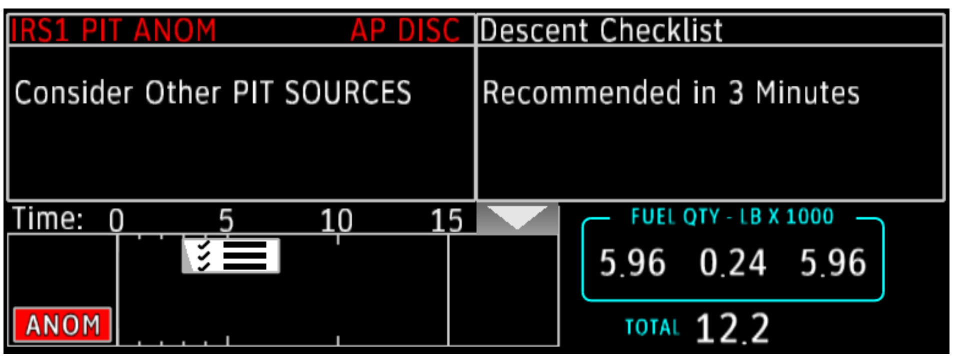


Figure 3. Example of Status Widget and Support Suite of Tools together. The SSOT provides additional information about the expected (descent checklist) and unexpected (pitch anomaly) events.

Design Philosophy

During the initial evaluation at Honeywell facility, pilot feedback indicated a preference for four types of information following an anomaly:

- What is wrong?
- What is impacted?
- Which source is good?
- What should the pilot do (guidance)?

Accordingly, we modified the design of the IDAD and SW/SSOT technologies to incorporate this information. Pilots also suggested that anomalous information should be removed from the display.

Based on these observations, the system was modified to perform the following actions following detection of a data anomaly:

1. Remove Anomalous Data from Impacted PFD(s) and Display Flag

The system removes anomalous data from impacted PFD(s) to improve crew diagnosis of ongoing anomalous autoflight behavior. Prior to being removed from the display, the anomalous data flashes for 10 seconds before a red flag appears near the anomalous data (e.g., SPD ANOM). The formatting and location of the red flag accommodates location of legacy PFD Failure flags.

The removal of anomalous content on the PFD worked as follows for the types of anomalies tested during flight simulations:

- Airspeed: airspeed tape elements removed: digits, increments, and indications
- Pitch: all Attitude Display components, horizon removed
- RA: radar altimeter indication removed

2. Display Explanation in Status Widget/Support Suite of Tools

Brief, explanatory information is displayed in the SW/SSOT to help the crew understand the impact of the data anomaly on autoflight systems (e.g., autopilot disconnect), identify which data source is anomalous, and if possible to predict when the anomaly could occur. This is detailed in the SW/SSOT section below.

3. Display EICAS Message

An EICAS message is displayed to harmonize with existing EICAS alerting philosophy.

Alerting philosophy

Trending toward an Anomaly

If the anomaly is initially detected as a trend, amber caution messages and the EICAS aural beeper tone (4 beeps) are used. Specifically, the SW depicts an amber icon on the timeline with a projection of when the system believes the anomaly will manifest, the SSOT indicates which source is anomalous and which autoflight systems may be impacted, and the corresponding amber EICAS caution message appears.

Current Anomaly

If the anomaly is detected at present, red warning messages and the Master Warning aural tone are used. Specifically, the SW depicts a red icon on the timeline at Time 0, indicating that the anomaly is happening at present, the SSOT indicates which source is anomalous, which autoflight systems are impacted, and provides guidance, and the corresponding red EICAS warning message appears.

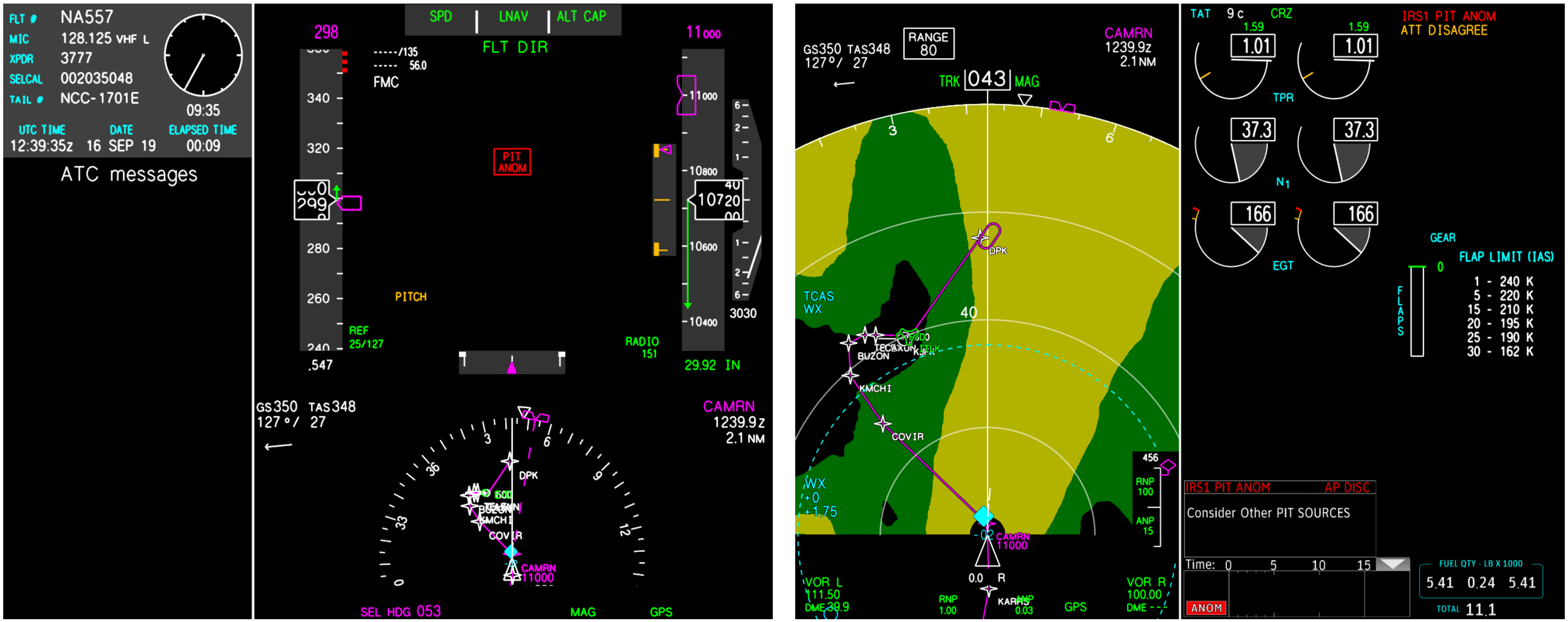
Example Images

Figures 4-7 depict examples of the PFD and MFD as implemented for the NASA flight simulator evaluation. In Figure 4 there are no IDAD/SW/SSOT indications; this is considered the 'baseline'. Figures 5-7 depict the PFD and MFD for airspeed, pitch, and radar altimeter anomalies, respectively. These images are for the affected side being fed by anomalous data, as the display elements are not changed for the unaffected side.

Figure 4. Baseline PFD and MFD for the left side displays.

IDAD detection of an airspeed anomaly triggers the following: 1) airspeed tape elements are removed from the PFD and an anomaly flag is added (“SPD ANOM”), 2) SW displays a red anomaly icon indicating that the issue is happening at present, 3) explanatory information is provided in the SSOT indicating that the ADRS1 source is anomalous, autopilot and autothrottle will disconnect, and other airspeed sources should be considered. EICAS indicates that an ADRS1 airspeed anomaly has occurred.

Figure 5. PFD and MFD during an airspeed anomaly.



IDAD detection of a pitch anomaly triggers the following: 1) Attitude Display elements and horizon are removed from the PFD and an anomaly flag is added ("PIT ANOM"), 2) SW displays a red anomaly icon indicating that the issue is happening at present, 3) explanatory information is provided in the SSOT indicating that the IRS1 source is anomalous, autopilot will disconnect, and other pitch sources should be considered. EICAS indicates that an IRS1 pitch anomaly has occurred.

Figure 6. PFD and MFD during a pitch anomaly.

IDAD detection of a radar altimeter anomaly triggers the following: 1) Radar altimeter is removed from the PFD an anomaly flag is added ("RA ANOM"), 2) SW displays a red anomaly icon indicating that the issue is happening at present, 3) explanatory information is provided in the SSOT indicating that the RA1 source is anomalous and that other RA sources should be considered. Additionally, the EICAS indicates that an RA1 anomaly has occurred.

Figure 7. PFD and MFD during a radar altimeter anomaly.

Information content for each anomaly type is summarized in Table 1.

Anomaly Notification Information Content:	Airspeed Anomaly	Pitch Anomaly	Radar Altimeter Anomaly
What is wrong?	ADRS1 SPD ANOM	IRS1 PIT ANOM	RA1 ANOM
What is impacted?	AP/AT DISC	AP DISC	--
Which source is good?	Other (Not ADRS1)	Other (Not IRS1)	Other (Not RA1)
What should pilot do (guidance)?	Consider Other SPD SOURCES	Consider Other PIT SOURCES	Consider Other RASOURCES

Table 1. Anomaly notification information content.

Transitioning from Predicted Anomaly to Current Anomaly

As discussed above, anomalous sensor behavior can be detected before crossing a threshold deemed to be anomalous. In such cases, the system can predict how much time it will take for the anomaly to reach the threshold. This duration is indicated on the SW timeline. In the figure below, a pitch anomaly is predicted to occur in approximately 30 seconds.

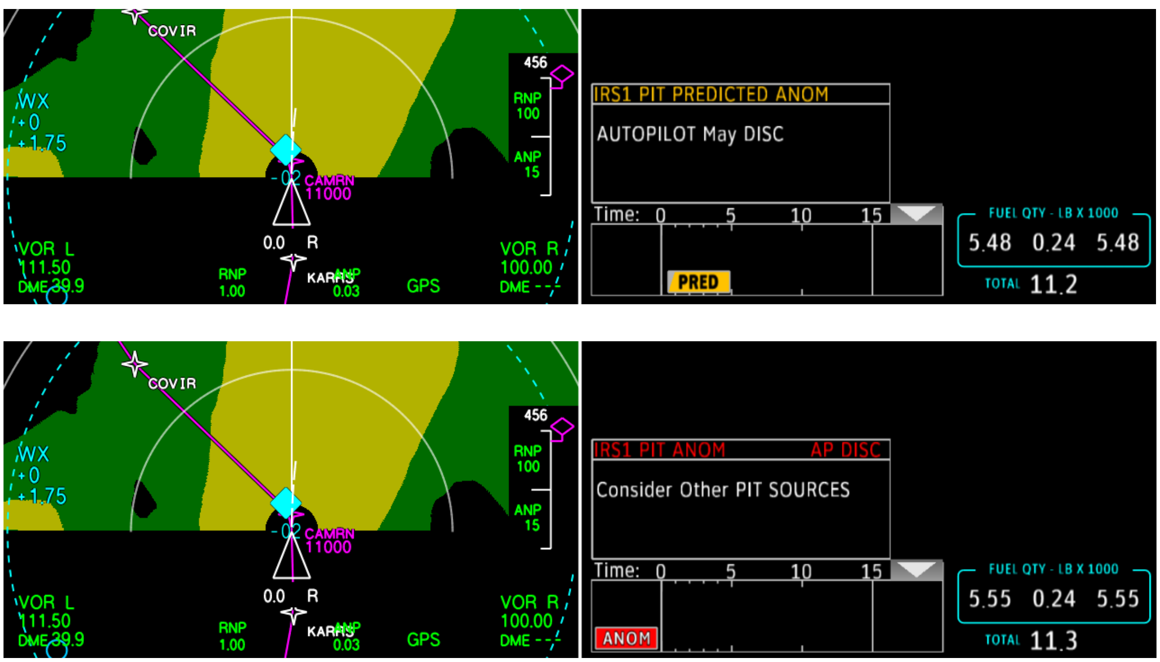


Figure 8. Predicted and Current Anomaly Indications. (Top) Predicted pitch anomaly indication in 30 seconds; (Bottom) Current anomaly indication at present.

If a predicted anomaly manifests and crosses the data threshold associated with an anomaly at present, the SW/SSOT and EICAS alerts change from amber to red and the master warning aural sounds.

Changes from Honeywell Evaluation to NASA LaRC Evaluation

The following images highlight the design changes from the 2018 Honeywell Evaluation to the 2019 NASA evaluation. After anomalous data is removed, the PFD no longer uses an amber box to indicate what has happened, which source is impacted, and provide guidance. Instead, the PFD shows a simple red flag near the anomalous information. The explanatory information is now provided in the SW/SSOT on the MFD. Airspeed, Pitch, and Radar Altimeter (RA) anomalies are depicted below.

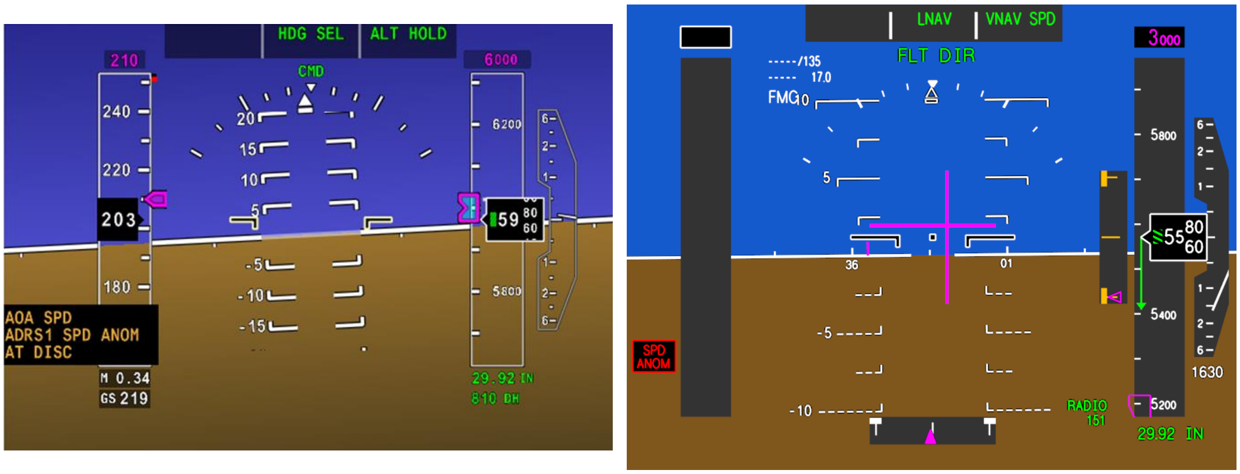


Figure 9. Airspeed Anomaly. (Left) Airspeed anomaly indication for Honeywell evaluation; (Right) Airspeed anomaly indication for NASA LaRC evaluation.

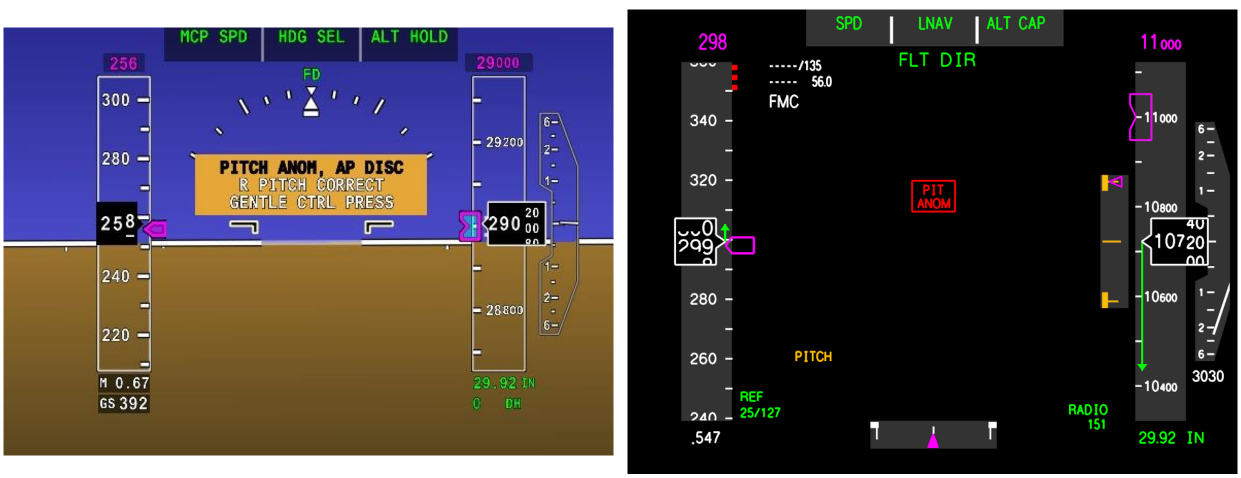


Figure 10. Pitch Anomaly. (Left) Pitch anomaly indication for Honeywell evaluation; (Right) Pitch anomaly indication for NASA LaRC evaluation.

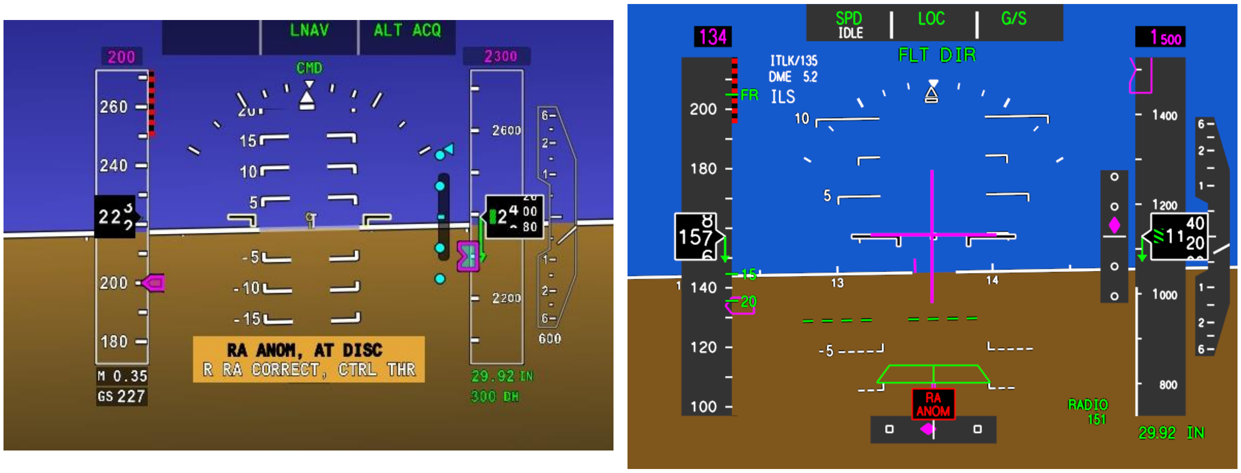


Figure 11. RA Anomaly. (Left) RA anomaly indication for Honeywell evaluation; (Right) RA anomaly indication for NASA LaRC evaluation.

