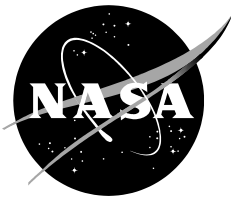


NASA/TM—2016—219396



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

*Eric R. Mueller
Ames Research Center, Moffett Field, California*

*Confesor Santiago
Ames Research Center, Moffett Field, California*

*Spencer Watza
Western Michigan University, Kalamazoo, Michigan*

October 2016

NASA STI Program ... in Profile

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

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

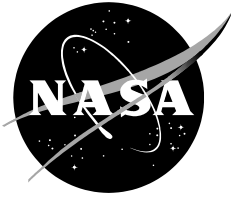
- **TECHNICAL PUBLICATION.** Reports of completed research or a major significant phase of research that present the results of NASA Programs and include extensive data or theoretical analysis. Includes compilations of significant scientific and technical data and information deemed to be of continuing reference value. NASA counterpart of peer-reviewed formal professional papers but has less stringent limitations on manuscript length and extent of graphic presentations.
- **TECHNICAL MEMORANDUM.** Scientific and technical findings that are preliminary or of specialized interest, e.g., quick release reports, working papers, and bibliographies that contain minimal annotation. Does not contain extensive analysis.
- **CONTRACTOR REPORT.** Scientific and technical findings by NASA-sponsored contractors and grantees.
- **CONFERENCE PUBLICATION.** Collected papers from scientific and technical conferences, symposia, seminars, or other meetings sponsored or co-sponsored by NASA.
- **SPECIAL PUBLICATION.** Scientific, technical, or historical information from NASA programs, projects, and missions, often concerned with subjects having substantial public interest.
- **TECHNICAL TRANSLATION.** English-language translations of foreign scientific and technical material pertinent to NASA's mission.

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

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

- Access the NASA STI program home page at <http://www.sti.nasa.gov>
- E-mail your question to help@sti.nasa.gov
- Phone the NASA STI Information Desk at 757-864-9658
- Write to:
NASA STI Information Desk
Mail Stop 148
NASA Langley Research Center
Hampton, VA 23681-2199

NASA/TM—2016—219396



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

Eric R. Mueller
Ames Research Center, Moffett Field, California

Confesor Santiago
Ames Research Center, Moffett Field, California

Spencer Watza
Western Michigan University, Kalamazoo, Michigan

National Aeronautics and
Space Administration

Ames Research Center
Moffett Field, CA 94035-1000

October 2016

Abstract

This study evaluated the performance of four prototype unmanned aircraft detect-and-avoid (DAA) display configurations, each with different informational elements driven by alerting and guidance algorithms. Sixteen unmanned aircraft pilots flew each combination of the display configurations, with half being given zero DAA surveillance sensor uncertainty 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 in altitude and heading bands had significantly fewer losses of well clear compared with alternative displays that lacked that information. This difference was significant from a statistical and practical perspective: 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. A modest level of DAA surveillance sensor uncertainty did not affect the proportion of losses of well clear or their severity. It is recommended that DAA traffic displays implement a band-type display in order to improve the safety of UAS operations in the National Airspace System. Finally, this report provides pilot response time distributions for responding to DAA alerts.

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. A DAA system consists of surveillance sensors that detect proximate air traffic, algorithms that determine whether and how to avoid those aircraft, and a traffic display that allows a pilot to evaluate, select and execute avoidance maneuvers. 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 (FAA 2013). A risk ratio is equal to the proportion of separation violations that occur with the DAA system relative to those that would have occurred without it. 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. The risk ratio of pilot-in-the-loop DAA systems is not known and is difficult to estimate. It is also unknown which parameters of the DAA system most improve the pilot’s ability to maintain separation, improving the risk ratio. The relationship between the DAA system’s design and its risk ratio must be determined before DAA operational and performance requirements can be finalized.

Accurate estimation of UAS DAA system risk ratio requires aircraft encounters that are representative of those likely to be encountered in actual operations. A detailed statistical analysis of such encounters is beyond the scope of this study and generally not used in pilot-in-the-loop studies due to the large number of encounters required to accurately replicate real-world encounter distributions. Instead, the performance of pilots employing DAA systems can be assessed and compared across experiment conditions. While observed mitigation levels cannot be directly employed to estimate risk ratios in actual operations, significant performance differences between experiment conditions are indicative of performance differences that should be expected in similar encounters in actual operations. Such observed differences may be used by standards developers and system builders to improve DAA system performance across a range of operational conditions, but are not in themselves sufficient to determine the risk ratio used in a safety case. To determine the risk ratio of a DAA system algorithm requires both statistical airspace encounter models (e.g., see Weinert 2013) and models of human pilot performance using DAA systems to remain well clear. To the extent that observed pilot performance in this study is again representative of what is expected in actual operations, models of pilot response derived from this study are a vital component in improving risk ratio estimates in Monte Carlo simulations.

Significant research effort has been devoted to understanding the technical and operational requirements for safely integrating UAS into non-segregated airspace (Dalamagkidis 2012, Ostwald 2007, Gillian 2012, Hesselink 2011, Fern 2015, Rorie 2015, Santiago 2015, and Rorie 2016). Many DAA system evaluations were part of a development program to produce a deployable system, so it is unclear how to generalize particular aspects of those designs to general requirements for future systems (Calhoun 2014). Other DAA systems have undergone human-in-the-loop testing to collect pilot-automation data that may generalize to new systems, but they have not been evaluated in the context of the development of an overall safety case for DAA-equipped UAS in the NAS (Bell 2012). Several studies preceding the research described in this paper directly contributed to the development of its displays and algorithms and influenced the experiment’s design. The first study indicated that a basic information display allowed more violations of the separation standard than a display that included maneuver guidance tools (Fern 2015, Santiago 2015).

The basic information display also produced longer pilot response times (Fern 2015). A follow-on study looked at the benefits of each display’s guidance features, finding that directive maneuver recommendations and an interactive “trial planning” capability significantly reduced the proportion of separation violations (Rorie 2015, Santiago 2015). Related research efforts focus on determining minimum DAA surveillance performance requirements and the effects of different sizes and types of separation standards (Park 2014 and Johnson 2015). Together with the research reported in this paper, these studies are expected to facilitate validation of the airspace safety case and provide a baseline set of DAA system performance standards.

The contribution of the present work is to quantify the performance of the piloted DAA system with respect to 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 airspace safety threshold, a set that may not be unique for all UAS platforms. The system performance is calculated through simulation of encounters between UAS, which are flown from ground control stations by professional pilots, and other aircraft that would result in close encounters. The pilot uses the DAA system to determine and execute a resolution, and metrics are recorded that relate to the success of the maneuver. 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, detailed statistical encounter modeling, and proper estimation of DAA system risk ratio in support of the safety case. This capability is necessary to demonstrate that a particular set of DAA requirements—including display features, surveillance sensor biases and errors, and alerting and resolution parameters—meets the safety threshold under all foreseeable operational conditions. Moreover, results related to the accuracy 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 (Fern 2015 and Rorie 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.

II. Methodology

The objective of the evaluation was to measure the self-separation efficacy of candidate DAA systems in which the primary distinguishing factor between the experimental conditions was the set of traffic display elements available to the pilot. Pilots flew several simulated UAS missions with each of the DAA systems, encountering six to eight aircraft per mission. The pilots used the DAA systems to determine whether a maneuver was necessary, coordinated that maneuver with an air traffic controller, and then followed the maneuver until it was safe to return to the mission flight plan. Objective metrics related to the degree of separation achieved by the UAS and the time required by the pilot to execute the maneuver were recorded, along with subjective feedback from the pilots on the DAA display designs. Half of the pilots were presented with perfect surveillance information and half the pilots observed intruder locations and velocities with modeled sensor uncertainties.

