PILOT NON-CONFORMANCE TO ALERTING SYSTEM COMMANDS DURING CLOSELY SPACED PARALLEL APPROACHES

Amy R. Pritchett  
Georgia Institute of Technology  
Atlanta, GA

R. John Hansman  
Massachusetts Institute of Technology  
Cambridge, MA

ABSTRACT

Pilot non-conformance to alerting system commands has been noted in general and to a TCAS-like collision avoidance system in a previous experiment. This paper details two experiments studying collision avoidance during closely-spaced parallel approaches in IMC, and specifically examining possible causal factors of, and design solutions to, pilot non-conformance.

INTRODUCTION

Closely Spaced Parallel Approaches

To reduce flight delays and increase airport capacity, methods of enabling closely spaced, independent parallel approaches in Instrument Meteorological Conditions (IMC) are being studied. For runways at least 4300 feet apart, the controller using today's radar can ensure aircraft separation; with the addition of the ground-based Precision Runway Monitor (PRM) this spacing can be reduced. [1] The use of new technologies to reduce this minimum separation would allow airports to effectively maintain their Visual Meteorological Conditions (VMC) capacity in IMC.

The task of ensuring adequate aircraft separation during parallel approach operations is problematic. The aircraft are closer together than during any other airborne phase of flight, severely limiting the warning time should one aircraft blunder into the other's approach path. New technologies such as the Cockpit Display of Traffic Information (CDTI), collision alerting systems and the crosslink of aircraft information may enable pilots, with cockpit systems, to maintain aircraft separation during approaches at close spacing.

Previous Flight Simulator Study Results

A previous simulator experiment had active airline pilots fly parallel approaches using the MIT part-task Advanced Cockpit Simulator to study the pilot effectiveness in avoiding encroaching traffic, both with and without the aid of an alerting system. While TCAS II may not be sufficient for closely-spaced parallel approaches [2], as an operational system it was available for use in this preliminary experiment as a baseline. The logic was modified to maintain a consistent threshold throughout the approach. [3] [4]

During each approach, the subject's response was recorded after traffic on a parallel approach blundered towards them. The study also examined several cockpit traffic displays.

The first measurement of interest is the resulting miss distance between aircraft. Overall, the intruder and subject aircraft came within 500 feet of each other 4% of the time, and within 1000 feet of each other 20% of the time. These percentages were found to be significantly lower when the approach was flown on autopilot and significantly higher when TCAS avoidance maneuvers were not displayed. These percentages are highly scenario-dependent and may not be indicative of pilot collision avoidance performance in all situations.

The presentation of the TCAS alerts and avoidance maneuvers correlated with a significant improvement in aircraft miss distance. However, the TCAS assumes that the pilot will react within five seconds, and then match or exceed the TCAS pitch command. [5] However, examination of the trajectories found the actual maneuvers flown by the pilots, when the TCAS maneuvers were shown, did not conform (i.e. meet the minimum commanded vertical speed) in 40% of the cases. As shown in Figure 1, an increased collision rate occurred when pilots did not follow the commanded maneuver.

Figure 1. Frequency of Incidents by Presentation of Alerts & Avoidance Maneuvers, and Pilot Conformance
No single causal factor of the low conformance rate can be isolated. Pilot reaction time alone does not show a strong effect. 66% of the pilots reacted within the five second allowance assumed by the TCAS system, and of these only 61% matched the displayed TCAS maneuver. Of the pilots who acted shortly before the alert or after the five second allowance (13% and 20% respectively), a significant number of pilots still matched what the TCAS guidance commanded (71% and 33% respectively).

Conformance to the (vertical) TCAS maneuver may be affected by the turning maneuvers that the pilots often performed at the same time. Overall, pilots did not turn in 32% of the approaches (i.e. the maximum bank angle after the alert was less than five degrees); 34% of the time the pilots turned away from the intruder, 11% of the time pilots turned toward the intruder, and 23% of the time pilots turned one way and then another. Pilots who did not follow the TCAS maneuver turned away significantly more often than those who followed the TCAS maneuver. This may suggest that the pilots, by executing a turn, felt a vertical maneuver was no longer required.

