JOURNAL OF AEROSPACE INFORMATION SYSTEMS
Vol. 18, No. 4, April 2021

Relating Sensor Degradation to Vehicle Situational Awareness for Autonomous Air Vehicle Certification

Donald H. Costello, III* and Huan Xu†
University of Maryland, College Park, Maryland 20742

https://doi.org/10.2514/1.I010905

Pilots use situational awareness (SA) to make appropriate aeronautical decisions. Autonomous vehicles will not have a human pilot (or operator) in the loop when off-nominal conditions present themselves, and will rely on sensors to build SA on their environment to make sound aeronautical decisions. As their sensors degrade, it is hypothesized that a point exists where the SA those decisions are based off will be inadequate for sound aeronautical decisions. It will be shown that this point can be identified through modeling and simulation of a simple sensor network to complete a task currently reserved for qualified pilots. This research highlights the process of determining an objective measure for the subjective end and relates it to a possible safety of flight certification for an autonomous system to perform tasks currently reserved for qualified pilots.

I. Introduction

MANY modern aircraft can, and are, operated through a set of pilot relief modes (i.e., autopilots) that allow the aircraft to complete nearly an entire flight without a pilot touching the controls and can be considered automation (which includes landing high-performance jet aircraft on the pitching deck of an aircraft carrier [1]). However, the pilot in command still has the responsibility for the aircraft and is expected to operate the vehicle under current certification standards. Pilots are trained to use their senses and experience to build their situational awareness (SA) while flying to enable them to safely accomplish their mission and make sound aeronautical decisions. The future of aviation is unmanned and ultimately autonomous. However, by eliminating the human pilot (or operator in the case of unmanned aircraft) we will be eliminating the SA that is currently required to safely accomplish a mission when off-nominal conditions present themselves. This paper defines an objective relationship between an autonomous vehicle SA when its sensors degrade and the ability to accomplish a task currently reserved for qualified pilots.

The first step in evaluating if the choices an autonomous system makes match that of a qualified pilot is to determine if the SA of the vehicle matches reality for the environment it is operating in. Since the dawn of aviation, many of the innovations we currently take for granted came from the military (some examples include radar [2], medevac air ambulance [3], jet engines [4], glow sticks [5], and advanced night vision technology [6]). Many military applications can transition easily to the civilian sector, as their functionality is similar. As both military and civilian pilots use SA, we elected to study SA of an autonomous vehicle completing a task currently reserved for qualified military pilots. In prior work we examined obtaining a safety of flight certification for a system that displays autonomous behavior. The underlying focus of our research relates to certifying an autonomous naval system to perform tasks currently reserved for qualified pilots. However, during our research we determined that in order for autonomous behavior to be certified it would need to demonstrate that it can make decisions similar to fully qualified pilots [7,8], to include situations where a fully qualified pilot makes decisions based on their SA when encountering off-nominal or unexpected conditions (such as degradation in the quality of information available to them).

Typically aircraft are designed to objective measures (i.e., maintain a desired speed at a desired altitude). During the certification process the system under test will be required to demonstrate that they can complete a subjective end (i.e., integrate with currently fielded systems). It is extremely difficult for designers to build an aeronautical system to accomplish a subjective end without an objective measure. This research focuses on developing a relationship between sensor performance degradation and vehicle SA in an attempt to establish an objective measure that can be provided to designers and certification officials for autonomous air vehicles to complete a task currently reserved for qualified pilots. This will enable certification officials to trust that an autonomous system has a clear understanding of the environment it is currently operating in, and will make appropriate aeronautical decisions (based off its programming) similar to those of a fully qualified pilot. We will develop an objective relationship between sensor degradation/error, through a modeling and simulation (M&S) environment, in a simple sensor network and sufficient SA for an autonomous vehicle to make a decision currently reserved for a qualified pilot. First, we will develop a scenario where an autonomous vehicle is reliant on its sensors to build its SA. The scenario will be built in such a way that the only factors affecting the SA of the vehicle are the accuracy of its sensors. We will then degrade those sensors to a point where the decisions it makes are no longer sound aeronautical decisions. And as a result of this work two inequalities (objective measure) are defined for when an autonomous vehicle has sufficient SA (subjective end) to make decisions currently reserved for qualified pilots.

The contributions of this paper are 1) the development of an objective measure for autonomous vehicle SA that accounts for sensor degradation; 2) the development of an scenario, within a Department of Defense (DoD)–recognized M&S environment, that specifically evaluates the effects of sensor degradation on error distance of a fused track of a threat aircraft; 3) the use of design of experiments (DOE) to determine the effects of sensor degradation and produce a predictive equations for the error distance of the fused track; and 4) use of subject matter expert (SME) opinion to define the point at which (within this scenario) the fused error distance is inadequate to make a decision currently reserved for qualified pilots.

This paper is structured as follows. In Sec. II, in addition to a review of related research in the area, we will discuss the issue of defining an objective measure for a subjective end to aircraft designers. We will discuss the evolution of handling qualities (HQ) specifications to include the use of the Cooper–Harper Rating (CHR) scale, which enables an objective value for a subjective task. We will also demonstrate how the scale was modified for the evaluation of a highly automated task and later used during the test and evaluation (T&E) of an autonomous aeronautical system. In Sec. III we will

Received 28 September 2020; revision received 3 December 2020; accepted for publication 21 January 2021; published online 10 February 2021.

*Ph.D. Candidate, Mechanical Engineering; Commander, U.S. Navy, U.S. Naval Air Warfare Center, Aircraft Division, Patuxent River, Maryland; donald.h.costello@navy.mil.

†Assistant Professor, Aerospace Engineering/Institute for Systems Research; mumu@umd.edu.

begin our use of DOE and an M&S environment to develop a quantitative relationship between sensor degradation and autonomous vehicle SA. In Sec. IV we will develop an objective measure for an autonomous vehicle's SA to accomplish a task, currently reserved for a qualified pilot, as its sensors degrade. In Sec. V, we summarize our findings as they relate to certifying autonomous systems to complete tasks currently reserved for qualified pilots. Directions for future research are also provided.

II. Background

This research focuses on developing a relationship between sensor performance degradation and vehicle SA (considered largely a subjective opinion). This is in an attempt to establish an objective measure that can be provided to designers, and certification officials, of an autonomous vehicle attempting to complete a task currently reserved for qualified pilots. Some related work is mentioned in Sec. II.A. Translating a subjective end into objective measures is not a new concept. Section II.B details how test pilots translate their opinion of the flying qualities of an aircraft into measures engineers can use to help improve the performance of the control laws via established rating scales. Section II.C details how the rating scales outlined in Sec. II.B have been transitioned to allow a test pilot to describe the behavior of an aircraft during a highly automated task (landing a high-performance jet aboard an aircraft carrier "hands-free"), and later for the evaluation of an autonomy demonstration vehicle. Similar to the use of ratings scales detailed in Sec. II, designers and certification officials will find an objective measures for subjective ends invaluable for evaluating the SA on an autonomous system.

A. Current Methods for Flight Certification and SA

This research focuses on the SA of an autonomous system as its sensors degrade. This will help build trust in autonomy, as without trust certification officials will be reluctant to grant a safety of flight certification for a system to operate without a human in the loop [9]. Currently a formalized/approved process does not exist for naval aircraft/systems that exhibit autonomous behavior (the system is able to respond to situations that were not explicitly preprogrammed) as there has never been a requirement for one to be developed. Several possible approaches have been proposed for autonomous control (dealing with the type of controller applied to the vehicle [10], updated the path planning based on sensor input [11], dynamically replan the flight path via adaptive controllers [12]) but none dealt with suboptimal sensor performance or were vetted through naval flight clearance authorities. Several issues have been identified for certificating autonomy (i.e., the complexity of autonomous systems results in an inability to test under all known conditions, difficulties in objectively measuring risk, and an ever-increasing cost of rework/redesign due to errors found late in the verification and validation (V&V) process [13]).