A. Well Clear Separation Requirement

It is typical for an aircraft separation function to have an associated separation requirement that it is trying to maintain and which, if violated, constitutes a “failure” of that function. The separation requirement for a DAA system is referred to as “well clear” (FAA 2013). The effects of different definitions of well clear have been evaluated in detail (Johnson 2015), and the final definition accepted by the community developing standards for DAA systems (RTCA 2013) is described in a separate paper (Cook 2015). In simplified terms, for any encounter in which the predicted horizontal miss distance (*HMD*) and current vertical separation of the aircraft (*ZTHR*) are less than the values given in Table 1, a loss of well clear (LoWC) has occurred if the value of modified tau (τ_{mod}) is less than 35 sec, where τ_{mod} is given by,

$$\tau_{mod} = \frac{r^2 - DMOD^2}{r\dot{r}} \quad (1)$$

where *r* is the range between aircraft, \dot{r} is the range rate between the aircraft, and *DMOD* is the distance modification parameter. This definition of modified tau is an approximation of the remaining time until the two aircraft will come within a distance *DMOD* of each other. The parameters specified in Table 1 are used as the basis for judging whether pilots using the DAA systems were successful in maintaining the appropriate separation and, in the cases they were not successful, the severity of the failure.

Table 1. Well clear parameter values

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

B. 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 (i.e. outside restricted airspace, operating IFR under a certificate of authorization or waiver). 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. The controllers were instructed not to decline the pilots' maneuver requests, however the interactions between the pilots and controllers provided realistic latency that could affect the success of the maneuvers. A companion paper presents additional details regarding pilot training procedures, instructions for the use of the DAA system, and metrics related to pilot performance (Rorie 2016).

C. Simulation Environment






The airspace simulator used in this experiment is composed of distinct hardware and software modules that provide the flexibility to test in flight, in batch simulations without human participants, or, as in this experiment, with simulated aircraft and real human pilots and air traffic controllers (Murphy 2015). The multi-aircraft control system (MACS) 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 (Prevot 2002). An additional ATC confederate would take control of intruder aircraft in MACS according to an experiment script in order to create predicted LoWCs with the subject UAS.

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

D. DAA Alerts

The alerting scheme used in this study, which has been refined over several previous studies, is shown in Table 2 (Santiago 2015 and Fern 2015). The separation criteria are based on the well clear definition given in Section A, along with a buffer to mitigate the impacts of uncertainty. All of the DAA display conditions used these alert levels and quantitative criteria. In order to determine whether an alert is appropriate, a dead reckoning trajectory prediction is made and all violations of the separation criteria along that prediction are recorded. The actual alert issued to the pilot is determined by the highest alert level for which the separation criteria are violated within the given time until loss of well clear. The aircraft symbol on the DAA system's tactical situation display is changed to the appropriate icon and an aural alert is issued.

Table 2. 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)	N/A	N/A		N/A

E. Display Designs

The pilots in the experiment flew four separate missions, each time with a different DAA display and associated algorithm. These DAA system designs built upon the findings of previous studies (Santiago 2015 and Fern 2015) and incorporated enhancements suggested by past participants. A detailed description of the display designs may be found in a companion paper (Rorie 2016).

The first display condition, called “information only,” provided basic information about each intruder, including relative altitude, bearing and range, along with alerts about intruders predicted to lose well clear. No maneuver guidance was provided to the pilots in this display configuration. The intruder state and aircraft-specific alerting information is common across all four display configurations.

The second display configuration incorporated all the informational aids contained in the information-only configuration, and it also included a “vector planner” feature (see Figure 1). That feature, which was represented as an arrow that the pilot could rotate around his aircraft symbol at the center of the tactical situation display, 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 most severe alert level predicted over a time horizon of 85 sec.



Figure 1. Vector planner horizontal and vertical guidance displays

The remaining two display configurations, called Stratway+ and the Omni Bands, showed bands of color around the UAS symbol at the center of the tactical situation display 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 displayed 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 (see Figure 2). No differentiation between the alert 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.

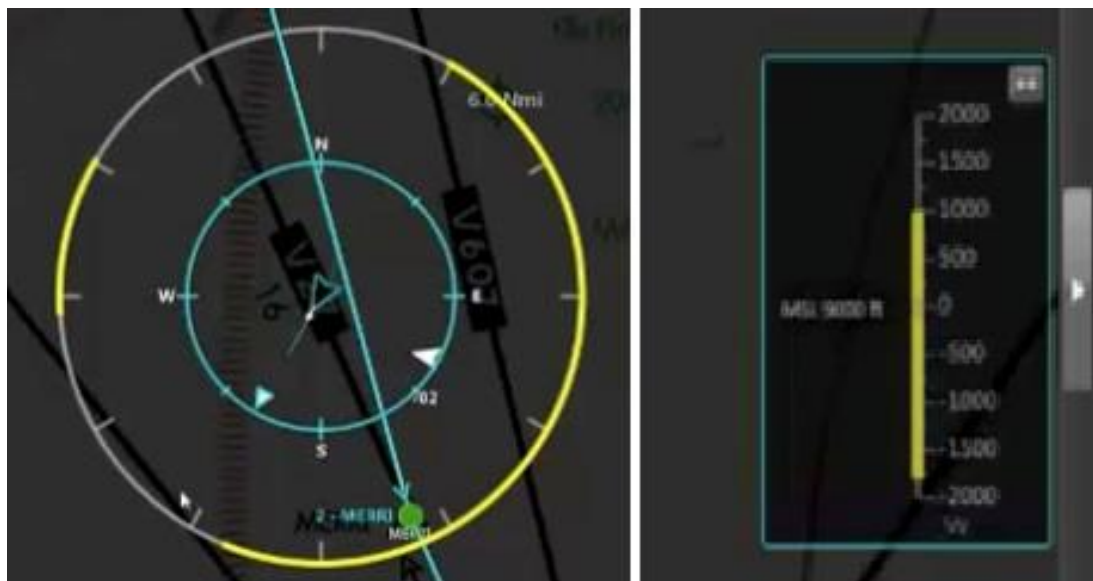


Figure 2. Stratway+ horizontal and vertical guidance displays

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 (see Figure 3). 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 algorithm underlying the alerts in all the displays, the guidance information in the vector planner, and the Omni Bands guidance may be found in a companion paper (Abramson 2016).

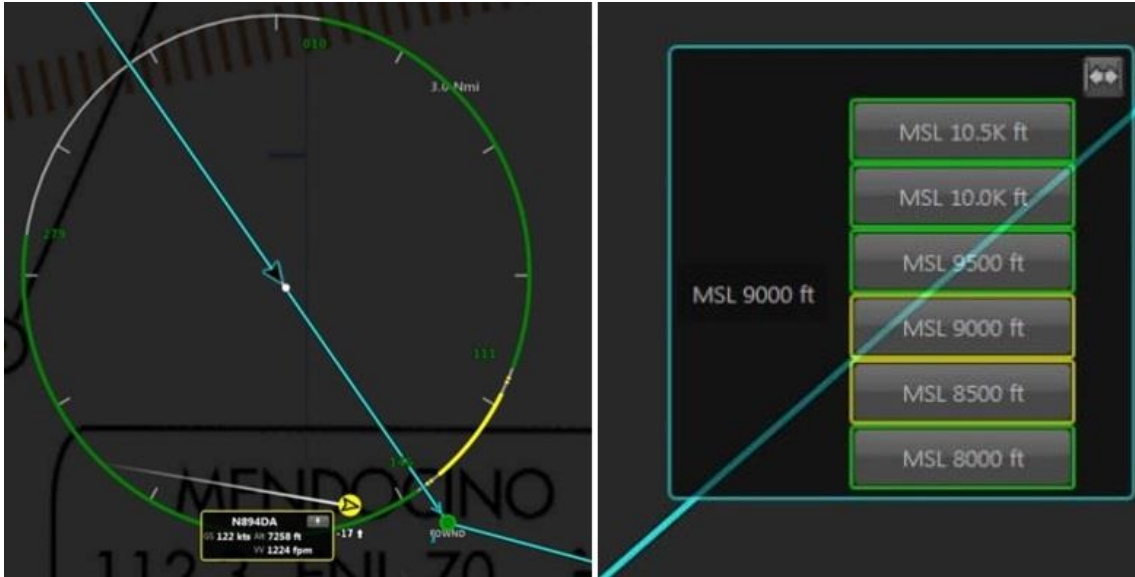


Figure 3. Omni Bands horizontal and vertical guidance displays

F. 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 certified DAA system. It was not designed to be high fidelity and did not capture complex dependencies between the relative bearing and orientation of the intruder aircraft, the available sensor types, and the sensor fusion algorithms. The objective in testing surveillance uncertainty was simply to determine whether significant differences in pilot acceptability and proportion of well-clear violations would occur with the introduction of a modest level of uncertainty.