Other possible factors for the low conformance rate have also been investigated. Examining the aircraft trajectories for the approaches where the pilots were not shown any TCAS alerts or maneuver guidance, the pilots' reactions only satisfied what the TCAS would have commanded in 25% of the approaches, suggesting that the TCAS maneuver is not what the pilots would do instinctively. Further analysis suggested that these problems may stem from the pilots' use of a different and less effective alerting algorithm for deciding when to generate alerts. The across-track deviation of the intruding aircraft appears to have been a major determinant in the decision to react, a conclusion also supported by pilot comments about their decisions to alert.

However, this type of alert generation logic -- based on intruder lateral deviation -- has been shown to be ineffective: it can generate a false alarm when the parallel traffic oscillate around their localizer during a normal approach, and it may not trigger an alert until the intruding aircraft has already established a high rate of convergence.

The enhanced displays tested in the previous experiment provided pilots with a fiducial marker indicating the cross-track position of a normal approach. All pilots indicated they liked this feature; some commented that it freed them from monitoring the convergence rate of the other aircraft. However, when presented with the enhanced traffic displays, pilots conformed significantly less often than when they were given the current TCAS II type traffic display. This may also suggest that pilots, given a more explicit traffic picture, may have felt a vertical maneuver was no longer required. This perception may have been erroneous, however, as more near-misses happened with these new displays. This may also indicate the display features providing a fiducial marker of the other approach path may have unintentionally encouraged a range-only or cross-track-only alerting logic.

Alerting & Avoidance Strategies During Parallel Approach

In order to avoid a potential collision hazard, two sequential steps must be executed. First, an alert must be triggered leaving sufficient time for an avoidance maneuver to be effective. Second, an avoidance maneuver must be selected executed. The minimum standards for these two steps are highly coupled; an avoidance maneuver which is more severe may allow for an alert later than that acceptable for a gentle avoidance maneuver.

For the task of parallel approaches, the alerting decision may be based on several different alert criteria. TCAS II alert decisions are based on a time-to-collision criteria, as projected from the ratio of measured range to range-rate. PRM uses a Non-Transgression Zone (NTZ) between the approach courses as an alerting criteria; should either aircraft enter the NTZ, an alert is given.

Other alerting criteria have also been examined. For example, the 'MIT Criteria' is based on contours of probability of collision. These contours were determined through numerical simulations which accounted for variability in sensors and which varied parameters of the intruding aircraft such as relative position, heading, bank and speed. One set of contours are shown in Figure 2; the contours change shape with relative heading, speed and bank. The relative position coordinates are drawn centered on the "own" aircraft. If an avoidance maneuver is triggered when an intruding aircraft enters the $p=0.001$ contour, then the probability of collision is at or less than $p=0.001$.  

![Figure 2. Probability Contours of Collision](image-url)
Each of these alerting criteria has benefits. Comparing the performance of the NTZ and MIT criteria, the MIT criteria is better at preventing both false alarms and late alerts (alerts given too late for an effective avoidance maneuver). These benefits are illustrated in Figure 3, with the 'own aircraft' at the origin. Intruder aircraft '1' will pass in front of the own aircraft; the NTZ criterion is still triggered, while the MIT criterion is not. Conversely, intruder aircraft '2' has a high closure rate; the MIT criteria triggers earlier, while the NTZ criteria, for closer runway spacings, risks generating a late alert.

However, the NTZ criteria has the benefit of being easy to visualize and monitor. The size and shape of the MIT alert contours, on the other hand, vary significantly with changes in the intruding aircraft's heading, speed and bank. This may make the MIT criteria difficult to explain and/or display to the pilots expected to act upon its alert.

Pilot Non-Conformance to Alerting System Commands

The assumption that pilots will conform to alerting system commands is not always valid. Studies of currently operational alerting systems have identified non-conformance situations where pilots have delayed in responding to automatic alerts, or have executed different resolutions to the hazard than commanded by the automatic system. For example, pilot questionnaires on the use of TCAS II reported pilots intentionally did not follow commanded avoidance maneuvers in 24.7% of the cases where alerts and commands were given. [7] A similar tendency was noted with the Ground Proximity Warning System (GPWS). [8]

However, the factors involved with non-conformance are not well understood. In addition to the potential benefits towards increasing airport capacity, the task of collision avoidance during closely spaced parallel approaches provides a useful case study into pilot non-conformance because of the measurable discrepancy between the types of alerting criteria consistent with subject reactions and the higher performance alerting criteria designed for alerting systems for this phase of flight.