An understanding of SA as it relates to aviation is critical to understanding how it will relate to the certification of autonomy. One of the most commonly accepted definitions for SA is "the perception of elements in the environment within a volume of time and space, the comprehension of their meaning, and the projection of their status in the near future" [14]. During flight school, student naval aviators (pilots) are taught that SA in aviation is being able to accurately diagnose what is happening around them and predict what will happen in the immediate future, thus enabling them to perform the assigned mission safely. Students with high SA are able to "stay ahead of the aircraft," whereas students with low SA tend to seem to be "holding onto the stab" during flight. From their first flight, aviators learn to use every available resource to develop their SA (e.g., radio calls, aircraft instruments, visually scanning outside of the aircraft, onboard radar, electro-optical/infrared (EO/IR) sensors, and seat of the pants feelings). Before obtaining full qualification, naval aviators will have proven to their commanding officer (CO) that they can develop their SA to an appropriate level that they can safely complete their assigned mission during off-nominal conditions [15]. The measurement of SA has proven to be intangible, and largely subjective. Pilots quickly learn that the only way to know exactly the level of their current SA is when they realize that they have none. When a pilot's SA is high (i.e., they have an accurate understanding of the environment they are operating in) they can make sound aeronautical decisions. However, when a pilot's SA is low (which they may or may not know at the time) their aeronautical decisions may not be sound.

Autonomous vehicles will use their sensors to build SA of their environment. When sensors are operating at 100% the SA they provide the vehicle should be adequate to make sound aeronautical decisions. However, at some point of sensor degradation the SA provided will no longer match reality. The advent of unmanned aerial vehicles (UAVs) has sparked an increase in research within the academic and flight test communities. When programming UAVs with automation (such as what actions to take in the case of lost link), or autonomous functionality (allowing the vehicle to make decisions based off the conditions they sense), it is vital for the system to be able to safety complete the assigned mission. Sensors are typically installed to inform the operator, or system, of the conditions the vehicle is operating in. These sensors could be as simple as a camera, or as complicated as a fusion of multiple sensors. Increases in processing power have enabled these vehicles to perform simple missions (e.g., collision avoidance and visual navigation), under fairly static conditions, providing they have access to sensor inputs. However, when human pilots realize that there may be an issue with their SA, they have the training and experience to rely on various inputs to diagnose their current interpretation of reality. Unless a system is programmed to react to sensor degradation, certification officials will hesitate to allow the system to make decisions based on the sensor input without a human in the loop to ultimately shoulder the responsibility for the air vehicle.

For pilots to make sound decisions, they need to have a clear understanding of the situation/environment they are operating in [16]. Teaching prospective pilots how to develop their SA and knowing when to question their perception are critical portions of flight training [15,17]. Researchers have spent decades developing models and methods for evaluating a pilots' SA (highly subjective) during flight and translating it into an objective measure [16–20]. Two methods that have provided ample data for research involve freezing a simulation and asking questions relating to the pilot's SA or asking questions of a pilot postmission [17,19]. Yet, neither of these methods allows pilots to rate their SA in real time to determine when it is lacking. One school of thought was to offer pilots more information to help build their mental picture. Modern aircraft can present a massive volume of data to the pilot. However, this overload of information has a tendency to detract from the pilot's SA, and work has been done to optimize how the information is presented [21,22].

As UAVs have become commonplace in aviation, the issue of sufficient operator SA has become a hot button issue. How can operators maintain appropriate SA to their air vehicle when they are not actually in the vehicle (as a pilot is for manned aviation)? Several papers have been published regarding increasing the SA of a detached operator as to the environment the vehicle/system is currently operating in (to include the status of the vehicles subsystems) on Earth [23–30] and space [31,32]. As vehicle-based computing power has increased, research has been accomplished to demonstrate that a vehicle can navigate via onboard sensors (without direct operator direction) [33–38]. It has been proposed that, as the level of autonomy increases, the required level of SA for the human operator will decrease and the required SA of the air vehicle will increase [39,40]. However, the current body of work lacks the ability to demonstrate to safety of flight clearance officials the ability of an autonomous system to maintain SA while completing its assigned mission as sensor performance degrades.

B. Development of an Objective Measure (Cooper–Harper Scale) for a Subjective End (Handling Qualities)

This subsection is used to illustrate how an objective measure (the dynamics of an aircraft, e.g., short period) can be used to accomplish a subjective end (CHR of the aircraft handling qualities). Throughout the 1920s and 1930s, despite meteoric advances in structures, aerodynamics, and propulsion, aircraft HQ languished under the conception that it would not be feasible to create objective design standards (satisfying black-and-white requirements) to achieve a subjective ends (satisfying pilots' needs) [41]. Aircraft designers did not have a clear direction for what equated to positive HQ. By the 1940s the first HQ specifications were established, enabling aircraft designers to build aircraft that would have satisfactory HQ for pilots. The specification dealt with both longitudinal and lateral characteristics for the full range of aircraft configurations. One example of an objective measure that led to favorable HQ (subjective end) was placing a quantitative upper limit on the absolute value of the stick-force gradient [41]. For further details on the establishment of objective measures for subjective ends for the first HQ specifications we refer the reader to Chapter 3 of Ref. [41].

Determining an aircraft's HQ is a daunting task, as different pilots may have different opinions on this subjective judgment. During test pilot school (TPS), future test pilots are trained on classical test techniques to evaluate aircraft. One of the corner stones of this training is the Cooper–Harper Handling Qualities Rating Scale as it forces a pilot to make a series of relatively unambiguous decisions to arrive at a rating of the current HQ of the aircraft [42]. CHR is the basis of the U.S. flying qualities military specification (Mil-F-8785B, later superseded by 8785C [43]), and divides the pilots opinion of the aircraft HQ into four levels. Level 1 is satisfactory. Level 2 is not satisfactory HQ, but performance is satisfactory. Level 3 includes maximum workload to get adequate performance (and deals with aircraft controllability). Level 4 is uncontrollable [42–44]. CHR 1–3 equate to Level 1 HQ. CHR 4–6 equate to Level 2 HQ. CHR 7–9 equate to Level 3 HQ. CHR 10 equates to Level 4 HQ. Figure 1 is from Mil-F-8785C and illustrates how an objective measure (aircraft characteristics, short-period dynamics) can be related to a subjective measure (flying quality level). For further details on aircraft HQ we refer the reader to Ref. [44].

C. Cooper–Harper Adjusted for Confidence in Automation

CHR allows the flight test community a method of achieving repeatable results for HQ evaluations. The scale was later used as the blueprint for a rating scale that measures a test pilot's confidence of a vehicle accomplishing a highly automated task, landing high-performance jet aircraft on the pitching deck of an aircraft carrier without pilot input [1]. The Precision Approach and Landing System (PALS) installed on United States Navy (USN) aircraft carriers allow a pilot to "couple" with the ship and land during adverse conditions (e.g., extreme weather, or when pilots are unable to perform an arrested landing on their own). Figure 2 is the PALS/Pilot Quality Rating (PQR) used during PALS certification testing. PQR allows test pilots to put their subjective opinion (confidence in the system at accomplishing a task) into an objective measure (PQR rating). For certification, a PALS system must return a PQR of 3 or less. The PQR scale gives PALS engineers an objective measure (PQR rating) for a subjective end (pilot confidence in the system) to use as