The DAA surveillance sensor model consisted of two steps: sampling of a Gaussian distribution and averaging of a limited time history of samples to achieve a current state estimate. The Gaussian sampling is characterized by a bias error and variance, with each sample uncorrelated to those before or after it. The moving average filter step ensures that the measurements are correlated to each other over time; see the diagram in Figure 4. 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.

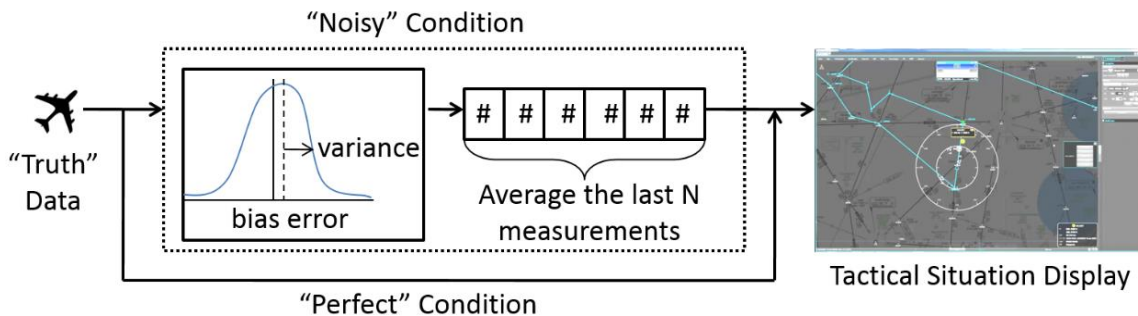


Figure 4. DAA surveillance sensor error model.

The sensor model was calibrated using ADS-B and airborne radar flight-test data. The differences between the state measurements as collected by the two sensors were used to calibrate the model. 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. The level of error tested in this simulation, which could be higher or lower than the error in a future, certified system, is quantified according to metrics contained in Figure 5 and Table 3 so that designers of future systems can estimate the level of impact that their system’s surveillance error will have on a pilot’s ability to successfully separate from

intruders using the DAA system. If such a future system has lower error than that tested here, then it is likely the impact on the DAA system’s effectiveness will also be lower than what was measured in this simulation.

The flight test and simulation errors were compared using two metrics: the actual absolute error in either the lateral or vertical estimate of an intruder’s position and 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. Because no alerting hysteresis was employed in this simulation, such errors cause alerts to appear and disappear as predictions move inside and outside the well clear boundary. The inconsistent movement of targets on the display makes it difficult for pilots 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 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 5 and Table 3.

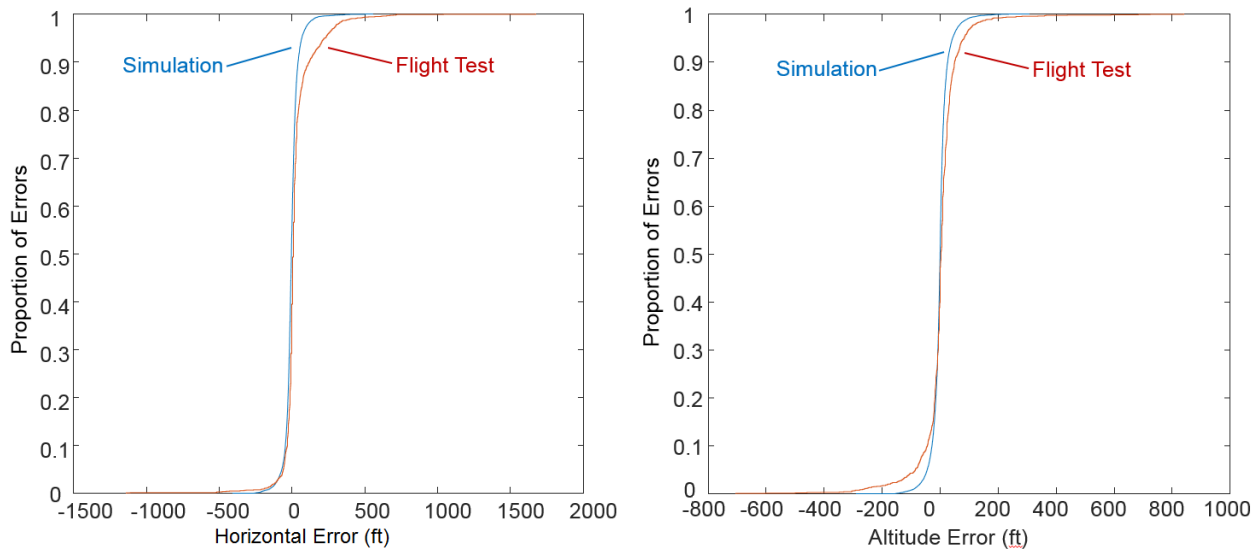


Figure 5. Comparison of non-cooperative surveillance system model and flight test errors.

Table 3. Comparison of non-cooperative DAA surveillance sensor accuracy between flight test and simulation

	<i>Flight Test</i>	<i>Simulation</i>
Mean Differential Horizontal Error (ft)	54.48	30.61
Median Differential Horizontal Error (ft)	20.08	19.03
Mean Differential Altitude Error (ft)	36.47	18.28
Median Differential Altitude Error (ft)	17.80	10.06

The error model used in the simulation was dependent on the equipment of intruder aircraft. Aircraft equipped with a Mode-C transponder are called “cooperative” intruders, while those lacking a transponder and therefore only detectable by a radar system are called “non-cooperative.” A key difference between the two levels of equipment is that a cooperative aircraft broadcasts its own altitude quantized in 100 ft bins, which makes the measured altitude of such aircraft relatively stable. The measured altitude of a non-cooperative aircraft is subject to the radar’s angular measurement accuracy and so can vary significantly from measurement to measurement. The accuracy of the two systems in the horizontal dimension is not significantly different.

G. Loss of Well-Clear Categories

Although all LoWCs are serious events, knowing the underlying reason a LoWC occurred is key to understanding how to reduce the rate of LoWCs in a deployed DAA system. For example, knowing that a LoWC occurred because an intruder changed altitude shortly before the LoWC will suggest a different remedy than a

LoWC that occurred because the pilot returned to course before the conflict had been resolved. For these reasons, six 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 estimate the overall pilot-DAA system performance in the real world by correcting for unavoidable simulation artifacts. LoWCs were assigned to one of six categories, and these categories are each classified according to whether or not they were the pilot's responsibility.

1. **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."
2. **Slow pilot response** - The losses in this category were due to slow creation and execution of encounter resolution maneuvers. They are distinguished from LoWCs in the "ineffective maneuver" category by the fact that the maneuvers 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 (pilots were instructed to maneuver without ATC coordination when necessary, but the point at which to maneuver without an amended clearance was left to pilot judgment). This category is assigned to "pilot's responsibility."
3. **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."
4. **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 probability that it 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."
5. **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 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 vertically separated by 500 ft would have appeared to have separation less than 450 ft at some instant during the encounter because of the surveillance system error. This category is *not* assigned to "pilot's responsibility."
6. **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.”

