

Piloted “Well Clear” Performance Evaluation of Detect-and-Avoid Systems with Suggestive Guidance

Eric Mueller,¹ Confesor Santiago²
NASA Ames Research Center, Moffett Field, CA, 94035

Spencer Watza³
Western Michigan University, Kalamazoo, MI, 49008

I. Introduction

REGULATIONS to establish operational and performance requirements for unmanned aircraft systems (UAS) are being developed by a consortium of government, industry and academic institutions (RTCA 2013). Those requirements will apply to the new detect-and-avoid (DAA) systems and other equipment necessary to integrate UAS with the United States’ National Airspace System (NAS) and will be determined according to their contribution to the overall safety case. That safety case requires demonstration that DAA-equipped UAS collectively operating in the NAS meet an airspace safety threshold (AST). Several key gaps must be closed in order to link equipment requirements to an airspace safety case. Foremost among these is calculation of the system’s “risk ratio”—the degree to which a particular system mitigates violation of an aircraft separation standard (Federal Aviation Administration 2013). The risk ratio of a DAA system, in combination with risk ratios of other collision mitigation mechanisms, will determine the overall safety of the airspace measured in terms of the number of collisions per flight hour. It is not known what the effectiveness is of a pilot-in-the-loop DAA system or even what parameters of the DAA system most improve the pilot’s ability to maintain separation. The relationship between the DAA system design and the overall effectiveness of the DAA system that includes the pilot, expressed as a risk ratio, must be determined before DAA operational and performance requirements can be finalized.

Significant research effort has been devoted to integrating UAS into non-segregated airspace (Dalamagkidis 2012; Ostwald 2007; Gillian 2012; Hesselink 2011; Santiago and Mueller 2015; Rorie, Fern, and Shiveley 2015, 2016). Several traffic displays intended for use as part of a DAA system have gone through human-in-the-loop simulation and flight testing. Most of these evaluations were part of development programs to produce a deployable system, so it is unclear how to generalize particular aspects of those designs to general requirements for future traffic displays (Calhoun 2014). Other displays have undergone testing to collect data that may generalize to new displays, but they have not been evaluated in the context of the development of an overall safety case for UAS equipped with DAA systems in the NAS (Bell 2012). Other research efforts focus on DAA surveillance performance and separation standards. Together with this work, they are expected to facilitate validation of the airspace safety case (Park 2014; Johnson 2015).

The contribution of the present work is to quantify the effectiveness of the pilot-automation system to remain well clear as a function of display features and surveillance sensor error. This quantification will help enable selection of a minimum set of DAA design features that meets the AST, a set that may not be unique for all UAS platforms. A second objective is to collect and analyze pilot performance parameters that will improve the modeling of overall DAA system performance in non-human-in-the-loop simulations. Simulating the DAA-equipped UAS in such batch experiments will allow investigation of a much larger number of encounters than is possible in human simulations. This capability is necessary to demonstrate that a particular set of DAA requirements meets the AST under all foreseeable operational conditions. Moreover, results related to the performance of the pilots’ use of displays and the time they needed to carry out different aspects of this task may be found in two companion papers (Rorie, Fern, and Shiveley 2015, 2016). This paper reports on a simulation that follows on work presented in Santiago’s study (2015) and builds directly on that study’s results.

¹ Research Aerospace Engineer, NASA Ames Research Center, M/S 210-10, Moffett Field, CA 94035

² Research Aerospace Engineer, NASA Ames Research Center

³ Student, Western Michigan University, Kalamazoo, MI, 49008

II. Methodology

A. Participants

The pilot subjects in this study were all active duty U.S. military pilots. They averaged 1100 hours of UAS flight time, with approximately 30 of those hours in domestic controlled airspace. Pilots did not have previous experience using a traffic display to conduct DAA operations in an air traffic control (ATC) environment. Two active duty air traffic controllers participated in the study as confederates, relaying standard clearances to the pilots and approving their requests for DAA maneuvers. For additional details regarding pilot training procedures, instructions for the DAA task, and metrics related to pilot response time, see Rorie, Fern, and Shiveley 2016.

B. Simulation Environment






The airspace simulator used to coordinate the subsystems used in this experiment was based on the Multi Aircraft Control System (MACS) (Prevot 2002). That system simulated all aircraft other than the UAS under study, provided a pilot interface for “pseudo-pilot” confederates to control the non-UAS aircraft, and provided an ATC interface that controller confederates used to respond to UAS pilot commands and direct overall airspace operations. An additional ATC confederate would take control of intruder aircraft in MACS according to an experiment script in order to create predicted losses of well clear (LoWC) with the subject UAS.

The UAS subject pilots controlled their aircraft using the vigilant spirit control station (VSCS), which was developed by the Air Force Research Laboratory (AFRL) (Feitshans et al. 2008). In addition to standard UAS interfaces, the pilots were provided with a tactical situation display (TSD) that was their primary interface to mission-specific information (e.g., the flight plan) and which contained all elements of the DAA human interface. For additional information on the pilot interface and airspace design see Rorie, Fern, and Shiveley 2016.

C. Experiment Design

The pilots in the experiment flew four separate missions, each time with a different combination of DAA display and algorithm elements. These DAA system designs built upon the findings of previous studies (Santiago 2015; Rorie, Fern, and Shiveley 2015) and incorporated enhancements suggested by past participants. An “information only” display provided only basic information about each intruder, including relative altitude, bearing and range, along with alerts about intruders predicted to lose well clear. The alerting scheme used in this study, which also builds on those previous studies, is shown in Table 1. A detailed description of the well clear definition and its parameters is given in Cook and Brooks (2015), and the values used in this study are given in Table 3. In the information-only display configuration no maneuver guidance was provided to the pilots.

Table 1. DAA alert definitions

ALERT LEVEL	SEPARATION CRITERIA	TIME UNTIL LOSS OF WELL CLEAR	ICON	AURAL ALERT VERBIAGE
DAA WARNING ALERT	DMOD = 0.75 nmi HMD = 0.75 nmi ZTHR = 450 ft	25 sec		“Traffic, Maneuver Now”
CORRECTIVE DAA ALERT	DMOD = 0.75 nmi HMD = 0.75 nmi ZTHR = 450 ft	75 sec		“Traffic, Separate”
PREVENTIVE DAA ALERT	DMOD = 0.75 nmi HMD = 1.0 nmi ZTHR = 700 ft	75 sec		“Traffic, Monitor”
DAA PROXIMATE ALERT	DMOD = 0.75 nmi HMD = 1.5 nmi ZTHR = 1200 ft	85 sec		N/A
NONE (TARGET)	Within surveillance field of regard	N/A		N/A

The second display configuration incorporated all the informational aids contained in the information-only configuration, and it also included a “vector planner” feature. That feature, which was represented as an arrow that the pilot could rotate around his aircraft symbol at the center of the TSD, would change color depending on the level of alert expected if the aircraft turned to the heading indicated by the arrow. A similar feature was available for altitude alert planning purposes: a set of altitude blocks near the UAS’s current altitude were color-coded according to the level of alert expected if the aircraft climbed or descended to and maintained the given altitude. The color-coding of these headings and altitudes was consistent with the highest alert level specified in Table 1 that was predicted over the subsequent 85 seconds.

The last two display configurations, called Stratway+ and the Omni Bands, showed bands of color around the UAS symbol at the center of the TSD at all times when a preventive threat level or higher was predicted. This represented the same information contained in the vector planner configuration; however, the pilot did not need to activate and slew the arrow symbol to see what threats lay in each direction; that information was always available as a band. In the Stratway+ implementation (Chamberlain 2015), a yellow band showed the pilots any heading or altitude that contained a preventive alert or higher; no differentiation between the levels was indicated. Pilots were also provided green bands that indicated the best heading or altitude to fly to in the event that a LoWC was unavoidable, but which would result in the fastest return to well clear status. The Omni Bands implementation was very similar to the Stratway+, except the level of alert for a given heading or altitude was indicated by changing the appearance of the bands: a preventive alert was represented by dashed yellow bands, a corrective alert by solid yellow bands, and a warning alert by solid red bands. Further information about the JADEM algorithm underlying the alerts in all the displays and the guidance information in the vector planner and Omni Bands displays may be found in Abramson et al. (2016).

D. Surveillance System Uncertainty

The surveillance system model created for this simulation was designed to provide the pilot with a realistic level of uncertainty expected in a minimum operational performance standards (MOPS)-compliant DAA system. It was not designed to be high fidelity, however, and did not capture complex dependencies between intruder geometry, sensor types, and sensor fusion algorithms. The objective in testing surveillance uncertainty was simply to determine whether major differences in pilot acceptability and rate of well clear violations would occur with a modest level of uncertainty.