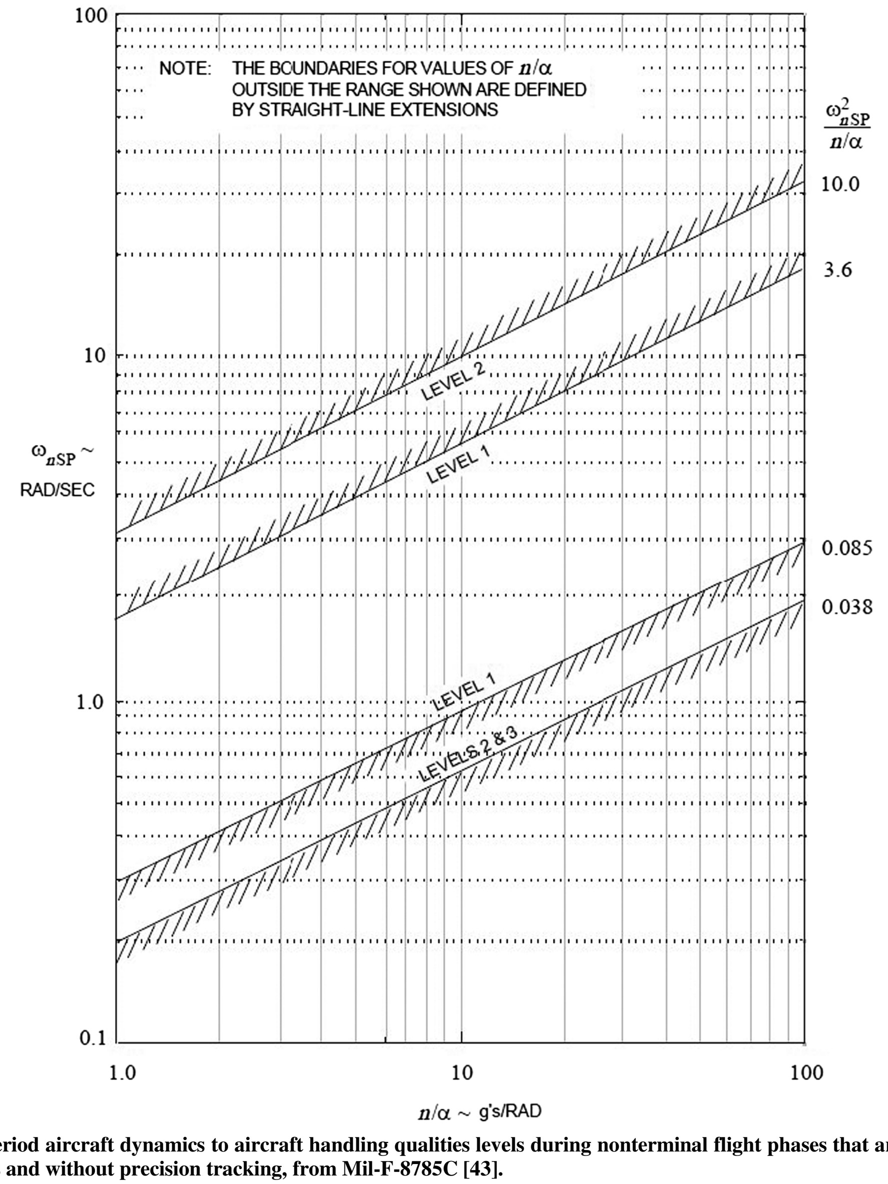

Fig. 1 Relating short-period aircraft dynamics to aircraft handling qualities levels during nonterminal flight phases that are normally accomplished using gradual maneuvers and without precision tracking, from Mil-F-8785C [43].

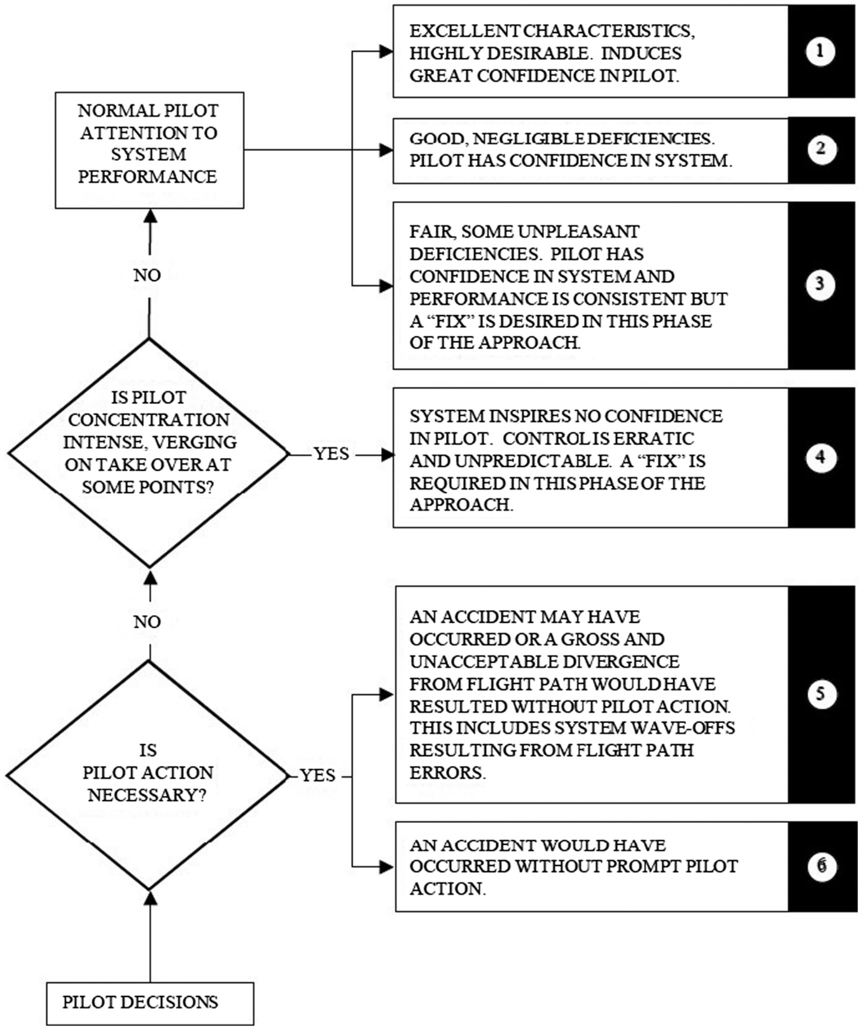

Fig. 2 PALS/Pilot Quality Rating Scale allows test pilots to objectively gage their confidence in the system under test [46].

they adjust the parameters within the system during certification testing [45].

PQR was later adopted by a flight test team evaluating an autonomous controller completing the United States Marine Corps (USMC) resupply mission in an optional piloted UH-1 helicopter during an autonomy demonstration program. The vehicle was able to use its onboard sensors (Global Positioning System [GPS], light detection and ranging [LiDAR], EO/IR cameras) to build SA and complete the resupply mission in both controlled and mission representative conditions [8]. However, on at least two occasions the safety pilot had to disengage the autonomous functionality due to safety of flight concerns when the systems SA did not match reality. The first instance required the safety pilot to take control when the vehicle was tracking dangerously close to trees (PQR-5). The second instance was on a later test flight and required the safety pilot to take control to avoid flying into trees (PQR-6). As these events occurred late in demonstration program, the test team decided to add a delay to the system before it started moving (to allow the onboard processors to spend extra time building SA of the environment). Once the delay was added to the system, further issues with path planning were not seen [47]. However, the fact that by simply adding a delay (giving the system more time to process its sensor data) to the system seemed to correct the inadequate vehicle SA implies that there may be a relationship between sensor degradation and SA in an autonomous system.

III. Problem Formation

In Sec. II.C we identified a possible relationship between sensor performance and vehicle SA in an autonomous system. In Sec. III we develop a relationship between sensor degradation and vehicle SA in an M&S environment through the use of DOE [48]. DOE has been used in the T&E of naval systems in the past. In a 2014 paper McCarley and Jorris used DOE during the investigation of an F/A-18 E/F strafing anomaly. In their work, DOE was used as a means of gaining the most statistical information from the fewest number of test points and ultimately generated a predictive equation that explained the strafing anomaly [49].

The United States Naval Test Pilot School (USNTPS) teaches DOE, and this section is structured to follow the steps of the process [50]. In Sec. III.A we will detail the M&S environment, give a statement of the problem, and detail the scenario we will be modeling. In Sec. III.B we will describe the choice of experimental factors (variables) and detail how we will measure them. In Sec. III.C we will discuss the measures of performance (MOP) for our experiment. Section III.D will detail how we plan to express the fused error distance as a function of the sensor errors.

A. M&S Environment and Statement of the Problem

As a truly autonomous system was not available for our evaluation, we elected to use a M&S environment for our research. Within the M&S environment we developed a scenario where an autonomous vehicle is reliant on its sensors to build its SA. The scenario was developed in such a way that the only factors affecting the SA of the vehicle are the accuracy of its sensors. Within the scenario the vehicle was required to make a decision, currently reserved for qualified pilots, based only on its degraded sensors.

For this experiment we used the Advanced Framework for Simulation, Integration and Modeling (AFSIM) environment. AFSIM is an engagement and mission level simulation environment written in

C++ originally developed by Boeing and now managed by the Air Force Research Laboratory (AFRL). AFSIM was developed to address analysis capability shortcomings in existing legacy simulation environments as well as to provide an environment built with more modern programming paradigms in mind. AFSIM can simulate missions from subsurface to space and across multiple levels of model fidelity [51]. As AFSIM has been used by both the USN and United States Air Force (USAF) to inform acquisition decisions and model aircraft system behavior, we elected to use it to generate evidence that may lead to certification of autonomous systems [52].