Comparison images are not provided for the SW/SSOT on the EICAS/MFD. The changes were mostly formatting related: larger font in SW, change “TRND” icon to “PRED” icon, remove graphical depictions of data anomaly from SSOT.

Methods

At this stage of technology maturity, our objective is to elicit pilot feedback on the technology to both validate the operational value and improve the design prior to further development. Accordingly, we conducted a formative evaluation consisting of short scenarios to demonstrate the technologies in simulated flight scenarios within a motion-based simulator that replicates the flight deck of a large civil transport aircraft. During this evaluation, we were primarily concerned with pilots' subjective impressions of the technologies and objective metrics related to their use of the technologies during flight operations.

Scenarios

To ensure that the crews understood the technologies of interest, scenarios were blocked together into Training, Baseline and Experimental blocks. The crews first received two training flights to familiarize themselves with the simulator, followed by three baseline scenarios in which an anomaly occurred but no technology interventions were provided, and then six experimental scenarios with the experimental technology. Within each block, the scenario presentation order was randomized. After the completion of all scenarios, a demonstration scenario illustrated one proposed behavior of returning display elements after an anomaly clears. Accordingly, all participants experienced the scenarios in the following table. Each scenario is mapped to a relevant aircraft incident/accident.

Run Order	Block	Trial Type (Training, Baseline, Experimental)	PF	Anomaly Type	Anomaly Detection	Anomaly Trigger	Accident/Incident
1	1	Training - baseline (no experimental tech)	Captain	Speed	--	Pitot block at 17000 ft	Icelandair 662 B757-200
2	1	Training - experimental (with experimental tech)	Captain	Speed	Current	Pitot block at 16000 ft	Icelandair 662 B757-200
3	2	Baseline	Captain	Pitch	--	Negative pitch rate error at 7 min	West Atlantic Sweden Cargo CRJ200
4	2	Baseline	Captain	Rad Alt	--	RA Fail at 2050 ft MSL	Turkish 1951 737-800
5	2	Baseline	Captain	Speed	--	Pitot-static block at 14000 ft	Icelandair 662 B757-200
6	3	Experimental	Captain	Rad Alt	Predicted then Current	Frozen RA at 1900 ft MSL	--
7	3	Experimental	Captain	Speed	Predicted then Current	Partial pitot block at 8 min	Midwest Express 717
8	3	Experimental	First Officer	Rad Alt	Current	RA error (-8 ft) at 1400 ft MSL	Turkish 1951 737-800
9	3	Experimental	Captain	Pitch	Predicted then Current	Pitch rate error at 9 min	West Atlantic Sweden Cargo CRJ200
10	3	Experimental	First Officer	Speed	Current	Partial pitot block at 6 min	Midwest Express 717
11	3	Experimental	First Officer	Pitch	Current	Pitch rate error at 7 min	West Atlantic Sweden Cargo CRJ200
12	4	Demonstration	Captain	Speed	Predicted then Current	Partial pitot block at 2 min, block persists for 4 min, then block clears	Air France 447

Table 2. Evaluation Scenario Details

To provide opportunities for crew familiarization with the flight deck, the first training trial consisted of runway takeoff from KDCA destined for KJFK under VMC conditions with climb, cruise, and descent phases before the anomaly occurred in the descent phase. All other scenarios initiated on the CAMRN FOUR STAR destined for KJFK under CAT II conditions, either at FL190 at SIE (approximately 5 minutes from top of descent) or at FL110 at CAMRN. With thunderstorms near the airport, these scenarios were completed with light to moderate turbulence. Weather and traffic were added to increase complexity which in turn increases workload; this increases the likelihood that participants experience the technology within a task context that approximates realistic workload.

Procedure

The following experimental procedure was approved by an external institutional review board (IRB) (Arclight, Inc IRB000087) on August 6, 2019 with approval number HON-2019-004 as well as being separately approved by NASA LaRC IRB (FWA00020089).

For all crews but one (Crew 4), two pilots participated together as a crew. The First Officer (FO) for Crew 4 could not attend due to illness so a qualified research pilot participated as a confederate FO. Before starting the experiment, participants provided informed consent, were briefed on the aims of the study and the configuration of the flight deck and agreed upon role of Captain and First Officer (FO). Captains sat in the left seat and First Officers sat in the right seat. The role of pilot flying (PF) and PM was prescribed per the scenario details provided above. A flight plan was pre-loaded for all scenarios and pilots were asked to manage altitude and speeds per the Flight Management Computer (FMC) guidance. PM was asked to perform routine monitoring tasks, respond to simulated ATC communications, and respond to all PF commands in accordance with standard operating procedures (SOP).

Pilots were instructed to verbalize their awareness and decision-making process in response to anomalous situations. Each scenario lasted approximately 10-15 minutes. Following the introduction of the anomaly, scenarios ended after the crew discussed the anomaly and decided upon a course of action. After all scenarios, experimenters would encourage pilots to comment on the technology design. The entire protocol lasted approximately 8-9 hours including lunch and breaks.

At least one secondary task was used during each scenario to keep the pilots focused on flying the aircraft rather than fixating on anomaly detection. Secondary tasks consisted of configuring the aircraft for landing, ANTI-SKID EICAS message, TCAS traffic alert, or ATC datalink messages to slow to a certain speed when passing through a specified altitude or being instructed to monitor a certain frequency. Also, as mentioned previously, weather conditions were poor (Cat II) with cells shown on the weather radar display and moderate turbulence. These conditions distracted participants such that they would not be able to fixate on sensor anomalies in a manner not representative of flight operations.

Design

A within-subjects design was employed such that all participants saw the same scenarios. After each trial, experimenters administered the Bedford and NASA TLX workload scales on a tablet computer. Following the end of the Baseline and Experimental blocks, experimenters administered the following on a tablet computer: Bedford and NASA TLX workload scales, Situation Awareness Rating Technique (SART), Acceptability Index, Complexity Index, System Usability Scale, and general comments. After completing all scenarios, experimenters followed-up with a post-evaluation questionnaire.

Experimenters captured pilot comments during and after the trials in written notes and later transcribed and categorized them for analysis.

Participants

Air Transport pilots were recruited to participate in the PITL evaluation. In total, 6 crews participated and completed all scenarios. The 11 pilots (10 men and 1 woman, all Part 121) represented 3 major U.S. passenger airlines and 1 cargo airline. Five of the six crews consisted of a 2-person crew, but one of the six crews utilized an experimenter as confederate FO due to illness. Pilot demographics are described in the table below:

<table>
<tr><th>Age</th><th>Total Flight Hours</th><th colspan="2">Common Aircraft Type Flown</th></tr>
<tr><td>Mean = 56.1 years</td><td>Mean = 13,324 hours</td><td rowspan="2">• B737
• B747
• B757
• B767</td><td rowspan="2">• B777
• A319
• A320
• MD-80</td></tr>
<tr><td>Range = 47-65 years</td><td>Range = 4,729-27,000 hours</td></tr>
</table>

Table 3. Pilot Demographics.

Flight Simulator Description

The pilot evaluation for this phase was conducted in the NASA Langley Cockpit Motion Facility's Research Flight Deck (NASA). The motion-based simulator is a hybrid design, mimicking aspects of a Boeing 757/767 – using the aerodynamic model and overhead panel, a Boeing 787 – incorporating four 17" LCDs and dual EFBs/HUDs, and an Airbus – in that it uses sidesticks. During this evaluation, experimenters manipulated the information sources and provided the new technology indicators and alerts both auditorily and visually (on the PFD and MFD/EICAS) as described previously. The HUDs were stowed and not used.

Figure 12. NASA LaRC Research Flight Deck.

As a reminder, erroneous display elements were removed the PFD on the affected side, and the SW/SSOT was presented at the bottom of the EICAS on the affected side.

Results

Subjective Workload

Both the Bedford Workload Scale (Roscoe, 1984) and the NASA Task Load Index (TLX) (Hart & Staveland, 1988) were administered. Scenarios were designed to ensure that pilots had the time and cognitive resources to consider and evaluate the technologies. Reported workload indicated that subjective workload was relatively low across all conditions, as indicated by the figures below. The experimental technology intervention did not contribute to an increase in subjectively rated workload as indicated by both Bedford ($mean_{Baseline}$ = 3.36, $mean_{Exp}$ = 3.15) and NASA-TLX responses. All graphed means include standard error bars.

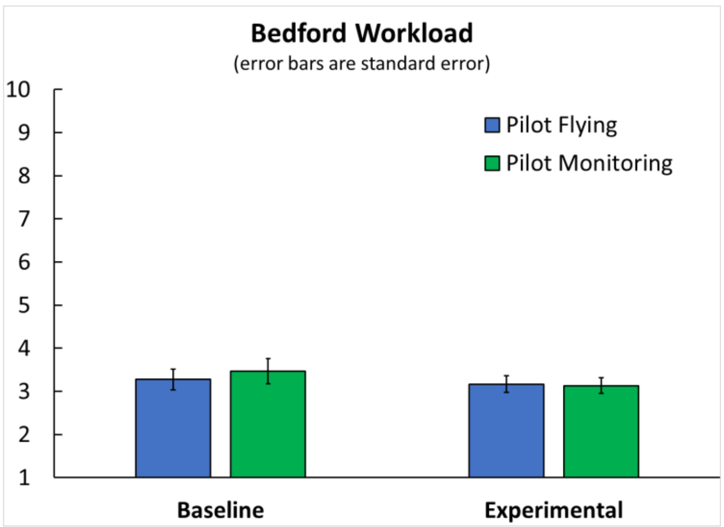


Figure 13. Bedford Workload responses averaged across participants.

With only six crews, statistical tests were not utilized to compare NASA-TLX responses, however, there appears to be a trend for lower workload responses when completing the trials with the experimental technology. This is more pronounced for the pilot monitoring (see right pane of Fig. 14).

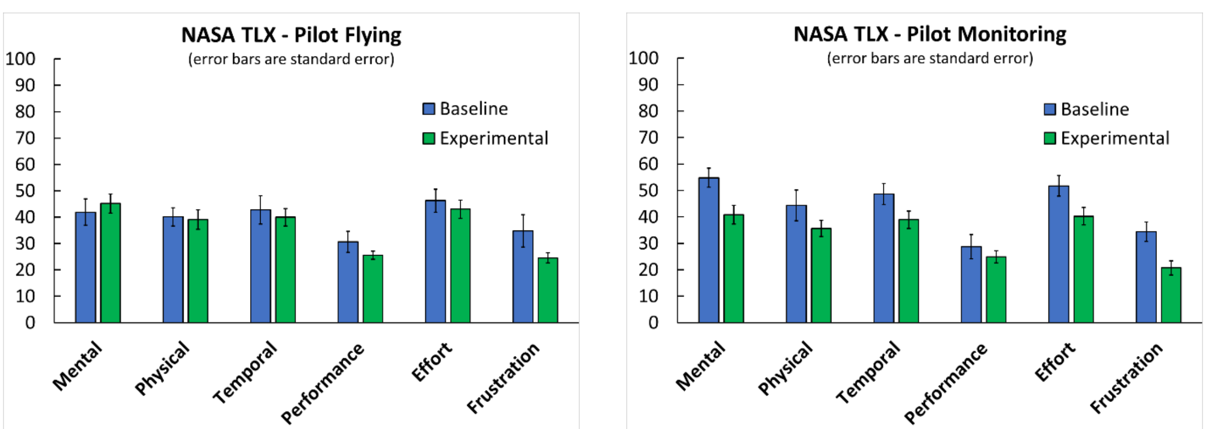


Figure 14. NASA TLX responses for Pilot Flying (Left pane) and Pilot Monitoring (Right pane) for the 6 subscales.

Acceptability

The following custom scale was used to assess perceived acceptability, as used by NASA teams on AIME experiments (Evans et al., 2016).

Rate the acceptability of the experimental technology (from 1 to 7):

1 = Very unacceptable. I did not like the technology and would not use it in normal operations.

4 = Average. I liked the technology and would use it in normal operations, but would like to see some improvements.

7 = Very acceptable. I like the technology very much, and would use it without any improvements.

The experimental technology intervention received high ratings of acceptability (mean = 5.82). This aligns with the narrative feedback (discussed below), indicating generally positive feedback for the anomaly detection concept, with specific topics for improvement.

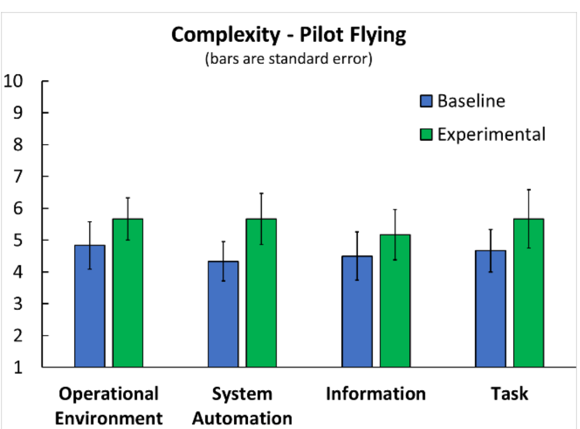
Complexity

The following custom scale was used to assess perceived complexity, as used by NASA teams on AIME experiments (Evans et al., 2016).

Complexity, 1-10, 1 = Not complex, 10 = Extremely complex

- Rate the complexity of the operational environment
- Rate the complexity of the system/automation
- Rate the complexity of the information
- Rate the complexity of the task

Complexity ratings were separated between Pilot Flying and Pilot Monitoring. As seen in Fig. 15, the Pilot Flying reported slightly higher complexity ratings for the scenarios completed with the experimental technology, while the Pilot Monitoring reported slightly lower complexity ratings for scenarios completed with the experimental technology.



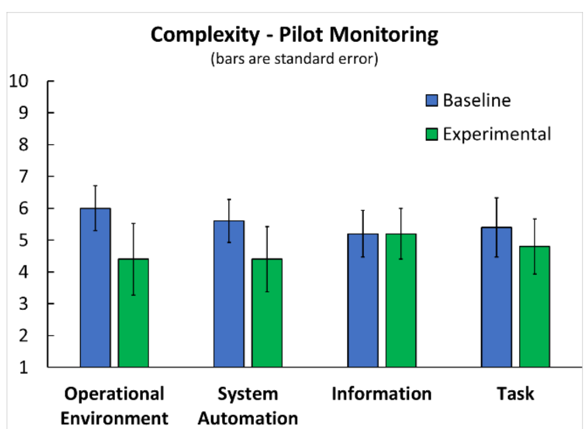


Figure 15. Complexity responses for Pilot Flying (Left pane) and Pilot Monitoring (Right pane).

The lower complexity ratings for the Pilot Monitoring may loosely explain the NASA-TLX ratings above, where the Pilot Monitoring reported slightly lower workload ratings across the subscales. This comports with the design philosophy of the SW/SSOT which is to support the monitoring function. An interpretation of this trend could be that the technology enabled monitoring such that the PM experienced the task and task environment as less complex.

Usability

After the block of baseline trials and experimental trials, the System Usability Scale (SUS) was administered (Brooke, 2013).

Technology Usability, SUS method, Rate 10 statements
With respect to the experimental technology, radio buttons 1 - 5 (1 = Strongly Disagree, 5 = Strongly Agree):
1. I think that I would like to use this system frequently.
2. I found the system unnecessarily complex.
3. I thought the system was easy to use.
4. I think that I would need the support of a technical person to be able to use this system.
5. I found the various functions in this system were well integrated.
6. I thought there was too much inconsistency in this system.
7. I would imagine that most people would learn to use this system very quickly.
8. I found the system cumbersome to use.
9. I felt very confident using the system.
10. I needed to learn a lot of things before I could get going with this system.

SUS items alternate between items where high scores are desirable ("Would Use Frequently") and items where low scores are desirable ("Unnecessarily Complex"). As is illustrated in Fig. 16 below, the experimental intervention received favorable usability ratings. Participant ratings suggest they would be confident using the technology, and would use it frequently; moreover, the technologies were not perceived as complex, cumbersome, nor requiring a particularly steep learning curve. SUS scores indicate that integration with the flight deck could be improved, which will be addressed in the Discussion section.

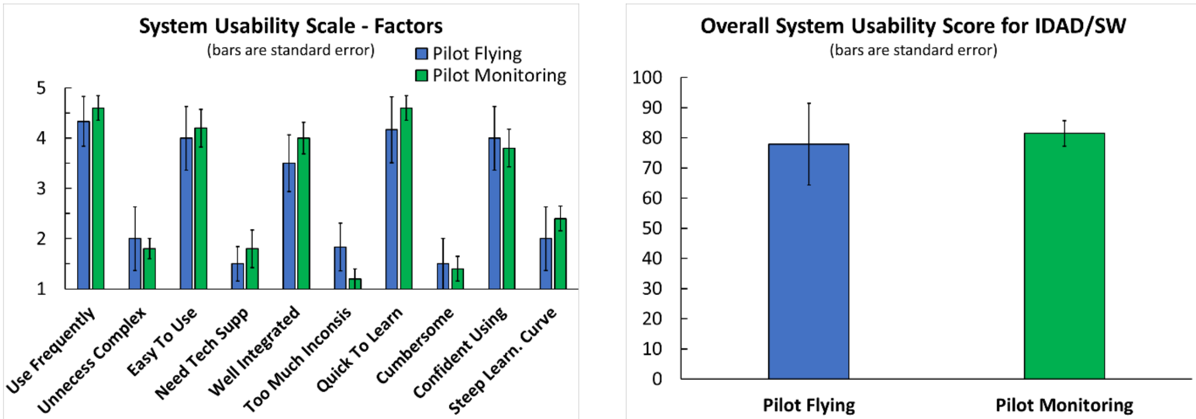


Figure 16. System Usability Scale responses. (Left pane) SUS Factors averaged across participants; and (Right pane) Overall SUS scores for Pilot Flying and Pilot Monitoring.

Situation Awareness Rating Technique (SART)

After the block of baseline experimental trials, the Situation Awareness Rating Technique (SART) was administered (Taylor, 1990). For the most part, there is not an appreciable difference between SART ratings provided on baseline trials and experimental trial with the technology.

Although there are no statistical conclusions to draw, it is worth noting that for Pilot Flying and Pilot Monitoring, there is a small trend for lower ratings of Attention Supply during Experimental Trials, which may contribute to the slightly higher rating of Understand Information and Situation during Experimental Trials.