The DAA surveillance sensor model combined sampling of a Gaussian distribution, characterized by a bias error and standard deviation, in which each sample was uncorrelated with those before or after it, with a moving average filter that averaged together the last N measurements (see the diagram in Figure 1). This second step allowed sensor measurements that were passed to the DAA system to be correlated in time, but they preserved the original variance of the Gaussian distribution. Half of the pilots flew all of their missions (and all of the display configurations) with this sensor model, while the other half of the pilots flew all missions with no sensor error (the “perfect” condition). This between-subjects comparison was designed to minimize the biasing effect that would occur if a subject saw both the perfect and imperfect conditions.

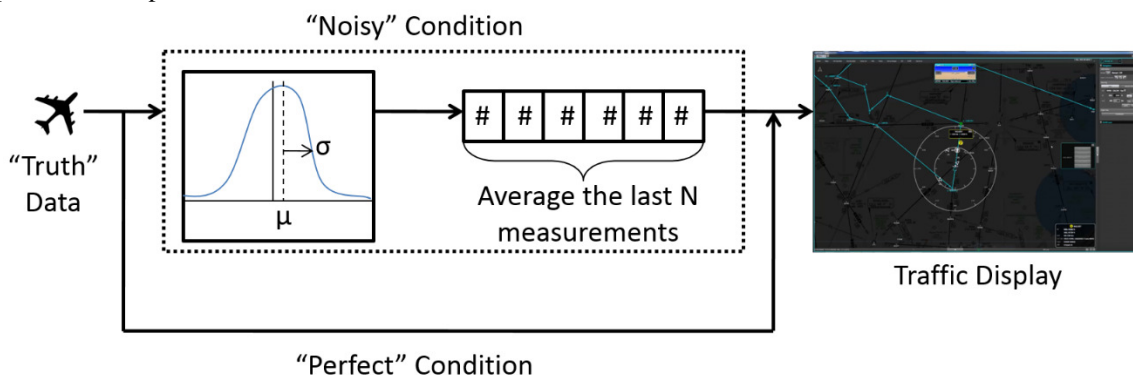


Figure 1. DAA surveillance sensor error model.

The sensor model was calibrated against flight-test data gathered at NASA’s Armstrong Flight Research Center in 2014. The goal of the calibration was not to precisely match the error characteristics of the prototype system under test there—an effort that would have been impossible given the simple design of the model. Instead, the goal was to provide pilots with a level of DAA surveillance error that was representative of errors they might experience with a future, certified DAA system. This level of error, which could be higher or lower than the error in a certified system,

is quantified according to metrics contained in Figure 2 and Table 2 so that designers of future systems can estimate the level of impact their system’s surveillance error will have on a pilot’s ability to perform DAA functions. If such a future system has lower error than that tested here, then it is likely the impact on the DAA function will also be lower than what was measured in this simulation.

Two primary metrics were used to compare the flight-test and simulation errors: (1) the actual absolute error in either the lateral or vertical estimate of an intruder’s position, and (2) the difference in error between consecutive sensor measurements. For this simulation, it was determined that the difference in consecutive measurements was the more important metric to match because it is the error that is experienced by the pilot. Such errors cause alerts to appear and disappear as predictions move inside and outside the well clear boundary, and the inconsistent movement of targets on the display makes it difficult to extrapolate the future position of the intruder. On the other hand, the absolute error between the displayed position of the intruder and the actual position of the (simulated) target is never apparent to the pilot. While such errors are critical in the actual performance of a DAA system, they do not directly degrade the performance of the pilot. The effect of such errors is better measured with a large number of realistic encounters between UAS and historical manned aircraft in batch simulations, not in pilot-in-the-loop simulations. For these reasons the surveillance model was calibrated against the measurement-to-measurement errors rather than the absolute errors. The results of the calibration are indicated in Figure 2 and Table 2.

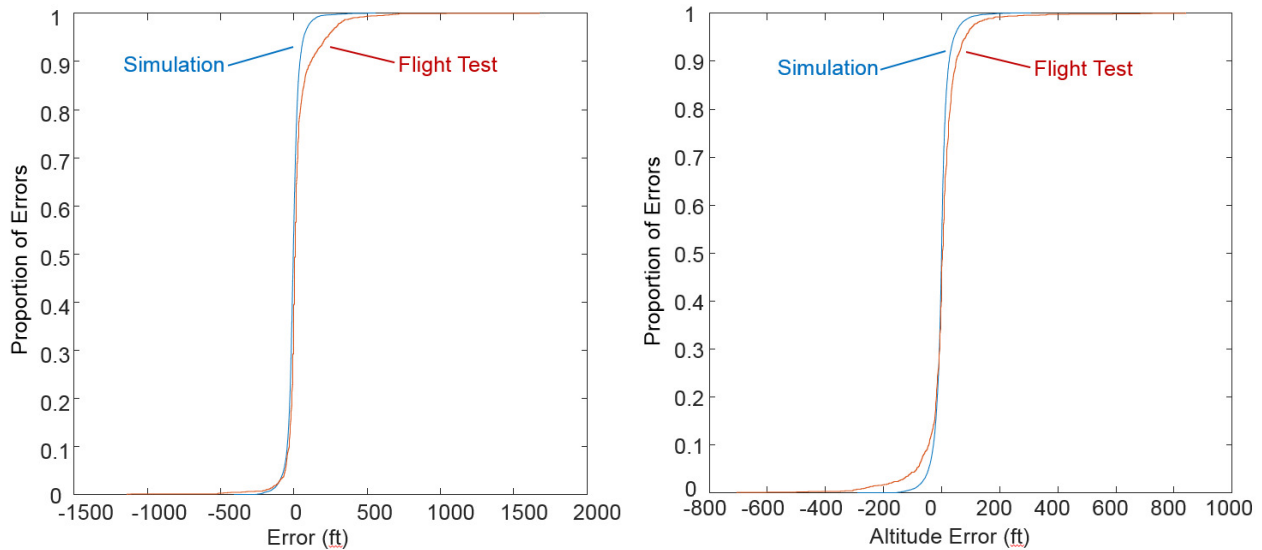


Figure 2. Comparison of surveillance system model errors (blue) to flight test error measurements (red) for lateral errors (left) and altitude errors (right).

Table 2. Comparison of DAA surveillance sensor accuracy between flight test and simulation

	Flight Test	Simulation
Mean Diff Horizontal Error (ft)	54.48	30.61
Median Diff Horizontal Error (ft)	20.08	19.03
Mean Diff Altitude Error (ft)	36.47	18.28
Median Diff Altitude Error (ft)	17.80	10.06

E. Loss of Well Clear Metrics

The separation standard that UAS must adhere to when they encounter VFR aircraft is referred to as “well clear,” and a failure to maintain the separation standard is called a “loss of well clear” (LoWC). The effects of different definitions of well clear have been evaluated in detail (Johnson, Mueller, and Santiago 2015), and the final definition accepted by the community developing standards for DAA systems (RTCA 2013) is described in a paper by Cook and Brooks (2015). The parameters that specify the separation standard are given in Table 3. A loss of well clear occurs when the following conditions are met at the current time:

$$[0 \leq \tau_{mod} \leq \tau_{mod}^* \text{ and } HMD \leq HMD^*] \text{ and } [|d_h| \leq h^*] \quad (1)$$

Table 3. Well clear parameter values

Parameter	Value
Modified tau (τ_{mod}^*)	35 sec.
DMOD	4000 ft.
HMD*	4000 ft.
Vertical threshold, d_h^*	450 ft

The fact that a LoWC occurred does not indicate the reason for the LoWC or how to reduce the rate at which it occurs. Although all LoWCs are serious events, knowing that a LoWC occurred because an intruder changed altitude shortly before the LoWC will suggest a different remedy than an LoWC that occurred because the pilot returned to course before the conflict had been resolved. For these reasons, seven different categories of LoWCs were created and each encounter that resulted in a LoWC was assigned to a single category. The analyses given in the results section use this categorization as appropriate in order to fairly judge the overall pilot-DAA system performance in the real world by correcting for unavoidable simulation artifacts. LoWCs were assigned to one of seven categories that were further classified according to whether or not they were the pilot’s responsibility:

- **Ineffective maneuver.** LoWCs in this category stemmed from the pilot’s selection of a maneuver that did not increase the inter-aircraft separation to more than the well clear standard despite being alerted to the conflict with sufficient time to respond (>10 sec). Losses in this category are typically caused by a lack of guidance information to the pilot (i.e., no assistance from a conflict resolution algorithm) or lack of pilot experience in the type or size of maneuver required to avoid the LoWC. This category is assigned to “pilot’s responsibility.”
- **Slow pilot response.** The losses in this category were due to slow creation and execution of encounter resolution maneuvers, and they are distinguished from LoWCs in the “ineffective maneuver” category by the fact that they would have created adequate separation if they had been executed earlier. Slow responses could result from pilots spending an excessive amount of time watching the progression of a conflict, or from repeated attempts to contact ATC over a busy frequency before attempting a maneuver. This category is assigned to “pilot’s responsibility.”
- **Early return to flight plan.** On several occasions pilots were able to successfully avoid LoWCs during the beginning of an encounter, but they turned back towards their original flight plan before that return path was free of conflicts with the intruder. This type of LoWC occurs only when specific guidance tools are unavailable and so it is important to distinguish this from other LoWCs. This category is distinct from the previous categories because the intruder was no longer an active threat, but the pilot, in an attempt to rejoin their route, directly caused a LoWC. This type of LoWC occurs when there is a lack of guidance information available regarding when a return is conflict-free. This category is assigned to “pilot’s responsibility.”
- **Intruder late maneuver/acceleration.** A complex choreography of intruder maneuvers was required to deliver the right sequence of encounters to each pilot and compensate for the unique resolution maneuvers performed by each pilot leading up to a given encounter. Occasionally, this complex series of maneuvers resulted in a first alert of a predicted LoWC with 10 seconds or less of warning time. Although such late maneuvers and resulting LoWCs will occur in the real world, the rate at which that will happen should be measured in separate simulations with realistic intruder trajectories, the consequences of which can be modeled with automated collision avoidance algorithms. The results of such encounters do not provide meaningful data on the performance of the pilot-DAA system because

the outcomes of the encounters are determined by the available warning time, not the system under test. This category is *not* assigned to “pilot’s responsibility.”

- **DAA surveillance system elevation error.** The elevation error modeled for the DAA surveillance system resulted in measurement-to-measurement altitude differences of up to hundreds of feet. Because the vertical separation requirement is only 450 ft and aircraft are routinely separated operationally by 500 ft, these surveillance errors can artificially indicate that a LoWC has occurred. In fact, almost all encounters that were in reality separated by 500 ft would have appeared at one point or another to have separation less than 450 ft because of the surveillance system error. The relatively small number of such LoWCs is due to the fact that pilots would proactively change altitude when they observed this intermittent alerting behavior in order to avoid such a loss. This category is *not* assigned to “pilot’s responsibility.”
- **Diverging loss.** An unanticipated and un-alerted type of LoWC occurred when an aircraft “clipped” the rear boundary of the well clear separation cylinder (4000 ft radius, ± 450 ft vertical distance) and the relative range rate between the aircraft was positive (diverging aircraft). When the range rate is positive the HMD was automatically set to infinity (rather than the current inter-aircraft range, which is the correct approach), so a LoWC was not predicted to occur and the pilot received no alert about that portion of the trajectory. Other elements of the DAA display did indicate that very close separation was predicted, so, although this assignment is certainly debatable, this category is assigned to “pilot’s responsibility.”
- **Bug.** On one occasion an error in the display of state and alert information to the pilot caused an LoWC. This error (“bug”) was immediately fixed and this category of LoWC is *not* assigned to “pilot’s responsibility.”

F. Well Clear Penetration Integral

The severity of LoWCs has been measured using several different metrics in a previous study (Santiago 2015). Typical methods include the time spent in a LoWC state and the geometric proximity of the two aircraft at their closest point of approach (CPA) normalized by the well clear separation cylinder (i.e., 4000 ft horizontally and 450 ft vertically). However, the UAS DAA community has not agreed to a single metric that relates LoWC severity to the elevated level of collision risk that accompanies the LoWC. The well clear penetration integral (WCPI) is a proposed method for incorporating both the geometric proximity of the aircraft and the time spent in violation in a single measure. The mathematical definition of the WCPI is given in Eq. (2), while an illustration of the difference in severity of two encounters with significantly different WCPI values is shown in Figure 3.

$$WCPI = \int \min \left\{ \frac{(4000 - HMD)}{4000}, \frac{(450 - DH)}{450} \right\} \frac{35 - \tau_{mod}}{35} dt \tag{2}$$

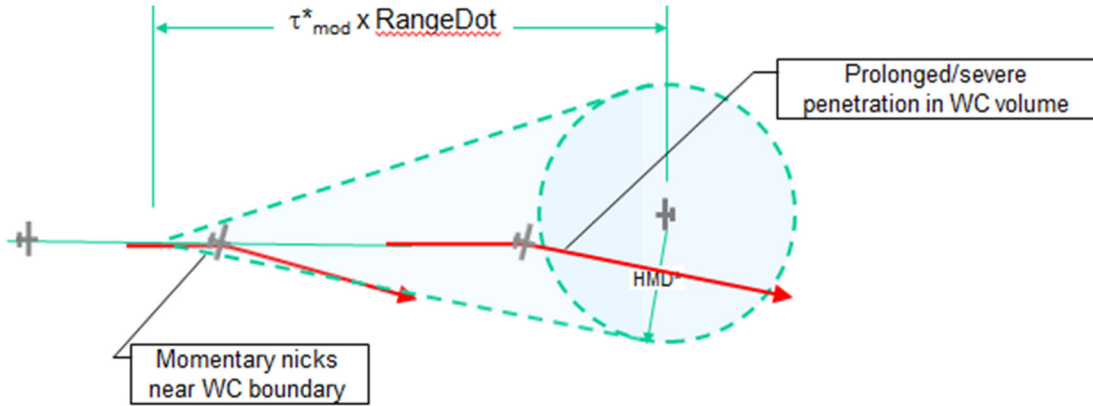


Figure 3: Differences in LoWC severity (credit: Bihrl Applied Research, unpublished).

The WCPI uses normalized separation terms inside the integrand in order to account for the differences between horizontal and vertical separations and the temporal term, modified tau. Equation 2 can be approximated with a summation ($j = 1$ to n) for all time steps of the LoWC, with horizontal miss distance, vertical miss distance, and modified tau values specified in Table 3. The purpose of the modified tau term is to reduce the contribution of the early stages of the LoWC, during which separation can still be quite large.

III. Results

The following sections present results related to the ability of pilots to avoid losses of well clear, the relationship between LoWCs and TCAS alerts for different encounters, and the behavior of pilots in reacting to and implementing avoidance maneuvers.

A. Loss of Well Clear Rate

The proportion of LoWCs out of all encounters that were predicted to become LoWCs without pilot intervention is shown in Figure 4. Only those losses of well clear that were deemed to have been pilot responsibility are shown. For example, if an intruder maneuvered without warning and caused an immediate LoWC, then that encounter would not be considered a LoWC and would be removed from the analysis for this chart entirely (though the data is retained and displayed as part of other analyses). The information-only display condition resulted in a LoWC rate two to four times higher than any other display, a result confirming the findings of previous studies (Santiago 2015). A pilot requires not just an alert that an encounter could result in a LoWC, they also need some level of guidance about how to respond to the alert. The two banding guidance displays achieved similar performance at around 5% LoWC, while the vector planner condition had approximately twice this rate of LoWCs. The banding displays provide a clear improvement in DAA performance by reducing the rate of LoWCs over the two alternative designs.

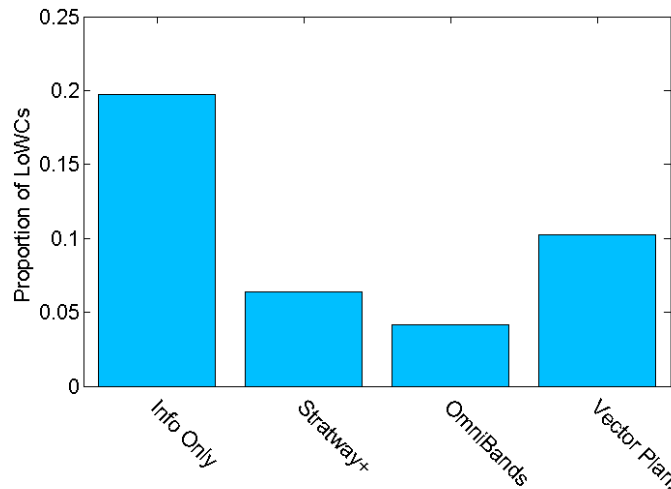


Figure 4. Proportion of losses of well clear by display. Only LoWCs that were the "pilot's responsibility" are included.

The results in Figure 4 appear to clearly indicate that the banding displays deliver better LoWC performance than the other two display conditions tested, but that conclusion requires a significance test. Because the number of LoWCs per scenario in each display condition are not normally distributed, the common analysis of variance (ANOVA) test is not appropriate. Instead, we apply the non-parametric Kruskal-Wallis test of significance (Kruskal and Wallace 1952; Daniel 1990). The significance matrix for the difference in number of LoWCs between each display configuration is given in Table 4.

