JOURNAL OF AIR TRANSPORTATION
Vol. 29, No. 2, April–June 2021

Autonomous Flight-Test Data in Support of Safety of Flight Certification

Donald H. Costello, III*
University of Maryland, College Park, Maryland 20742
Jason Jewell†
Aurora Flight Sciences Corporation, Manassas, Virginia 20110
and
Huan Xu‡
University of Maryland, College Park, Maryland 20742

https://doi.org/10.2514/1.D0220

The current safety of flight clearances for unmanned aircraft requires a qualified operator who can make decisions and ultimately bears the responsibility for the safe operations of the vehicle. The future of aviation is unmanned, and ultimately autonomous. Yet, a method for certifying an autonomous vehicle to make decisions currently reserved for qualified pilots does not exist. Before we can field autonomous systems, a process needs to be approved to certify them. This paper analyzes the flight-test data (both developmental and operational) of an autonomous decision engine selecting an appropriate landing site for a large rotorcraft in an unprepared landing zone. In particular, this paper focuses on using legacy test and evaluation methods to determine their suitability for obtaining a safety of flight clearance for a system that possesses autonomous functionality. We will show that the autonomous system under test was able to complete a mission currently reserved for qualified pilots under controlled conditions. However, when confronted with conditions that were not anticipated (or programmed), the software lacked the judgment a pilot uses to complete a mission under off-nominal conditions.

I. Introduction

THE use of unmanned aircraft in aviation is expected to increase over the next decade because they can operate far beyond the limits of human endurance [1]. However, current safety of flight certification standards require a qualified operator in the loop. This operator, who controls the vehicle and makes decisions, is ultimately responsible for the safe operations of the vehicle [2]. Many modern aircraft can, and are, operated through a set of pilot relief modes (i.e., autopilots) that allow the aircraft to complete nearly the entire flight without a pilot touching the controls (which includes landing high-performance jet aircraft on the pitching deck of an aircraft carrier [3] or preparing for a landing on another planet during the Mars entry, descent, and landing phase of a mission [4]). However, the pilot in command still has the responsibility for the aircraft and is expected to operate the vehicle under current certification standards. Federal Aviation Administration (FAA) certification for unmanned vehicles only deals with small vehicles (referred to as quadcopters or similar small drones), and requires the operator to be within line of sight of the vehicle [5]. Future systems are expected to allow vehicles to operate autonomously. Autonomous aircraft will not have an operator in the loop, and they will ultimately require a new process for certifying an autonomous vehicle to accomplish tasks that are currently reserved for qualified pilots [2,6–8].

Many military applications can and have transitioned easily to the civilian sector (e.g., radio detection and ranging (known as radar) [9], medevac air ambulance [10], jet engines [11], glow sticks [12], and advanced night vision technology [13]). Therefore, we choose to examine a safety of flight certification for the unprepared (i.e., not an aerodrome or helipad) confined area landing (CAL)/landing zone (LZ) mission currently carried out by H-1 and H-60 variant helicopters by the U.S. Navy (USN) and U.S. Marine Corps (USMC) [14]. In an attempt to provide a path forward for certifying autonomy in aviation, this paper provides insight into the final portion of the certification process: flight-test (both developmental and operational). We examine flight-test data of an autonomous controller as installed on a FAA certified (experimental certification) UH-1 attempting to accomplish the unprepared CAL/LZ mission to determine if the current process can lead to a safety of flight clearance of autonomous behavior. We examined data through the lens of a developmental test (DT) program, which is used to determine if the vehicle can satisfy the requirements of the contract for which it was acquired (normally a set of objective measures). Following the DT evaluation, we examined data through the lens of an operational test (OT) program, which is used to determine if the vehicle is suitable for the mission for which it was designated when operating under mission representative conditions (normally a subjective opinion of the OT team). Both the DT and OT are designed to examine the possible corners of the operational envelope or the edge cases in the software verification [15].