We propose the following scenario for analyzing the effects of sensor error on the SA of an autonomous vehicle: An autonomous UAV (we will refer to it as the Bucket Fighter) is operating over hostile territory. It is in a stationary orbit to provide intelligence, surveillance, and reconnaissance (ISR) information to ground forces. The information it provides is essential for the overall mission to be accomplished. However, the Bucket Fighter can be considered a high-value airborne asset (HVAA) that is unable to defend itself. As the platform is considered HVAA there is a set range it is required to maintain from threat aircraft. A fully qualified pilot is expected to take in the information available to them (both from communications with other assets and onboard systems) to determine when an aircraft reaches one of these prebriefed limits. When a threat aircraft reaches a defined range, the Bucket Fighter will be required to RETROGRADE (withdraw from station in response to a threat, continue mission as able). Once the threat is no longer a factor, the vehicle can RESET to its orbit. During a RETROGRADE, the ISR platform can continue to complete its assigned mission. When a threat aircraft reaches a defined range, the Bucket Fighter will be required to SCRAM (egress for defensive or survival reasons). If the UAV were to execute a SCRAM, it will no longer be able to provide support for ground forces, as a RESET is not authorized after a SCRAM. A description of these terms, and others used by the DoD, can be found in Ref. [53].

For the sake of this hypothetical scenario the researchers set the RETROGRADE and SCRAM ranges to 20 and 10 nautical miles (nm). An autonomous UAV's ability to accurately identity when a threat aircraft has reached its RETROGRADE and SCRAM range is critical for it to perform its mission. If it were to RETROGRADE or SCRAM too early, it may lead to an unacceptable degradation to the assigned mission (ISR support for ground forces). If it were to RETROGRADE or SCRAM too late, it may lead to a situation where a threat aircraft would engage the defenseless HVAA.

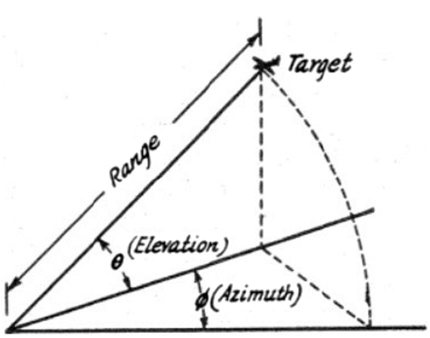

Fig. 3 Graphical depiction of the three possible error parameters of the sensors installed on the Bucket Fighter [54].

B. Experiment Factors (Variables)

Within the M&S environment, we installed two sensors on the Bucket Fighter: a generic infrared search and track (IRST) and a generic air-to-air radar. Both sensors were given an unlimited field of view and had the ability to track the threat aircraft for the duration of the simulation. In the M&S environment, we had the ability to apply errors into each sensor in the form of a σ (standard deviation [SD]) value. The error can be applied to the azimuth, elevation, and range of the track. Figure 3 is a pictorial of these parameters. It can be assumed that the only factors (environmental, mechanical, or other) that can cause degradation to the individual sensor can be illustrated by the errors detailed above.

During the scenario, the M&S package generates a random number to determine where on the normal distribution to pull the error value for each sensor component. This error shifts each time the individual sensor performs a sweep. The errors are constant at each point in the simulation of the same scenario to enable repeatable results (i.e., at 20 nm [on every run of the simulation] the simulation may pull a value of $0.73 \times \sigma$ regardless of what the σ value is).

The Bucket Fighter has the ability to fuse the tracks provided by the IRST and radar. This fused track is based not only on the raw sensor data, but it also uses velocity measurements and any past detection to build a predictable model for the track. This enables the autonomous UAV to more accurately track the target the longer it has been tracked by the sensors. It can be assumed that the fused track would match reality if there were zero error within the sensors. Figure 4 contains two AFSIM screen captures from a test run from different angles depicting the threat aircraft location based on the IRST (red), radar (green), and fused track (white triangle). The threat aircraft is approximately 20 nm west of the UAV [55].

For DOE we chose the factors to be azimuth error, elevation error, and range error as resident in the radar and the IRST. This will give a total of six factors in the experiment with one level each (six variables). For each of the six factors, we use the following null hypothesis: No statistical significance can be found between the "error value" (IRST/radar azimuth, elevation, range) and the error distance (distance between the fused track and the threat aircraft).

C. Measures of Performance

We are attempting to measure the SA provided as sensors degrade. Therefore, we elected to use error distance as the MOP in this research. In particular we will measure the error distance at 20 and 10 nm (corresponding to our hypothetical RETROGRADE and SCRAM range). Based on the errors inherent in the sensors (the six error σ s), we hypothesized that we could provide a predictive equation that would give the error distance at 20 and 10 nm. We will use SME opinion (four senior naval officers who have extensive

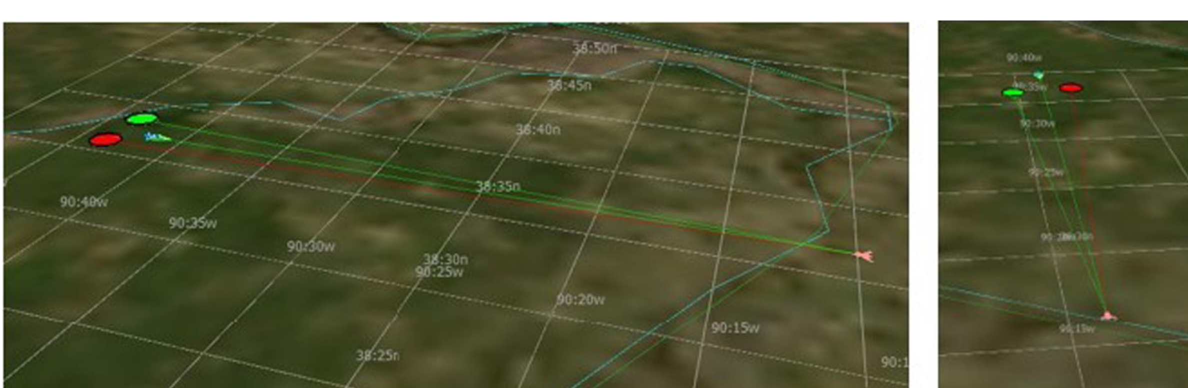

Fig. 4 Two AFSIM screen captures from a test run [55].

Fig. 5 Screen capture from the start of a test run. The threat aircraft is in the west and the Bucket Fighter (UAV) is in the east [55].

experience in dealing with RETROGRADE and SCRAM situations) to determine what error distance corresponds to inadequate SA to make a decision normally reserved for qualified pilots.

D. Fused Error Distance as a Function of Sensor Error

With the assistance of researchers from AFRL (Dayton, OH) and analysts from the Naval Air Warfare Center Aircraft Division (NAWCAD) (Patuxent River, MD) we adjusted a demonstration simulation from the standard unclassified AFSIM training software to meet the needs of our research. All output data from AFSIM used in this research were approved for public release [55].

We started with the Bucket Fighter providing ISR information to notional ground forces from a static location. We then elected to place a threat aircraft 60 nm from the Bucket Fighter. Both platforms were placed at 20,000 ft MSL and the threat aircraft tracked directly at the Bucket Fighter at 300 knots. For this hypothetical scenario, it can be assumed that the Bucket Fighter's only method of building an air picture (its SA of what is around it while airborne) is through its onboard sensors (a generic IRST and generic radar). Figure 5 is a screen capture depicting a top view from the start of the scenario with the Bucket Fighter in the east and the threat aircraft tracking inbound from the west. We studied the effect of sensor error on error distance (distance between the fused track and the actual location of the threat aircraft) at 20 and 10 nm in an attempt to quantify the SA level of the Bucket Fighter at critical decision points (RETROGRADE and SCRAM range) to determine at which point the SA provided to the Bucket Fighter was sufficient to make a sound aeronautical decision.

