Limitations of Commercial Aviation Safety Assessment Standards

Rodrigo Rose

STAMP Workshop 2024
Introduction

• The processes used to assess the safety of commercial aircraft were developed throughout the 20th century and formalized into standards in the 1990s

• Modern commercial aircraft are highly automated and rely on complex interactions between hardware, software and humans

• The Boeing 737 MAX accidents have highlighted that commercial aircraft are not immune to severe design flaws
  – Government agencies, academics and the standards community were aware of this before the accidents
  – Impetus to address these deficiencies before another major accident
Common View of the Limitations in Boeing 737 MAX Safety Assessment

“When all flight deck effects are considered, the introduction of the MCAS function invalidated aircraft-level assumptions for flight crew responses related to erroneous AOA failures under certain conditions”
– Joint Authorities Technical Review Report

“Based on the incorrect assumptions about flight crew response and an incomplete review of associated multiple flight deck effects, MCAS’s reliance on a single sensor was deemed appropriate and met all certification requirements”
– Lion Air 610 Final Report

“Boeing made fundamentally faulty assumptions about critical technologies on the 737 MAX, most notably with MCAS”
– House Committee on Transportation & Infrastructure Report

Most analyses identify the flawed assumptions, but don’t systematically question the safety assessment methods that allowed the assumptions to slip through
Limitations in Safety Assessment Standards

- CAST analysis was performed on JT610 and ET302

- Four main limitations identified:
  - Limited integration of human considerations in the safety assessment process
  - Absence of a systematic methodology that supports the identification of unrecognized assumptions
  - Limited guidance for uncovering and managing non-failure based causal scenarios leading to losses
  - Limited framework for understanding non-linear (e.g., circular or balancing) causal relationships
Limitation 1: Human Considerations

“The safety assessment process described in this document assumes that flight crews, cabin crews, maintenance crews, and other individuals participating in the operation of the aircraft follow documented procedures in foreseeable operating conditions…”
### Limitation 1: Human Considerations

<table>
<thead>
<tr>
<th>Table A-7 - AFHA Format Example</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>ID #</strong></td>
</tr>
<tr>
<td>---------</td>
</tr>
<tr>
<td>4.1.1.T1</td>
</tr>
</tbody>
</table>

- **Assumptions about flight crew response are used to make decisions about severity classifications.**
- **Severity classifications are used to make design decisions.**
Limitation 2: Identification of Assumptions

In traditional safety assessments, assumptions are listed because there is some level of doubt about their validity.

"Assumptions should be captured and formally communicated to the appropriate development information sources. The assumption may then be confirmed, or corrected based on new development information. In the latter case, a design change or a revision of the AFHA may be required."
Limitation 2: Identification of Assumptions

Documented Assumption: Continuous unintended nose down stabilizer trim inputs would be recognized as a Stab Trim or Stab Runaway failure and procedure for Stab Runaway would be followed

---

“Condition: Uncommanded stabilizer trim movement occurs continuously.”

Reality:
• MCAS stabilizer movement not continuous
• MCAS commands bounded by 2.5° authority
• Pilots can counter nose-down movement with manual electric trim inputs
• No MCAS command for 5 seconds after reset

---

Undocumented Assumption: Erroneous MCAS activations always result in “continuous unintended nose down stabilizer trim inputs”
Limitation 3: Capturing Non-Failure Cases

- Developmental factors
  - Unsafe interactions between intended functions/behavior
  - Unsafe combinations of failures and intended behavior

- Non-developmental factors
  - Maintenance error
  - Manufacturing error
  - Operational error
  - etc.

Difficult to obtain meaningful probabilities for
Autoflight malfunction at low altitude which results in an unsafe flight path in an autopilot OFF, single channel or fail passive configuration,

- Autopilot Malfunction in the Pitch Axis at Low Altitude.

  - Erroneous Runaway/oscillatory stab output un-arrested by column cutout

    - AND

      - Undetected stab trim runaway
      - Column Trim Cutout Fails to Interrupt Stab Motion

        - FCC-730 produces undetected erroneous MCAS or Flaps Up/Dn discrete

          - P < 10^{-9} Input failures cause FCC to produce an undetectable erroneous MCAS engage discrete
Misleading Air Data from the Left and Right ADIRU – Airspeed / Altitude.

- Misleading Air Data from the Left ADIRU
- Misleading Air Data from the Right ADIRU

  - Erroneous AOA-L data from the Captain’s side

    - Failure of AOA-L vane / annunciation
    - Incorrect AOA output from the ADIRU-L output

      - OR

        - Erroneous AOA-L Sensor
        - Incorrect AOA output from ADIRU-L output
        - Loss of Power to AOA-L Heater
Limitation 3: Capturing Non-Failure Cases

- ARP4754 seeks to minimize development errors
- ARP4754 requires “Correctness Checks” to be conducted through its recommended “Validation Methods”

“Does the requirement contain errors of fact?”

“Is the requirement verifiable?”

“Is the source of the requirement identified and correct?”

“Are all requirements from safety assessments included? Are all system failure conditions identified and classified correctly? Is the impact of unsafe design or design errors considered?”

- ARP4761’s Common Mode Analysis (CMA) qualitatively considers how aspects like software error, pilot training, or manufacturing defects can invalidate logic in FTA

These processes are not step-by-step methodologies to interrogate and challenge what you think is true about the system.
Limitation 4: Capturing Complex Non-linear Causality

- Non-linear causality often involves behavior that reinforces itself or cancels itself out.

- Capturing non-linear causality requires being able to capture repeated actions, appropriate timing of decisions, sequences of crew and automated actions, etc.
Limitation 4: Capturing Complex Non-linear Behavior

“Operational events should be added to the relevant failure condition statements, creating new combined failure conditions. When considering the combination, it is important to ensure that the operational event is independent from the original failure condition.”
Moving Forward

Can STPA help address some of these gaps?
Thank you!

Questions, Comments, Feedback?

rlrose@mit.edu