Before certification of an autonomous system to complete the CAL/LZ mission, officials need to be provided certification evidence that the system can complete tasks currently reserved for fully qualified helicopter aircraft commanders (HACs) [16]. As a truly autonomous system has never been subjected to formal flight testing to support a safety of flight certification, exercising the existing process to evaluate a single mission set will provide significant lessons learned as we transition to more autonomous functionality within aviation. We demonstrate that the autonomous system under test was able to perform the CAL/LZ mission under controlled conditions. However, when confronted with conditions that were not anticipated/programmed (e.g., obstacle types that were not anticipated; compound malfunctions on the vehicle; or changing environmental conditions), its software lacked the judgment a pilot uses to complete a mission under off-nominal conditions.

This paper is structured as follows. In Sec. II, in addition to a review of related research into certifying autonomy, we discuss certifying the CAL/LZ mission, the flight-test process, and the system under test (to include a brief overview of the available flight-test data). In Sec. III, DT methods and results are summarized for the system under test. In Sec. IV, OT methods and results are

---

Received 26 June 2020; revision received 13 October 2020; accepted for publication 15 October 2020; published online 12 November 2020.

*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.

†Experimental Test Pilot; jewell.jason@aurora.aero.

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

summarized, and a system suitability for the mission is provided. In Sec. V, we decompose the results of the flight-test data for lessons learned regarding flight testing of autonomous systems for safety of flight certification. In Sec. VI, we summarize our findings as they relate to certifying autonomous systems to complete missions currently reserved for qualified pilots. Directions for future research are also provided.

II. Background

This research focuses on the flight-test of an autonomous system to complete the CAL/LZ mission to determine if it is suitable for a safety of flight certification. This will help build trust in autonomy because without trust certification officials will be reluctant to grant a safety of flight certification [17]. A simplified version of the steps leading to a safety of flight clearance for an autonomous system to complete the CAL/LZ mission is presented in Fig. 1. While the flowchart may appear to be a workflow diagram, it is actually a simplified version of the critical path leading to a safety of flight certification. The first step is to determine the requirements the system must complete to accomplish the mission for which it was acquired. Step two involves awarding a contract to a vendor to develop a system that can complete the mission requirements. The vendor will then need to validate the software (ensure the software meets the requirements from the contract), and perform modeling and simulation (M&S) as a risk mitigation step before flight testing. DT will then be performed to ensure the system has completed the requirements of the contract. Finally OT will be performed to ensure the system can complete the mission under mission-representative conditions. Once the system under test has accomplished all the steps, it will be granted a safety of flight clearance. This paper focuses on steps 5 and 6 of the simplified safety of flight certification process. Some related work is mentioned in Sec. II.A. Our proposed method for certifying autonomy for the CAL/LZ mission is covered in Sec. II.B. A review of the flight-test process (both DT and OT) is provided in Sec. II.C. Finally, the system under test [autonomous aerial cargo/utility system (AACUS)] and the available data are detailed in Sec. II.D.

A. Current Methods for Flight Certification

Currently, when an aircraft is certified as safe for flight (when operated safely, they will not break down or cause a danger to the general public), it is implied that they will be operated by a qualified pilot [or operator in the case of large unmanned aerial vehicles (UAVs) such as Global Hawk or Predator]. As an example of a currently fielded system, the USN currently operates the MQ-8 Fire Scout UAV. The Naval Air Systems Command (NAVAIR) has certified the large rotorcraft to fly without a qualified HAC. However, an air vehicle operator (AVO) is ultimately responsible for the safe operation of the vehicle. During preflight mission planning, the AVO programs the vehicle to complete parts of the mission without operator input (similar to an autopilot). In the event of lost link, the system will fly to a preplanned point and land. The system does not perform any evaluation of the landing point; it simply executes a preplanned route to a LZ and autolands [18].

Currently, a formalized/approved process does not exist for naval aircraft/systems that exhibit autonomous behavior (i.e., a system that is able to respond to situations that were not preprogrammed) because there has never been a requirement for one to be developed. Parallel paths are being taken around the world and by other organizations to achieve this goal [19]. However, this paper focuses on the achievement of a safety of flight clearance for a naval autonomous system. Several possible approaches have been proposed, but none have been vetted through the military, or civilian, flight clearance authorities [15,20–23]. The decision space for certifying a vehicle to complete all tasks assigned is extremely complex, which is why this work focused on flight testing in support of a safety of flight clearance of an autonomous controller completing a specific mission: to execute a safe landing of a large rotorcraft (capable of transporting passengers) within an unprepared CAL/LZ. This will enable an exercise of the flight-test process for just one mission normally reserved for a fully qualified HAC (other missions/tasks would include power line avoidance, see-and-avoid, formation flying, and visual navigation), thus limiting the complexity and scope of flight-test.