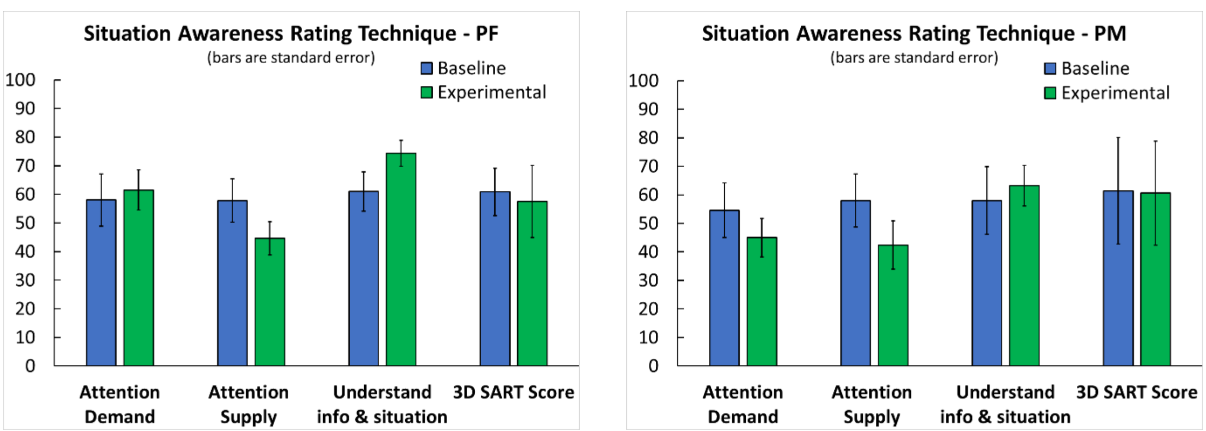


Figure 17. Situation Awareness Rating Technique responses for Pilot Flying (Left pane) and Pilot Monitoring (Right pane).

Behavioral Responses

Time-stamped video recordings were used to quantify: 1) the time required for the crew to detect that an anomalous situation had occurred, and 2) the time required to decide on a course of action (COA) after detecting the anomaly.

Pilot recognition of anomalous data before any aural indication

Both baseline and experimental trials included anomalies that triggered multiple aural alerts. For example, both triggered legacy aural alerts for "ATT Disagree" and "IAS Disagree." Given the commonality of aural indications, we looked at all trials to assess how frequently crews recognized an anomalous situation prior to the aural alert. Of the 54 trials (6 crews × 9 trials), crews recognized anomalies prior to any aural alerts on three trials, or 5.5% of the trials. This suggests that detecting anomalies in modern aircraft is based only on visual indications is difficult, corroborating results of accident and incident analyses. There are likely many reasons for this, chief among them is that pilots are trained to trust displayed information and high demands on visual attention; pilots can miss or overlook important cues, even if attentive and vigilant.

Predicted and Current anomaly timing

Before considering timing results, it should be noted that Predicted anomalies effectively start earlier than Current anomalies, as subtle perturbations in the affected data source escalate to a higher magnitude and transition to a Current Anomaly. Fig. 18 illustrates how the timing works for a Predicted and Current anomaly.

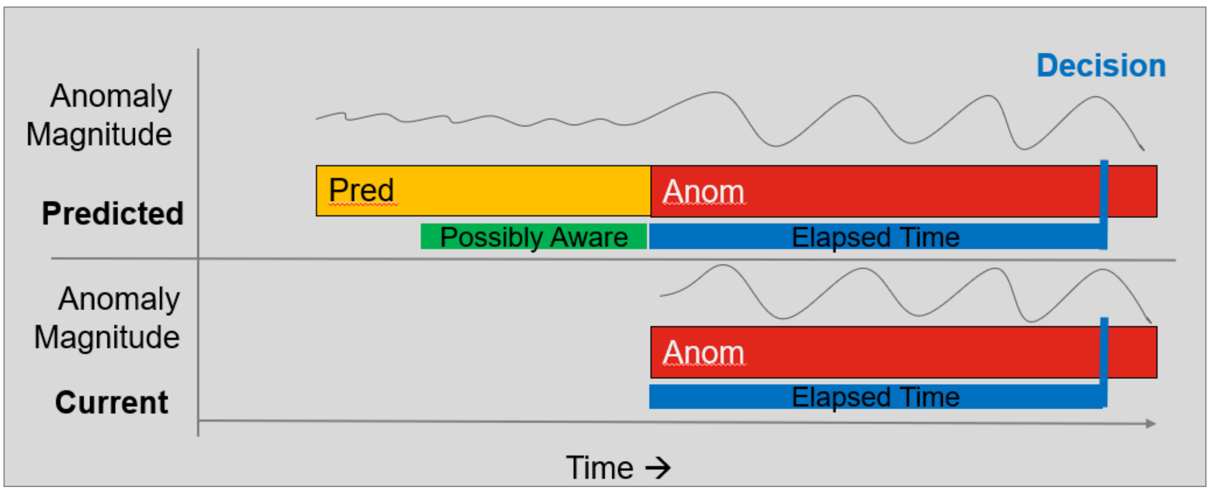


Figure 18. Explanation of anomaly timing for predicted and current anomalies.

Given the difficulty in determining crew awareness of predicted anomalies, we chose to calculate the "start" time for crew awareness for Predicted trials from the onset of the current anomaly, not the onset of the initial perturbations. This aligned Predicted trials with the timing of Current Anomalies in both Baseline and Experimental trials, enabling comparisons. Crew debriefing comments indicated that the predicted alerts started orienting them to the situation and priming them for a decision.

Time to recognize anomalous elements

Experimenters reviewed the audio-video logs for all trials to identify when, if ever, crews recognized that some displayed element was anomalous. Elapsed time to recognize any anomalous elements was calculated from the onset of the anomaly until the crew verbalized recognition of a specific anomaly. There were a few trails where experimenters could not definitively identify a moment of recognition.

Radar altimeter (RA) trials were unique in that anomalies occurred late in the approach (1400-2100 ft AGL), so crews were under time pressure and, in all but one of the RA trials, nearly immediately initiated a Go-Around. Crews reported that they responded to the aural alerts and presence of any Caution or Warning EICAS, by deciding to Go-Around, without recognizing the issue was with the PF's RA indication.

As noted above, Baseline and Experimental Current trials were similar in the rapid onset of the anomaly, enabling a direct comparison. Predicted anomalies, however, were slowly developing situations that started with very subtle perturbations. Consequently, "start" time for Baseline, Experimental Current, and Experimental Predicted trials was defined from the onset of the Current anomaly for timing calculation.

Experimenters were unable to discern a definitive moment of recognition for 1 Baseline and 1 Experimental Current Airspeed Trial, thus there are only 5 samples for each instead of 6. As depicted in Fig. 19, although there are limited samples, there is certainly a trend with non-overlapping standard error bars for crews to be faster to recognize the anomalies in Experimental trials (both Current and Predicted) than Baseline trials. This suggests that the EICAS notification and PFD display mitigation, 10-seconds of flashing before the anomalous elements was removed and an anomaly flag was presented, was effective in orienting crew attention to the anomalous indication.

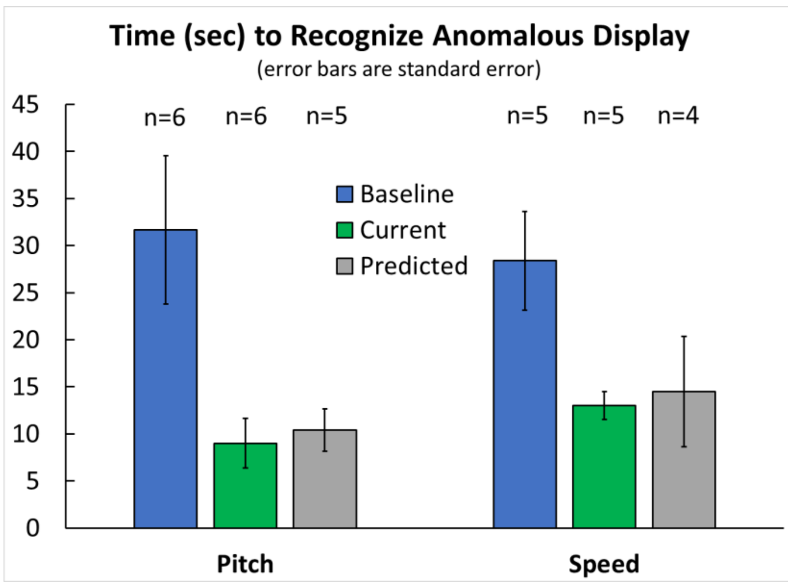


Figure 19. Average time required to recognize anomalous display elements across Baseline and Experimental Trials (Current and Predicted).

Decision time after anomaly onset

For all trials, experimenters reviewed the audio-video logs to identify when crews decided on a final course of action in response to an anomalous situation. A common decision was to discontinue approach into JFK which has Cat II conditions and go to an airport with better weather. One of the impacts of the anomalies was that the flight was no longer LAND-3 capable, meaning it was not legal to continue the published Cat-2 approach procedure. Elapsed decision time was calculated from the onset of Current anomalies, defined a priori in experimental scripts as time or altitude triggers, to the identified decision point. Experimenters could confidently identify the decision point for all trials, so results depicted in Fig. 20 are for all 6 crews.

There is a trend, with non-overlapping standard error bars, indicating that decisions were made faster during Experimental trials (Current and Predicted) than Baseline trials. This suggests that the Experimental Display interventions can streamline the process toward deciding on a course of action.

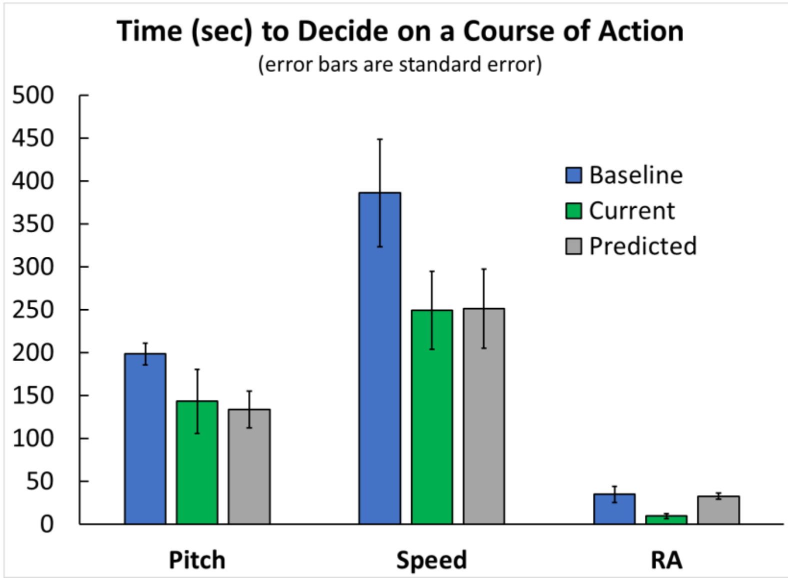


Figure 20. Average time for crew to decide on a course of action for each type of anomaly and condition.

Pilots commented that removing the anomalous indication simplified their determination of which side of the aircraft had correct information. Several pilots also commented that they would be distracted by a displayed anomalous indication, continuously referencing it to assess whether it was valid again. While the results are preliminary, the magnitude of the time difference for the Airspeed anomaly suggests crews could decide over 2 minutes faster with anomaly alerting, possibly supporting more operational flexibility and less disruption.

As mentioned previously, crews responded very quickly to the aural and visual indication associated with the RA anomalies. With such fast responses, a floor effect likely precludes discerning and interpreting any meaningful differences.

Course of action analysis

For all trials, experimenters reviewed the audio-video logs to identify actions the crews performed in response to an anomaly. After reviewing all the logs, experimenters identified the following categories of actions:

- Executed a Go-Around (Go-Around)
- Ask for ATC Accommodation (Declare Emergency, Ask for Vectors and/or Block Altitude)
- Divert for Better Weather
- Continue Approach
- Transfer Control to PM
- Transfer Control back to PF
- Check Minimum Equipment List (MEL) for Continuing Approach
- Experimenter Prompt (Experimenter prompted the crew: remind Cat II weather at destination, offer help with bringing up standby display and/or switching to ALTN)

After categorization, experimenters then revisited the annotated log notes and identified which actions were performed during each trial. Figure 21 depicts the frequency of these categories of actions broken out by Anomaly Type (Current Pitch, Current RA, etc.) and Trial Type (Baseline, Experimental).

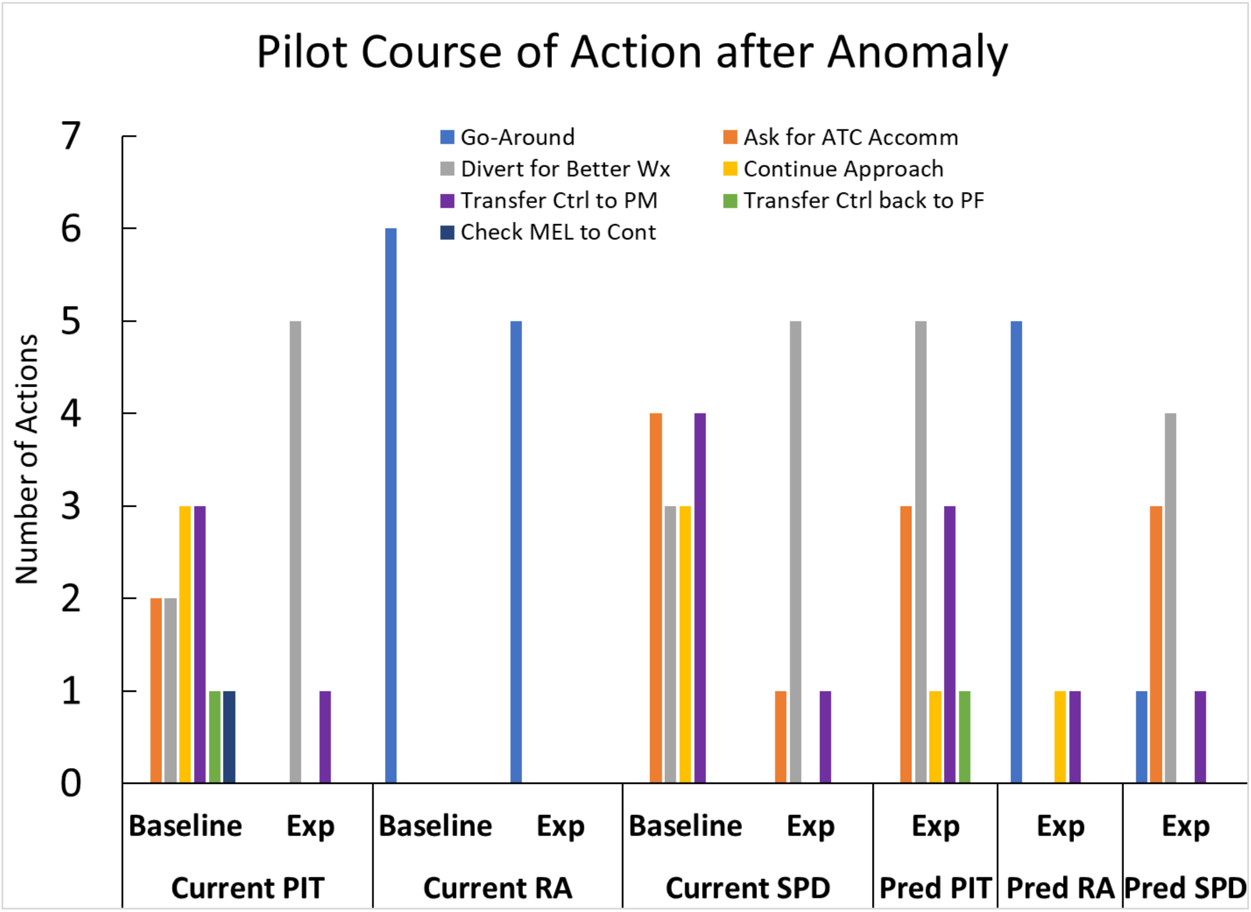


Figure 21. Actions Taken by Anomaly/Trial Type. Experimenter prompts are excluded from the figure.

The figure reinforces the other observations about RA trials - that crews mostly (16 of 18 trials) performed an immediate Go-Around, regardless of Trial Type. Execution of a Go-Around precluded the performance of any other categorized action, save one crew who Transferred Control to PM and Continued the Approach during the Predicted RA Trial.

When examining Current Airspeed and Pitch trials across Trial Type (Baseline, Experimental) as depicted in Table 4, a pattern emerges of action profiles excluding Experimenter Prompts: crews performed more actions across more action categories during Baseline trials. Although from a small sample, it is encouraging that the Experimental Mitigation seems to support more consistent and fewer action steps. One interpretation is that the additional information about the anomalies focuses decision making which results in resolving the situation with fewer and more consistent actions.

		Current Pitch	Current Airspeed	Current RA
Baseline	Actions	12	14	6
	Types	6	4	1
Experimental	Actions	6	7	5
	Types	2	3	1

Table 4. Actions Performed by Trial/Anomaly Type. Experimenter prompts are excluded from the table.

Demonstration Scenario

One demonstration scenario was administered after the experimental trials, without the use of turbulence, to solicit feedback about returning anomalous data display elements after the underlying data source anomaly cleared. This is a documented phenomenon, as several accidents and incidents involved an intermittent blockage that cleared (e.g., Air France 447 blockage cleared as they lost altitude). In terms of the IDAD detection algorithm, it is possible to detect the return of anomalous data to the expected range.

In the demonstration scenario, a partial pitot blockage initially caused an airspeed anomaly. After four minutes, the partial pitot blockage clears. The IDAD algorithm detected the return of valid airspeed and triggered the return of airspeed display elements to the affected PFD. As with the removal of PFD elements, the airspeed blinks while returning to the PFD.

Across the six crews, pilots generally supported returning the display elements, provided a high level of confidence in its accuracy, but had differing opinions on what determines whether the data returns (e.g., return when data is valid, return when data is valid only after a dampening period, have an anomaly on new data source for original data source to return). However, crews did not want the data source to be removed and returned repeatedly, as this would become a nuisance and with each onset would likely drive the crew to repeat the associated checklist again. In terms of how the data should return when the anomaly clears, three of the crews indicated a preference for a notification that the anomaly cleared to allow pilots to initiate the return of display elements and switching data sources.

Discussion

The objective of this evaluation was to elicit pilot feedback on the combined IDAD + SW/SSOT technology to both validate the operational value, understand its performance impact, and improve the design prior to further development. Along with NASA LaRC colleagues, we conducted a formative evaluation consisting of short scenarios to demonstrate the technologies in simulated air transport flight scenarios within a motion-based simulator.

Overall Findings

Across the subjective ratings, pilots rated the combined IDAD + SW/SSOT technology intervention favorably and there was clearly consensus on the operational value of the proposed technology.