Overview

In order to examine pilot non-conformance during closely-spaced parallel approaches, two simulator experiments were conducted. The first examined the alerting and avoidance maneuver decisions made by subjects with reference to various CDTI features in the absence of an alerting system. The second experiment examined possible design considerations for increasing pilot conformance to higher-performance alerting criteria which may not match their normal alerting decisions. These two experiments will be discussed, followed by a general discussion.

SIMULATOR STUDY OF ALERTING CRITERIA AND AVOIDANCE MANEUVERS PREFERRED BY SUBJECTS IN THE ABSENCE OF AN ALERTING SYSTEM

Based on the results from the preliminary experiment, it was hypothesized that the traffic display features can, and should, support a more sophisticated mental model for pilots to use in generating alerts and selecting avoidance maneuvers. This should provide for better pilot confidence in, and following of, automatically displayed avoidance maneuvers (when available), and reduce erroneous pilot reactions. To test this hypothesis, a follow-on flight simulator experiment was conducted. This experiment had the following two objectives: 1) provide a preliminary study of how the display features of a cockpit traffic display affect a person's mental 'alert generation logic', used to assess when an avoidance maneuver is necessary and what the avoidance maneuver should be, and 2) ascertain how display features affect a user's ability to detect a conflict. An alerting system was not used.

The experiment runs each consisted of three sequential parts:

- The Flight The subjects were told they were flying an approach, and should press a red button on the sidestick as soon as they thought the aircraft on the parallel approach was blundering towards them, as evidenced by the traffic display.
- The Maneuver Selection Once the subject indicated the parallel approach traffic was deviating towards them, the traffic display was
blanked and six possible maneuvers were graphically shown to the subjects. The subjects were asked to select the maneuver considered best for maintaining inter-aircraft separation.

- **Numerical Simulation** The simulator then predicted the miss distance resulting from the selected avoidance maneuvers, providing a first order measurement of the subjects' decision making.

The simulator used a Silicon Graphics Indigo 2 workstation for the displays and aircraft dynamics computations. A sidestick was connected for the flying task. The aircraft dynamics used point-mass calculations with performance constraints representative of air transport aircraft. The pitch and heading acquisition models used a critically damped controller, while the localizer acquisition controllers were slightly under damped, modeling the actual wavering about the approach path of the aircraft.

In total, nineteen subjects flew the experiment. The basic characteristics of the subjects varied widely. Two were airline flight crew, four were Certified Flight Instructors (CFI) in general aviation aircraft (one with jet fighter experience), two held Private Pilot Licenses, and the remaining eleven were students without piloting experience.

Five displays were tested. All were based on a moving map type display, with a top-down view, heading-up orientation, iconic presentation of the other aircraft's positions and a text presentation of the other aircraft's altitudes. Traffic information was updated once per second, a technically feasible rate with current datalink systems. These five displays were:

- **Baseline Display**: emulated the current TCAS display.
- **Fiducial Mark Display**: added the reference indication of the parallel approach path, emulating the enhanced EHSI display tested in the preliminary experiment.
- **Heading Display**: added a graphic indication of the other aircraft’s heading.
- **Noisy Projection Display**: added a graphic indication of heading rate and projected position for the next 15 seconds; the projection was based on the noisy measurement of the other aircraft’s bank that sensors can provide.
- **Smooth Projection Display**: added a graphic indication of heading rate and projected position within the next 15 seconds; the position projection used theoretical knowledge of heading rate to give a more smooth projection.

Subject workload was also varied to test its effect on subject's decisions. The subjects were told their primary task was to keep their wings level despite turbulence, using a side-stick. To do this, bank angle was shown on an artificial horizon drawn approximately three inches away from the edge of the traffic display. The turbulence was set to two different levels, generating two different levels of workload. The subjects were not briefed on these qualities.