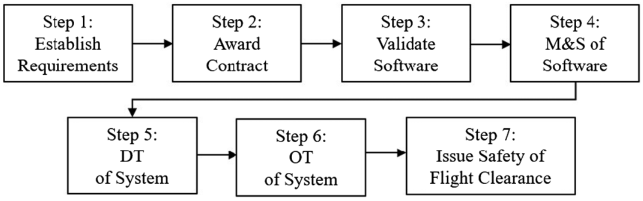

Fig. 1 Simplified flowchart detailing the steps leading to a safety of flight clearance for an autonomous system to accomplish the CAL/LZ mission. This paper focuses on steps 5 and 6.

There have been several proposed approaches for certification of unmanned/autonomous systems. A majority of the work deals with small unmanned aerial vehicles or theoretical methods for certifying large vehicles. One common theme is to identify errors in the software early in the design cycle since the later a defect is found, the more resources (both in time and money) are required to correct the issue [24–28]. Many of the approaches involve M&S to determine if the software is adequate for the system requirements [26,29–36]. Another common approach involves employing formal methods for safety-critical software verification and validation (V&V) (e.g., run-time verification [37–48], model checking [22,49–60], and theorem proving [49,60–66]). Some papers have detailed methodologies for V&V for the unmanned see-and-avoid requirement, but only for a two-dimensional problem [67,68]. Other proposals highlight the limitations of programming at simulating a pilot's ability to sense and accurately build their situational awareness (SA) during flight [69–75] and then make decisions based on changing situations [39,76].

One drawback of these approaches is the limited scope of their work. As an entire approved methodology does not exist, previous work has been limited to one or two pieces of the V&V process; and most did not consult aviation certification officials. One notable exception is the work done by the Formal Methods Group at National Aeronautics and Space Administration's (NASA's) Langley Research Center. Currently, NASA has published several papers on obtaining flight clearances for unmanned aerial systems to operate within the national airspace [77–79]. Their work focuses on formally defining the specification from the requirements of operation within the national airspace, and then V&V via theorem provers. This is designed to give certification officials confirmation that the software will perform per the requirements. However, their work focuses on an objective standard (such as maintain 1,000 ft separation), and not a judgment task (such as interpret the environment and make the best decision). As the current safety of flight certification process is designed to approve a system to be used by a fully qualified pilot, it has been hypothesized that before a safety of flight certification will be granted for an autonomous system, the system under test needs to demonstrate that it can perform as a qualified pilot would [80,81]. One issue with this plan is the complexity of accomplishing it. 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 V&V process [24].

The idea of autonomy and automation in transportation is not new. The automobile industry is one example of the increasing use of automation in our everyday lives. Modern automobiles have several capabilities that may be considered "driver relief modes" or automation. These capabilities include, but are not limited to, cruise control, brake assist, and hands-free parallel parking. While self-driving cars have been studied for decades, it was not until the Defense Advanced Research Projects Agency Grand Challenge (in 2005) that major advances were seen in the practical application of self-driving cars [82]. Tesla vehicles have had the hardware and software installed for truly autonomous operation since the 2016 model year. However, to operate the vehicle in autonomous mode, a licensed driver has to be at the wheel, ready to take over at all times, for it to be legally operated [83–85].

B. Certifying Autonomy for the CAL/LZ Mission

When certification officials grant a safety of flight clearance, they are certifying that if the system is used by a qualified pilot, it will be safe and can complete the mission that it was designed for [16]. However, the process of certifying a pilot is a trust process. When certifying a pilot, the commanding officer is putting his or her stamp of approval on that pilot; and they are designating that they trust their judgment when unplanned events occur [86]. By eliminating the pilot from the equation, certification officials need to be able to justify a safety of flight clearance without the benefit of a human when off-nominal conditions occur. For the purposes of tractability, we narrow the scope of the problem to a particular flight envelope (i.e., a box) in which the decision engine can exhibit autonomous behavior. This approach will allow certification officials to grant a safety of flight clearance, providing the decision engine will not violate one of the limits of the box. We used the size, slope, wing, elevation, escape route, and power (SWEEP) procedure executed by qualified HACs in the USN and USMC [14] to define the box for the proposed flight clearance of an autonomous system.