Table 1 Summary of the terms in the multiple regression model [Eq. (1)]

Term	Definition
Y	Error distance/independent variable
X_1	IRST azimuth σ value
X_2	IRST elevation σ value
X_3	IRST range σ value
X_4	Radar azimuth σ value
X_5	Radar elevation σ value
X_6	Radar range σ value
β_0	Y intercept
β_1	Weight of the X_1 variable
β_2	Weight of the X_2 variable
β_3	Weight of the X_3 variable
β_4	Weight of the X_4 variable
β_5	Weight of the X_5 variable
β_6	Weight of the X_6 variable
ε	Error that exists within the model

Equation (1) is the multiple regression model that explains the relationship between Y (error distance/the independent variable) and multiple X_X values (the six error σ s/dependent variables): $X_1 =$ IRST azimuth σ value, $X_2 =$ IRST elevation σ value, $X_3 =$ IRST range σ value, $X_4 =$ radar azimuth σ value, $X_5 =$ radar elevation σ value, and $X_6 =$ radar range σ value. The corresponding β_x values are the relative weights of each variable, and β_0 is the Y intercept. The ϵ term represents the error that exists within the model that cannot be accounted for and will drop out when we develop our predictive equation ($\hat{Y}$). Table 1 summarizes the various terms in Eq. (1).

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \beta_5 X_5 + \beta_6 X_6 + \epsilon \tag{1}$$

IV. Experimental Results and Analysis

In Sec. III we developed a multiple regression model where we express the fused error distance as a function of the various sensor errors in the sensor network. In Sec. IV.A we describe how we will gather data at various error σ levels to characterize the system. In Sec. IV.B we perform multiple variable regression analysis on the data gathered in the M&S environment to populate the variables in Eq. (1) at 20 and 10 nm. In Sec. IV.C we develop inequalities that define sufficient SA for an autonomous vehicle to make a decision that is currently reserved for qualified pilots.

A. Conduct of the Experiment

In an attempt to limit the scope of possible errors, and provide useful data to analyze, we limited the error σ to between three and seven (nm or deg). We programmed in the ability to introduce three error variables into each of the sensors (azimuth [deg], elevation [deg], and range [nm]) in the form of defining one σ for each variable. For this research we varied the six variables between three, five, and seven at the start of each test run and recorded the observed error distance (distance between the fused track generated by the autonomous UAV and actual location of the threat aircraft) within the M&S environment. By manually updating the six σ values with all 729 combinations between each run, we hope to provide enough data to generate predictive equations through multiple variable regression analysis. Equations (2) and (3) are the predictive equations (at 20 and 10 nm) we plan on population with the results of our regression analysis. Table 2 summarizes the various terms in Eqs. (2) and (3). The completed equations will be used to provide a quantitative evaluation of an autonomous systems SA to complete a task currently reserved for qualified pilots. Table 3 is a 10-run subset of the 729 combinations we plan on evaluating.

$$\hat{Y}_{20} = b_{0-20} + b_{1-20}X_1 + b_{2-20}X_2 + b_{3-20}X_3 + b_{4-20}X_4 + b_{5-20}X_5 + b_{6-20}X_6 \tag{2}$$

Table 2 Summary of the terms that are in the predictive equations [Eqs. (2) and (3)]

Term	Definition
$\hat{Y}_{20}$	Predictive error value at 20 nm
$\hat{Y}_{10}$	Predictive error value at 10 nm
X_1	IRST azimuth σ value
X_2	IRST elevation σ value
X_3	IRST range σ value
X_4	Radar azimuth σ value
X_5	Radar elevation σ value
X_6	Radar range σ value
b_{0-x}	Y intercept for the x equation (20 or 10 nm)
b_{1-x}	X_1 variable in the x equation (20 or 10 nm)
b_{2-x}	X_2 variable in the x equation (20 or 10 nm)
b_{3-x}	X_3 variable in the x equation (20 or 10 nm)
b_{4-x}	X_4 variable in the x equation (20 or 10 nm)
b_{5-x}	X_5 variable in the x equation (20 or 10 nm)
b_{6-x}	X_6 variable in the x equation (20 or 10 nm)

Table 3 Ten of the 729 data points

Run no.	Y_{20} (m)	Y_{10} (m)	X_1 (deg)	X_2 (deg)	X_3 (nm)	X_4 (deg)	X_5 (deg)	X_6 (nm)
50			3	3	5	7	5	5
141			3	5	7	3	5	7
248			5	3	3	3	5	5
339			5	5	3	5	5	7
397			5	5	7	7	3	3
469			5	7	7	5	3	3
554			7	3	7	5	5	5
594			7	5	3	7	7	7
656			7	7	3	3	7	5
723			7	7	7	7	3	7

Y_{20} is the 20 nm error distance in meters. Y_{10} is the 10 nm error distance in meters. X_1 is the value of one σ error in IRST azimuth in degrees. X_2 is the value of one σ error in IRST elevation in degrees. X_3 is the value of one σ error in IRST range in nm. X_4 is the value of one σ error in radar azimuth in degrees. X_5 is the value of one σ error in radar elevation in degrees. X_6 is the value of one σ error in radar range in nm.

$$\hat{Y}_{10} = b_{0-10} + b_{1-10}X_1 + b_{2-10}X_2 + b_{3-10}X_3 + b_{4-10}X_4 + b_{5-10}X_5 + b_{6-10}X_6 \tag{3}$$

B. Analysis of the Data

As discussed in Sec. IV.A, we planned on evaluating 729 different combinations of the six variables (error σ s). Table 4 is a subset of 10 (the same 10 as Table 3) of the runs with the observed error distance (measured in meters) at 20 and 10 nm.

Table 4 Ten of the 729 data points

Run no.	Y_{20} (m)	Y_{10} (m)	X_1 (deg)	X_2 (deg)	X_3 (nm)	X_4 (deg)	X_5 (deg)	X_6 (nm)
50	467.9	342.9	3	3	5	7	5	5
141	325.8	181.5	3	5	7	3	5	7
248	331.2	243.3	5	3	3	3	5	5
339	496.7	381.7	5	5	3	5	5	7
397	617.9	481.6	5	5	7	7	3	3
469	434.0	320.5	5	7	7	5	3	3
554	567.3	412.1	7	3	7	5	5	5
594	818.1	600.9	7	5	3	7	7	7
656	932.4	765.5	7	7	3	3	7	5
723	813.4	615	7	7	7	7	3	7

Y_{20} is the 20 nm error distance in meters. Y_{10} is the 10 nm error distance in meters. X_1 is the value of one σ error in IRST azimuth in degrees. X_2 is the value of one σ error in IRST elevation in degrees. X_3 is the value of one σ error in IRST range in nm. X_4 is the value of one σ error in radar azimuth in degrees. X_5 is the value of one σ error in radar elevation in degrees. X_6 is the value of one σ error in radar range in nm.

Table 5 Twenty and 10 nm regression data obtained through Microsoft Excel multiple regression analysis

Twenty-mile regression data			Ten-mile regression data		
$\hat{Y}_{20}$ b term	Coefficient	p	$\hat{Y}_{10}b$ term	Coefficient	p
b_{0-20}	−528.337	4.02E − 80	b_{0-10}	−441.693	5.90E − 73
b_{1-20}	57.427	3.80E − 123	b_{1-10}	49.665	9.80E − 119
b_{2-20}	54.105	2.31E − 113	b_{2-10}	46.854	2.10E − 109
b_{3-20}	4.653	1.91E − 02	b_{3-10}	−4.607	9.01E − 03
b_{4-20}	54.649	5.75E − 115	b_{4-10}	47.578	8.30E − 112
b_{5-20}	61.384	9.58E − 135	b_{5-10}	53.085	4.70E − 130
b_{6-20}	−10.208	3.31E − 07	b_{6-10}	−15.638	4.84E − 18

We then used multiple variable regression analysis resident in Microsoft Excel to preform regression analysis on the 729 data points to determine the effects each independent variable (the six σ values) has on the two dependent variables (error distance at 20 and 10 nm). The 20 nm data adjusted R^2 value (indicates the percentage of the variance in the dependent variable that the independent variables explain collectively) was 0.822, and the 10 nm adjusted R^2 value was 0.818. R^2 describes levels of predictive accuracy with 0.75, 0.50, and 0.25, respectively, describing substantial, moderate, or weak [56]. The analysis of variation (ANOVA) significance F value was 6.831E-267 for 20, and 8.674E-263 for 10 nm (both of which show an extremely high statistical significance for the respective model). Table 5 details the relative coefficients for the predictive equation and the individual p values. All of the p values are well below 0.05. Therefore, we must reject the six null hypotheses as there is a significant relationship between each sensor error value and the fused track error distance.