Four scenarios were flown, in random order, within each test block. These scenarios were designed to represent a variety of collision trajectories, with high and low convergence rates. One of the four was not hazardous; in another scenario, the ‘other’ aircraft never varied from its approach path.

The complete test matrix was three dimensional, with five displays, four types of scenarios and two workload levels being varied. Most subjects had 40 experiment runs, fully combining all types of displays, workload levels and scenarios, allowing for within-subject comparisons; four subjects did not have runs with the smooth predictor display. The scenarios were flown in blocks of four; each included all runs for each display-workload combination.

The collision avoidance system available in the preliminary experiment, TCAS II, uses convergence rate to estimate time remaining to collision as a basis for generating an alert. The subjects' reactions, in contrast, did not have a consistent time to point of closest approach at their reactions, as shown in Figure 4. The time to point of closest approach ranged from -13.39 seconds (the subject reacted after the point of closest approach) to 34.32 seconds, with a mean of 14.37 seconds. The wide spread suggests the subjects' alerting criteria does not take into account convergence rate, differing from the alerting criteria used by TCAS.

The subjects' reactions were instead consistent with a criteria based on range or lateral separation. The distribution of the lateral separation between the
aircraft at the time of the reaction is shown in Figure 5. A Chi-Squared goodness-of-fit test found its distribution approximates a normal distribution with a high probability (p > 99%). The mean lateral separation at the time of the reaction is 1346 feet, with a standard deviation of 345 feet. These statistics were similar for both high and low convergence rate scenarios. For comparison, in the high convergence rate blunders, the aircraft lateral separation could decrease 200 feet between every one second update of information about the other aircraft. Therefore, the variance of this distribution is comparable to that expected from a standard deviation of 1.75 seconds in reaction time around an alerting criteria based purely on lateral separation.

Although the newer displays were purposefully designed to give indications of relative convergence rate and trend before an abnormal lateral position was reached by the intruder, no differences can be found in the method used by the subjects to generate alerts with each of the different displays.

The largest determinant of predicted collision avoidance performance was the convergence rate of the intruding aircraft. In scenarios with a high convergence rate, subject's reactions were too late for any of the six simulated avoidance maneuvers to be effective in 42% of the cases, highlighting the need for a collision avoidance system or for subjects to use a more effective alerting strategy.

In addition to the timing and validity of the subject's alerting decisions, the performance of the subjects in selecting a safe direction of flight for an avoidance maneuver was measured. The most popular maneuvers were Turn Away and Climb (55%), and Turn Away while maintaining altitude (36%). Each maneuver appeared to be selected the same amount, regardless of display. However, these maneuvers were not always effective.

This experiment suggests that subjects may disagree with the alerts and avoidance maneuvers made by more efficient alerting systems, such as the TCAS II system used in the preliminary experiment or alerting systems developed specifically for parallel approach such as the 'MIT' logic discussed in the previous section. Pilot conformance has been demonstrated to have a significant effect on pilot performance at collision avoidance. Therefore, methods of encouraging pilot conformance may require explicit consideration in the design, evaluation and implementation of collision avoidance systems.

EXPERIMENTAL EVALUATION OF PILOT CONFORMANCE TO COLLISION AVOIDANCE COMMANDS

This experiment served as a preliminary investigation of methods of promoting conformance to alerting system commands through the explicit display of the criteria underlying automatically generated alerts. Two criteria were tested: the Non-Transgression Zone (NTZ) criteria consistent with subject reactions in the previous experiment -- but with a low performance -- and a higher performance MIT criteria intended for use in alerting systems. In some cases, alert criteria were explicitly displayed to the pilot which supported the timing of automatically generated alerts, creating consonance between the display and the alerting system. In other cases, the explicitly displayed alert criteria contradicted the timing of the automatically generated alerts, creating dissonance between the alerting system and displays. For comparison, baseline conditions with no automatic alerts and/or no display of alert criteria were also tested.

Each run consisted of three sequential parts:

1. **The Flight** The subjects were told they were flying an approach, and should indicate when the aircraft on a parallel approach was blundering towards them, as evidenced by a traffic display. In some cases, automatic alerts were given. Subjects were asked to use their best judgment; conformance to the automatic alerts was not mandated.