Table 4. Significance value for rate of LoWC by display condition

Display Condition	Info Only	Stratway+	OmniBands	Vector Planner
Info Only	1.0	0.0159	0.0011	0.0866
Stratway+		1.0	0.2067	0.5569
OmniBands			1.0	0.0902
Vector Planner				1.0

The values in the table are color-coded according to the level of significance, with green being significant at the 0.05 level, beige being marginally significant between 0.05 and 1.0, and the rest not significant. Compared to the information-only display, both Stratway+ and Omnibands are significantly different in terms of the number of LoWCs per scenario (0.0159 and 0.0011, respectively). The vector planner shows a marginally statistically significant reduction in the rate of LoWCs (0.0866) compared with information only. The Omnibands also have a marginally statistically significant reduction in number of LoWCs compared with the vector planner (0.0902). There is no measureable difference between the two bands displays (0.2067). These results confirm the observations from Figure 4 that the banding displays provide an important improvement in pilot performance of the DAA function.

B. Losses of Well Clear by Alert Time and Range

Previous studies have shown that the rate at which LoWCs occur is related to the time between first alert and the predicted LoWC, which is in turn related to the inter-aircraft range at this first alert (Santiago 2015). This section presents LoWC results as a function of these temporal and geometric parameters. For the analysis in this section, LoWCs that a pilot could reasonably have avoided with the information and interfaces at their disposal were retained, while those that were completely beyond a pilot’s control were filtered out. See Section II.E for a description of these categories. The effects of factors beyond a pilot’s control will be determined through separate large-scale, fast-time simulations involving tens of thousands of aircraft. This section reports on the effectiveness of the combined pilot-DAA system in remaining well clear as a function of the time and distance at which a “valid” alert is received, defined as the time at which three (not necessarily consecutive) corrective or warning alerts are provided for the same intruder.

The outcomes of all encounters across the four display conditions as a function of the first valid alert is shown in Figure 5. The green sections of each bar indicate encounters that were predicted (and scripted as part of the scenario) to result in a LoWC, but were successfully resolved by the pilot using the DAA system. As observed in prior simulations (Santiago and Mueller 2015), the rate of LoWCs is relatively constant when more than 20 sec of warning is provided to the pilot, but the rate increases to 44% when less than 20 sec is available and is 60% when the time is under 10 sec. What is novel in this chart is the assignment of a reason for the LoWC, and these causes vary significantly between the short and long alert time LoWCs. Most of the LoWCs at low alert times are due to late intruder maneuvers, which is precisely the reason the alert time was short, and so this is an expected result. Although intruder maneuvers continue to cause some LoWCs at larger alert times, the reasons for the LoWCs shift towards factors that are the pilot’s responsibility: ineffective maneuvers, slow responses, and early returns to the flight plan. While the first two of those factors can be reduced by increasing the alert time available to the pilot, the early return to flight plan is in no way affected by the original alert time. Instead, as will be clear from the charts that partition these results by individual display conditions, better DAA guidance can eliminate this cause of LoWC entirely.

The LoWC rate when alert times are under 10 seconds is 60%, which may appear low given that the pilot would have almost no time to formulate and uplink a maneuver before it was too late for the aircraft’s dynamics to physically avoid a LoWC. However, pilots did notice such intruders approaching and would be ready to execute a resolution maneuver immediately at the first alert. The banding displays were particularly effective in this regard; pilots could observe the red warning alert bands grow towards their current heading and altitude and therefore were primed to act when the alert was issued.

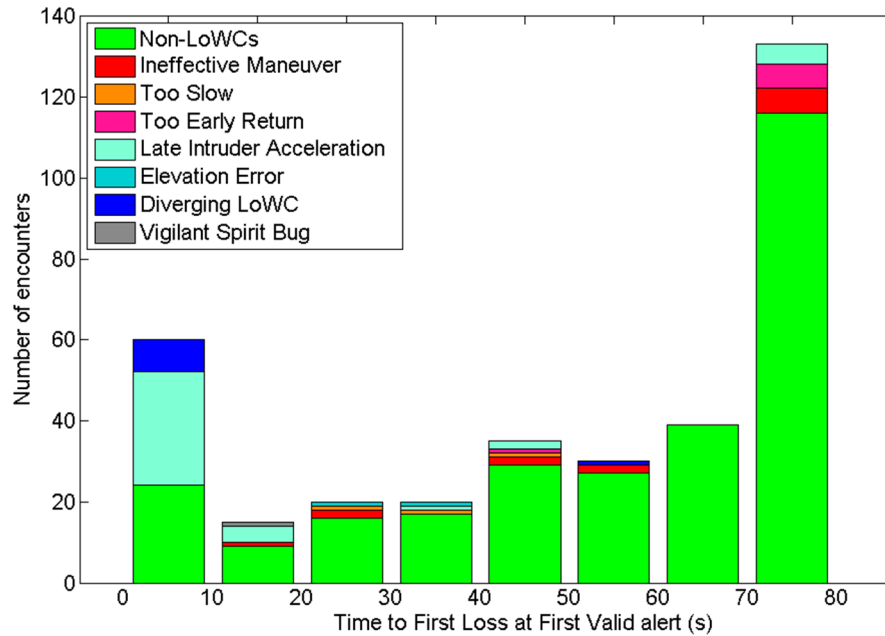


Figure 5. Outcomes of encounters by time to first loss of well clear at which an alert was provided. All displays.

The outcomes of encounters in the information-only display condition are shown in Figure 6; just under 20% of the total number of encounters became LoWCs that were the pilot's responsibility. In addition to a higher overall rate of LoWC, the rate at which LoWCs were avoided when less than 10 sec of warning time was available increased to 73% compared with 45% for the other three displays (60% overall). This indicates that with the information-only display pilots are less able to anticipate potential LoWCs when alerts are not timely, either because of surveillance sensor uncertainty or because of late-maneuvering intruders. The rate of LoWCs when the full alert time was provided to the pilot (between 70 and 80 sec to LoWC) in this condition is very high at 25% (9 of 36) compared with the rate of the other three displays combined (8.3%, 8 of 96). These high LoWC rates at both ends of the alert time spectrum, along with the observation that the causes of most LoWCs at the long alert times are prevented with DAA guidance, indicates that the information-only display is likely insufficient to meet the safety requirements for DAA systems.

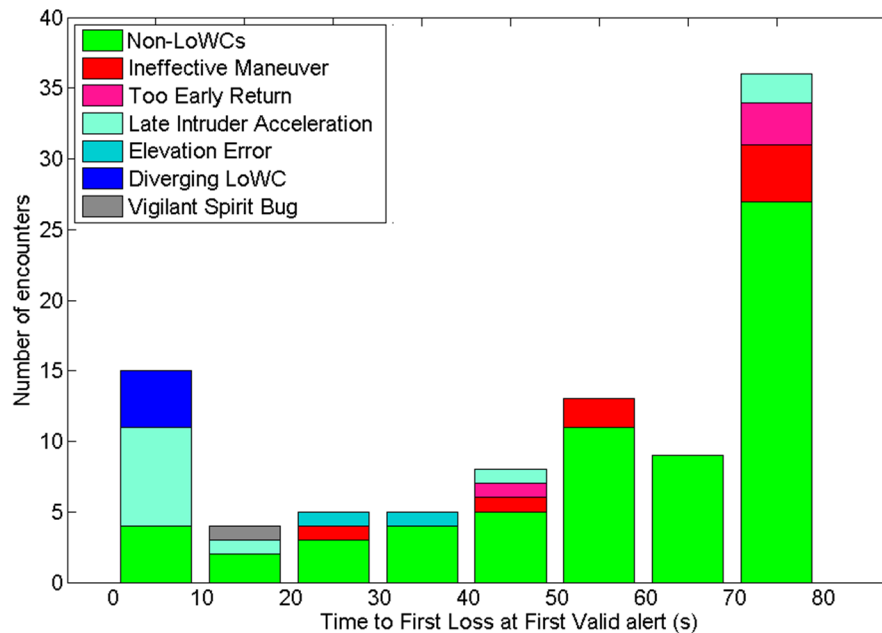


Figure 6. Outcomes of encounters by time to first loss of well clear at which an alert was provided. Information-only display.

The encounter outcomes for the vector planner display are shown in Figure 7. The overall pilot-responsible LoWC rate was intermediate between the information-only display and the two banding displays at about 10%. Ineffective maneuver LoWCs dropped significantly due to the additional information provided to the pilot about alerts that would occur if specific headings or altitudes were selected. Interestingly, this display configuration had the most LoWCs attributed to slow execution of the maneuver (i.e., two), but the mean pilot response time for an alert was faster than for the information-only display (Rorie, Fern, and Shively 2016). For some encounters, pilots may have spent too much time using the vector planner tool to search for a successful resolution maneuver, though on average it did not slow them down by more than three to five seconds compared with the banding displays. Although the vector planner display did improve on the information-only display in terms of the LoWC rate at both ends of the alert time range and provided shorter pilot response times, it made the altitude and heading alerting information directly available on the banding displays more difficult to access. For these reasons the vector planner display should not be implemented for a DAA system.