Equations (4) and (5) are predictive equations ($\hat{Y}$) that depict an anticipated fused error distance (dependent variable) based on the various error σ s (independent variables) internal to the system at 20 and 10 nm, respectively (X_1 = IRST azimuth σ value, X_2 = IRST elevation σ value, X_3 = IRST range σ value, X_4 = radar azimuth σ value, X_5 = radar elevation σ value, X_6 = radar range σ value). The corresponding b_X values are the relative weights of each variable, and b_0 is the Y intercept from Table 5.

$$\hat{Y}_{20} = -528.337 + 57.427X_1 + 54.105X_2 + 4.653X_3 + 54.649X_4 + 61.384X_5 - 10.208X_6 \tag{4}$$

$$\hat{Y}_{10} = -441.693 + 49.665X_1 + 46.854X_2 - 4.607X_3 + 47.578X_4 + 53.085X_5 - 15.638X_6 \tag{5}$$

Next, we used a random number generator (integers between three and seven) to populated 25 test points for the evaluation of the predictive equations. We elected to limit out evaluation of the regression analysis to σ s between three and seven, as that was the population of the data that we used for the regression analysis. Table 6 details these test points and their observed error distances at 20 and 10 nm. Table 7 then compares the predicted error distance and observed error distance from the M&S environment. The predicative equations generated error distances across the 25 points with less than a 10% average error at both 20 and 10 nm (distance between the observed range and fused track).

C. DOE Conclusions

Based on this output and SME (four senior naval officers who have extensive experience in dealing with RETROGRADE and SCRAM situations) opinion, we determined that if the system could generate an error distance less than 800 m at 20 miles, and 400 m at 10 miles, then the SA provided by its sensors is accurate enough for it to make the RETROGRADE or SCRAM decision normally reserved for qualified pilots. As the error distance from the predictive equation is within 10% of the observed error distance, we used 727 m for 20 nm (worst case: $727 + (727 * .1) = 799.6$), and 363 for 10 nm (worst case: $363 + (363 * .1) = 399.3$). Equations (4) and (5) were then translated to be inequalities, Eqs. (6) and (7). When Eq. (6) is true, the SA provided by the onboard sensors is sufficient to make a sound RETROGRADE decision at 20 miles. When Eq. (7) is true, the SA provided by the onboard sensors is sufficient to make a sound SCRAM decision. If Eq. (6) or Eq. (7) were to be false, the SA provided by the onboard sensors is not adequate for making a sound RETROGRADE or SCRAM decision.

$$727 > -528.337 + 57.427X_1 + 54.105X_2 + 4.653X_3 + 54.649X_4 + 61.384X_5 + -10.208X_6 \quad (6)$$

$$363 > -441.693 + 49.665X_1 + 46.854X_2 - 4.607X_3 + 47.578X_4 + 53.085X_5 - 15.638X_6 \quad (7)$$

Table 6 Results from 25 test runs of randomly generated σ values

Test run no.	Y_{20} (m)	Y_{10} (m)	X_1 (deg)	X_2 (deg)	X_3 (nm)	X_4 (deg)	X_5 (deg)	X_6 (nm)
T—1	476.5	362.9	7	5	4	3	3	4
T—2	588.5	458.0	5	6	4	6	4	5
T—3	537.4	411.0	5	5	4	6	5	6
T—4	307.0	195.0	3	5	5	4	4	6
T—5	451.1	338.6	4	5	4	6	4	5
T—6	396.5	271.7	7	3	6	3	4	5
T—7	517.5	384.6	3	5	5	4	7	5
T—8	481.0	367.9	5	5	3	4	5	7
T—9	291.6	210.8	5	3	4	3	4	3
T—10	493.6	387.0	6	6	4	4	4	3
T—11	394.0	291.6	6	4	4	3	4	4
T—12	859.2	687.7	6	7	5	5	7	4
T—13	750.7	571.3	3	7	7	5	7	5
T—14	824.5	645.3	7	7	6	5	6	5
T—15	628.6	463.5	7	3	6	5	6	6
T—16	470.0	345.3	4	5	5	6	4	5
T—17	425.8	289.8	5	3	7	6	3	5
T—18	463.8	322.0	5	3	5	6	6	7
T—19	742.9	586.2	4	7	7	4	7	3
T—20	571.8	400.6	7	3	6	4	6	7
T—21	874.1	710.4	7	6	3	7	5	7
T—22	394.7	279.7	3	6	4	5	3	7
T—23	540.5	418.4	7	4	5	6	3	3
T—24	471.1	345.9	5	6	5	4	4	5
T—25	384.8	262.7	5	5	5	3	4	6

Y_{20} is the 20 nm error distance in meters. Y_{10} is the 10 mn error distance in meters. X_1 is the value of one σ error in IRST azimuth in degrees. X_2 is the value of one σ error in IRST elevation in degrees. X_3 is the value of one σ error in IRST range in nm. X_4 is the value of one σ error in radar azimuth in degrees. X_5 is the value of one σ error in radar elevation in degrees. X_6 is the value of one σ error in radar range in nm.

V. Conclusions

This paper demonstrated that a relationship (objective measure) can be defined for autonomous vehicle SA (subjective end) and sensor degradation. Section II details how defining an objective measure for a subjective end is not a new idea within the flight test community and highlighted inadequate vehicle SA in an autonomous technology demonstration vehicle. Sections III and IV dealt with M&S of a hypothetical simplified sensor network to define the relationship. Future work within AFSIM that uses multiple Monte Carlo runs with random seed values for where on the σ curve to pull the error value on each run could build a more accurate error equation. Additionally, future work focusing on defining this relationship on a mature system during flight test would give vehicle designers the ability to program a vehicle to complete tasks currently reserved for qualified pilots under off-nominal conditions and eventually obtain a safety of flight certification.

References

[1] Lisa, M., and Staub, J., "EA-18G Sea Trials, Flying Qualities, Performance and Precision Approach and Landing Systems," *2008 Society of Experimental Test Pilots Annual Symposium*, Soc. of Experimental Test Pilots, Lancaster, CA, 2008.

Table 7 Results from 25 test runs Y_x, the corresponding results from of predictive equations $\hat{Y}_x$, and the absolute distance between the two in meters and percentage

Test run no.	Y_{20} (m)	$\hat{Y}_{20}$ (m)	Delta (m)	Delta (%)	Test run no.	Y_{10} (m)	$\hat{Y}_{10}$ (m)	Delta (m)	Delta (%)
T—1	476.5	460.1	16.4	3.44	T—1	362.9	361.2	1.7	0.46
T—2	588.5	614.5	26.0	4.41	T—2	458.0	488.9	30.9	6.76
T—3	537.4	611.5	74.1	13.80	T—3	411.0	479.5	68.5	16.68
T—4	307.0	327.0	20.0	6.52	T—4	195.0	227.4	32.4	16.60
T—5	451.1	502.9	51.8	11.49	T—5	338.6	392.4	53.8	15.90
T—6	396.5	405.1	8.6	2.16	T—6	271.7	295.8	24.1	8.86
T—7	517.5	521.4	3.9	0.75	T—7	384.6	402.3	17.7	4.59
T—8	481.0	491.0	10.0	2.09	T—8	367.9	373.4	5.5	1.48
T—9	291.6	308.6	17.0	5.84	T—9	210.8	236.9	26.1	12.39
T—10	493.6	583.0	89.4	18.11	T—10	387.0	474.7	87.7	22.67
T—11	394.0	409.9	15.9	4.05	T—11	291.6	317.8	26.2	8.99
T—12	859.2	866.7	7.5	0.87	T—12	687.7	708.2	20.5	2.98
T—13	750.7	686.2	64.5	8.59	T—13	571.3	534.3	37.0	6.47
T—14	824.5	853.5	29.0	3.52	T—14	645.3	684.5	39.2	6.08
T—15	628.6	626.9	1.7	0.27	T—15	463.5	481.5	18.0	3.87
T—16	470.0	503.9	33.9	7.22	T—16	345.3	387.8	42.5	12.31
T—17	425.8	393.8	32.0	7.52	T—17	289.8	281.5	8.3	2.87
T—18	463.8	555.5	91.7	19.77	T—18	322.0	418.7	96.7	30.02
T—19	742.9	709.4	33.5	4.51	T—19	586.2	567.7	18.5	3.16
T—20	571.8	562.1	9.7	1.70	T—20	400.6	418.2	17.6	4.40
T—21	874.1	823.9	50.2	5.74	T—21	710.4	662.3	48.1	6.78
T—22	394.7	363.2	31.5	7.99	T—22	279.7	257.7	22.0	7.87
T—23	540.5	581.1	40.6	7.52	T—23	418.4	468.1	49.7	11.89
T—24	471.1	506.2	35.1	7.44	T—24	345.9	389.2	43.3	12.51
T—25	384.8	387.2	2.4	0.63	T—25	262.7	279.1	16.4	6.25
Average error over 25 test runs				**6.24**	**Average error over 25 test runs**				**9.31**

The error percentages are also summarized.