In terms of workload, Bedford and NASA-TLX ratings were no higher for the experimental technology. Interestingly, there is a trend for the PM to experience lower workload for Mental Demand, Temporal Demand, and Effort during experimental trials. In terms of complexity, PF reported slightly higher complexity ratings for the scenarios completed with the experimental technology, while PM reported slightly lower complexity ratings for the experimental technology. In terms of usability, the experimental technology received favorable usability ratings. Participants indicated they would be confident using the technology. Moreover, the technology was not perceived as complex, cumbersome, nor requiring a particularly steep learning curve. Usability scores do indicate that integration with the flight deck could be improved.

In terms of behavioral observations, crews were both faster to recognize anomalous events in Experimental trials than Baseline trials and faster to decide on a course of action in Experimental trials than Baseline trials. This was true both for Current and Predicted Experimental trials. While there was no difference between Current and Predicted trials, Crew debriefing comments indicated that the predicted alerts started orienting them to the situation and priming them for a decision.

Pilots also provided many suggestions to improve the design and expand the functionality to further improve the technologies. Most importantly, they were supportive of integrating anomaly alerts with the existing crew alerting philosophy and removing erroneous display elements from the PFD.

The following observations are noted based on the subjective data, objective data, video recordings, and experimenter notes. The training and certification implications are also discussed for each section.

Predicted Anomalies

The concept of predicting the onset of an alert was well supported by the crews. During the post-evaluation survey period, crews noted specifically that it addresses startle/surprise by providing an early indication that an upcoming situation could impact autoflight. As expected, the crews noted that judgments regarding data validity must have a high level of confidence.

A predicted onset time interval was provided to crews by way of the predicted anomaly icon on the SW. None of the crews utilized the precise indication of time to onset, instead noting that they were aware that they had some amount of time to discuss and prepare prior to anomaly onset.

Consequently, we believe the implementation could be changed so that no SW/SSOT is required, and predicted anomalies are indicated in the EICAS when they are below some threshold (e.g., 90 secs from present) and on the PFD with a predicted anomaly flag adjacent to impacted display element. A predicted EICAS energy alert was implemented and evaluated on previous NASA evaluations (Duan et al., 2016). This would provide consistency and predictability in the case of receiving a predicted anomaly and removes the burden from the prediction algorithm to be completely precise in terms of the detection interval; furthermore, by picking a relatively short interval threshold, such as 90 seconds, the time estimate is more likely to be accurate than a longer interval.

Additionally, a minimum prediction window could also be established, in terms of proximity to onset at current. If a predicted anomaly is detected at 3 seconds, there is questionable value in providing a predicted anomaly followed by a current anomaly in rapid succession. Such a situation could confuse the crew, therefore we believe a minimum threshold should be established as well (e.g., 5 seconds).

It should be noted for future investigations that there is a time/accuracy tradeoff for anomaly detection in that the longer a relationship is evaluated, the higher the confidence in the rating. However, this comes at the cost of timeliness. If the evaluation period lasts too long, the utility of providing an alert to the crew could be lost.

Certification implications – high; predictive alerts represent a new type of alert and additional work is needed to specify precisely how pilots should handle them in the context of other operations. The logic and real-time operation of the algorithms being used to detect these new anomalies would need to be detailed, but it does not necessarily require any advanced analytics like machine learning; moreover, it is more transparent than a black box approach in that it explainable in terms of common statistical methods and known time and magnitude thresholds.

Training implications – high; predictive alerting would need to be trained and integrated into the pilot workflow. While the notion of predictive alerting seemed intuitive when briefing pilots, they would need to be trained on how to react to and handle predictive alerts.

Procedural implications – high; care should be taken to define how crews respond to predicted alerts. Currently, most alerts reflect current situations so there is less ambiguity in how to respond than an alert for something that could happen in the near future. Procedures should support a consistent response to predicted alerts that is aligned with their intended function and operational utility. For example, they could be treated as an "awareness" item where pilots should communicate and brief how they would response, but not take any actions until it escalates to a current anomaly alert.

EICAS/ECAM/Crew Alerting Integration

The familiarity of the crew alerting system of each target platform should be leveraged. This dictates that each anomaly will have its own checklist or be integrated into existing checklists. Such an approach utilizes a familiar format and workflow when the crew is handling abnormal situations, leverages pilot training and expectations, and reduces certification risk while not compromising utility of the new anomaly detection and alerting.

Certification implications – low to medium; a significant effort would have to be undertaken to harmonize the new anomaly alerts with existing checklist procedures and verify their effectiveness; however, one mitigating factor is that the checklist actions would be similar to legacy checklists (e.g. IAS Disagree).

Training implications – low; any new anomaly alert would have to be evaluated whether an associated checklist procedure was required. Regardless of whether it required a checklist or not, crews would need additional training on how to respond to anomaly alerts.

PFD Element Removal

All crews were receptive to the notion of removing erroneous information from the PFD. While we did not set out to determine if this drove behaviors, sped up detection time, or improved decision making, we suspect that it did not compromise decision making as crews would reference the absence of information during problem solving. Furthermore, the crews indicated in the post-evaluation survey that removing erroneous PFD elements simplified decision making and discouraged their use of erroneous data.

Certification implications – medium; by aligning PFD element removal with that described in the Boeing 787 FCOM, we hope to minimize the certification impact of such a proposition. However, we have also proposed additional flags, that would be needed to be integrated as well. Additionally, we proposed a period of blinking prior to removal.

Training implications – medium; again, by aligning with other PFD element removal being replaced by flags, we hope to minimize the training requirements, but pilots would need to be made aware of and train with this change.

Status Widget

While there was some positive feedback regarding the SW/SSOT, the majority of the crew alerting can be achieved by integrating the SW/SSOT information in the impacted PFD and the existing crew alerting systems for each platform. This conclusion is framed by the current reality of the PM role. If the PM role were to be reconsidered in a future cockpit, with redesigned displays, the SW/SSOT should be considered as a novel display element likely incurring increased certification risk. However, if the SW/SSOT were to be considered as a retrofit option, we believe it should be placed on the PM side instead of the PF.

In terms of specific findings in the present evaluation, the font size was possible too small and red text needed additional contrast. This, combined with the SW/SSOT being located outside of their normal scan, made it more difficult to quickly reference. Additionally, the predicted time estimate was not utilized, as crews only noted that they had some amount of time before the predicted anomaly occurred. The expected items (e.g., Descent checklist Due in 4 mins) received variable feedback. For the RA scenarios, anything beyond the aural alert, including SW/SSOT information is questionable as the crews did not have time to view it and opted to initiate a go around. Alert suppression under certain conditions (e.g., below certain altitudes) should be investigated.

Certification implications – high; the SW/SSOT represents a new display concept leveraging time and as of the evaluation was placed on a flight critical display. Accordingly, we expect the certification of this for retrofit, or for a future PM display concept, would require significant work to be certified.

Training implications – high; the SW/SSOT represents a new display concept leveraging time. Training pilots to monitor this display element would require extensive training.

Retrofit Recommendations

After considering the evaluation results and certification risk, we provide the following recommendations in terms of retrofit implementations.

Anomaly Detection

In addition to the existing data validation logic, failure annunciations, and legacy EICAS alerts, we believe that an independent monitoring system, akin to the IDAD monitoring technology, should be introduced to increase and expand validation of avionics sensor data.

As demonstrated in the present evaluation, predicted and current anomalies can be detected with redundant sources (ADRS1 vs. ADRS2). Additionally, predicted and current anomalies could be detected with independent sources (ADRS1 vs IRS1), although we have not yet developed this more advanced technology.

Crew Alerting

To align predicted and current anomalies with the existing alerting philosophies, we believe predicted anomalies should be Memo messages (white) and current anomalies should be Caution messages (amber). This also aligns with conclusion of previous predicted alert evaluation (Duan et al., 2016). In the present evaluation, predicted anomalies were amber Cautions and current anomalies were red Warnings. Consequently, crews sometimes took immediate action in the form of automation changes upon seeing the predicted anomaly (e.g., crew 2, scenario 50, run 6). While this is encouraging in terms of crew ASA, we do not necessarily want the crew to take immediate action upon receiving a predicted anomaly. Rather, the aim is for the crew to discuss the situation and prepare for their next action, which could be better facilitated with a Memo message. An example of this is crew 1 (scenario 51, run 12) where the FO puts his hand on the throttles upon seeing the predicted alert.

Additionally, new checklists should be created for novel anomalies (e.g., sensor exhibiting too little variance), even if it is awareness items with no associated actions like the "IAS DISAGREE" alert and checklist, which was also triggered as a legacy alert for high variance airspeed anomalies. We realize that this creates a certification risk, but it is needed for new anomalies, so whenever possible the anomaly messages should be incorporated into the indenting or categorization scheme of related legacy alerts, like "IAS DISAGREE" in the crew alerting system. From this evaluation, an example of a novel anomaly would the Radar Altimeter since there was no related disagree alert like there is for Airspeed and Pitch.

For aircraft that do not have an EICAS, it has display space for information on the Engine Display, below Engine Alerts, with master caution/master warning. For example, on the A320/330/350, the ECAM/Warning Display could be used for integration along with the checklists.

PFD Indications

In terms of PFD indications, based on the evaluation feedback and precedent in instrument training, we believe anomalous PFD elements should be removed from the display and replaced with an anomaly flag along with a corresponding EICAS message, as tested in the evaluation.

There was unanimous support for this in the evaluation, but this may come with high certification risk. We do, however, think there is precedent from the current way PFD Failure Flags are depicted on modern aircraft, and from general aviation instrument training where pilots cover erroneous display elements with a sticky note or instrument display cover during training. Information removal is a familiar concept for all pilots with an instrument rating.

If removal of PFD information, as tested in the evaluation, was deemed too risky to implement, we would recommend using the PFD flags, but leaving the anomalous information on the PFD. One crew indicated that pilots may want to compare the anomalous data against other sources. Ultimately, we believe removal of this information supports safer operation, in the event that pilots do not notice the failure flag and trust the anomalous information.

Another alternative to removal of PFD elements is to provide some indication on the impacted element, in addition to the PFD flag and EICAS message. This could be in the form of a yellow line or cross (i.e., "X") through the affected element. However, we believe these values could potentially be employed by the pilots or misread due to the occluding element.

Status Widget

Given that data anomalies are ultimately quite rare, we believe the SW/SSOT concept is potentially too risky to retrofit, and instead believe it should be considered as a forward fit option with a fully reconsidered PM role and displays. The infrequent utilization may not support the certification and training cost for retrofit.

The current location of the SW/SSOT, at the bottom of the EICAS, is an underutilized space, but it is not currently a part of the pilot scan pattern. The location of the SW/SSOT could be reconsidered in a forward fit.

Themes and Precursors

In the Contract's Year 2 Report (Whitlow and Dillard, 2018), we assessed how well the mitigation technologies addressed the themes and precursors common in incidents/accidents involving compromised pilot state awareness identified in Year 1 (Whitlow et al., 2017). We do not repeat those conclusions here but do reference them as the mappings are the same for the present NASA LaRC evaluation. Please see Appendix A for the Year 2 Themes and Precursors conclusions.

Next Steps

The present evaluation findings are helpful in guiding future iterations of the anomaly detection methodology and the pilot interface. In future iterations, the most impactful changes will likely be: i) expanding the anomaly detection to multiple, independent sensors as well as ii) creating platform-specific interfaces that employ PFD elements removal with flags plus fully integrated alerts with tailored checklists.

Detection Methodology

The IDAD detection technology used in the present evaluation monitored a single sensor to detect each type of anomaly (e.g., high amounts of variability in IRS1 pitch). The next step is to mature the technology to compare redundant sensors (IRS1 pitch vs. IRS2 pitch), then to compare independent sources (IRS1 pitch vs ADRS1 airspeed). Our partner, AT CORP, developed several proofs of concept angle of attack (AOA) estimators based on independent sources, then evaluated their efficacy for capturing ground truth AOA variance. They documented their methodology and some promising early results in a report that is in Appendix B. Additionally, there is no limit to the dimensionality of multi-source detection. For example, airspeed, pitch, and AOA could be included in a three-dimensional relationship, and so on. This will increase the confidence in the assessment, could support additional types of anomaly detection, and may provide additional computational tractability.

Further testing of the detection methodology would best take place on an aircraft bench, such as the Honeywell Embraer E2 Bench, where the sensor inputs directly reflect the aircraft configuration and data can be simulated to produce anomalous situations.

Pilot Interface Design

Based on the present evaluation results, the pilot facing interface could see some significant changes that would benefit from being evaluated, some that are platform-specific considerations (Airbus, Boeing, Embraer). For such an evaluation, the recommended scope is removal of information from the PFD plus flags, combined with fully integrated EICAS/ECAM alerting with tailored checklists (no status widget). For the anomalies being tested, we propose drafting a checklist for use in a future flight simulator evaluation.

References

Brooke, J. (2013). SUS - a retrospective. Journal of Usability Studies.

Commercial Aviation Safety Team (CAST) (2014a). SE208: Airplane State Awareness – Airplane Systems Awareness, Detailed Implementation Plan, Commercial Aviation Safety Team. December 4, 2014.

Commercial Aviation Safety Team (CAST) (2014b) Airplane State Awareness – Airplane Systems Awareness, Joint Safety Analysis Team: Final Report Analysis and Results. June 17, 2014.

Duan, P., M. Uijt de Haag, T. Etherington and L. Smith-Velazquez, "Energy State Prediction Methods for Airplane State Awareness," Proceedings of the IEEE/AIAA 35th Digital Avionics Systems Conference (DASC), Sacramento, CA, September 2016.

Evans, E., Young, S. D., Daniels, T., Santiago-Espada, Y., & Etherington, T. (2016). Analysis of pilot feedback regarding the use of state awareness technologies during complex situations.

Hart, S. G., & Staveland, L. (1988). Development of the NASA task load index (TLX): Results of empirical and theoretical research. In P. A. Hancock & N. Meshkati (Eds.), Human Mental Workload (pp. 239-250). Amsterdam: North Holland.

NASA. NASA Langley Research Center. NASA Space Environments Testing Management Office. https://www.nasa.gov/sites/default/files/atoms/files/larc_flight_simulation.pdf

Roscoe, A. H. (1984). Assessing pilot workload in flight: Flight test techniques (No. AGARD-CP373). Neuilly-sur-Seine, France: NATO Advisory Group for Aerospace Research and Development (AGARD).

Taylor, R. M. (1990). Situational Awareness Rating Technique (SART): The development of a tool for aircrew systems design. AGARD-CP-478, Situational Awareness in Aerospace Operations, NATO-AGARD, Neuilly Sur Seine, France, 3/1 –3/17.

Whitlow, S., & Dillard, M. (2018). Honeywell Enhancing Airplane State Awareness (EASA) Project: Year 2 Report on Design Reviews and Evaluation Report (NASA Technical Report).

Whitlow, S., Holder, Barbara, Gannon, A., & Lancaster, Jeff. (2017). Report on Conceptual Designs and Recommendations for Further Development (NASA Technical Report).

Young, Steven D., Maarten Uijt De Haag, Taumi Daniels, Emory Evans, Kimberlee H. Shish, Stefan Schuet, Timothy Etherington, and Daniel Kiggins. (2016a). Evaluating Technologies for Improved Airplane State Awareness and Prediction. In AIAA Infotech@ Aerospace, 2043, 2016.

Young, Steven D., Taumi Daniels, Emory Evans, Evan Dill, Maarten Uijt de Haag, and Tim Etherington. (2016b). Flight Simulation Study of Airplane State Awareness and Prediction Technologies. In 2016 IEEE/AIAA 35th Digital Avionics Systems Conference (DASC), 1–11. IEEE, 2016.

Appendix A: Themes and Precursors Conclusions from Year 2 Report

To review the impact of the technologies, we assessed how well the technologies addressed the themes and precursors common in incidents/accidents involving compromised pilot state awareness identified in Year 1 (Whitlow et al., 2017). After reviewing the results, we concluded that these technologies addressed the following themes and precursors (bold text), with rationale provided.

- Flight path Management (FPM)
  - Unstabilized approaches
  - Problem with ground-based navigation support--- monitor altitude
  - **Slowly developing situation**
    - **Rationale: Predicted anomaly detection would identify a slow degradation of airspeed or slow bank that could likely be missed by crew.**
  - Distracted by non-flight critical situations
- Aircraft Systems
  - **Bad Data**
    - **Rationale: Anomaly detection, in general, provides another safeguard against data irregularities that are not associated with an annunciated failure by the sensing system.**
  - **Lack of specific alerts**
    - **Rationale: Predicted and current alerts are an addition that would enhance pilot state awareness beyond the current state of practice.**
  - Uncommanded control maneuvers
  - Complex mode logic
  - Procedural Support
- Flight Crew Interface (FCI)
  - **Incorrect response to alert**
    - **Rationale: Current and predictive alert could identify impacts to better inform how the crew responds to the situation.**
  - Cognitive tunneling
  - **Startle/surprise to unexpected situation**
    - **Rationale: By providing alerts to current and predicted anomalies with explanations, crew will be less likely to succumb to startle/surprise; as a result, they should develop a more accurate understanding of the actual**

airplane state and be more likely to make the correct decision on how to respond.

- **Lack of shared awareness of crew**
  - **Rationale: Both technologies encourage crew to discuss the situation and compare different information sources, likely developing an improved and more accurate shared awareness of airplane state.**
- **Monitoring challenges**
  - **Rationale: Anomaly detector provides an automated means to monitor for possible situations related to critical flight data then alert the crew; the timeline would provide a new means to help structure the crew's monitoring behavior, including expected and unexpected events.**
- **Automation awareness**
  - **Rationale: Predicted anomaly notification includes likely system impact so that should the autopilot (AP) or autothrottle (AT) disconnect, crew should recognize these disconnects better and understand the reason behind them.**
- **Poor CRM**
  - **Rationale: By providing new information relevant to airplane state, crew are encouraged to discuss; furthermore, the proposed central location of SW/SSOT (EICAS) physically and cognitively brings crew together to discuss situations, potentially improving CRM.**
- Pilot fatigue

This page intentionally left blank.