We define a suitable landing as one that satisfies the SWEEP checks performed by qualified HACs. While not all of the steps were specifically programmed into the tactical aerial logistics system (TALOS) (the decision engine that controls AACUS), it is important to understand each component of SWEEP as it relates to the system under test (AACUS/TALOS). In Ref. [81], Costello and Xu describe how the SWEEP checklist can be used to define a clearance envelope where a system can be allowed to exhibit autonomous behavior. This can be considered run-time verification, the system would be allowed to exhibit autonomous behavior providing it remained in the clearance envelop defined by the SWEEP checklist. If it were to reach a limit of the clearance envelope it would revert to known behavior. The components of SWEEP are described as follows:

1) The first component is size. The TALOS used light detection and ranging (LiDAR) to build a three-dimensional image to help determine a landing point free from obstructions and large enough for the vehicle. It was programmed to use a 10 m diameter as a clear zone for landing. That diameter needed to be an additional 10 m clear of obstacles (a total of 20 m from obstructions).

2) The second component is slope. Although the TALOS did not specifically determine the slope of a LZ, it used a rough approximation (similar to what a pilot would do) to determine if the slope of the LZ posed an unsafe condition. The slope limits allowed by the controller were more restrictive than the actual limits of the test vehicle.

3) The third component is wind. The TALOS was programmed to continuously evaluate the wind based on the control inputs and the deviations in ground track (Global Positioning System based). This is a standard technique for the test and evaluation (T&E) of helicopters. On approach, it would continue to update its local wind model until it reached 50 ft above ground level (AGL). It then used that wind speed and direction for approach. Before landing, the system would maneuver the nose of the aircraft into the wind to minimize crosswind and maximize headwind.

4) The fourth component is elevation. Elevation had a negligible effect on the available flight-test data, and it was not evaluated. The system under test did not possess a health monitoring system for elevation data. Although not evaluated during the test period, the elevation would have a dramatic impact on power available. Providing the data were accurate, it would be a variable for the power portion of the SWEEP checks.

5) The fifth component is the escape route. The TALOS used the situational awareness obtained by processing the sensor data available to build an escape route. Although none of the LZs evaluated required a complicated escape route, one was displayed to the safety pilot/flight-test engineer for each approach. During approach, the TALOS would monitor the LZ to ensure SWEEP remained valid. If SWEEP became invalid, the TALOS would initiate a wave off and fly the escape route back to a hold point.

6) The sixth component is power. All of the evaluated test LZs and aircraft configurations accommodated a power margin greater than 5% (a nominal safety buffer the AACUS/TALOS test team put in place). Although not evaluated during this test period, it would be a simple limit to place on an autonomous controller.

C. Flight-Test Overview

Flight-test is performed on a naval system before granting a safety of flight clearance. It is important to understand the purpose of the two types of flight-tests (DT and OT) as they pertain to granting a flight clearance. The FAA, NASA, and each of the three branches of the U.S. military have an airworthiness certification process for aircraft. For naval aviation, airworthiness certification authority is delegated to the NAVAIR. When a new capability/software/weapon/air frame is acquired, and before naval personnel operate it, the NAVAIR must grant a flight clearance (also referred to as a safety of flight certification). Aircraft subsystems, software, components, and ultimately the aircraft itself are certified through an established risk mitigation process; the final portion of the process is the flight-test [16]. Flight-test can be further broken down to either DT or OT. The qualification process for naval aviators (pilots) is considered to be a trust process. Unlike the civilian sector, military pilots are trusted by their commanding officers to complete missions critical to national interests. While each pilot is required to log a minimum amount of flight time and show competency in aircraft procedures before qualification, a commanding officer will not designate them as fully qualified until the individual has earned the trust of the commanding officer in their decision-making abilities in off-nominal conditions [86].