H. Well Clear Severity Metrics

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, no single metric has been shown to equate LoWC severity with the elevated level of collision risk that accompanies the LoWC. The well clear penetration integral¹ (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).

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

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 HMD, ZTHR, and τ_{mod} values specified in Table 1. The purpose of the τ_{mod} 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 severity of the LoWCs, and the time required by pilots to respond to traffic alerts and to coordinate and implement avoidance maneuvers.

A. Loss of Well Clear

The proportion of LoWCs out of all encounters that were predicted to become LoWCs without pilot intervention is shown in Figure 6. This metric is equivalent to the risk ratio of the DAA self-separation function (the ratio of losses of well clear with the DAA system to those without it) for the set of encounters modeled in this study. Only those losses of well clear that were deemed to have been the pilot’s 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 (though the data is retained and displayed as part of other analyses). The information-only display condition resulted in a proportion of LoWCs two to four times higher than display conditions with additional pilot aids, a result confirming the findings of previous studies (Santiago 2015). It appears that to reduce the proportion of LoWCs, 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 banding guidance displays (Stratway+ and Omni Bands) had similar performance at around 5% LoWC, while the vector planner condition had approximately twice this proportion of LoWCs. The proportion of LoWCs that occurred in the banding display conditions is clearly lower than the two alternative designs, suggesting improvement in DAA performance may be achieved by using these displays concepts.

¹ The WCPI was originally proposed by Bihle Applied Research, but it has not been published.

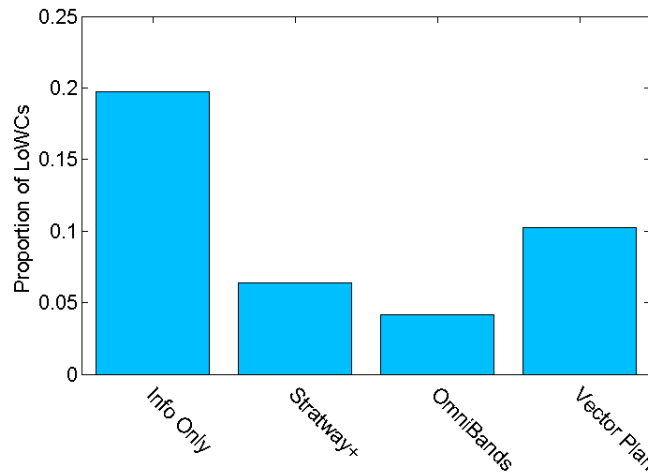


Figure 6. Proportion of losses of well clear by display condition. Only LoWCs that were the pilot's responsibility are included.

Although the results in Figure 6 appear to indicate that the banding displays are responsible for reducing the proportion of LoWCs compared with the other two display conditions tested, it is important to apply a test for statistical significance to confirm this observation. Because the proportions of LoWCs per scenario in each display condition are not normally distributed (i.e. the frequency of LoWCs does not follow a Gaussian distribution, partly because it is impossible to have a negative proportion of LoWCs in a scenario), the common analysis of variance test is not appropriate. Instead, we apply the non-parametric Kruskal-Wallis test of significance (Kruskal 1952 and 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 (p-value) for proportion of LoWCs by display condition

Display Condition	Info Only	Stratway+	Omni Bands	Vector Planner
Info Only	1.0	0.0159	0.0011	0.0866
Stratway+		1.0	0.2067	0.5569
Omni Bands			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 $p=0.05$ level, beige being marginally significant between $p=0.05$ and 0.10 , and the rest not being significant. Compared to the information-only display, both Stratway+ and Omni Bands are significantly different ($p=0.0159$ and $p=0.0011$, respectively) in terms of the number of LoWCs per scenario. The vector planner shows a marginally statistically significant reduction in the proportion of LoWCs ($p=0.0866$) compared with information only. The Omni Bands also have a marginally statistically significant reduction in number of LoWCs compared with the vector planner ($p=0.0902$). There is no significant difference between the two bands displays ($p=0.2067$). These results confirm the observations from Figure 6 that the banding displays significantly improve pilot performance in the use of the DAA function compared with the alternatives tested in this and previous simulations (Santiago 2015).

B. Losses of Well Clear by Alert Time and Range

Previous studies have shown that the likelihood of LoWCs is related to the time between first alert and the predicted LoWC, which is in turn related to the intruder aircraft's 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, only LoWCs caused by the LoWC types categorized as the pilot's responsibility are included (ineffective maneuver, slow pilot response, early return to the flight plan, diverging encounter; see Section II.G for a description of these categories). The effects of the other factors, which were beyond the pilot's control, on DAA system performance is determined in separate work through large-scale, fast-time simulations involving tens of thousands of encounters (Lee 2016). 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 any combination of 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 time to first LoWC at the first valid alert is shown in Figure 7. In the legend, reasons for LoWCs that were assigned to pilot responsibility are annotated with a “P” and those that were not the pilot’s responsibility have an “NP.” 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 2015), the proportion of LoWCs is relatively insensitive to alert time when more than 20 sec of warning is provided to the pilot. However, encounters that (without pilot action) would result in LoWC and whose first valid alert occurs with less than 20 sec remaining until first predicted LoWC result in an actual LoWC at a rate of 53% (40 of 75 encounters became LoWCs). When the first valid alert occurs within 10 sec of the predicted LoWC, the proportion of encounters that proceed to a LoWC is 60% (36 of 60 encounter became LoWCs). Figure 7 also depicts 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 (first valid alert received with less than 20 sec to LoWC) are due to late intruder maneuvers, which is precisely the reason the alert time was short, and so this is an intuitive 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 mitigated by providing increased alert time to the pilot, the early return to flight plan is not affected by the original alert time. Instead, as will be clear from the charts that partition these results by individual display conditions, improved DAA guidance was observed to eliminate this latter cause of LoWC in this simulation.

The proportion of predicted LoWCs that materialize when first 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 often were 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 growing towards their current heading and altitude and therefore were primed to act promptly when an alert was issued.

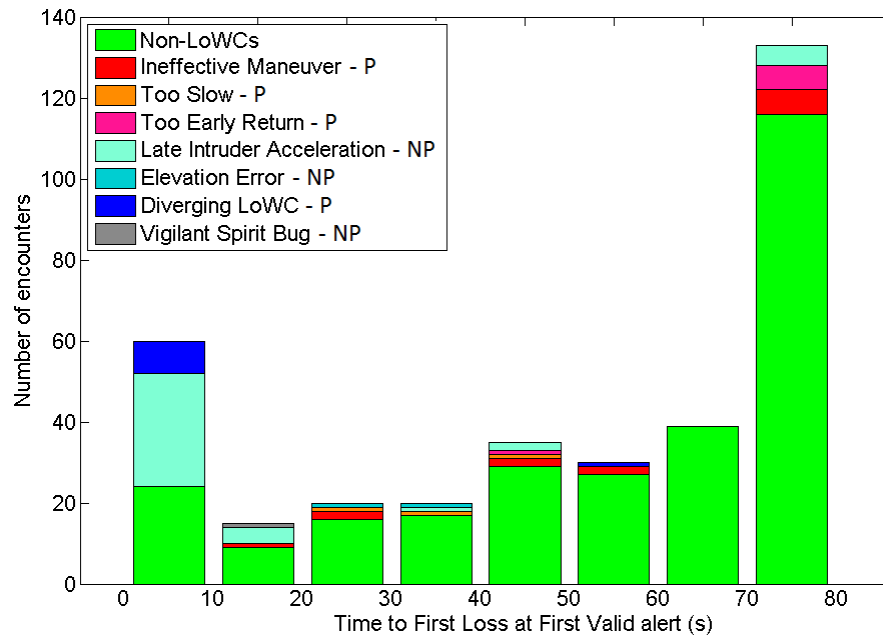


Figure 7. Outcomes of encounters by time to first loss of well clear at which an alert was provided. All displays.