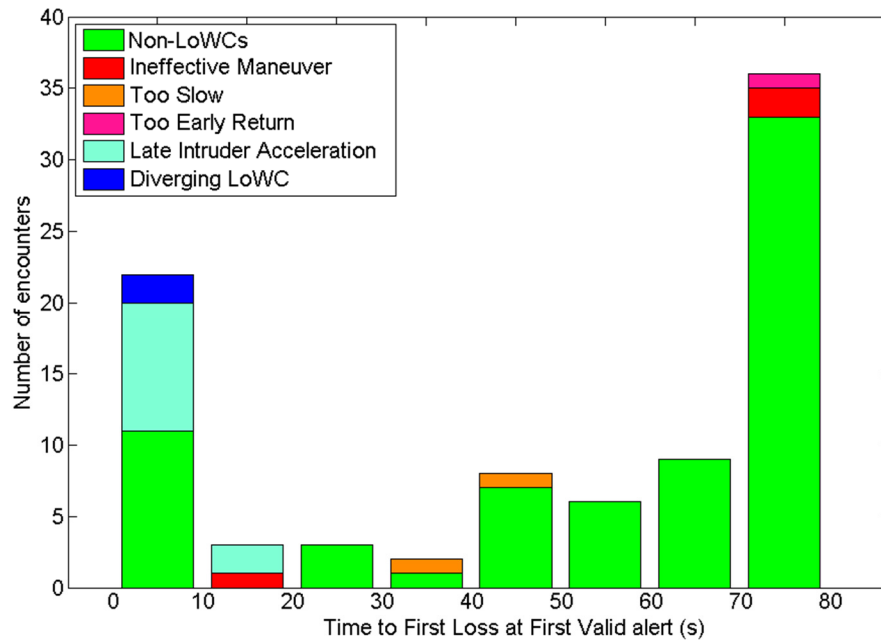


Figure 7. Outcomes of encounters by time to first loss of LoWC at which an alert was provided. Vector planner display.

The banding displays showed the lowest LoWC rates among the four display conditions, and the specific encounter outcomes for each are shown in Figure 8 and Figure 9. Pilots were unable to avoid a LoWC during several encounters using the no-fly bands, though the three slow and ineffective maneuver LoWCs all occurred with less than half of the normal alerting time. On two occasions the pilots returned to their original flight plan too early and caused a LoWC, possibly because of a mismatch between the aircraft dynamics used to calculate the band guidance and the dynamics used in the simulation itself. However, the overall pilot-responsibility LoWC rate of 6% is excellent compared with the non-banding displays, and is statistically identical to the 4% LoWC rate provided by the Omni Bands display. As shown in Figure 9, the only pilot-responsibility LoWCs that occurred in this display condition were under diverging circumstances—the aircraft were in fact getting farther apart when the LoWC was recorded. These LoWCs can likely be avoided simply by changing the alerting logic so that it permits alerts even when aircraft are predicted to be separating. Within the limits of simulation fidelity and encounter realism, the Omni Bands display performed nearly flawlessly. The authors recommend using a band-type display that indicates at least the heading and altitude commands that would lead to corrective or warning alerts. The alerts should be calculated by simulating the trajectory of the aircraft until it reaches the new heading or altitude rather than instantaneously setting the aircraft’s state to the commanded state and building the trajectory from there. The additional trajectory fidelity this provides reduces the rate of LoWCs and provides better information to the pilot. A DAA system based on this type of display and algorithm is likely to meet the safety requirements necessary to operate UAS in the NAS.

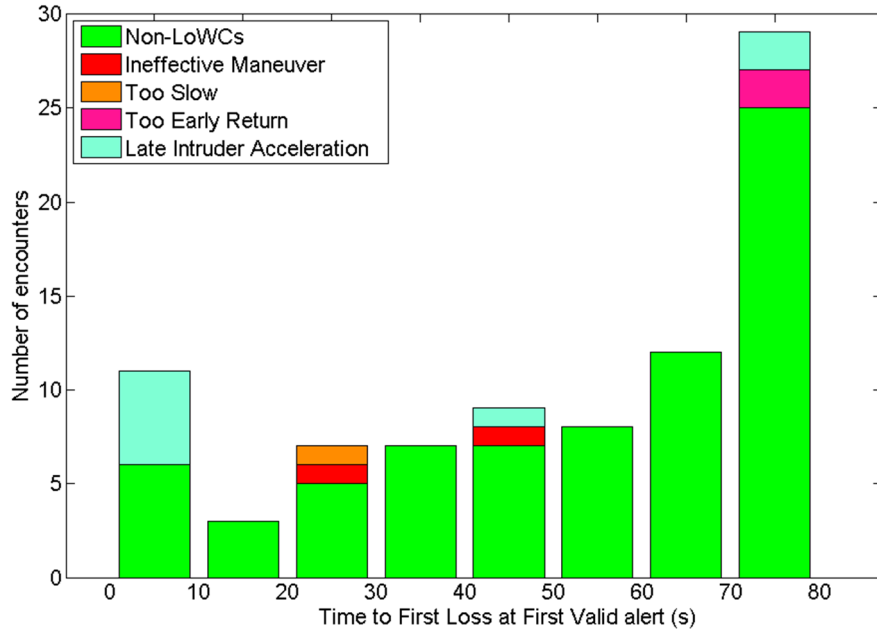


Figure 8. Outcomes of encounters by time to first loss of LoWC at which an alert was provided. Stratway+ display.

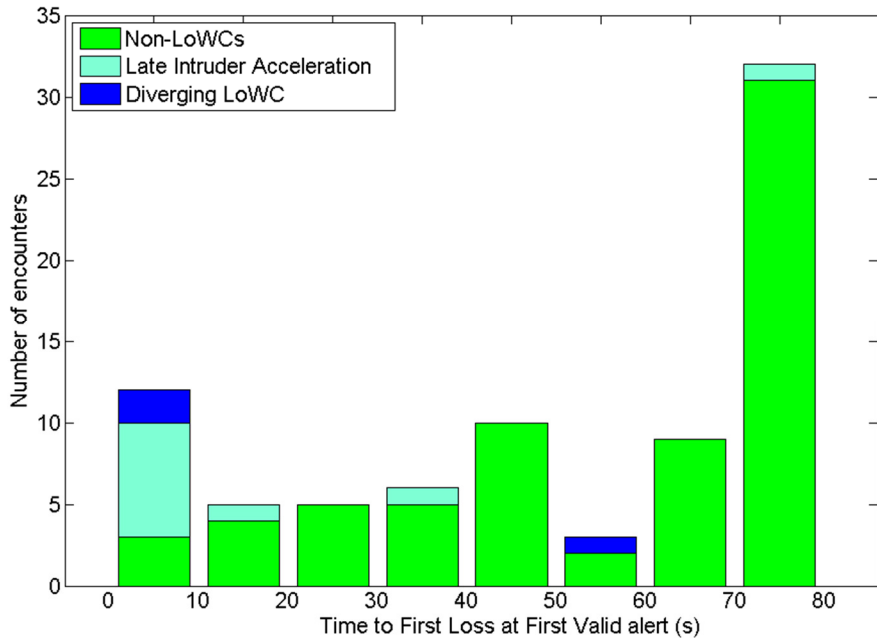


Figure 9. Outcomes of encounters by time to first loss of LoWC at which an alert was provided. Omni Bands display.

The time until well clear is first lost is a useful metric from a pilot’s perspective because it indicates roughly how quickly they must respond to an alert; however, other aspects of the DAA system and NAS operate on distance rather than time. For example, the detection range of a DAA surveillance sensor is largely a function of range and does not depend strongly on the direction or velocity of the intruder. Similarly, air traffic controllers use the current range between aircraft in deciding whether or not to issue traffic advisories to pilots regarding proximate traffic (Mueller 2015). The relationship between the time to CPA at first alert and the range at first alert depends on the velocity of each aircraft and their relative encounter geometry. The outcomes of encounters as a function of the range at first valid alert across all four display conditions are shown in Figure 10. An interesting difference between the range and time-to-CPA charts is that in the range chart the LoWCs are concentrated at ranges of 4 nmi and below, with few LoWCs

at larger ranges and many non-LoWCs at these larger ranges. In contrast, the time-to-CPA charts have a large number of LoWCs at the long and short ends of the alerting time range. This difference suggests that many of those long-alert-time LoWCs occurred at short range, implying slow closure rates. These acute angle encounters can be the most difficult to resolve and can also transition from a non-threat to a threat status very quickly. Although pilots should continue to be alerted according to time-to-CPA criteria, they should also pay particular attention to those intruders that are closer than 4 nmi away and request separation from such vehicles if they have not already received a traffic advisory.