Investigation of Methods for Validation of Air Data Measurements

Version: 1

October 31, 2019

NASA Contract: NNL17AA31T

NASA ATD 2.2 Enhancing Airplane System State Awareness

Honeywell PO: 6400421466

Prepared for:

Dr. Stephen Whitlow

Honeywell Aerospace Advanced Technology

HONEYWELL INTERNATIONAL INC

1985 DOUGLAS DRIVE

GOLDEN VALLEY MN 55422

Prepared by:

ARCHITECTURE TECHNOLOGY CORPORATION

Architecture Technology Corporation

9971 Valley View Road

Eden Prairie, MN 55344-3552

Table of Contents

1 Introduction

Under a 2015 NASA Research Announcement in support of the Airspace Technology Demonstration (ATD) Project, Honeywell is performing a multi-year study to investigate technologies for indicating airplane system status and dependencies during complex, non-normal situations. In particular, Honeywell is researching and developing technologies to attempt to mitigate the Loss of Airplane System State Awareness (LASSA) including: enhanced displays, novel information management approaches, and new alerting functions. In a 2017 interim report, Honeywell identified 26 candidate solutions related to display, alerting, and information management that could help mitigate and support pilots in recovering from loss of airplane state awareness. One of the 26 candidate solutions proposed the development and application of data source validation algorithms and alerting to notify the crew of non-failure sensor anomalies.[1] In their report Honeywell describes the problem and proposed solutions as follows:

> *Source Data Validation & Indication*
>
> *When an input source fails, such as air data, radar altimeter, or an angle of attack sensor, systems handle the situation rather elegantly by de-coupling from autoflight, automatically switching the data source, and providing an alert to the crew. However, most aircraft systems do not handle non-failure anomalous data elegantly— they present the erroneous data to the crew (e.g., West Air Sweden Cargo CRJ200), and allow autoflight to be driven by it (e.g., Turkish 1951 B737-800). An intermittent anomaly can be particularly insidious, since the crew can have difficulty determining what display source to trust. Most input data are validated with voting methods of like-sensors, which has contributed to accidents. In two separate examples (Lufthansa LH1829 A321 (05-Nov-14); XL Airways End of lease Functional Check Flight (FCF) A320 (27-Nov-200), where the system invalidated a correct AOA sensor because it differed from the other two AOA sensors that had been fouled by precipitation and washing without protection, respectively, and froze at a fixed value once at altitude.*
>
> *The Team proposes to augment existing voting methods with robust statistical methods with respect to how flight critical input data sources are validated. It would be rather straightforward to develop a statistical model of sensor input that, once developed, could be used to evaluate whether current values deviate. If a deviation is detected, then the system would alert the crew to help improve their diagnosis of ongoing anomalous autoflight behavior, or help them to decide whether to disengage autoflight proactively. A sensed dimension would be still be presented, but it would transition to grayscale as a representation of diminishing trust in its accuracy.*

[1] Report on Conceptual Designs and Recommendations for Further Development (Deliverable Item 6.4), December 5, 2017, Dr. Stephen Whitlow, Dr. Barbara Holder, Dr. Aaron Gannon, and Dr. Jeffrey Lancaster, Honeywell Aerospace Advanced Technology, Honeywell, Inc.

Based on discussions with Honeywell, ATCorp proposed and performed a study of methods by which the angle of attack of an aircraft can be validated without relying on measurements from an airplane's air data system (ADS) or air data computer (ADC). The purpose of this document is to describe and evaluate proposed methods by which the angle of attack of an aircraft can be validated without relying on air-data measurements. These methods are based on well-established mathematical formulations describing the flight dynamics of a conventional fixed-wing aircraft.

This document explores three methods: one based on flight path angle, one based on vertical acceleration, and one based on elevator deflection. The three methods were evaluated using a basic 6 DOF (Degree of Freedom) desktop research simulator. Due to the limited scope of the project, only simple methods were explored; but these methods provide useful insight into the nature of the problem. Based on the observations obtained in this study, further research into signal comparator systems should consider more sophisticated methods such as those involving Kalman filtering and state observers.

1.1 Background

Pilot Airplane System State Awareness is an interactive process between pilots, their aircraft and the environment in which the aircraft is operating. Often, one of the reasons for loss of state awareness is erroneous or conflicting information being given to the pilots. This can occur when sensors fail and aircraft systems fail to properly identify the problem. One particularly pernicious problem is the failure of the air data system. Since the ADS relies on direct interaction with the ambient environment to establish its parameters, it is exposed to atmospheric hazards such as moisture, extreme temperatures, and vibration, depending on the sensor/probe location.

In particular, the measurement of Angle-Of-Attack (AOA), symbolized by the Greek letter α(Alpha), is particularly vulnerable to failures. As such, high assurance systems have triple redundancy AOA sensors which are then compared by signal voting algorithms to determine the correct state. Most of the time, this level of redundancy is sufficient to ensure that the appropriate AOA is received by the ADS. However there have been several instances, where multiple AOA sensors have been damaged simultaneously in a fashion such that a simple voting algorithm is unable to determine the appropriate signal. In these cases, the best result is for the system to determine the unreliability of the signal and either warn the air crew or engage reversionary modes to compensate. In rare instances, the failures have been pernicious enough that the voting algorithm chose the wrong AOA signal when two sensors failed at a similar measurement state. The result was that the voting algorithm failed to recognize the failure and provided erroneous information to the rest of the flight deck.

2 Estimating Angle Of Attack

There are multiple methods that can be used to indirectly determine the angle of attack of an aircraft from other state data. This document explores three methods, one based on the flight path angle, a second based on the vertical acceleration, n_z, and a third based on the elevator deflection.

2.1 Estimating AOA Using Flight Path Angle

The flight path angle of an aircraft is the angle its vertical path makes with the inertial horizontal plane. In steady level flight, the relationship between the pitch angle θ, the flight path angle γ_a, and the angle of attack α, is shown in Figure 2.1 and Equation (2.1) The pitch angle defines the angle the aircraft body axis makes with the horizontal plane, and the flight path angle is generally slightly lower. The difference is the angle of attack.

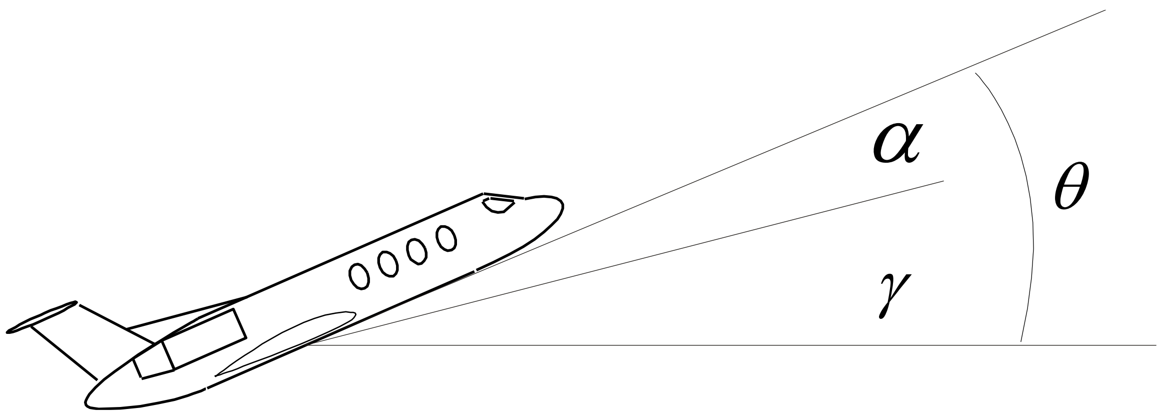


Figure 2.1. High level flow for flight planner

$$\theta = \gamma_a + \alpha \tag{2.1}$$

The flight path angle term carries the *'a'* subscript denoting that the flight path angle is measured with respect to the air-mass.

2.1.1 Cruise Flight and Other Considerations

One important observation to make initially is that in cruise flight, where the flight path is level, the flight path angle will be zero. Hence, the angle of attack will equal the pitch angle. Since this condition characterizes a large majority of the flight, the angle of attack can be easily monitored since it should equal the pitch angle exactly. It can also be seen that for cruise flight no extensive calculation for γ_a is required. It is sufficient to determine that the vertical speed of the aircraft is close to zero within some tolerance.

Beyond level cruise flight, it should be emphasized that this method is best employed only during steady operations. The calculation of AOA using this method is inherently problematic if the aircraft Phugoid dynamics are sufficiently excited. As will be explored further in Section 3, the Phugoid mode dominates the Pitch signal and the Flight Path Angle (FPA) signal (a combination of speed and altitude). Therefore the method is essentially relying on the differencing of these two signals to eliminate the Phugoid to result in a steady AOA response. Numerical inconsistencies render this method largely useless during aggressive maneuvering.

2.1.2 Calculating Flight Path Angle

In cases where the aircraft is climbing or descending, a reasonable approximation of the flight path angle must be calculated. This can be obtained either from air data measurements or from inertial measurements providing a decent wind approximation is available. If wind measurements are available the flight path angle can easily be determined from basic trigonometry, where V_v is the vertical speed of the aircraft, and V_a is the true airspeed of the aircraft as measured along the flight path.

$$\gamma_a = \sin^{-1}\left(\frac{V_v}{V_a}\right) \tag{2.2}$$

However, it is likely that if angle of attack data is compromised, it may be that other air data is also compromised, so a method of determining flight path angle from inertial or navigational data is desired. In most cases inertial data can give the groundspeed and ground track of the aircraft as well as the vertical speed of the aircraft with respect to the inertial frame. These measurements are of course inertial, and therefore not measured with respect to the air-mass. Furthermore, the ground-speed represents the horizontal speed of the aircraft, not the speed of the aircraft along its flight path. To estimate the air data from the inertial data an estimate of the winds is needed. Often this is available to the aircraft either from a datalink, or due to prior on-board calculations when both air-data and inertial data were available. Generally the conversion is as simple as subtracting the wind component from the ground speed as shown in Equation (2.3). Equation (2.3) shows an implicit vector subtraction of the respective velocities. The format of the wind data on the aircraft will determine the exact formulation needed. There is likely no good way to account for vertical gusts.

$$\vec{V}_a = \vec{V}_g - \vec{V}_w \tag{2.3}$$

Since the groundspeed is a horizontal measurement rather than the total measure, the flight path angle is slightly different, as shown in Equation (2.4). A subscript *'h'* is added to the true airspeed term to denote it is a horizontal measurement. In most cases this distinction matters very little to the overall speed because the cosine of angles in the normal range for flight path angle are very nearly unity (i.e., $V_a \approx V_{ah}$) and either Equation (2.2) or (2.4) will suffice.

$$\gamma_a \approx \tan^{-1}\left(\frac{V_v}{V_{ah}}\right) \tag{2.4}$$

In modern aircraft inertial or navigation data is usually quite good. Modern GPS systems in the United States and Canada make use of the Wide Area Augmentation System (WAAS) which improves the accuracy of the GPS signals. The WAAS specification requires it to provide a position accuracy of 25ft or less (for both lateral and vertical measurements), at least 95% of the time. Actual performance measurements of the system at specific locations have shown it typically provides better than 5ft in most parts of North America. With these results, WAAS is capable of providing the data needed to approximate air-data.

2.1.3 Maneuvering Flight

In maneuvering flight, the calculation of flight path angle requires a more thorough estimate of angle of attack beyond what is provided in Equation (2.1). The formal definition of flight path angle is shown in Equation (2.5).[2] The nomenclature C_θ and S_θ are simplified notation for $\cos\theta$ and $\sin\theta$. This is done for all trigonometric manipulations to simplify the ultimate expression.

$$\gamma_a = \sin^{-1}\left(C_\alpha C_\beta S_\theta - \left(S_\phi S_\beta + C_\phi S_\alpha C_\beta\right)C_\theta\right) \tag{2.5}$$

Equation (2.5) must be rearranged in terms of γ_a to solve for α. Due to the transcendental functions, it can only be done approximately, however a numerical solver could be used to solve it exactly if needed. First, it will be assumed that the sideslip angle of the aircraft, β, is small. This eliminates several terms. Unfortunately, this leaves the expression still in terms of both the sine and cosine of α, so another simplification is made. Since the cosine of small angles is nearly unity, the cosine of α is assumed to be one. The cosine of 15^o is 0.965, which is an AOA value that the aircraft is unlikely to exceed in normal flight. The Euler angles θ and ϕ are known and therefore constant for this calculation.

$$\gamma_a = \sin^{-1}\left(C_\alpha S_\theta - \left(C_\phi S_\alpha\right)C_\theta\right) \tag{2.6}$$

$$\sin\gamma_a = S_\theta - \left(C_\phi C_\theta\right)S_\alpha \tag{2.7}$$

$$\alpha = \sin^{-1}\left(\frac{S_\theta - S_{\gamma_a}}{C_\phi C_\theta}\right) \tag{2.8}$$

2.2 Estimating AOA Using Vertical Acceleration

The vertical acceleration of the aircraft, n_z, often can be used as a direct stand-in for angle of attack since n_z tracks angle of attack closely. The vertical acceleration is a function of the lift of the aircraft divided by its mass. In steady level flight, when the vertical acceleration is pointed down and the lift vector is mostly pointed up, the relationship is straightforward as shown in Equation (2.9).

$$n_z \approx \frac{L}{mg} \approx \frac{qS_w C_L}{mg} \approx \frac{qS_w\left(C_{L_o} + C_{L_\alpha}\alpha\right)}{mg} \tag{2.9}$$

The dynamic pressure is defined as follows:

[2] Aircraft equations of motion can be found in reference texts such as: Etkin, Bernard, and Lloyd D. Reid. Dynamics of Flight: Stability and Control. New York: Wiley, 1996. Print.

$$q = \tfrac{1}{2}\rho V_a^{\,2} \tag{2.10}$$

Based on Equation (2.9), for a given combination of weight, altitude, and airspeed, n_z varies linearly with α . (see Equation (2.11) and (2.12))

$$n_z \approx \frac{qS_w C_{L_o}}{mg} + \frac{qS_w C_{L_\alpha}}{mg}\alpha \tag{2.11}$$

$$\alpha \approx \frac{n_z - \dfrac{qS_w C_{L_o}}{mg}}{\dfrac{qS_w C_{L_\alpha}}{mg}} \tag{2.12}$$

Using n_z to approximate angle of attack requires considerable information about the aircraft, including the altitude, the airspeed, its current weight, and some basic lift curve information. Also, it is the most sensitive to errors in the airspeed measurement due to the squaring of speed in the dynamic pressure term. However, it will be the most accurate for capturing the dynamic response of the signal, and it will work during maneuvering. Furthermore, it is just a linear approximation, and may not yield perfect answers in the upper ranges of the alpha curve. However, its primary advantage is the response speed.

2.3 Estimating AOA Based on Elevator Deflection

For a given aircraft configuration, including flap position, gear position, center of gravity, etc.., there is generally a fixed theoretical relationship between the elevator deflection (or other pitch controlling device) and the angle of attack. This relationship can be seen in the non-dimensionalized pitching moment Equation (2.13), where δ_e is the elevator deflection. In a steady state condition the pitching moment is zero, (i.e. $C_m = 0.0$) and the unsteady terms (i.e. $\dot{\alpha}, q$) are small. This leaves Equation (2.14) which can be rearranged to solve for α (Equation (2.15).

$$C_m = C_{m_0} + C_{m_\alpha}\alpha + \frac{\bar{c}_w}{2u_o}C_{m_{\dot{\alpha}}}\dot{\alpha} + \frac{\bar{c}_w}{2u_o}C_{m_q}q + C_{m_{\delta_e}}\delta_e \tag{2.13}$$

$$0 = C_{m_0} + C_{m_\alpha}\alpha + C_{m_{\delta_e}}\delta_e \tag{2.14}$$

$$\alpha = \frac{-\left(C_{m_0} + C_{m_{\delta_e}}\delta_e\right)}{C_{m_\alpha}} \tag{2.15}$$

This is a truly interesting result because it suggests that the angle of attack can be determined directly without any information about the aircraft state. All dimensional terms drop out and all inertial measurements are irrelevant. The disadvantage however is that the estimate is only valid for smooth air. Any external disturbance that creates a shift in angle of attack, would not be predicted by elevator deflection. However, the estimate gives a really good estimate of what the steady state angle of attack should be. This estimate should be used with discretion. For instance, if no alpha response is observed with the deflection of the elevator, the alpha senor (or elevator) should be considered suspect. Furthermore, the signal should never be considered as a substitute for any signal being used in a feedback control capacity, because it doesn't contain any real information about the aircraft dynamics. It is only good as a reference estimate.

Implicit in this approximation is that the dynamics between elevator and angle of attack is instantaneous. As will be demonstrated in Section 3, the dynamics are governed by the short period. During high frequency manipulation of the controls, this results in a lag between this approximation and the AOA signal. To make up for this lag, a simple first order lag is added to the AOA approximation, where α_c is the calculated value from Equation (2.15), and α_l is the lagged output.

$$\alpha_l = \frac{1}{\tau s + 1}\alpha_c \tag{2.16}$$

3 Modal Properties Analysis

The modal properties of the aircraft dynamics impact how various measurement signals should be expected to respond. This information can be used to determine whether a particular signal is correct or faulty.

3.1 Longitudinal Modes

There are two second order (oscillatory) modes that describe the longitudinal dynamics of a fixed wing aircraft. The two modes are typical of the control-fixed longitudinal characteristics of stable airplanes over a wide range of conditions.[3] It is beyond the scope of this effort to provide a detailed mathematical development of these modes, however a brief discussion of the modes and the relevance to our analysis are provided below.

3.1.1 Short Period

The short period mode is named because it is the faster of the two modes. It is the mode which defines the aircraft's pitching about its center of gravity (see Figure 3.1). Generally, the short period mode is over ten times faster than the other longitudinal mode. The short period mode controls the dynamics between elevator deflection and the aircraft's angle of attack. The behavior of the angle of attack is dominated by the short period mode. Similarly, the n_z signal, since it is largely a function of angle of attack, mostly reflects the short period dynamics.

[3] Etkin, Bernard, and Lloyd D. Reid. Dynamics of Flight: Stability and Control. New York: Wiley, 1996. Print.

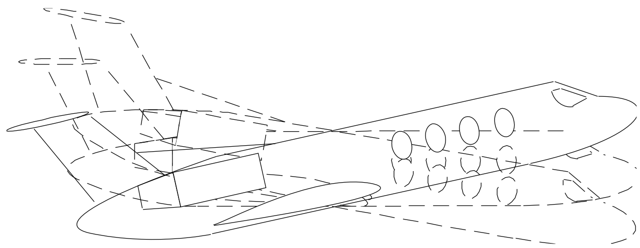


Figure 3.1. Illustration of Short Period Dynamics

3.1.2 Phugoid

The Phugoid mode is the slower of the two longitudinal modes. The Phugoid reflects a gradual interchange between potential and kinetic energy about the equilibrium altitude and airspeed (see Figure 3.2). The Phugoid mode is characterized by changes in pitch attitude, altitude, and velocity at a nearly constant angle of attack. The Phugoid will have a strong influence over the behavior of the pitch angle, but is nearly non-existent in the angle of attack. In fact, pitch angle is influenced by both dynamic modes, so the transient response from each mode can be observed in pitch signals.