Information-Only Display Condition

The outcomes of encounters in the information-only display condition are shown in Figure 8. The results inferred from this figure will be compared with aggregate results from the other displays that may be seen in subsequent figures. Just under 20% of all encounters in this condition became LoWCs (considering only LoWCs that were assigned to the pilot’s responsibility). In addition to a higher overall proportion of LoWCs, the proportion of LoWCs that occurred when less than 10 sec of warning time was available was higher (73%) as compared with the other three displays (45%). This indicates that with the information-only display, pilots are less able to anticipate potential LoWCs when

alerts are not timely. The failure to predict LoWCs is likely due to surveillance sensor uncertainty or a lack of guidance that shows alerts would occur with only small changes to relative heading or altitudes. The proportion of LoWCs when the full alert time was provided to the pilot (between 70 and 80 sec to LoWC) is very high at 25% (9 of 36) compared with the aggregate proportion of LoWCs for the other three displays (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 notably inferior to displays that provide explicit guidance.

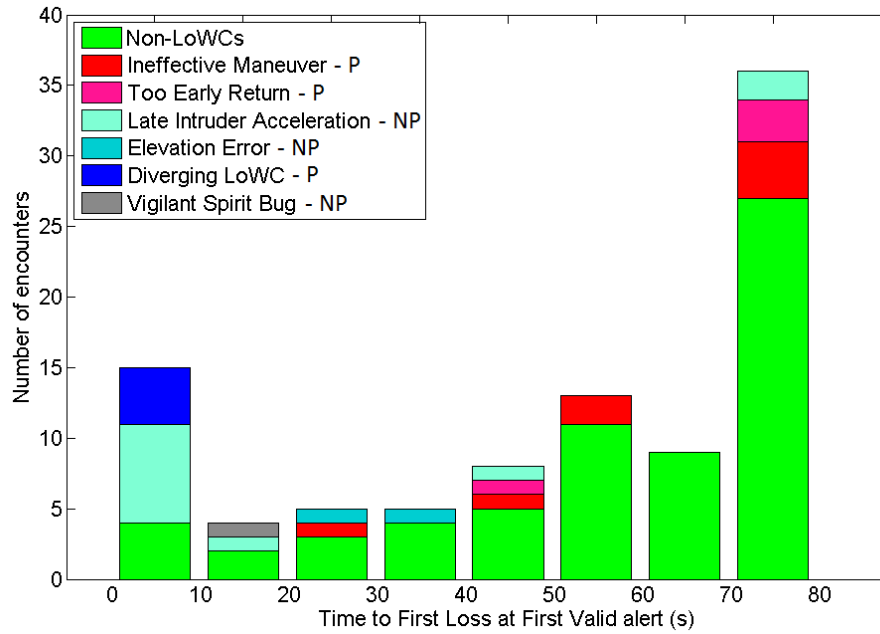


Figure 8. Outcomes of encounters by time to first loss of well clear at which an alert was provided. Information-only display.

Vector Planner Display Condition

The encounter outcomes for the vector planner display are shown in Figure 9. Recall that the overall proportion of LoWCs the pilot was responsible for was intermediate between the information-only display and the two banding displays at about 10% (reference Figure 6 for this result). The number of LoWCs due to ineffective maneuvers dropped, likely 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 two LoWCs attributed to slow execution of the maneuver, more than any other display, but the mean pilot response time for an alert was faster than for the information-only display (Rorie 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 proportion of LoWCs at both ends of the alert-time range and provided shorter pilot response times, the altitude and heading alerting information was not shown at all times the way it was on the banding displays and subjective feedback indicated this detracted from the display's effectiveness.

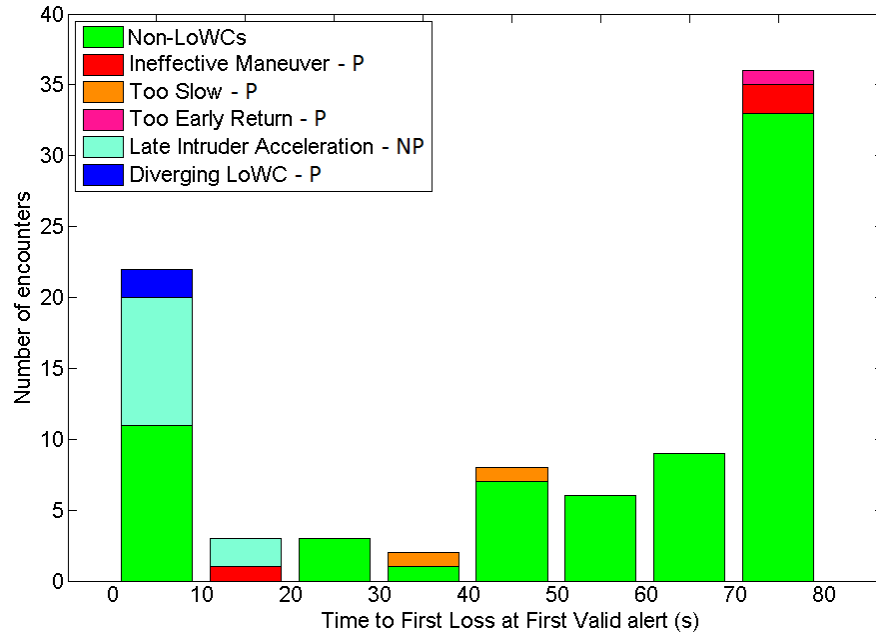


Figure 9. Outcomes of encounters by time to first loss of LoWC at which an alert was provided. Vector planner display.

Banding Display Conditions

The banding displays showed the lowest LoWC rates among the four display conditions, and the specific encounter outcomes for each are shown in Figure 10 and Figure 11. 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. This situation could happen during flight if the DAA trajectory prediction is not sufficiently well matched to the actual aircraft's dynamics, and could occur in simulation because the vigilant spirit control station used a different aircraft model than the DAA system. However, the overall proportion of pilot-responsibility LoWCs of 6% for the Stratway+ display is not significantly different from the 4% proportion of LoWCs provided by the Omni Bands display. Both rates of LoWC improve upon the non-banding displays, in some cases in a statistically significant way (see Table 4). As shown in Figure 11, the only pilot-responsibility LoWCs that occurred in the Omni Bands display condition were under diverging circumstances. These LoWCs can likely be avoided by changing the alerting logic so that it permits alerts even when aircraft are predicted to be diverging. Overall, the banding displays have lower LoWC proportions than all other displays tested in this or earlier simulations that used the same environment (Santiago 2015).

The determination of headings and altitudes that will lead to alerts, which is the core element of the banding displays, may be calculated using a variety of aircraft modeling approaches (Suarez 2012, Theunissen 2014, Consiglio 2015). A common approach is to use the current aircraft location and velocity with a set of candidate headings to predict a new trajectory, in effect commanding an instantaneous change in heading. However, this method has the potential to lead to predictions that headings are alert free when, in reality, by the time the aircraft has turned to that particular heading, it will no longer be alert free. This potentially inaccurate guidance could be avoided by simulating the turn segment of the aircraft trajectory until it reaches the new heading and subsequently continuing along a straight-line trajectory. The most severe alert at any step along this trajectory (whether in the turn or straight segments) would be used to indicate the alert level of that heading command. This higher-fidelity trajectory prediction should provide better information to the pilot, particularly in the close-range encounters when the highest alert level occurs on the turn segment.