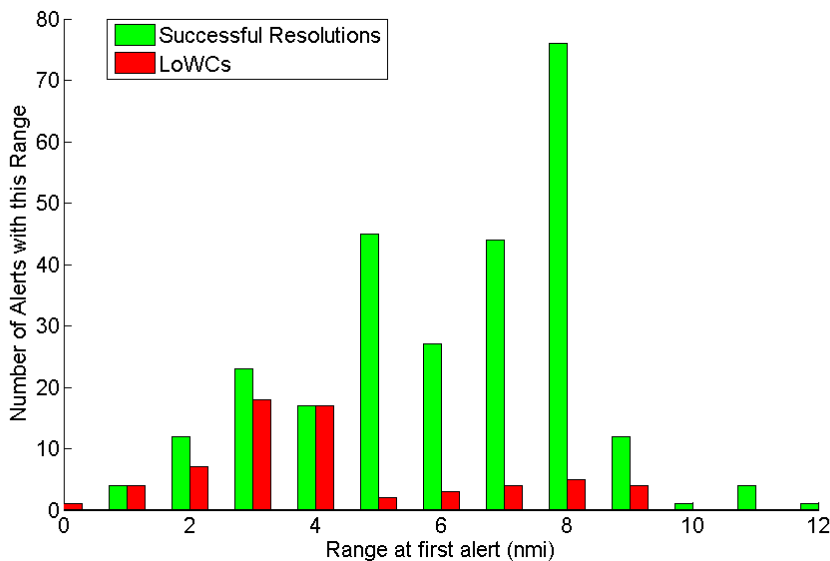


Figure 10. Outcome of encounter by range at which first alert was received.

C. Severity of Losses of Well Clear

The severity of violations for those encounters that resulted in losses of well clear was evaluated in three ways: geometric proximity of the closest point of approach, duration of the violation, and a combined integral that incorporated both geometric and temporal aspects of the violation into the single “well clear penetration integral” (WCPI) defined by Eq. 2.

The severity index is a measure of the geometric proximity of two aircraft that compensates for the different distance thresholds used in the vertical and horizontal domains. It does not use any temporal measures to evaluate severity. It is defined as

$$S_{index} = \min_t \left\{ \max \left[\frac{horz.range(t_i)}{h_{sep_{CA}}}, \frac{vert.range(t_i)}{v_{sep_{CA}}} \right] \right\} \quad (3)$$

where the $h_{sep_{CA}}$ and $v_{sep_{CA}}$ are the geometric portions of the well clear definition, with values of 4000 ft and 450 ft, respectively. The maximum normalized separation is computed at each time step, and the minimum value over the entire trajectory becomes the separation index. For predictions rather than the actual encounter trajectory, the index is computed in the same way but is done over the prediction made at each time step. The lowest predicted separation index over all these predictions is what is reported in Figure 11. Over all the displays, only 30% of encounters actually penetrated the 4000 ft horizontal and 450 ft vertical separation cylinder; the rest violated it with a modified tau under 35 sec but maneuvered away before reaching the separation cylinder. The DAA systems overall raised the median predicted separation index from 0.59 to 1.06, and none of the LoWCs became near mid-air collisions (NMACs). The small number of LoWCs that were the pilot’s responsibility when using the banding displays makes comparison of this metric among the display conditions challenging. No conclusions are drawn regarding the severity of LoWCs between the displays using this metric.

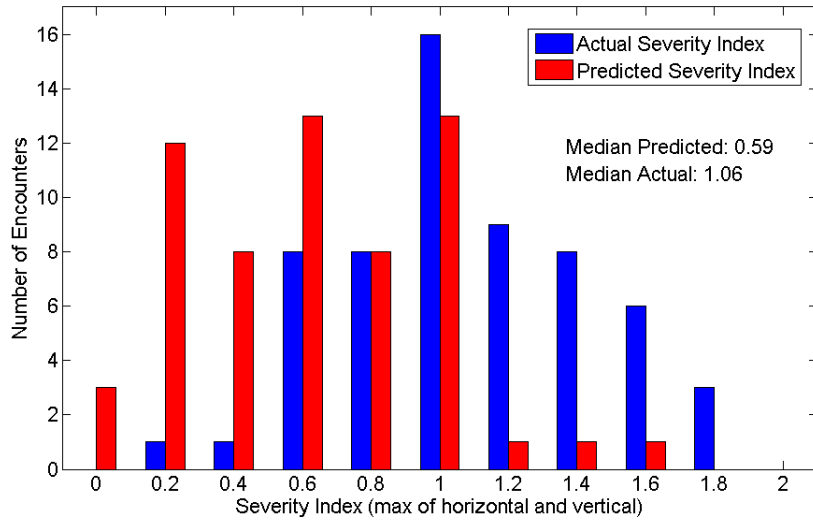


Figure 11. Predicted and actual mitigated geometric severity of all encounters.

The second method of evaluating LoWC severity is to examine the amount of time each aircraft spent in an LoWC or warning alert status. This temporal metric is useful because minor violations of well clear will also be of short duration, while longer durations represent extended durations in which the two aircraft are in a state of heightened risk. Although the functional relationship between duration and risk is not known, the two variables are positively correlated. In addition, pilots are instructed to avoid LoWCs and warning alerts, and to maneuver immediately to regain well clear or remove the alert when these situations occur, so the pilots are to some degree attempting to minimize this metric.

All encounters that resulted in some time spent in a warning alert were divided into their respective display conditions. They were then placed in 15 sec-wide bins and the total number of encounters in each bin plotted in the histogram in Figure 12. The results largely parallel the LoWC rate results, with the information-only display having the largest total time in warning alert status and the most unique encounters in each bin, except for the 15–30 sec bin. The vector planner had the second-most time in warning alert status, and the banding displays had the least time in warning alert status. It is important to avoid warning alerts because pilots will be permitted to deviate from their ATC clearance when these alerts occur, potentially impacting normal ATC operations and degrading safety. The banding displays are likely to cause less disruption to normal NAS operations and therefore should be strongly considered for use by DAA systems.

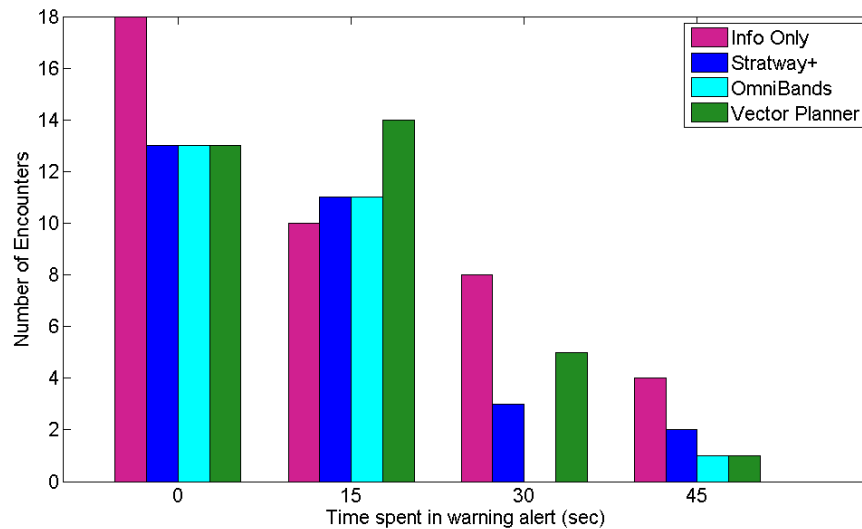


Figure 12. Number of encounters that spent the given amount of time in a warning alert status.

A histogram of the time spent in LoWC according to display condition is shown in Figure 13. Again paralleling previous results on the rate of LoWC and time spent in warning alert status, the information-only display had the most cumulative time in LoWC, followed by the vector planner, and finally Stratway+ and the Omni Bands. Although this result does not add insight into the relative strengths of the different displays, it does reinforce the conclusion that a display based on the Omni Bands is likely to provide the safest DAA system.