[2] Eldridge, C., "Electronic Eyes for the Allies: Anglo-American Cooperation on Radar Development During World War II," *History and Technology*, Vol. 17, No. 1, 2000, pp. 1–20.
https://doi.org/10.1080/07341510008581980
[3] Bradley, M., Nealiegh, M., Oh, J. S., Rothberg, P., Elster, E. A., and Rich, N. M., "Combat Casualty Care and Lessons Learned from the Last 100 Years of War," *Current Problems in Surgery*, Vol. 54, No. 6, 2017, pp. 315–351.
https://doi.org/10.1067/j.cpsurg.2017.02.004
[4] Giffard, H., "Engines of Desperation: Jet Engines, Production and New Weapons in the Third Reich," *Journal of Contemporary History*, Vol. 48, No. 4, 2013, pp. 821–844.
https://doi.org/10.1177/0022009413493943
[5] Rodd, S. A., "Government Laboratory Technology Transfer: Process and Impact Assessment," Ph.D. Thesis, Virginia Polytechnic Inst. and State Univ., Blacksburg, VA, 1998.
[6] Alexander, C., "Development of the Combiner-Eyepiece Night-Vision Goggle," *1990 Technical Symposium on Optics, Electro-Optics, and Sensors*, Oct. 1990.
https://doi.org/10.1117/12.20951
[7] Costello, D. H., and Xu, H., "Generating Certificaion Evidence for Autonomous Aerial Vehicles Decision-Making," *Journal of Aerospace Information Systems*, Vol. 18, No. 1, 2021, pp. 3–13.
https://doi.org/10.2514/1.I010848
[8] Costello, D. H., Xu, H., and Jewell, J., "Autonomous Flight Test Data in Support of a Safety of Flight Certification," *Journal of Air Transportation*, 2020 (article in advance).
https://doi.org/10.2514/1.D0220
[9] Hoff, K. A., and Bashir, M., "Trust in Automation: Integrating Empirical Evidence on Factors that Influence Trust," *Human Factors: The Journal of Human Factors and Ergonomics Society*, Vol. 57, No. 3, 2015, pp. 407–434.
https://doi.org/10.1177/0018720814547570
[10] Ashokkumar, C. R., and York, G. W., "Trajectory Transcriptions for Potential Autonomy Features in UAV Maneuvers," *AIAA Guidance, Navigation, and Control Conference*, AIAA Paper 2016-0380, 2016.
https://doi.org/10.2514/6.2016-0380
[11] Humphreys, C. J., Cobb, R., Jacques, D. R., and Reeger, J. A., "Dynamic Re-Plan of the Loyal Wingman Optimal Control Problem," *AIAA Guidance, Navigation, and Control Conference*, AIAA Paper 2017-1744, 2017.
https://doi.org/10.2514/6.2017-1744
[12] Avram, R., Zhang, X., Muse, J. A., and Clark, M., "Nonlinear Adaptive Control of Quadrotor UAVs with Run-Time Safety Assurance," *AIAA Guidance, Navigation, and Control Conference*, AIAA Paper 2017-1896, 2017.
https://doi.org/10.2514/6.2017-1896
[13] Clark, M., Alley, J., Deal, P. J., Depriest, J. C., Hansen, E., Heitmeyer, C., Nameth, R., Steinberg, M., Turner, C., Young, S., Ahner, D., Alonzo, K., Bodt, B. A., Friesen, P. F., Horris, J., Hoffman, J. A., Gross, K. H., Humphrey, L., Childers, M., and Corey, M., "Autonomy Community of Interest (COI) Test and Evaluation, Verification and Validation (TEVV) Working Group: Technology Investment Strategy 2015–2018," Office of the Assistant Secretary of Defense for Research and Engineering TR, 2015.
[14] Endsley, M. R., "Design and Evaluation for Situation Awareness Enhancement," *Human Factors and Ergonomics Society Annual Meeting Proceedings*, Vol. 32, No. 2, 1988, pp. 97–101.
https://doi.org/10.1177/154193128803200221
[15] "Naval Air Training and Operations Procedures Standardization (NATOPS) General Flight and Operating Instructions," Commander Naval Air Forces, 2016.
[16] Endsley, M. R., "Toward a Theory of Situation Awareness in Dynamic Systems," *Human Factors*, Vol. 37, No. 1, 1995, pp. 32–64.
https://doi.org/10.1518/001872095779049543
[17] Nguyen, T., Lim, C. P., Nguyen, N. D., Gordon-Brown, L., and Nahavandi, S., "A Review of Situation Awareness Assessment Approaches in Aviation Environments," *IEEE Systems Journal*, Vol. 13, No. 3, 2019, pp. 3590–3603.
https://doi.org/10.1109/JSYST.2019.2918283
[18] Wickens, C. D., Goh, J., Helleberg, J., Horrey, W. J., and Talleur, D. A., "Attentional Models of Multitask Pilot Performance Using Advanced Display Technology," *Human Factors*, Vol. 45, No. 3, 2003, pp. 360–380.
https://doi.org/10.1518/hfes.45.3.360.27250
[19] Endsley, M. R., "Situation Awareness Misconceptions and Misunderstandings," *Journal of Cognitive Engineering and Decision Making*, Vol. 9, No. 1, 2015, pp. 4–32.
https://doi.org/10.1177/1555343415572631
[20] Snow, M. P., and Reising, J. M., "Comparison of Two Situation Awareness Metrics: SAGAT and SA-SWORD," *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, Vol. 44, No. 13, 2016, pp. 49–52.
https://doi.org/10.1177/154193120004401313
[21] Frey, T., "Cooperative Sensor Resource Management for Improved Situation Awareness," *Infotech@Aerospace 2012*, AIAA Paper 2012-2467, 2012.
https://doi.org/10.2514/6.2012-2467
[22] Ackerman, K., Xargay, E., Talleur, D. A., Carbonari, R. S., Kirlik, A., Hovakimyan, N., Gregory, I. M., Belcastro, C. M., Trujillo, A., and Seefeldt, B. D., "Flight Envelope Information-Augmented Display for Enhanced Pilot Situational Awareness," *AIAA Infotech @ Aerospace*, AIAA Paper 2015-1112, 2015.
https://doi.org/10.2514/6.2015-1112
[23] Hogan, J., Pollak, E., and Falash, M., "Enhanced Situational Awareness for UAV Operations over Hostile Terrain," *1st UAV Conference*, AIAA Paper 2002-3433, 2002.
https://doi.org/10.2514/6.2002-3433
[24] Jaenisch, H., Handley, J., and Bonham, L., "Data Modeling for Unmanned Vehicle Situational Awareness," *2nd AIAA "Unmanned Unlimited" Conf. and Workshop & Exhibit*, AIAA Paper 2012-2467, 2003.
https://doi.org/10.2514/6.2003-6523
[25] Reichard, K., Banks, J., Crow, E., and Weiss, L., "Intelligent Self-Situational Awareness for Increased Autonomy, Reduced Operational Risk, and Improved Capability," *1st Space Exploration Conference: Continuing the Voyage of Discovery*, AIAA Paper 2005-2692, 2005.
https://doi.org/10.2514/6.2005-2692
[26] Velagapudi, P., Owens, S., Scerri, P., Lewis, M., and Sycara, K., "Environmental Factors Affecting Situation Awareness in Unmanned Aerial Vehicles," *AIAA Infotech@Aerospace Conference*, AIAA Paper 2009-2057, 2009.
https://doi.org/10.2514/6.2009-2057
[27] Jones, R., Strater, L., Riley, J., Connors, E., and Endsley, M., "Assessing Automation for Aviation Personnel Using a Predictive Model of Situation Awareness," *AIAA SPACE 2009 Conference & Exposition*, AIAA Paper 2009-6790, 2009.
https://doi.org/10.2514/6.2009-6790
[28] Bateman, A. J., DeVore, M., Richards, N. D., and Wekker, S. D., "Onboard Turbulence Recognition System for Improved UAS Operator Situational Awareness," *AIAA Scitech 2020 Forum*, AIAA Paper 2020-2196, 2020.
https://doi.org/10.2514/6.2020-2196
[29] Endsley, M. R., "From Here to Autonomy: Lessons Learned From Human–Automation Research," *Human Factors: The Journal of Human Factors and Ergonomics Society*, Vol. 59, No. 1, 2017, pp. 5–27.
https://doi.org/10.1177/0018720816681350
[30] Ruano, S., Cuevas, C., Gallego, G., and García, N., "Augmented Reality Tool for the Situational Awareness Improvement of UAV Operators," *Sensors (Basel, Switzerland)*, Vol. 17, No. 2, 2017, p. 297.
https://doi.org/10.3390/s17020297
[31] Hanson, M., and Gonsalves, P., "Space Situational Awareness Using Intelligent Agents," *AIAA Space 2003 Conference & Exposition*, AIAA Paper 2003-6286, 2003.
https://doi.org/10.2514/6.2003-6286
[32] Scharringhausen, J.-C., Kolbeck, A., and Beck, T., "A Robot on the Operator's Chair—The Fine Line Between Automated Routine Operations and Situational Awareness," *SpaceOps 2016 Conference*, AIAA Paper 2016-2508, 2016.
https://doi.org/10.2514/6.2016-2508
[33] Collins, B., Kessler, L., and Benagh, E., "An Algorithm for Enhanced Situation Awareness for Trajectory Performance Management," *AIAA Infotech@Aerospace 2010*, AIAA Paper 2010-3465, 2010.
https://doi.org/10.2514/6.2010-3465
[34] Wagter, C. D., and Mulder, J., "Towards Vision-Based UAV Situation Awareness," *AIAA Guidance, Navigation, and Control Conference and Exhibit*, AIAA Paper 2005-5872, 2005.
https://doi.org/10.2514/6.2005-5872
[35] Shim, D., and Sastry, S., "A Situation-Aware Flight Control System Design Using Real-Time Model Predictive Control for Unmanned Autonomous Helicopters," *AIAA Guidance, Navigation, and Control Conference and Exhibit*, AIAA Paper 2006-6101, 2006.
https://doi.org/10.2514/6.2006-6101
[36] Lu, B., Coombes, M., Li, B., and Chen, W.-H., "Aerodrome Situational Awareness of Unmanned Aircraft: An Integrated Self-Learning Approach with Bayesian Network Semantic Segmentation," *IET Intelligent Transport Systems*, Vol. 12, No. 8, 2018, pp. 868–874.
https://doi.org/10.1049/iet-its.2017.0101
[37] Zhu, L., Cheng, X., and Yuan, F.-G., "A 3D Collision Avoidance Strategy for UAV with Physical Constraints," *Measurement*, Vol. 77,