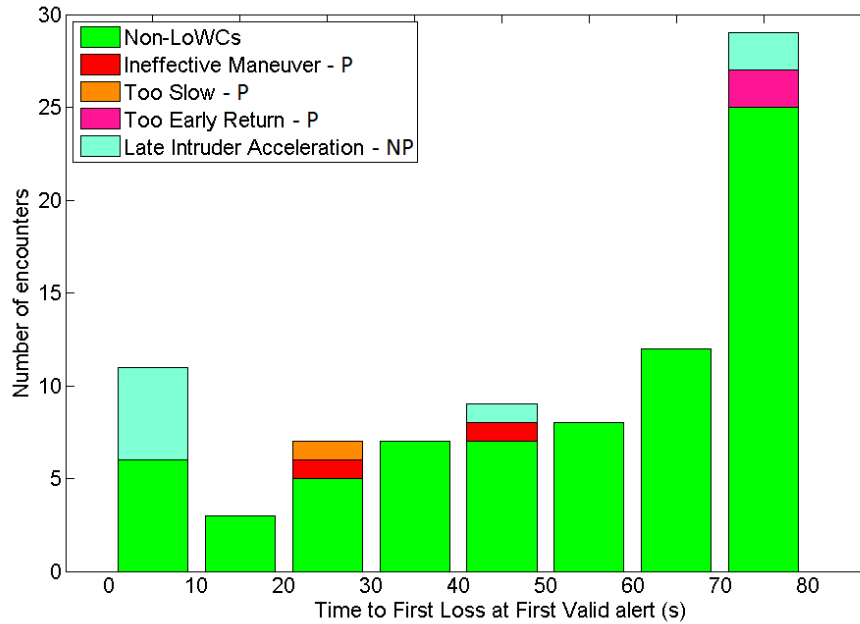


Figure 10. Outcomes of encounters by time to first loss of LoWC at which an alert was provided. Stratway+ display.

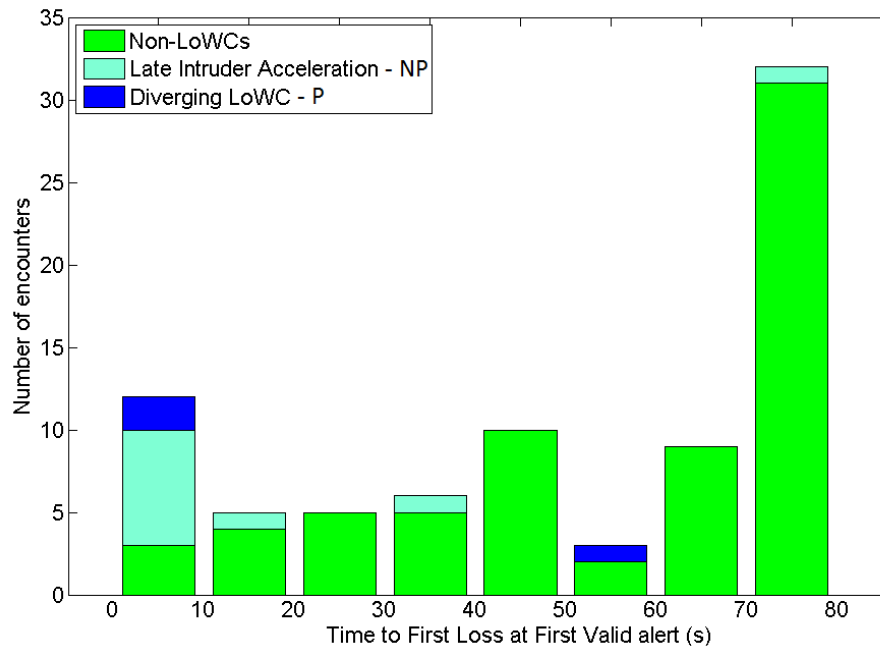


Figure 11. Outcomes of encounters by time to first loss of LoWC at which an alert was provided. Omni Bands display.

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. Other separation-related components of the DAA system and the NAS operate on distance rather than time. For example, the probability of detecting an intruder with an airborne radar is largely a function of range and does not depend strongly on the relative bearing or velocity of the intruder (the relative bearing, velocity and range together determine the time to LoWC). 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 12. An interesting difference between the range and time-to-CPA charts is that the highest concentration of LoWCs occur at and below 4 nmi range at first alert, with few

LoWCs at larger distances. In contrast, the time-to-CPA charts (Figure 7 - Figure 11) 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 (Mueller 2015).

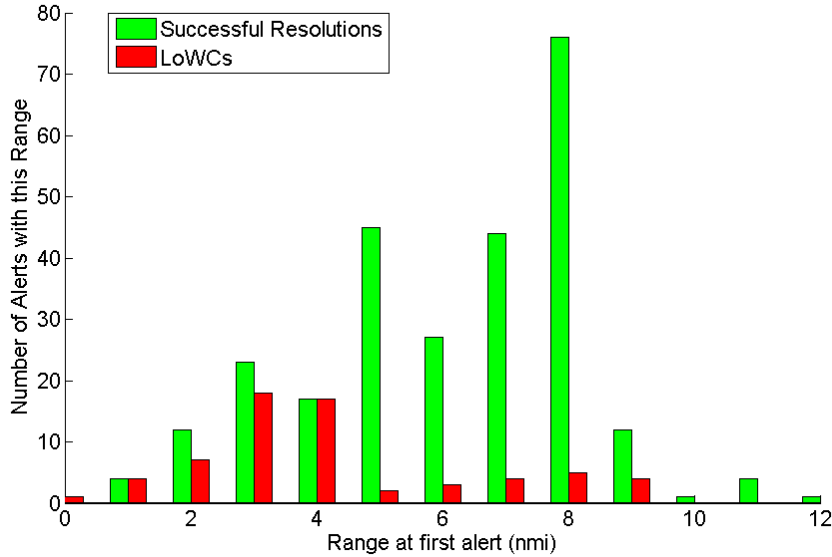


Figure 12. Outcome of encounter by range at which first alert was received.

C. Severity of Losses of Well Clear

The severity of violations for those encounters that resulted in LoWCs was evaluated in three ways: a severity index of the geometric proximity of the closest points 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 thresholds of the well clear definition, with values of 4000 ft and 450 ft, respectively. The maximum normalized separation is computed at each time step, t_i , and the minimum value over the entire trajectory becomes the separation index. Lower values of the index are more severe, with $S_{index} = 0.0$ being a direct collision. To compute the index for a trajectory prediction rather than an actual flown trajectory, the same equation is used but the minimization is done over every time step of every prediction. The lowest predicted separation index among all these predictions for encounters that resulted in LoWC is reported in Figure 13. Over all the displays, only 30% of LoWCs actually penetrated the 4000 ft horizontal and 450 ft vertical separation cylinder; the rest violated it with a modified tau under 35 seconds but maneuvered away before reaching the separation cylinder. The DAA systems overall raised the severity index from a median predicted value of 0.59 to a median actual value of 1.06, and none of the LoWCs became near-mid-air collisions. 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.

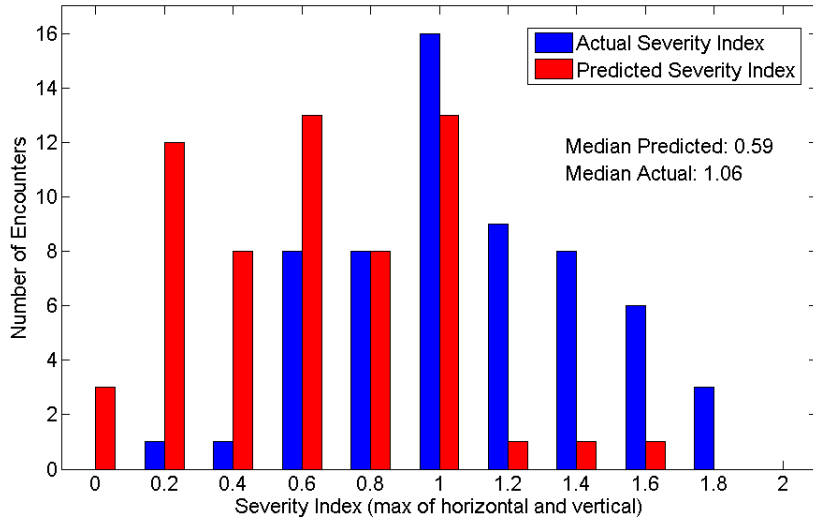


Figure 13. Predicted and actual mitigated geometric severity of all LoWCs.