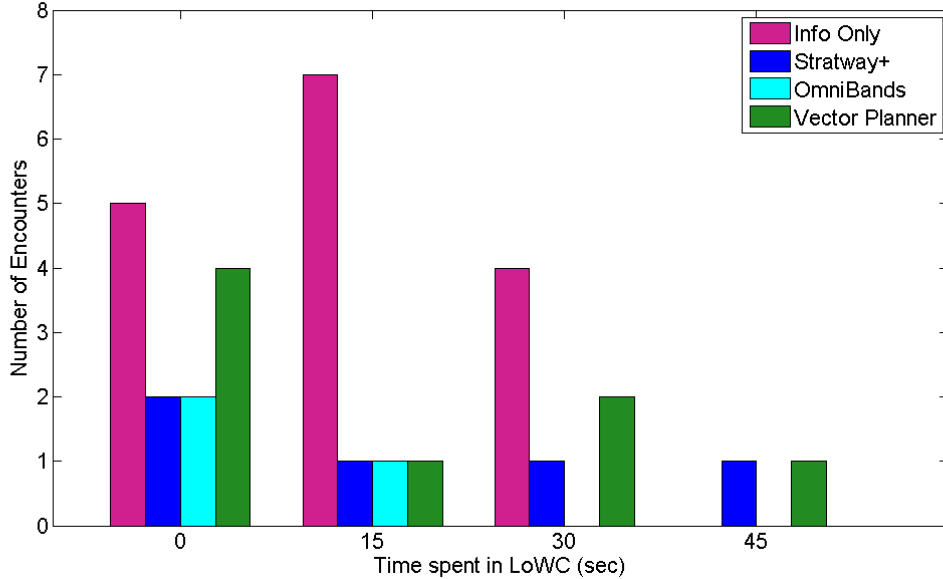


Figure 13. Number of encounters that spent the given amount of time in a LoWC status.

The WCPI (see Section II.F) was also calculated for each trajectory that resulted in a LoWC and the results binned by metric magnitude and separated by display condition. The results of this process are shown in Figure 14. The order of display conditions according to this metric are the same as the previous metrics, though the relative magnitudes of them tell a slightly different story: information-only and vector planner displays were quite close in terms of the total WCPI (34.4 and 30.9), primarily because the vector planner condition had more long duration LoWCs. Stratway+ was intermediate (18.8) while the Omni Bands barely registered on this metric (3.7). That latter display condition benefited from both short durations of LoWC and small penetrations of the separation cylinder. The WCPI metric, consistent with all other metrics presented in this paper, indicates that the Omni Bands performed the most successfully in helping the pilot avoid LoWCs and reducing the severity of LoWCs when they did occur.

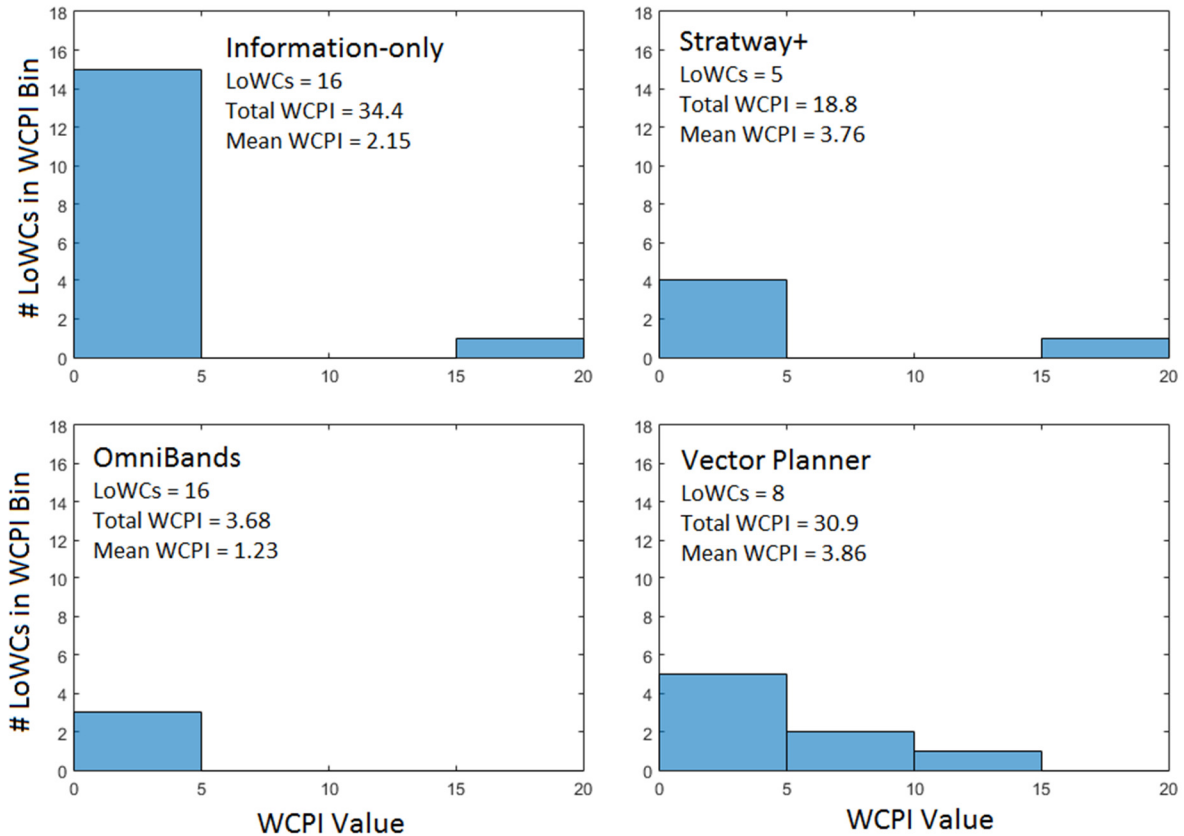


Figure 14. Histogram of the WCPI values by display

D. Pilot Response Time

Gathering data to create a model of pilot response time is a key goal of these pilot-in-the-loop simulations. Such models allow more realistic and accurate fast-time simulations that can be used to determine the frequency and outcomes of encounters simulated in this experiment. A detailed analysis of each step in a pilot’s response sequence is presented in Rorie, Fern, and Shively 2016, and a UAS pilot model is being developed at MIT Lincoln Laboratory using the identical alerting scheme and Omni Bands guidance designed for this experiment (Kuffner, Guendel, and Darrah 2016). Those detailed models are necessary for simulations matching their level of fidelity, but for simpler simulations a straightforward distribution of response times may be more appropriate.

The total response time between first alert and upload of a resolution maneuver was analyzed as a function of several independent variables: display type, time to first loss of separation, level of surveillance sensor uncertainty, and intruder type (cooperative vs. non-cooperative). The latter independent variables showed little difference in response time between conditions, but the banding display response times were similar to each other and shorter than the other two displays. Because of this similarity between the two banding displays and their overall good LoWC performance, the response-time data for those displays was combined into a single set. That set was then analyzed for any effect of time to LoWC at first alert. It was found that when the first alert was within the warning threshold, the response times were significantly shorter than when the first alert was corrective. This difference is consistent with the instructions provided to the pilots not to coordinate or seek ATC approval for well clear maneuvers when a warning alert is present.

The response-time data for the two banding displays was divided into those encounters in which the first alert was less than 25 seconds to LoWC, and those in which the first alert was more than 25 seconds to LoWC. Several distributions were fit to this data: the gamma distribution was the best fit for both sets. Cumulative distributions of both the simulation data and the models are shown in Figure 15, along with the shape and scale parameters necessary to recreate those models. Although the response-time models do not take some important factors into account when calculating a response time (e.g., follow-up maneuvers after the initial upload), these distributions should be useful

for lower-fidelity simulations that require an estimate of delay between the time when a threat is detected and the time a resolution maneuver is executed.

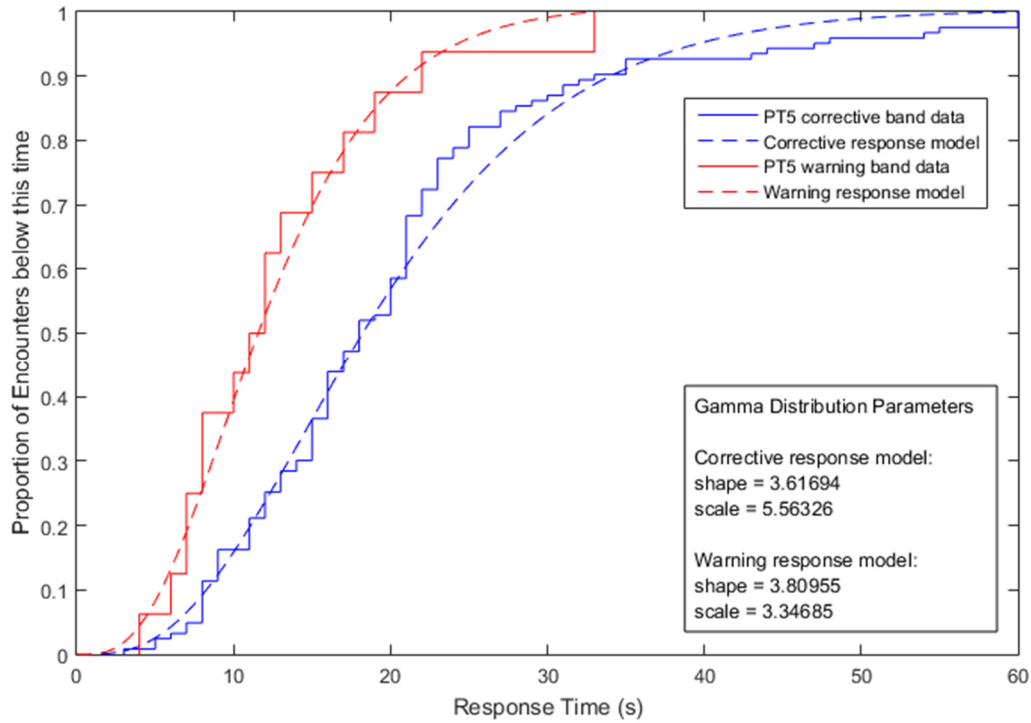


Figure 15. DAA maneuver response time data and models.