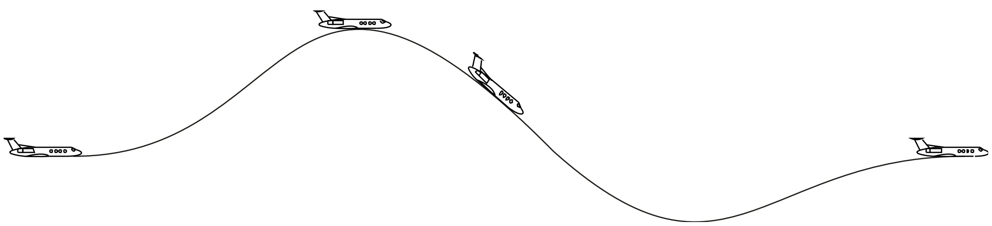


Figure 3.2. Illustration of Phugoid Dynamics

3.2 Simulation Examples

Using a desktop simulation (described in Section 6), example data tracks were developed to demonstrate the impact of these dynamic modes on the aircraft state. Figure 3.3 shows the excitement of the Phugoid dynamics with an elevator doublet. The Figure shows a single elevator doublet in the first 5 seconds of the signal. This doublet immediately causes a response in AOA, N_z, and pitch. Both the AOA and N_z response to the elevator doublet are governed by the short period. The response is nearly immediate, with an overshoot prior to returning to an equilibrium value. After the initial transient decays, both AOA and N_z remain relatively stable, however the pitch signal shows a slower, more pronounced oscillation. This oscillation, the Phugoid, shows that the pitch is governed nearly equally by both of the longitudinal modes, essentially acting together.

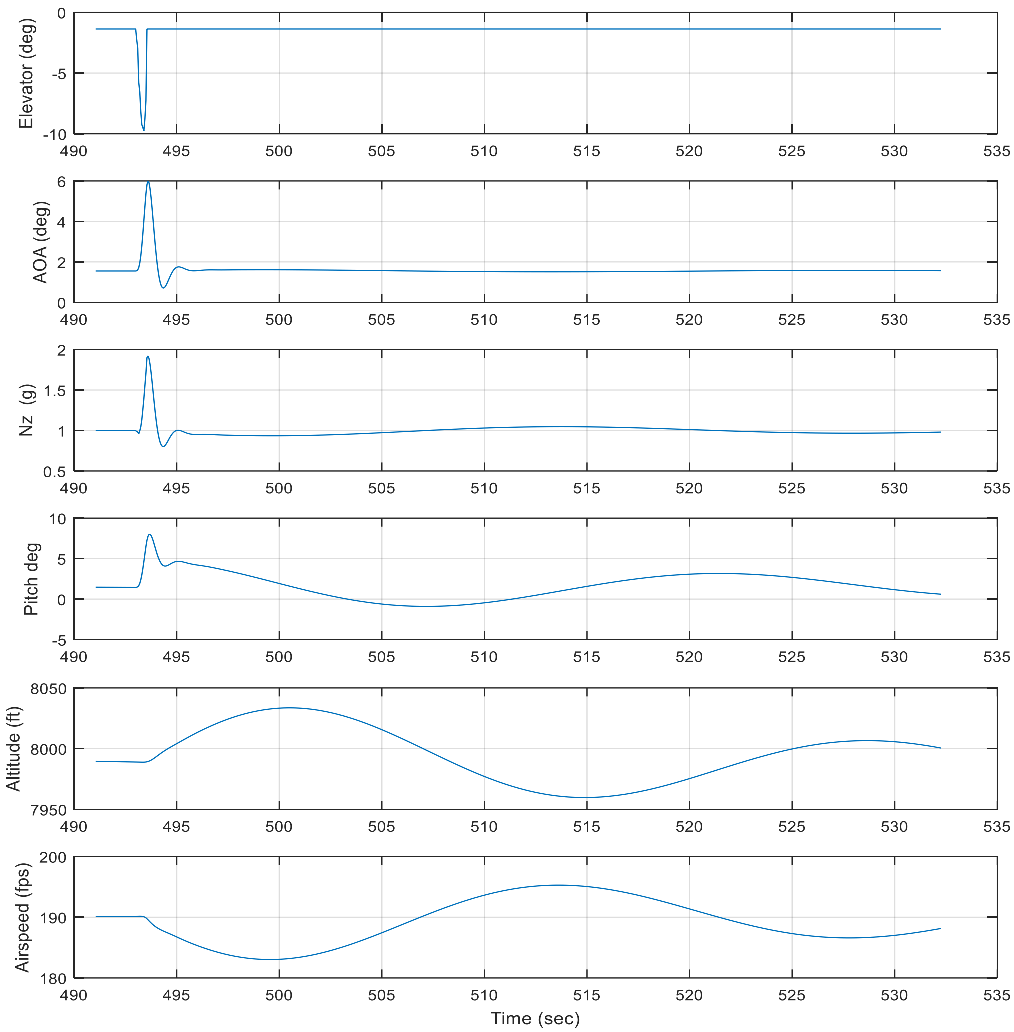


Figure 3.3. Phugoid excitement after single elevator doublet

The important conclusion is that from a dynamic response, while both pitch and AOA respond to elevator deflections, pitch angle generally cannot serve as a good surrogate for AOA, whereas, the N_z signal, governed much more exclusively by the short period, can serve as a stand-in for AOA. Notice also that airspeed and altitude are dominated by the Phugoid mode, and show nearly no short period response. The conclusion is that AOA and N_z responses are dominated by the short period, speed and altitude are dominated by the Phugoid, and the pitch angle is influenced by both modes. This phugoid influence illustrates the difficulty of using pitch information to reconstruct AOA.

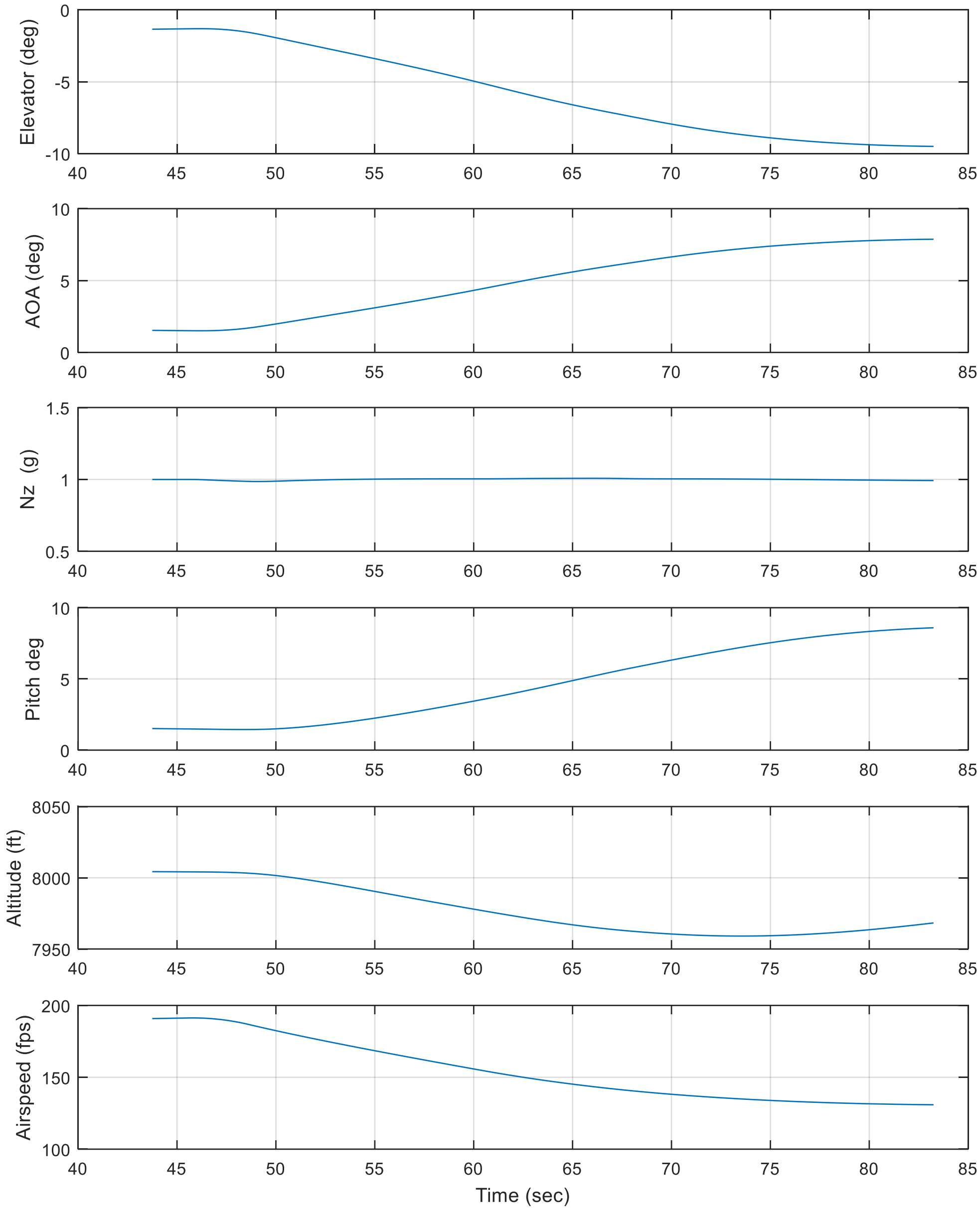


Figure 3.4. Speed change in level flight under autopilot control

Figure 3.4 shows a constrained response where the autopilot is used to facilitate a speed change from 100 KIAS to 70 KIAS (the plot shows speed in ft/sec true airspeed). In this case, where the aircraft is in level flight, both the Phugoid and short period mode are completely damped out of the system by autopilot inputs. In this case, pitch serves as a good surrogate for AOA. In fact it should equal AOA exactly.

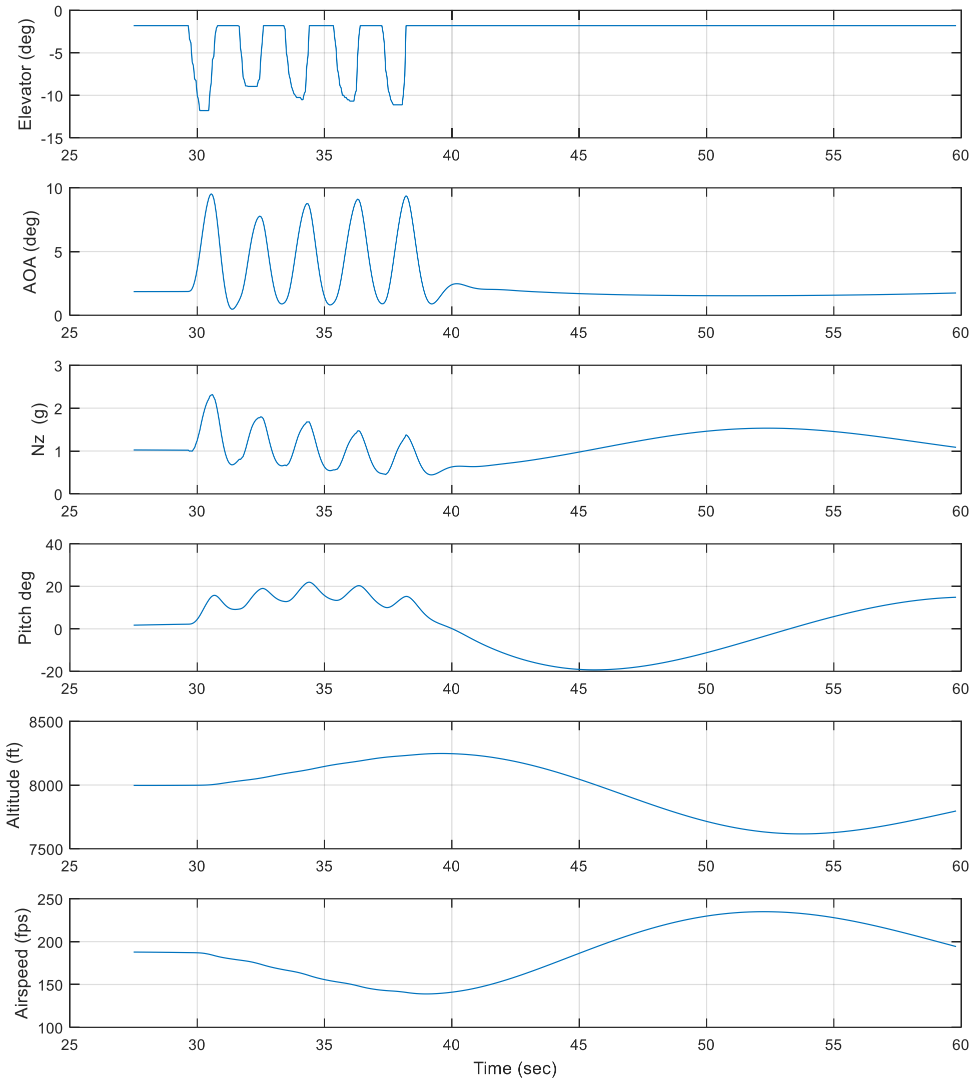


Figure 3.5. Aggressive Pitch Maneuvering

Also note that the N_z response is level, showing that for equilibrium conditions, it makes a poor surrogate for AOA. The elevator deflection tracks the AOA very well. In cases where the short period is sufficiently faster than the commanded input, the AOA response to elevator appears instantaneous. In steady conditions, both elevator[4] and pitch can be used to estimate AOA. Figure 3.5 shows more dramatic longitudinal maneuvering to further highlight the short period response in the signals. The response shows how closely AOA will track the elevator. The AOA will lag elevator, showing that a better

[4] In flight mechanics nomenclature 'up-elevator' from a pilot's perspective is denoted as negative to maintain a consistent sign convention

approximation for AOA is obtained by the elevator response fed through a 1st order lag. The lag doesn't capture the overshoot of the second order short period, but it will tend to improve peak alignment.

4 Signal Comparison

From the modal property analysis of the prior section, it can be seen that each method may have its individual strengths. The pitch signal is good for determining angle of attack exactly, but it is only valid in steady level flight. The n_z signal is good for determining if the AOA signal properly reflects the dynamics of the short period, but otherwise doesn't directly enable any value-by-value comparison.

4.1 Signal Similarity

The fact that n_z is similar to AOA in terms of its dynamic properties, but differs by some unknown (difficult to calculate) linear relationship, led to the conclusion that perhaps signal similarity techniques could be employed between the two signals, which would forgo the need for sophisticated processing using other state data. Two methods were tried. The first was a traditional method based on the convolution of a sampled part of the signal. This worked well in a static Matlab environment, but did not perform well in the simulation. The main problem is that the short period dynamics just are not observed often enough to make the comparison. Most of the flight trajectory is smooth as observed in Figure 3.4. For the convolution algorithm to work reliably, the signal would have to possess dynamic properties more like those in Figure 3.5. Indeed, the algorithm did work when the operator continuously cycled the elevator. Another attempt was made to develop an ad-hoc method that would identify the peaks of signals based on the derivative of the respective signals. This method worked reasonably well, but small minority of peaks were missed, which led to some false positives. And the main problem still remained, that if there was not sufficient short period excitement, the algorithm could not render an opinion.

Therefore, for comparison, a simple value-by-value comparison was employed. A moving average of the signal errors was maintained, and if a mean signal error exceeded a specified tolerance, the signal was rejected. This method worked adequately well, but also meant that for n_z, accurate representations of the linear relationship between AOA and n_z needed to be maintained.

4.2 Moving Average

The moving average method was based on a simple moving average of the value-by-value difference between the measured signal f_m and reference signal f_r, where the measured signal represents the data from the sensor, and the reference signal represents the approximation to the signal from one of the methods from section 2. The relationship is shown in Equation (4.1). The term *k* indicates the location in the main sample. The moving average considers the prior *m* points. If the error E_k exceeds a tolerance, the signal is considered to have failed.

$$E_k = \sum_{i=(k-m)}^{k} \frac{\left|\left(f_m(i) - f_r(i)\right)\right|}{m} \tag{4.1}$$

$$E_k > E_{tol} \rightarrow \text{Failed Condition} \tag{4.2}$$

A typical error tolerance used in the simulation was 0.5 deg.

5 Signal Comparison Algorithms

The overall algorithm consists of a series of signal comparators that work together to determine which of the three AOA sensor signals is accurate. There are five different comparators. The first is the traditional simple voting comparator, which just differences the three input signals and looks for an outlier signal. Then there are four special signal comparators that compare each sensor signal to a calculated reference signal. The comparators are named for the primary contributing input to its respective reference signal.

1. Simple Voting
2. Pitch-Based Comparator
3. FPA-Based Comparator
4. Nz-Comparator
5. Elevator-Comparator

Each comparator, has limits placed on its availability for use, based on the state of the aircraft. These limits are referred to as the operating envelope of the comparator, and reflects the boundaries outside of which, the comparator cannot produce an accurate assessment. The comparator operating envelopes are shown below:

1. Simple Voting (No envelope limits)
2. Pitch-Based Comparator (Vertical Speed +/- 250 ft/min, +/- 5 deg)
3. FPA-Based Comparator (Vertical Speed +/- 1000 ft/min, +/- 20 deg)
4. Nz-Comparator (Vertical Speed +/- 5000 ft/min, +/- 90 deg)
5. Elevator-Comparator (Vertical Speed +/- 5000 ft/min, +/- 90 deg)

The pitch and FPA comparators are essentially the same comparator, using the measured pitch and approximated flight path angle in conjunction to estimate AOA. A distinction is made between them based on operating envelope. Since steady level conditions can yield an exact figure for AOA using the FPA method, under very steady conditions the FPA comparator essentially gets two votes.

The limits are of most concern for the FPA-based algorithms, which are the most accurate, but also the most susceptible to error. Deviations from level flight can lead to large errors in the reference approximation signal. For the *Nz* and elevator systems, limits are established, but they are arbitrary. It is not clear what limits, if any, are needed. Figure 5.1 shows the basic algorithm. Each comparator is provided with the three AOA signals, and performs an evaluation of the signal validity based on the respective reference signal associated with the particular comparator.