2. **Certainty and Timeliness Ratings** The traffic display was blanked and subjects were asked to rate their certainty in their decision and, if an automatic alert had been given, the timeliness of the automatic alert.

3. **Numerical Simulation of Avoidance Maneuvers** The simulator then projected the resulting miss distance between the intruder and of the subject aircraft resulting from avoidance maneuvers triggered by the subject's reaction, by the NTZ alert criteria, and by the MIT alert criteria. These numerical simulations were transparent to the subject.
The simulator used a Silicon Graphics Indigo 2 workstation for the displays and aircraft dynamics computations. A sidestick was connected for the flying task, and a mouse for the avoidance maneuver selection. Subjects controlled their progress, selecting further practice or commencement of the experiment runs.

In total, twelve subjects participated. Three held Certified Flight Instructor (CFI) ratings; six had some flight experience, and the remaining three were students without flight experience. No subjects were airline flight crew.

Three displays were tested. All were based on a moving map display, with a top-down view, track-up orientation, iconic presentation of the other aircraft’s positions, and a text presentation of the other aircraft’s altitude. All features of the traffic display were updated once per second, an update rate feasible with current technology.

- **Baseline Display**: Emulated the current TCAS CDTI, with an additional indication of the other aircraft’s heading, as shown in Figure 6.
- **NTZ Criteria Display**: Added a graphic indication of a Non-Transgression Zone between the approaches, as shown in Figure 7. This criteria is consistent with subjects’ reactions in previous experiments.
- **MIT Criteria Display**: Added a graphic indication of the alert criteria used by the prototype MIT alerting logic to the baseline display, as shown in Figure 8. The shape of this alert criteria changes with each update of information about the other aircraft, making it a potentially distracting feature.

Three different automatic alerting conditions were used in the experiment:
- No automatic alerts were given to the subjects.
- Automatic alerts based on an NTZ criteria were given. This underlying criteria was the same as that shown explicitly on the NTZ Alert Criteria display.
- Automatic alerts based on the MIT prototype alerting logic were given. This underlying criteria was the same as that shown explicitly on the MIT Alert Criteria display.

Four scenarios were flown, in random order, within each test block. These scenarios were designed to test a variety of conditions. Half of the time, the NTZ criteria would generate a false alarm or trigger before the MIT criteria; in the other half of the cases the MIT criteria would trigger before the NTZ criteria.

The test matrix for this experiment was three dimensional, testing all combinations of displays, alerts and traffic conflict scenarios. Altogether, subjects completed 36 experiment runs, allowing for within-subject comparisons. The scenarios were flown in 9 blocks of four, where each block included all the runs for each particular display-workload combination. Paired-comparison statistical tests were used to analyze differences between conditions.
When no automatic alerts were given, the subject's reactions appeared to be strongly correlated with criteria shown explicitly on the display, as measured by the time difference between the subject's reactions and when each of the alert criteria would have triggered. The mean values of these differences are shown in Figure 9. The average difference between the subject's response time and the time the NTZ criteria triggered is significantly different when the NTZ criteria is shown compared to when the baseline display is shown (p < 0.01). A similar effect is found for the MIT alert criteria, with a statistically significant difference between subject's reactions with the baseline display available and with the display of the MIT criteria (p < 0.05).

Combined display and automatic alert effects were also found. In general, consonance between the criteria on the display and the criteria used to generate the automatic alert reduced the difference in time between the subject's reactions and the time when each type of automatic alerts were given, as shown in Figure 10. In other words, responses to automatic alerts based on the higher-performance MIT criteria were the quickest when the MIT criteria was explicitly shown on the display. In contrast, subjects' reactions varied the most from the time of the MIT criteria-based automatic alerts when the dissonant NTZ criteria was explicitly displayed. However, because subjects' reactions to automatic alerts based on the MIT criteria were variable, statistical significance of these trends can not be proven. Subjects' reactions were significantly closer to automatic alerts based on the NTZ criteria when either alert criteria was explicitly shown on the traffic display. The mean difference in time between the subject's reactions and the time of NTZ-based automatic alerts drops significantly from the runs with the baseline display (p < 0.01).