E. Surveillance Uncertainty

The level of sensor uncertainty evaluated in this simulation did not cause significant differences in most of the metrics used to evaluate the displays. The displays that performed best in terms of the number and severity of LoWCs with perfect information (the banding displays) also did the best in the surveillance uncertainty condition. The major difference between the perfect and noisy conditions occurred when non-cooperative intruders were separated vertically by 500 ft: when uncertainty was present, the intruders would appear to change altitude from measurement step to measurement step, which would cause intermittent, spurious corrective and warning alerts between the correct preventive alerts. In the perfect condition only preventive alerts were displayed.

Pilots in the experiment occasionally maneuvered to avoid an intruder that was currently at the preventive alert stage; nearly every example of this type of maneuver was due to uncertainty in the vertical separation of the intruder from the ownship as described previously. This fact is crucial to consider for the deployed DAA system because operationally the typical vertical separation between IFR and VFR aircraft (500 ft) is only slightly larger than the vertical separation standard (450 ft). The closeness of these two standards means that UAS are likely to request lateral maneuvers from ATC to separate themselves from non-cooperative intruders even if those intruders have more than the minimum required separation. This additional maneuvering compared with cooperative aircraft is not likely to impact the efficiency of the NAS because encounters between UAS and non-cooperative aircraft are rare (less than one encounter for every 5 flight hours, [RTCA 2015]) and because under current operations manned aircraft pilots do not know the precise altitude of non-cooperative intruders and so are more likely to request vectors from ATC to avoid them.

IV. Conclusions

This study evaluated four DAA display configurations, each with different informational and guidance elements. Two different DAA algorithms drove the display elements for these systems. Sixteen UAS pilots flew each combination of the display configurations, with half being given “perfect” DAA surveillance sensor errors and the other half experiencing errors that were comparable, and in some cases slightly better than, errors that were measured in DAA system flight tests.

The displays that showed intruder alert information for altitudes and headings near the current UAS flight state had significantly fewer losses of well clear, from a statistical as well as from a practical perspective. Compared with the information-only display, which provided alerting but no guidance information at all, the rate of losses of well clear was less by a factor of four. In addition to fewer losses, those losses that did occur lasted for shorter periods and did not penetrate as far into the geometric “separation cylinder” as those in the non-banded displays. In particular, the Omni Bands display, which indicated the level of alert in its banding symbology, had slightly better performance than a similar display, which showed only the presence of an alert and not its severity. It is recommended that DAA traffic displays implement a band-type display based on the characteristics reported in this and companion papers (Rorie, Fern, and Shively 2016) in order to improve the safety of UAS operations in the National Airspace System.

A modest level of DAA surveillance sensor uncertainty did not affect the rate of losses of well clear or their severity, but it did cause pilots to make maneuvers when their UAS was separated vertically from an intruder by 500 ft, which is acceptable under today’s airspace procedures. Finally, pilot response time distributions were calculated for cases when the first alert from an intruder was a (less imminent) corrective alert versus a (more imminent) warning alert. These response-time distributions may be used to improve the fidelity of non-human-in-the-loop simulations.

References

Abramson, M., Santiago, C., Refai, M., Mueller, E., and Snow, J., “Java Architecture for Detect-and-Avoid Extensibility and Modeling,” NASA/TM-2016-XXXX, in publication.

Bell, S., Drury, J., Estes, S., and Reynolds, C., “GDTI: A ground station display of traffic information for use in sense and avoid operations,” IEEE, 31st Digital Avionics Systems Conference, Williamsburg, VA, October 2012.

Calhoun, G., Miller, C., Hughes, T., and Draper, M., “UAS sense and avoid system design and evaluation,” Proceedings of the Human Factors Ergonomics Society 58th Annual Meeting, October 2014.

Chamberlain, J. P., Consiglio, M. C., Comstock, J. R., Ghatas, R. W., & Munoz, C. "NASA controller acceptability study 1 (CAS-1) experiment description and initial observations." NASA/TM-2015-218763, 2015.

Cook, S. P., and Brooks, D., “A Quantitative Metric to Enable Unmanned Aircraft Systems to Remain Well Clear,” Air Traffic Control Quarterly, Vol. 23, No. 2/3, pp 137-156, 2015.

Dalamagkidis, K., Valavanis, K., and Piegl, L. A., *On Integrating Unmanned Aircraft Systems into the National Airspace System: Issues, Challenges, Operational Restrictions, Certification and Recommendations*, 2nd ed., Springer-Verlag, Dordrecht, 2012. Daniel, W. (1990). "Kruskal–Wallis one-way analysis of variance by ranks". *Applied Nonparametric Statistics* (2nd ed.). Boston: PWS-Kent. pp. 226–234. ISBN 0-534-91976-6.

Federal Aviation Administration, “Sense and Avoid (SAA) for unmanned aircraft systems (UAS),” SAA Workshop Second Caucus Report, 2013.

Feitshans, Gregory L., Allen J. Rowe, Jason E. Davis, Michael Holland, and Lee Berger. "Vigilant spirit control station (VSCS)—‘The face of COUNTER’." In *Proceedings of AIAA Guidance, Navigation and Control Conf. Exhibition*, 2008.

Gillian..., M., and Cox, V., “Integration of unmanned aircraft systems into the national airspace system. Concept of operations v2.0.” American Society of Testing & Materials, Tech. Rep., September 2012.p

Hesselink, H. H., and Schmitt, D. R., “UAS air traffic insertion starts now,” National Aerospace Laboratory, NLR, Tech Report, 2011.

Johnson, M., Mueller, E., and Santiago, C., “Investigating the impacts of a separation standard for UAS operations in enroute and transition airspace,” 2015, unpublished.

Kruskal, W., and Wallis, A., "Use of ranks in one-criterion variance analysis". *Journal of the American Statistical Association*, Vol. 47, No. 260, 1952, pp 583–621. doi:10.1080/01621459.1952.10483441.

Kuffner, M., Guendel, R. E., and Darrah, S., "Investigating Traffic Avoidance Maneuver Decisions of Unmanned Aircraft Pilots," AIAA Aviation 2016 Forum, Washington D.C., June 2016.

Mueller, E., Isaacson, D., and Stevens, D., "Air Traffic Controller Acceptability of Unmanned Aircraft System Detect-and-Avoid Thresholds," NASA UAS Integration into the National Airspace System Project, Report Number XXX-XX, 2015.

Ostwald, P., and Hershey, W., "Helping global hawk fly with the rest of us," in ICNS Conference, 2007.

Park, C., Lee, S.M., and Mueller, E., "Investigating detect-and-avoid surveillance performance for unmanned aircraft systems," Proceedings of 14th AIAA Aviation Technology, Integration, and Operations Conference, AIAA 2014-2413, June 2014.

Prevot, T. "Exploring the many perspectives of distributed air traffic management: The Multi Aircraft Control System MACS." In Proceedings of the HCI-Aero, pp. 149-154, 2002.

Rorie, C., Fern, L., and Shively, J., "An evaluation of detect and avoid displays for unmanned aircraft systems: the effect of information level and display location on pilot performance," AIAA Aviation, June 2015.

Rorie, R. C., Fern, L., and Shively, R. J., "The Impact of Suggestive Maneuver Guidance on UAS Pilots Performing the Detect and Avoid Function," AIAA SciTech 2016 Conference, San Diego, CA, January 2016. DOI: 10.2514/6.2016-1002.

RTCA, Inc. Special Committee 228, "Detect and avoid (DAA) Whitepaper," RTCA-SC228-WG1-WP01-01, 2013.

RTCA Inc., "Minimum Operational Performance Standards for Unmanned Aircraft Systems (UAS) Detect and Avoid (DAA) Systems Appendix B: Safety Assessment for Performance, Functional, and Operational Requirements for the DAA MOPS," RTCA Paper No. 197-15/SC228-027, Washington D.C., July 2015.

Santiago, C., and Mueller, E., "Pilot Evaluation of a UAS Detect-and-Avoid System's Effectiveness in Remaining Well Clear," Eleventh UAS/Europe Air Traffic Management Research and Development Seminar, Lisbon, Portugal, 2015.