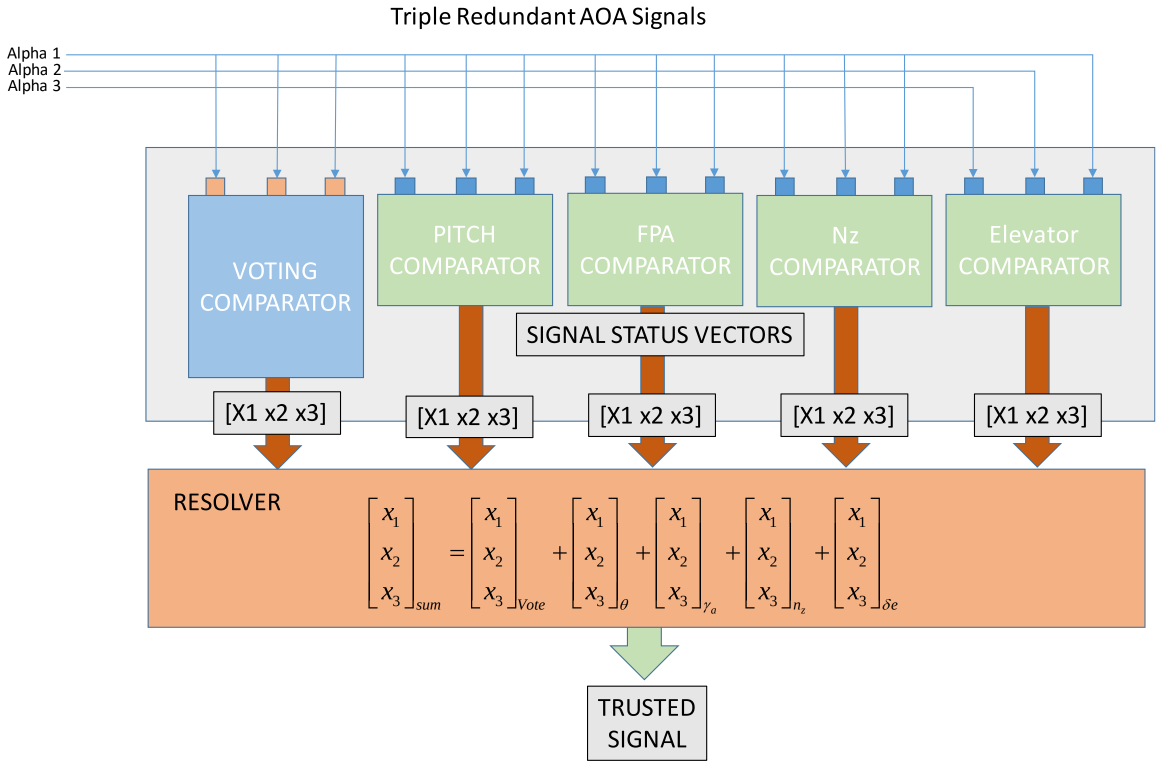


Figure 5.1. Algorithm for resolving triple redundant AOA signals

The resolver evaluates the signals from each of the comparators and determines the signal most likely to be correct. For now, the resolver sums the result from each of the comparators (see Equation (5.1)). In the ideal scenario, where all sensors are operating perfectly, or when a single sensor failure is encountered, all the comparators should agree with the basic voting comparator on which signals are valid. In the case where the voting comparator has simultaneous failed signals that present nearly the same value and hence rejects the one likely correct value, the 4 reference-based comparators will be able to override the voting algorithm. It may be desirable to establish a minimum score any that signal must achieve to be believable. In the case that where all three sensors fail, the minimum score would prevent the resolver from returning a good evaluation to a signal just because it is the 'best' of the failed signals.

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}_{sum} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}_{Vote} + \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}_{\theta} + \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}_{\gamma_a} + \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}_{n_z} + \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}_{\delta e} \tag{5.1}$$

6 Desktop Simulation

The aircraft simulation used for the research is a 6 DOF (Degree of Freedom) research simulation originally developed to test handling characteristics of light aircraft. The aircraft model is of a Ryan Navion, which is a light civil aircraft capable of roughly 150kts. The model is capable of supporting any type of aircraft, if the proper data is available. Due to constraints on scope of this current effort, the Navion model was used rather than modifying the simulation to a high-speed civil transport.

6.1 Main Interface

The main interface enables the operator to start and pause the simulation. An image of the interface is shown in Figure 6.1. The simulation time is established when the simulation is first started. The start-pause button is a toggle button that toggles the operation of the simulation. It initially starts the simulation and will pause it at any time during operation. The main interface window shows three times. The first is the simulation time. This time represents the time since the first initialization. It is based on original clock time when the simulation was initialized plus all the time that the simulation has run. The middle time is the simulation time with an offset added to correspond with the current clock time. It is normal for the simulation to have an offset between clock time and simulator time, especially after numerous start-pause cycles. The third time is the current computer clock time. The simulation is maintaining a proper frame rate (maintaining real-time operations) when the offset time and the clock time are identical. The simulation will provide a warning if the simulation breaks a frame, and the soft-real time executive will attempt to make up the lost time on successive frames. The main interface also allows the simulation to record data in a '.CSV' (comma separated format), that can easily be imported into other tools. Pressing the 'Record Data' button starts the process, and the button will turn orange while data is being recorded. Toggling the button starts and stops data collection.

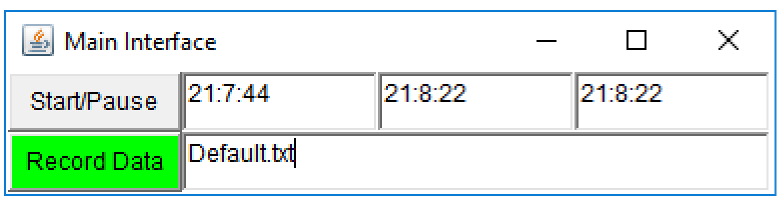


Figure 6.1. Main interface controls the operation of the simulation and allows the use to record data

Stopping and restarting the data recording function without changing the file name will cause the system to overwrite the data in the file, so it is important to keep track of the filenames in use. Table 6-1 shows the data collected in the data file.

Table 6-1. Description and of data members collected in data files

Value	Units	Description
t	seconds	Simulation Time
δ_e	degrees	Elevator Deflection
α	degrees	Angle of Attack (truth)
n_z	g	Vertical Acceleration (body frame)
θ	degrees	Pitch Angle

Value	Units	Description
ϕ	degrees	Roll Angle
α_{s1}	degrees	Angle of Attack Sensor Model 1
α_{s2}	degrees	Angle of Attack Sensor Model 2
α_{s3}	degrees	Angle of Attack Sensor Model 3
V_{IAS}	ft per sec	Indicated Airspeed / Calibrated Airspeed
h	ft	Altitude
V_{TAS}	ft per sec	True Airspeed
α_f	degrees	Fused angle of attack from sensors
V_{GS}	knots	Ground speed
$\dot{h}$	ft per min	Vertical Speed
$\alpha_{r_{FPA}}$	degrees	Angle Of Attack Reference Signal from FPA method
$\alpha_{r_{nz}}$	degrees	Angle Of Attack Reference Signal from Nz method
$\alpha_{r_{\delta e}}$	degrees	Angle Of Attack Reference Signal from δe method

6.2 Control Surface Window

The control surface window is shown in Figure 6.2. The control surface window indicates the position of the control surfaces and shows their motion in real time. The control surface pointers on the left cannot be used to fly the aircraft. They are strictly position indicators. A joystick is required to manipulate the control surfaces directly. On the right side of the screen, there is a longitudinal trim control and a throttle control. These elements can be manipulated by a mouse to control the inputs. The longitudinal trim can also be manipulated by the point of view control that is on most modern joysticks.

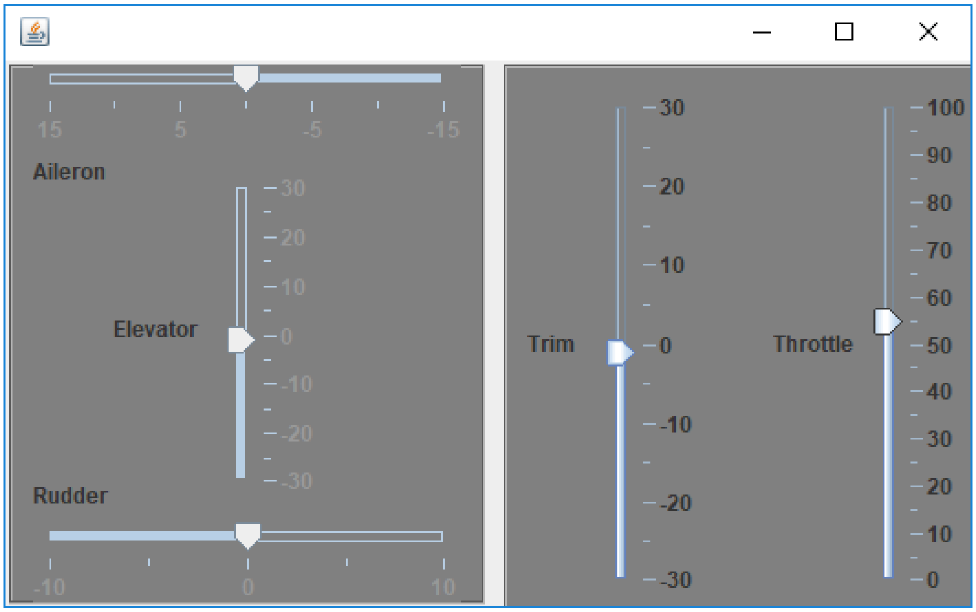


Figure 6.2. The Control Surface window gives a real-time display of control surface and throttle position

6.3 Mode Control Panel

The Mode Control Panel is shown in Figure 6.3. The mode control panel is the primary interface for the autopilot. The autopilot functions similarly to an air-transport system. The white knobs allow the value in the settings window to be changed. Clicking on the right side of the knob increments the value up and the left side increments the value down. Engaging a button enables a particular mode. A green indicator light illuminates when a mode has been activated. For the vertical speed dial, clicking in the uppermost part causes a reduction in vertical speed, and clicking in the lower part causes an increase in vertical speed.

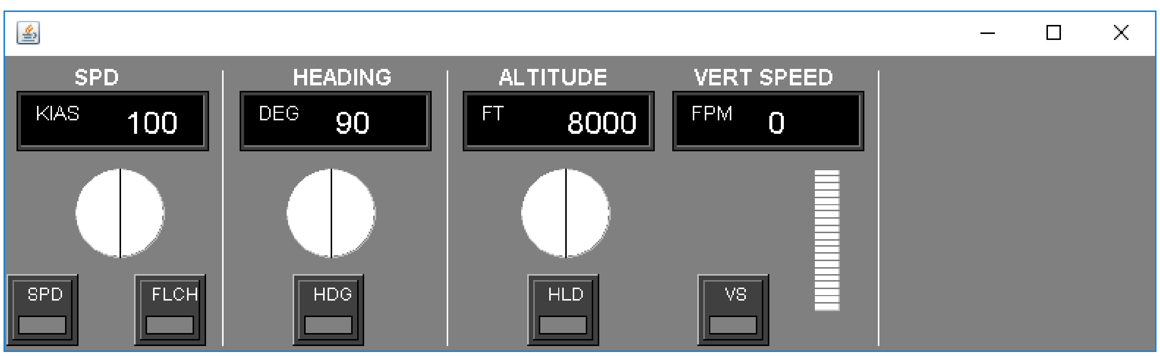


Figure 6.3. The Mode Control Panel controls autopilot functionality

HDG button. The HDG (Heading) button engages the heading capture mode. The aircraft will turn towards the selected heading using a standard rate turn and capture the heading.

HLD button. Selecting the HLD (Hold) button will engage altitude hold, which will command the aircraft to capture the current altitude (not the altitude displayed in the window, which is used for altitude capture purposes).

VS button. Engaging the VS (Vertical Speed) button will command the aircraft to climb or descend at the prescribed vertical speed in the Vertical Speed window. The aircraft will climb/descend towards the altitude displayed in the altitude capture window. At the appropriate time, the autopilot will transition to altitude hold and capture the altitude in the adjacent window.

FLC button. Engaging the FLC (Flight Level Change) button will command the aircraft to climb or descend at the prescribed indicated airspeed in the speed window. The aircraft will climb/descend towards the capture altitude displayed in the altitude capture window. At the appropriate time, the autopilot will transition to altitude hold and capture the altitude in the adjacent window.

SPD button. Selecting the SPD (Speed) button will engage speed capture, which will command the aircraft to capture the speed shown in the speed window.

6.4 Flight Instruments Window

The Flight Instruments window is shown in Figure 6.4. Instrument window includes the Standard-6 instruments as well as a power level indicator and a true airspeed indicator. The indicators give the ability for the operator to view the aircraft behavior from a pilot's perspective. The vertical speed indicator show instantaneous vertical speed rather than the heavily lagged vertical speed that would normally be seen when using a traditional VSI (Vertical Speed Indicator) instrument. The ground-speed function has not been enabled, so the Ground Speed indicator will read true airspeed.

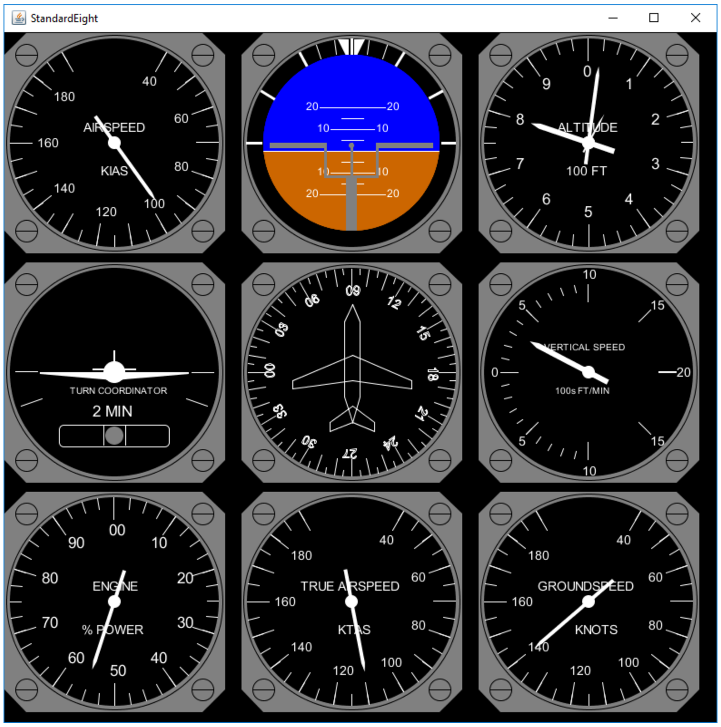


Figure 6.4. Traditional flight instruments provide a pilots view of the aircraft state

6.5 Alpha Sensor Control Panel

For this project, a simplified alpha sensor modeling capability was developed. The Sensor model is based on a 1st order Gauss Markov process which uses truth data as the primary input and then allows the operator to vary the standard deviation of the noise (sigma) and also to add a bias to the signal. The maximum standard deviation is 5.0 degrees (the display does not show the decimal point) and the bias can be (+/- 5.0 degrees). The damping ratio of the Gauss Markov process is fixed at 0.5. Three instances of the sensor models included in the simulation to model the three redundant sensors that are present on most high assurance systems. The Alpha Sensor Control Panel (Figure 6.5) gives the operator control over the behavior of each of the sensors, and can control the bias and noise standard deviation independently. A freeze toggle button also allows the operator to freeze a signal (and to subsequently unfreeze it). The control panel allows the operator to test the functionality of the comparators under various circumstances very quickly.

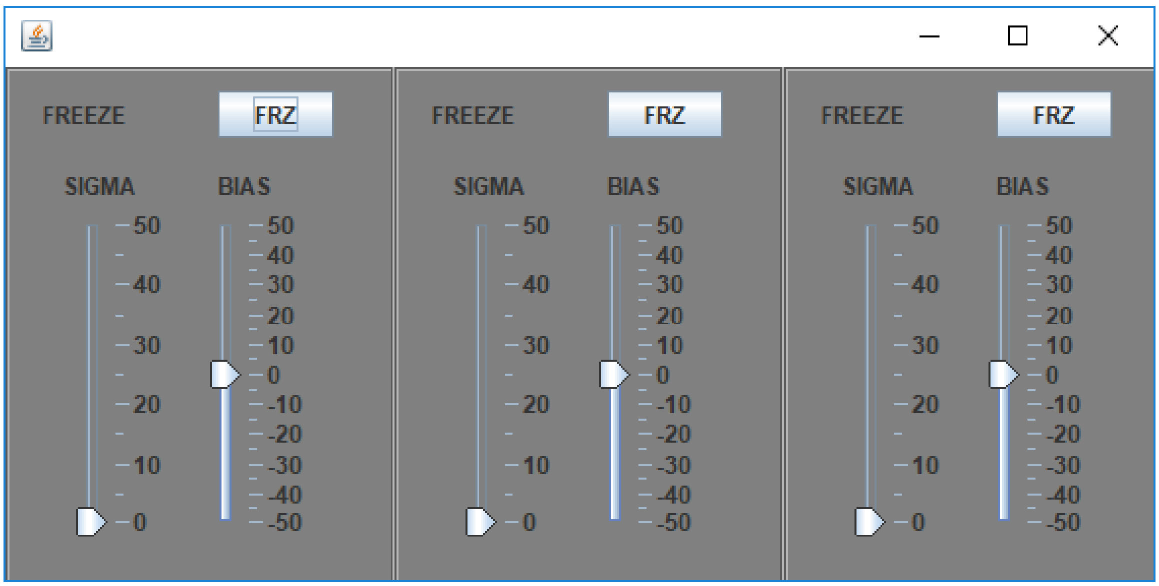


Figure 6.5. The Alpha Noise/Bias window enables the operator to control the error properties of the Alpha Sensors

6.6 Comparator Window

A comparator window (Figure 6.6) was developed to indicate the status of all three simulated alpha-sensors, when compared to the reference values associated with each of the reference comparators. The five reference comparators as discussed in Section 5 are shown in the window. For each comparator, there is an indicator light that illuminates either green when a signal is perceived to be good, amber if the signal is perceived to be bad, or off (grey) if the comparator cannot render a judgement based on the current aircraft state. A final resolver shows the overall best estimate of the alpha sensor status.

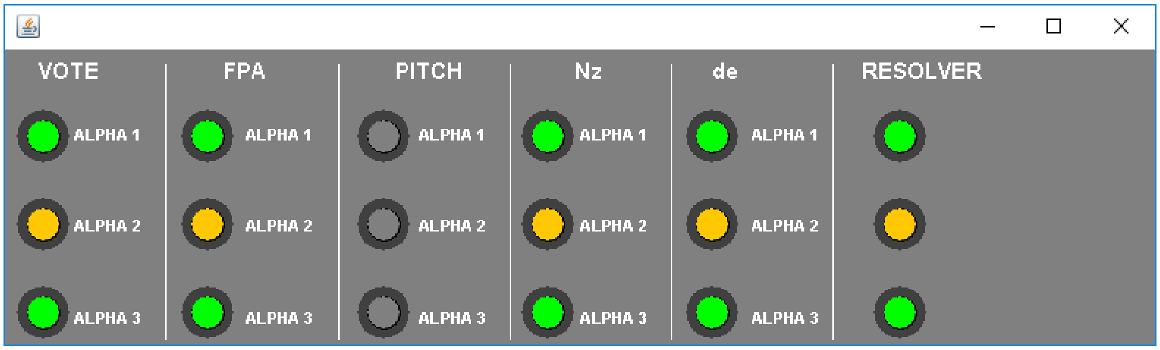


Figure 6.6. The Comparator window shows the status of the Alpha signals

6.7 Stripcharts

The simulation is configured to have three strip-chart windows that enable monitoring of simulation data in real-time. Currently there are two windows with six charts each and one single chart window for comparator signals.