The second method of evaluating LoWC severity is to examine the amount of time each aircraft spent in a LoWC or warning alert status. This temporal metric is useful because minor LoWCs are usually of short duration, while longer durations represent extended periods in which the two aircraft are in a state of heightened collision risk. Although the functional relationship between duration and collision 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 mitigate the alert when these situations occur, so the pilots are to some degree attempting to minimize this metric. 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 overall safety. In contrast warning alerts, pilots are told that action *will be* necessary when their aircraft is in a corrective alert state so that the encounter does not progress to a warning alert or LoWC; they are not told to initiate a maneuver to avoid a corrective alert before it occurs. For this reason it would be improper to evaluate a display according to the time spent in a corrective alert.

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 bins and plotted in the histogram in Figure 14. The results according to this metric largely parallel the LoWC proportion results from the previous section, 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.

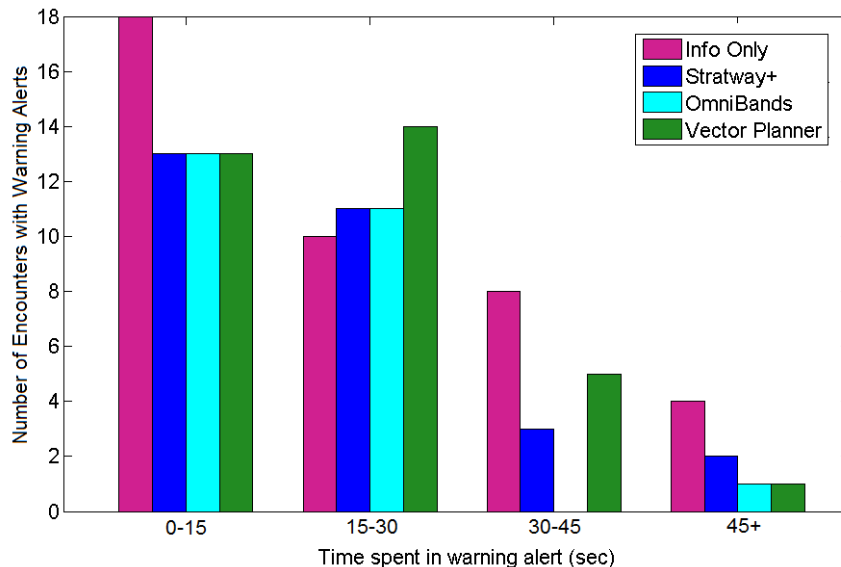


Figure 14. Number of encounters that spent the given amount of time in a warning alert status.

A histogram of the time spent in LoWC according to display condition is shown in Figure 15. This metric provides the same observations indicated by the previous results: 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 banding display is likely to provide the safest DAA system.

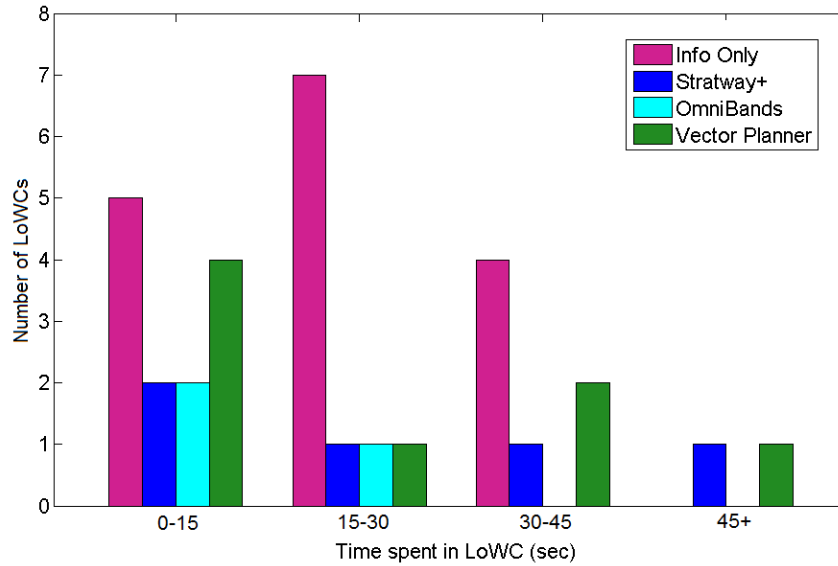


Figure 15. Number of encounters that spent the given amount of time in a LoWC status.

The WCPI (see Section II.H) was calculated for each trajectory that resulted in a LoWC. The results were then separated by display condition and binned by metric magnitude. The results of this process are shown in Figure 16. The performance of the display conditions according to this metric are similar to the previous metrics, though the relative magnitudes of the conditions' WCPI metrics 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 had a very low value of WCPI (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 banding displays performed the most successfully in helping the pilot avoid LoWCs and reducing the severity of LoWCs when they did occur.

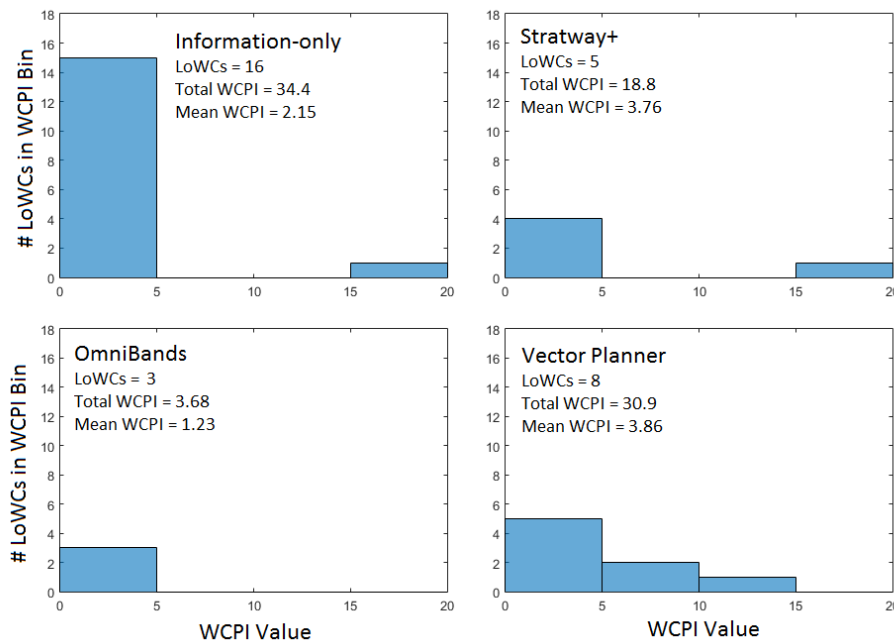


Figure 16. Histogram of the WCPI values by display

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 how often close encounters will occur in real operations and predict whether they would result in LoWCs. A detailed analysis of each step in a pilot's response sequence is presented in a companion paper (Rorie 2016), and a UAS pilot model is being developed using the identical alerting scheme and Omni Bands guidance designed for this experiment (Kuffner 2016). Those detailed models are necessary for simulations at higher levels 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 last two 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, and the gamma distribution was the best fit for both sets. Cumulative distributions of both the simulation data and the models are shown in Figure 17, 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.