Jan. 2016, pp. 40–49.
https://doi.org/10.1016/j.measurement.2015.09.006

[38] You, S., Gao, L., and Diao, M., "Real-Time Path Planning Based on the Situation Space of UCAVs in a Dynamic Environment," *Microgravity Science and Technology*, Vol. 30, No. 6, 2018, pp. 899–910.
https://doi.org/10.1007/s12217-018-9650-5

[39] Adams, J., "Unmanned Vehicle Situation Awareness: A Path Forward," *Human Systems Integration Symposium*, Annapolis, MD, 2007, https://pdfs.semanticscholar.org/e4a0/55d5de1de0321b974922498d1b316134305d.pdf.

[40] Wickens, C. D., "Situation Awareness and Workload in Aviation," *Current Directions in Psychological Science: A Journal of the American Psychological Society*, Vol. 11, No. 4, 2002, pp. 128–133.
https://doi.org/10.1111/1467-8721.00184

[41] Vincenti, W., *What Engineers Know and How They Know It: Analytial Studies from Aeronautical History*, Johns Hopkins Univ. Press, 1990, Chaps. 1–3.

[42] Cooper, G., and Harper, R. P., "The Use of Pilot Rating in the Evaluation of Aircraft Handling Qualities," NASA TN-D-5143, 1969, https://ntrs.nasa.gov/archive/nasa/casi.ntrs.nasa.gov/19690013177.pdf.

[43] "Military Specificaion: Flying Qualities of Piloted Airplanes, MIL-F-8785C," Nov. 1980, http://www.mechanics.iei.liu.se/edu_ug/tmme50/8785c.pdf.

[44] Hodgkinson, J., *Aircraft Handling Qualities*, AIAA Education Series, AIAA, Reston, VA, 1999.

[45] Jolly, J., and Clark, C., "CVN Precision Approach and Landing Systems (PALS) Certification Tests," NAVAIR, Test Plan Number SA17-02-0721B02, NAS Patuxent River, Maryland, 2018.

[46] Jewell, J., "What's it Doing Now? Testing Autonomy on a Manned Planform," *Society of Experimental Test Pilots 2018 East Coast Symposium*, Soc. of Experimental Test Pilots, Lancaster, CA, 2018.

[47] "Deliverables for Contract N00014-12-C-0671, Aurora Flight Sciences," 2018. Released by NAVAIR Center for Autonomy.

[48] Montgomery, D. C., *Design and Analysis of Experiments*, 4th ed., Wiley, New York, 1997, Chaps. 1–3.

[49] McCarley, Z., and Jorris, T., "Design of Experiments Used to Investigate an F/A-18 E/F Strafing Aunomaly," *11th AIAA SoCal Aerospace Systems & Technology (ASAT) Conference*, Santa Ana California, 2014.

[50] O'Conner, J., "Introduction to Aircraft and Systems Test and Evaluation: Design of Experiments Screening/Characterizing Example," United States Naval Test Pilot School, SYS Short Course Module SY04, 2020.

[51] Clive, P. D., Johnson, J. A., Moss, M. J., Zeh, J. M., Birkmire, B. M., and Hodson, D. D., "Advanced Framework for Simulation, Integration and Modeling (AFSIM) (Case Number: 88ABW-2015-2258)," *Proceedings of the International Conference on Scientific Computing (CSC)*, Las Vegas Nevada, July 2015, pp. 73–77.

[52] Hanlon, N., Garcia, E., Casbeer, D., and Pachter, M., "AFSIM Implementation and Simulation of the Active Target Defense Differential Game," *2018 AIAA Guidance, Navigation, and Control Conference*, AIAA Paper 2018-1582, 2018.
https://doi.org/10.2514/6.2018-1582.

[53] "Multiservice Brevity Codes," Air Land Sea Applaction Center, FM 3-987.18, MCRP 3-25B, NTTP 6-02.1, AFTTP(I) 3-2.5, Feb. 2002, http://www.fermentas.com/techinfo/nucleicacids/maplambda.htm.

[54] "Radar Techniques—Primer Principles, April 1945 QST Article," 2019, https://www.rfcafe.com/references/qst/radar-techniques-primer-principles-apr-1945-qst.htm.

[55] Dickerson, B. M., "NAVAIR Public Release 2020-643," NAWCAD, 2020.

[56] Hair, J., Joe, F., Sarstedt, M., Hopkins, L., and Kuppelwieser, V. G., "Partial Least Squares Structural Equation Modeling (PLS-SEM): An Emerging Tool in Business Research," *European Business Review*, Vol. 26, No. 2, 2014, pp. 106–121.
https://doi.org/10.1108/EBR-10-2013-0128

D. Casbeer
Associate Editor