- One strip chart window shows the primary longitudinal states, which are: Elevator Deflection, AOA, N_z, Pitch, Airspeed, and Altitude.
- A second strip chart window shows the three AOA sensor models, a fused version of AOA from the voting algorithm (average of all good sources), and the truth AOA.
- The comparator strip-chart has four angle of attack estimates plotted on the same chart. These lines are as follows: Truth AOA (yellow), FPA estimate (Orange), Elevator Estimate (red), and the N_z Estimate (green).

7 Example Simulation Results

This section provides a summary of the results obtained during the simulation.

- Three charts are shown that indicate the performance of the reference signals.
- Each chart shows elevator deflection followed by the 'truth' AOA.
- Then each of the reference signals are presented on their own respective graphs.
- Finally, a single chart is shown with both truth AOA and the 3 reference signals.

Overall, the simulation results show that the *Nz* solution is the most reliable in all cases. While it may not always be the most numerically precise, it can be expected to reliably track the AOA signal in most all maneuvers. Therefore, the *Nz* solution, notwithstanding the additional data needed to calculate it, may be the best method overall for reliably checking AOA.

Figure 7.1 shows the aircraft in a gentle speed reduction from 100 KIAS to 70 KIAS. In this case all of the reference signals are nearly identical to the actual angle of attack. This plot demonstrates that during gradual maneuvers in non-turning flight, all of the reference signals can be expected to provide good results.

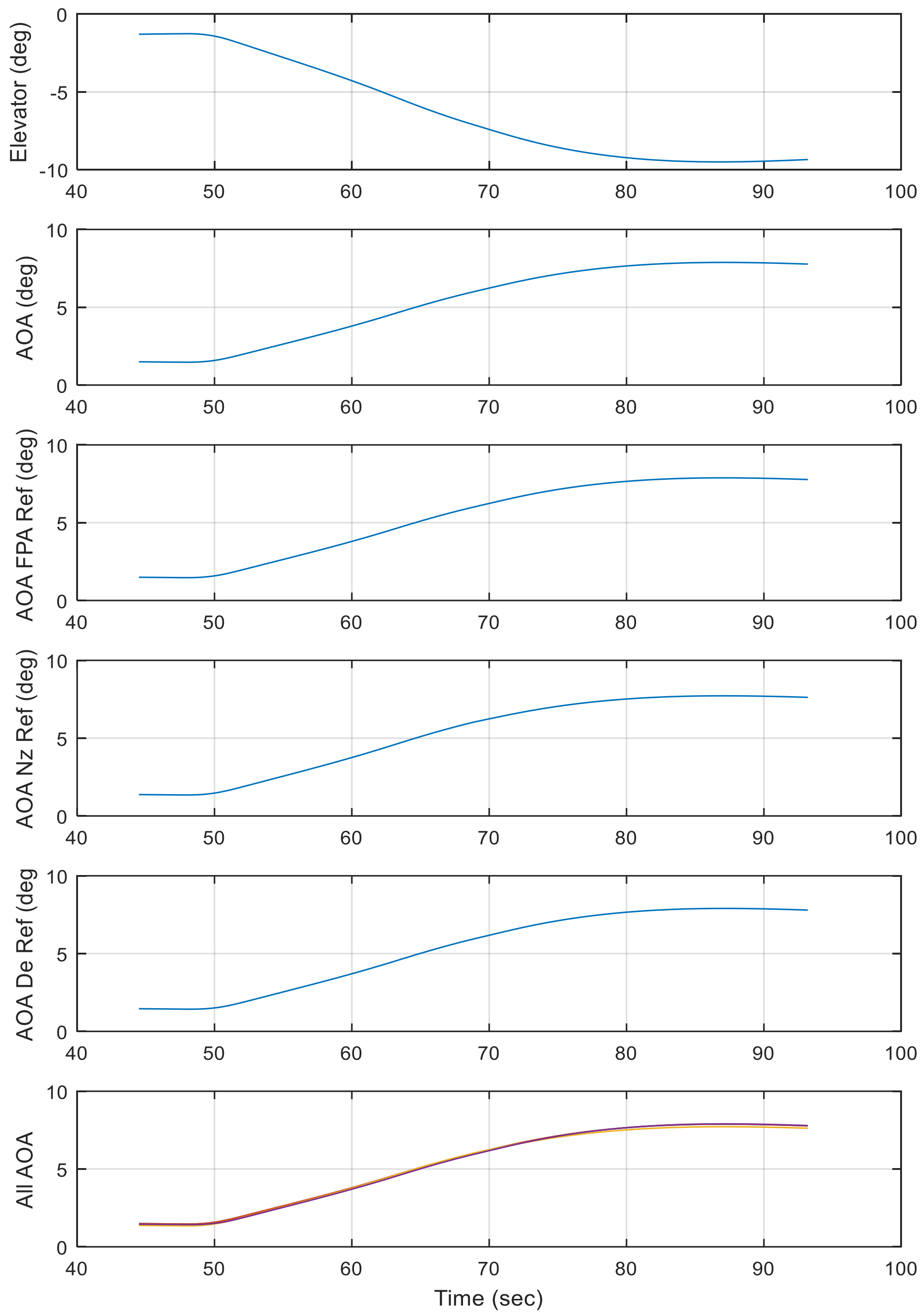


Figure 7.1. AOA Reference signals during a speed reduction

Figure 7.2 shows the aircraft in a 180 degree turn in level flight at 100 KIAS. The aircraft starts in level flight and then establishes a 30 degree bank angle for the turn. At the end of the turn, the aircraft reduces its bank angle and captures the new heading. The maneuver shows the increase in angle of attack needed to maintain level flight in the 30 degree bank. It also shows that during the establishment of the bank angle, the FPA reference solution shows some unsteady behavior. This unsteady behavior is largely reduced when the bank angle stabilizes at 30 deg suggesting that the bank rate has more impact on FPA-reference unreliability than the bank angle itself. It is not clear why this is the case and will require further research to fully understand. The elevator signal shows some bias in the signal more so during the turn than during level flight, but it amounts to only about 0.25 degree. The reason for the bias may be due to the omission of unsteady terms, which may have a greater contribution in the turn. Further study is required to understand the observation. Overall, the Nz reference signal tracks the entire maneuver the most reliably.

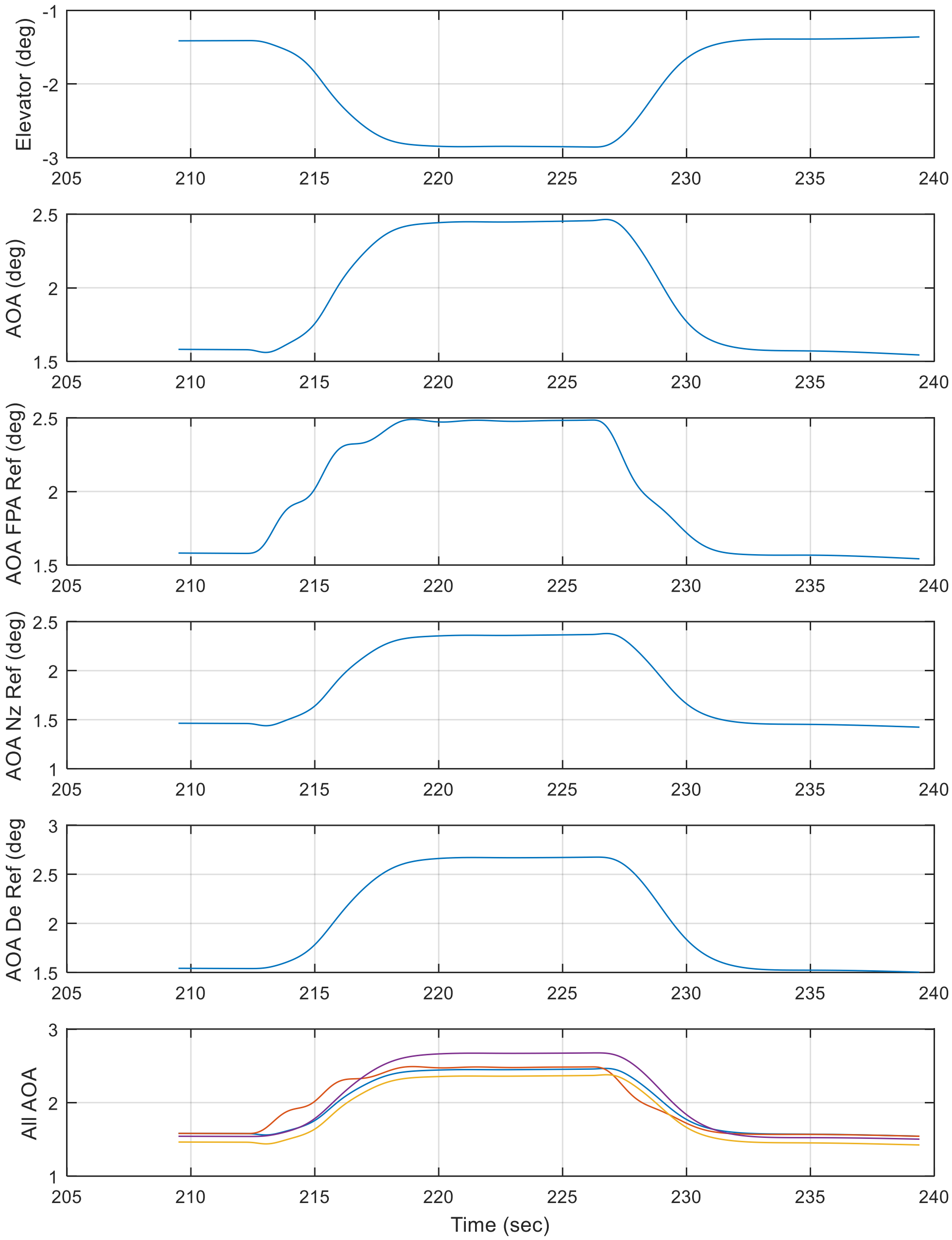


Figure 7.2. AOA Reference signals during a 180 degree turn

Figure 7.3 shows the reference signals subjected to multiple elevator doublets while in level flight. Here all the signals track the actual AOA quite well with the exception of the elevator reference. This is to be expected. Since the elevator reference AOA is just a scaled and lagged elevator signal, the short period dynamics that characterize the de/AOA relationship are missing. Therefore the overshoot that is seen in the other signals in the lower portion of the sinusoids is missing in the elevator signal. It emphasizes that the elevator reference signal can estimate a steady-state value but is not good at capturing the dynamics of higher-frequency maneuvering.

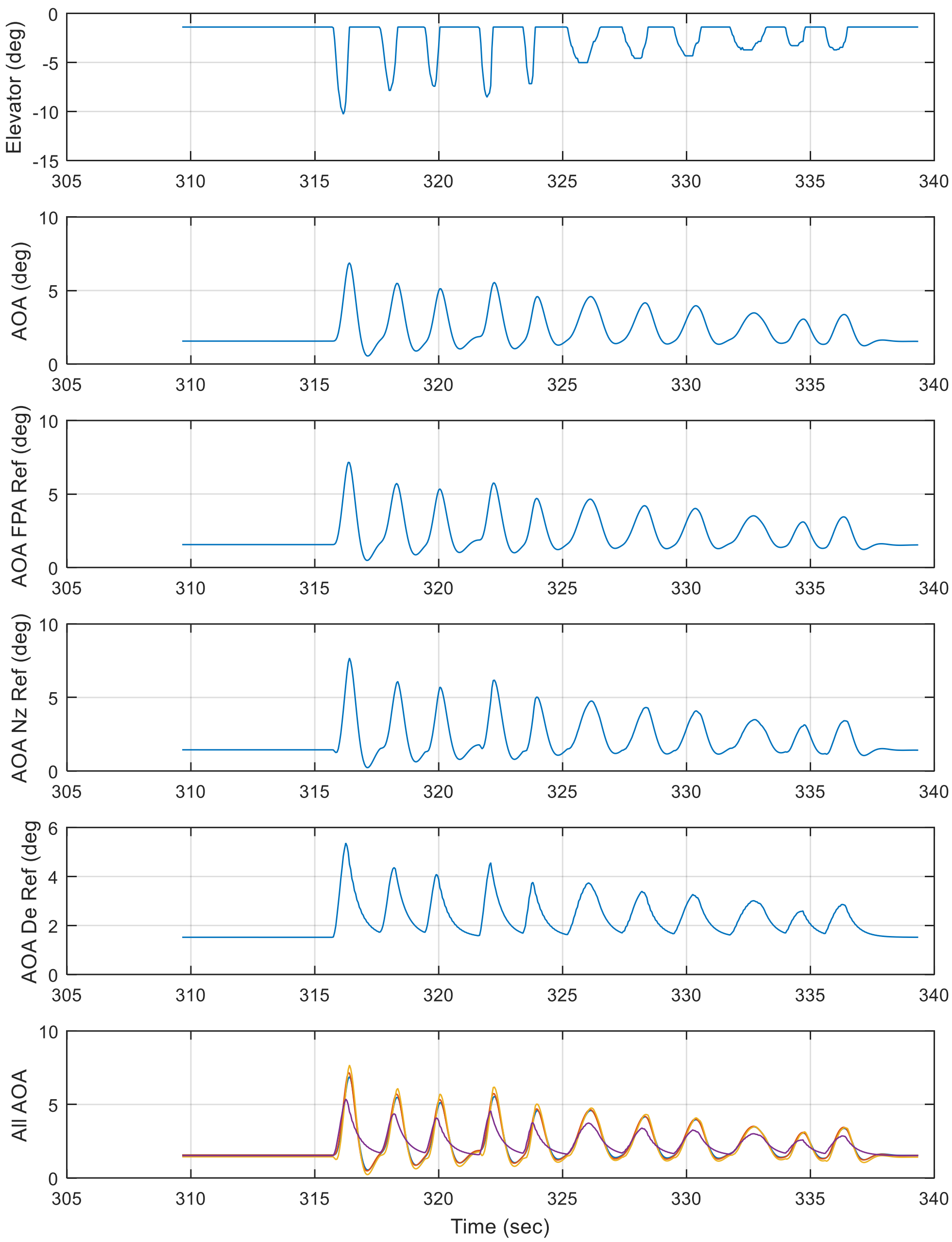


Figure 7.3. AOA Reference signals during elevator doublets

8 Conclusions

In this effort a basic set of reference signals, which can be assembled from aircraft state information, were constructed to demonstrate how AOA can be estimated when the actual sensor AOA data is not reliable. Three signals using independent methods were developed: The FPA reference solution, the *Nz* reference solution, and the elevator reference solution. Overall, the *Nz* reference solution proved the most reliable for capturing both steady-state and higher-frequency maneuvering. The FPA reference solution works well in level flight, but loses precision in turns especially when the bank rate is high. The elevator reference solution gives a reasonable approximation for steady state AOA values, but cannot reproduce the dynamics associated with the short period. Therefore, it is inadequate for any high-frequency operations.

This work is a first step in determining methods for evaluating failed states. This work provides a proof-of-concept demonstration performed with limited resources. The work does demonstrate concept viability and illustrates some basic concepts that must be understood when making such calculations. Further research is recommended in the following areas:

- **Expanded Scope**: The scope of this effort was to focus on angle of attack exclusively, because it is perhaps the most vulnerable sensor measurements to atmospheric contamination or damage. However all aircraft states ideally would have some means of reference based validation. It would be good to expand the scope of the project to include all air data.
- **More Sophisticated Analysis**: The mechanisms presented in this effort represent the simplest methods by which an angle-of-attack reference signal could be created. More sophisticated methods should be considered, such as methods involving some combination of State Observers and Kalman Filtering.
- **Error and Uncertainty Analysis**: With the current analysis, each reference signal has been created using ideal inputs without any errors. A rigorous analysis of available data is needed to determine the likely noise and error properties of input signals, as well as the likely error bounds of the reference signals.
- **Aircraft Performance**: The current analysis was performed on a low subsonic aircraft model. While the basic physics of aircraft dynamics is the same for all fixed wing aircraft, it would be best if the aircraft model more accurately reflected the target aircraft class: the high-speed civil transport aircraft. The performance differences may impact the error analysis and identify specific anomalies that aren’t present with low speed aircraft.

REPORT DOCUMENTATION PAGE			Form Approved OMB No. 0704–0188

1. REPORT DATE *(DD-MM-YYYY)*
01-02-2020

2. REPORT TYPE
Contractor Report

3. DATES COVERED *(From - To)*

4. TITLE AND SUBTITLE
Honeywell Enhancing Airplane State Awareness (EASA) Project: Final Report on Refinement and Evaluation of Candidate Solutions for Airplane System State Awareness

5a. CONTRACT NUMBER

5b. GRANT NUMBER
NNL16AA06B

5c. PROGRAM ELEMENT NUMBER

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER
340428.02.40.07.01

6. AUTHOR(S)
Whitlow, Stephen; Dillard , Michael

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
NASA Langley Research Center
Hampton, Virginia 23681-2199

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
National Aeronautics and Space Administration
Washington, DC 20546-0001

10. SPONSOR/MONITOR'S ACRONYM(S)
NASA

11. SPONSOR/MONITOR'S REPORT NUMBER(S)
NASA/CR-2020-220441

12. DISTRIBUTION/AVAILABILITY STATEMENT
Unclassified-
Subject Category 03
Availability: NASA STI Program (757) 864-9658

13. SUPPLEMENTARY NOTES
Langley Technical Monitor: Steve D. Young

14. ABSTRACT
The loss of pilot airplane state awareness (ASA) has been implicated as a factor in several aviation accidents identified by the Commercial Aviation Safety Team (CAST). These accidents were investigated to identify precursors to the loss of ASA and develop technologies to address the loss of ASA. Based on a gap analysis, two technologies were prototyped and assessed with a formative pilot-in-the-loop evaluation in NASA Langley's full-motion Research Flight Deck. The technologies address: 1) data source anomaly detection in real-time, and 2) intelligent monitoring aids to provide nominal and predictive awareness of situations to be monitored and a mission timeline to visualize events of interest. The evaluation results indicated favorable impressions of both technologies for mitigating the loss of ASA in terms of operational utility, workload, acceptability, complexity, and usability. The team concludes that there is a feasible retrofit solution for improving ASA that would minimize certification risk, integration costs, and training impact.

15. SUBJECT TERMS
Anomaly detection; Angle of attack; Checklist; Flight deck; Prediction

16. SECURITY CLASSIFICATION OF: a. REPORT	b. ABSTRACT	c. THIS PAGE	17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
U	U	U	UU	70	STI Help Desk (email: help@sti.nasa.gov)
					19b. TELEPHONE NUMBER *(Include area code)* (757) 864-9658