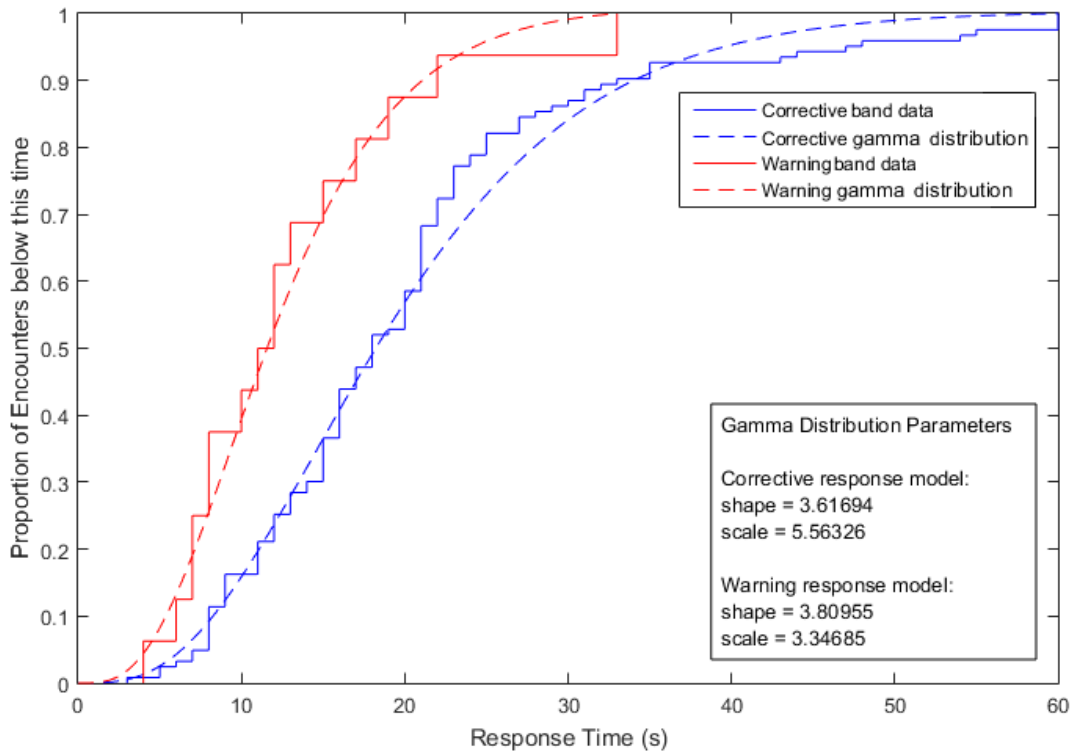


Figure 17. DAA maneuver response time data and models.

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 corrective and 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 altitudes of non-cooperative intruders and so are more likely to request vectors from ATC to avoid them.

IV. Conclusions

This study evaluated four UAS DAA display configurations, each with different informational and guidance elements. Sixteen UAS pilots flew each combination of the display configurations, with half being given zero DAA surveillance sensor uncertainty 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 results of the piloted simulation evaluation should allow DAA system designers to select combinations of alerting and guidance algorithms along with display configurations that reduce the airspace risk ratio relative to competing designs.

The displays that showed intruder alert information in altitude and heading bands had significantly fewer losses of well clear compared with alternative displays that lacked that information. This difference was significant from a statistical as well as a practical perspective. Compared with the information-only display, which provided alerting but no guidance information at all, the proportion of losses of well clear for the banded displays was lower 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. It is recommended that DAA traffic displays implement a band-type display based on the characteristics reported in this and companion papers (Rorie 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 proportion 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 an acceptable degree of separation 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. In *IEEE Digital Avionics Systems Conference*, 2012.

Calhoun, G., Miller, C., Hughes, T., and Draper, M. UAS sense and avoid system design and evaluation. In *Human Factors Ergonomics Society Annual Meeting*, 2014.

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

Consiglio, M., Munoz, C., Hagen, G., Narkawicz, A., Upchurch, J., Comstock, J., Ghatas, R., Vincent, M., and Chamberlain, J. Human-in-the-loop Experimental Research for Detect and Avoid. In *IEEE/AIAA Digital Avionics Systems Conference*, 2015. DOI 10.1109/DASC.2015.7311423.

Cook, S. P., and Brooks, D. A Quantitative Metric to Enable Unmanned Aircraft Systems to Remain Well Clear. In *Air Traffic Control Quarterly*, 23(2/3):137-156, 2015.

Dalamagkidis, K., Valavanis, K., and Piegler, L. A. On Integrating Unmanned Aircraft Systems into the National Airspace System: Issues, Challenges, Operational Restrictions, Certification and Recommendations, 2nd ed. Dordrecht: Springer-Verlag, 2012.

Daniel, W. "Kruskal–Wallis one-way analysis of variance by ranks". *Applied Nonparametric Statistics* (2nd ed.), Boston: PWS-Kent, 1990, pp 226-234.

Federal Aviation Administration (FAA). Sense and Avoid (SAA) for unmanned aircraft systems (UAS). In *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 *AIAA Guidance, Navigation and Control Conference*, 2008.

Fern, L., Rorie, R. C., 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. In *AIAA Aviation Conference*, June 2015.

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 Technical Report*, 2012.

Hesselink, H. H., and Schmitt, D. R. UAS Air Traffic Insertion Starts Now. National Aerospace Laboratory Technical Report, 2011.

Johnson, M., Mueller, E., and Santiago, C. Characteristics of a well clear definition and alerting criteria for encounters between UAS and manned aircraft in Class E airspace. In *Air Traffic Management Research and Development Seminar*, 2015.

Kruskal, W., and Wallis, A. Use of ranks in one-criterion variance analysis. *Journal of the American Statistical Association*, 47(260):583-621, 1952. DOI: 10.1080/01621459.1952.10483441.

Kuffner, M., Guendel, R. E., and Darrah, S. Investigating Traffic Avoidance Maneuver Decisions of Unmanned Aircraft Pilots. In *AIAA Aviation 2016 Forum*, 2016.

Lee, S. M., Park, C., Thippavong, D., Isaacson, D., and Santiago, C. Evaluating Alerting and Guidance Performance of a UAS Detect-and-Avoid System. NASA TM-2016-219067, 2016.

Mueller, E., Isaacson, D., and Stevens, D. Air Traffic Controller Acceptability of Unmanned Aircraft System Detect-and-Avoid Thresholds. NASA TM-2015-219392, 2015.

Murphy, J., Jovic, S., and Otto, N. Message Latency Characterization of a Distributed Live, Virtual, Constructive Simulation Environment. In *AIAA SciTech Conference*, 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. In *AIAA Aviation Technology, Integration, and Operations Conference*, 2014.

Prevot, T. Exploring the many perspectives of distributed air traffic management: The Multi Aircraft Control System MACS. In *International Conference on Human-Computer Interaction in Aeronautics*, 2002, pp. 149-154.

Rorie, R. C., Fern, L., and Shively, R. J. The Impact of Integrated Maneuver Guidance Information on UAS Pilots Performing the Detect and Avoid Task. In *Human Factors and Ergonomics Society Annual Meeting*, 59(1):55-59, 2015. DOI: 10.1177/1541931215591012.

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

RTCA, Inc. Special Committee 228, "Detect and avoid (DAA) Whitepaper," RTCA-SC228-WG1-WP01-01, Washington D.C., 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. In *Air Traffic Management Research and Development Seminar*, 2015.

Suarez, B., Kirk, K., and Theunissen, E. Development, Integration and Testing of a Stand-alone CDTI with Conflict Probing Support. In *AIAA Infotech@Aerospace Conference*, 2012. DOI 10.2514/6.2012-2487.

Theunissen, E., Suarez, B., and Uijt de Haag, M. The Impact of a Quantitative Specification of a Well Clear Boundary on Pilot Displays for Self Separation. In *IEEE Integrated Communications, Navigation and Surveillance Conference*, 2014. DOI 10.1109/ICNSurv.2014.6819987.

Weinert, A. J. , Harkleroad, E. P., Griffith,, J. D., Edwards, M. W. M., and Kochenderfer, M. J. Uncorrelated encounter model of the national airspace system version 2.0. Project Report ATC-404, Massachusetts Institute of Technology, Lincoln Laboratory, 2013.