Several statistically significant effects were noted between cases with no automatic alerts and with each of the types of alerts. These effects correlate with subjects' responses to "How did the (automatic) alerts affect your decisions?" These responses and measures indicate a tendency for the decision-making process to be affected in three ways:

- The automatic alerts may have been used as additional input to the subjects' reasoning.
- The automatic alerts may have served as a cue for the subjects to evaluate the situation.
- The automatic alerts may have given the subjects greater trust in their reactions when they coincided.

This results provide insight into the relative effects of automatic alerts and the explicit display of alert criteria, and highlight the importance of consonance between the displays and the automatic alerts. Practical considerations for the task of closely spaced parallel approaches require further study, however. Although benefits were found with the display of the MIT criteria, it did not completely meet the ultimate objective of enabling the subjects to consistently use strategies good enough to ensure collision avoidance. In addition, the display of the MIT criteria -- or a similar criteria -- may not be the final or best display to provide to pilots because of its potentially distracting changes in shape and size on a critical area of the traffic display.

**SUMMARY**

The first experiment found subjects' reactions to alerting system commands were consistent with an NTZ-type alerting criteria. Subjects also selected turn-away type maneuvers 91% of the time. These decisions can vary from the automatic alerts given during parallel approaches. Displaying rate did not noticeably affect subjects' reactions.

The second experiment identified a strong effect of the display of alert criteria on subjects' reactions. When the display was consonant with the automatic alert, subjects' reactions were the closest to the automatic alerts. Conversely, dissonant display information from automatic alerts generated by the MIT criteria created sizeable differences between subject reactions and automatic alerts.
DISCUSSION

These results raise broader issues about pilot interaction with executive alerting systems. Alerting systems with executive roles are designed with the implicit assumption that pilots will execute the commands quickly and precisely. In cases of non-conformance, pilots instead elect to examine the situation, and execute a resolution to the hazard which may not resemble the commands. The pilots may also consider information not used by the alerting system. Their actions effectively change the role of the alerting system and of the pilots. In doing so, the anticipated benefits of the alerting system may not be fully realized, and alerting system outputs to the pilots are treated as information sources instead as executive commands.

The frequency with which pilots perform this re-evaluation may be higher than the measured non-conformance rate. When pilots follow commands, it is unknown whether they trust them completely, or take on the extra workload of independent analysis, and then accept the commands.

Two factors may contribute to pilot non-conformance: pilots may perceive a need to confirm the alerting system’s commands, and then they may disagree with the commands and not conform.

The pilots’ perceived need to confirm the commands may involve several factors, including:
• The pilot may be concerned that the alerting system will fail to act as it should, both by giving false alarms and by failing to give timely alerts.
• The pilot may feel the alerting system can not consider all information or has other objectives.
• The pilots may place greater confidence in their own decisions than in the alerting system’s.

If the pilots do not have confidence in the alerting system, they may attempt to confirm its alerts and commands. This confirmation process alone can cause a delay in the pilots’ responses. If the pilots’ assessments do not agree with the alerting system’s commands, they may additionally execute different resolutions to the hazard. A resulting mismatch between pilot decisions and alerting system commands may contribute to non-conformance.

In addition, involvement of pilots in the decision making removes the ability to analyze the system behavior with the same degree of certainty. This variability may limit the extent to which the performance of the combined pilot-alerting system can be predicted during design and certification.

It may be possible to encourage informed decisions by the pilots by incorporating consonance between the pilot’s displays and the alerting system’s commands. In situations where the commands are valid, this method promotes pilot conformance, while maintaining the benefits of a pilot in the loop when pilots have better reasoning.

While these experiment results suggest that alerting system consonance may improve conformance, it requires that the underlying logic of the alerting system be communicable in a quickly understood form. Alerting systems are being proposed for operations which are very complex, require very specific types of performance, or involve many computations. In these situations rigorously encouraging pilot conformance through display consonance may not be possible because of the complexity of the alerting system’s functionality. Such cases may represent a limit on the use of alerting systems.

ACKNOWLEDGMENTS

This work was supported by the National Aeronautics and Space Administration/ Ames Research Center under grant NAG 2-716. The authors would like to thank the students, flight instructors and airline pilots who donated time to fly the simulator as subjects.

REFERENCES