The purpose of DT is to ensure that the system under test can meet the requirements for which it was acquired under (normally a contract). DT is performed by trained test pilots: graduates of an internationally recognized test pilot school. The DT points (individual data points required to characterize the system under test during the evaluation) are controlled and designed to determine if the capability meets the individual specifications/requirements from the contract and must be flown by trained test pilots. An example of a developmental test requirement might be “the aircraft will achieve a level accelerated speed of 300 kt at 10,000 ft Mean Sea Level (MSL).” This requirement has a clear condition (300 kt at 10,000 ft MSL) and a clear method to achieve the specification (level acceleration). DT also offers an iterative approach to expanding a safety of flight clearance (envelope) by providing data to compare to other types of analyses (such as M&S or wind-tunnel data). DT is considered a black-or-white evaluation of an aircraft against the contract specifications. The test points for DT are typically objective. Once a new capability (i.e., full aircraft, new software, or weapon) has successfully demonstrated that it meets the requirements of DT, it can transition to OT.

The purpose of OT is to ensure that the new capability is suitable for the mission it is expected to complete. For a new capability to be deemed suitable (and pass OT), it must be able to perform the mission under mission-representative conditions by fleet-representative aircrew. An example of an OT requirement may include “the aircraft must be able to integrate into a multiplane strike verses a remote target in a contested environment.” Modern OT differs from DT in several ways beyond simply the training required for its aircrew. DT is designed to ensure the capability matches the requirements of the contract. OT is designed to ensure that the end user can use the capability to complete its designated mission. It is possible for a capability to successfully pass DT but fail during OT. This is one of the reasons that U.S. federal law only requires OT [87]. Unlike the objective evaluation of DT, OT is mainly a subjective evaluation of the system under test’s suitability for the mission it is designated for.

D. System Under Test (AACUS/TALOS) Overview

To evaluate current certification methods for the possible safety of flight certification of autonomy, we required a system that possessed autonomous functionality. In 2017, Aurora Flight Sciences (AFS) developed the TALOS decision engine for the AACUS program under an Office of Naval Research (ONR) contract [88]. AFS installed the TALOS on a modified UH-1 that flew under a FAA experimental certificate. The FAA granted the safety of flight clearance for the vehicle with the stipulation that any time the vehicle flew (autonomously or not), a HAC was required to be on board. The TALOS used the data available from the onboard sensors combined with the onboard processing power and data buses to build SA of the environment the decision engine would be operating in. The safety pilot (who was a trained experimental test pilot and fully qualified HAC) was required to monitor the systems decisions while the vehicle completed its mission autonomously, and the HAC was ultimately responsible for the safety of flight. The same test pilot acted as the HAC during all test flights in this research.

The AACUS/TALOS was designed to execute the U.S. Marine resupply mission. We used the available data to analyze the systems performance during the CAL/LZ mission (a submission of the resupply mission). Although AFS has published papers within the flight-test community, their work focused on how the system was designed, operated, and tested [89,90]: not on how the flight-test results can be used for safety of flight certification of autonomy. Similar work was done by the U.S. Army in modifying a Black Hawk for field navigation and landing site selection [91]. However the flight-test data available from AFS are diverse enough that they can be evaluated under current U.S. Department of Defense processes [87] for a potential flight clearance of the autonomous controller to complete the CAL/LZ mission. During the test program, the safety pilot monitored the system under test while it performed autonomous flight. By using a safety pilot, AFS and the ONR were able to examine autonomous functionality despite the lack of certification standards for autonomous vehicles. The 21 flight-test events occurred between 11 December 2017 and 23 May 2018. These events were chosen based on the fact that the software controlling the TALOS had reached a maturity point where future modifications did not have an effect on how it chose its LZ. The test events also concentrated on the actual landing portion of the demonstration and not the other aspects of the contract. The flights can be broken down to DT- and OT-like conditions.

The flights supporting the AACUS/TALOS final demonstration, rehearsals, and follow-on technology maturation assessment (December 2017 through January 2018) can be seen as DT events. The dataset, consisting of six flights concentrated on the system requirements from the contract, and the test points were scripted as such. The LZs were located on Quantico Marine Corps Base in Virginia, and they were designed to demonstrate the autonomous functionality of the AACUS/TALOS. During the DT period, all of the flights were choreographed by the test team to demonstrate the system's ability to satisfy the requirements of the ONR demonstration contract.

The follow-on events supporting a large-scale field training exercise at Twentynine Palms (a USMC base in California) can be seen as OT events. During operations in California, 15 flights were flown in the spring of 2018 in preparation for, and in support of, an integrated training exercise (ITX) with actual Marines [89]. The USMC uses Twentynine Palms to simulate real-life conditions Marines may find once deployed. The LZs were chosen by actual Marines to support conditions that can be considered as mission representative. During the OT period, all of the test flights were designed to evaluate the systems capability to complete the assigned task under mission-representative conditions.

III. Developmental Flight-Test of AACUS/TALOS

In this section, we further discuss the aspects of the DT (step 5 from Fig. 1) of the system under test. The evaluation of the objective requirements from the contract is covered in Sec. III.A. The various test points that will be tracked during the DT period, as well as how the system under test will be characterized, are outlined in Sec. III.B. A summary of the DT program is provided in Sec. III.C. Furthermore, in order for a system to pass DT and move on to OT, a positive DT/OT transition recommendation (to include a documentation of any deficiencies found during DT) is required. We provide a notional positive recommendation for the system under test in Sec. III.D.

A. Requirements of AACUS/TALOS for the Autonomous CAL/LZ Mission

For an autonomous system to obtain a safety of flight certification for the CAL/LZ mission, it will need to demonstrate that it can accurately complete SWEEP checks. As the only parts of SWEEP that were programmed into the TALOS were size (to include obstacle detection), slope, wind, and escape route, the DT flight-test data will evaluate those requirements (elevation and power margin were not evaluated during this test program).

1) The first requirement is LZ size. The contract set the requirement for a 10 m radius (UH-1 rotor arc is 24 ft, 1.6 in.); this radius must be an additional 10 m from any obstacles. The system was required to scan the possible LZ from altitude (approximately 200 ft AGL) and determine if the LZ is large enough for the vehicle. A human pilot uses experience to judge the size of a LZ, but using onboard sensors has the potential of being more exact.

2) The second requirement is LZ slope. The contract set the requirement for less than 3 deg of slope (actual UH-1 limit is 6 deg). The system was required to scan the possible LZ from altitude (approximately 200 ft AGL) and determine if the LZ is within limits. Slope is the most difficult parameter for a pilot to determine from altitude. Often, on approach, a HAC will abort a landing when the slope is not as anticipated from altitude.

3) The third requirement is obstacle detection. The contract requirement was for the system to detect and avoid an obstacle the size of an 18 in. pelican case (depicted in Fig. 2 [94]). If a helicopter were to land on an obstacle, the risk of dynamic rollover would be real. Similar to excessive slope, a dangerous situation can develop if only one skid were to touch down during a normal landing. During the CAL/LZ mission, a crew chief actively looks out the side of the helicopter, clearing the LZ for the pilots from when the aircraft is over its landing spot through landing. The system was required to scan the possible LZ from altitude (approximately 200 ft AGL) and determine if the LZ is clear of obstacles. The system under test was required to continuously monitor the touchdown point for possible obstructions during approach through touchdown.

4) The fourth requirement is wind. The system under test was able to continuously evaluate the local wind conditions by comparing the ground track of the vehicle against the control inputs. As this is a standard technique for developmental testing of helicopters, it is not part of this research. As the vehicle began its approach to landing, it stopped evaluating the winds at 50 ft AGL. It then used that wind direction and magnitude to determine if the winds were within limits. Before landing, the system would maneuver the nose of the aircraft into the relative wind to limit crosswinds and maximize headwind.

5) The fifth requirement is escape route. The system under test was required to scan the area around the LZ and determine a safe route to a hold point before starting its approach for landing. The AACUS/TALOS used the rapidly exploring random tree algorithm [92,93] and the information available through its sensors to build the escape route. If the LZ were to become fouled (something moves into the previously cleared space) or SWEEP were no longer valid during approach (such as an obstacle were to be detected during approach), the system would wave off and fly the escape route to a hold point. In the field, a

Fig. 2 Photograph of a Marine carrying a 24 × 20 × 16 in. pelican case during the AACUS ONR final demonstration [94].

Table 1 Completed DT matrix of AACUS/TALOS for autonomous CAL/LZ mission (P = pass, F = fail, and N/A = not applicable)

Flight number: date	Size	Obstruction	Slope	Fouled LZ	No. of landings	No. aborted
F096: 11 Dec. 2017	P	P	P	P	7	0
59F097: 12 Dec. 2017	P	P	P	P	6	0
59F098: 13 Dec. 2017	P	P	P	P	5	0
59F100: 22 Jan. 2018	P	P	P	P	3	0
59F101: 23 Jan. 2018	P	P	P	N/A	7	0
59F102: 24 Jan. 2018	P	P	P	P	5	0

ground vehicle or wildlife may foul the LZ. Or, once the sensor package is closer to the landing zone, it may detect a condition that violates the requirements for a valid LZ.

B. Developmental Flight-Test Matrix

When preparing for a flight-test program, military T&E leadership develop a list of specific test points required to accomplish a test program. Typically, these test points are laid out in an easy to follow test matrix. As developmental flight-test is resource intensive, leadership will develop test points that are designed to evaluate the edge cases of the system under test. These edge cases typically define the edges of the envelope that will be in a safety of flight certification. These edge cases are typically first identified during risk mitigation M&S before flight-test (step 4 from Fig. 1). The test matrix offers the flight-test community a simple to understand status of the test program, as well as a method to annotate flight-test results. To pass the DT, the system under test will need to accomplish a minimum of 25 autonomous landings (nominal value we selected for this research) with no safety of flight issues. During the landings, the system must demonstrate that it can select a LZ that is not obstructed and has a slope that meets the requirements of the test program. In addition, the system must demonstrate it can identify an 18 in. pelican case in a possible LZ. Finally, during approach to landing, the system must be able to identify an interloper if it were to enter the LZ, abort the landing, and fly an escape route to the hold point. The flight-test matrix, in addition to daily flight reports prepared after each flight, are used by the flight-test community to characterize the system under test when they evaluate the systems compliance with the requirements of the contract for which it was acquired. Using the CAL/LZ mission as the foundation for evaluation, the flight-test community can help inform certification officials decisions for certifying autonomous behavior. The test matrix for this evaluation can be found in Table 1. The columns for Table 1 can be described as follows:

1) “Flight number: date” specifies the flight-test number and the date of flight.

2) “Size” tracks the system’s ability to select a LZ that meets the minimum size requirement. During DT, this was evaluated by placing obstacles (the test team used pelican cases described in Sec. III.A) in known locations in the test LZ area to determine if the system can accurately choose a valid landing point (both by the safety pilot in real time and by postflight analysis). Figure 3 depicts two LZs. Both photographs were taken from the pilot’s perspective in a UH-1, 200 ft AGL over Naval Air Station (NAS) Patuxent River [95]. The left image does not meet the requirements of the contract; the right does.

3) “Obstruction” tracks the system’s ability to select a LZ that meets the obstacle clearance threshold (no obstacles larger than an 18 in. pelican case). During DT, this was by examining the selected LZ to determine that the LZ was not obstructed (both by the safety pilot in real time and by postflight analysis). This and the first column of the test matrix will be accomplished by placing test pelican cases around a known location to test the system’s ability to choose a valid LZ. Figure 4 depicts two LZs that are obstructed by vehicles.

4) “Slope” tracks the system’s ability to select a LZ that meets the maximum slope requirement. During DT, this will be evaluated by examining the selected LZ to verify that it meets the slope requirement (both by the safety pilot in real time and by postflight analysis). Figure 5 depicts a LZ at NAS Patuxent River used by the DT community for slope landing evaluation [96]. The photograph was taken from the pilot’s perspective in a UH-1, 200 ft AGL over NAS Patuxent River. The three surveyed LZs have different slopes that test pilot’s use during flight testing.

5) “Fouled LZ” tracks the system’s ability to sense an interloper that fouls the LZ during approach. During DT, this will be evaluated by driving a golf cart into the LZ while the system is on approach to landing. Upon sensing the LZ is fouled, the system will execute the escape route (which is displayed to the safety pilot before approach) and return to the hold point.

6) “No. of landings” and “No. aborted” track safe autonomous landings and aborted approaches by the safety pilot for violation of requirements. To successfully pass DT, we stipulated that the system must complete 25 autonomous landings and have zero approached aborted by the safety pilot for violation of the requirements.

Each DT flight was recorded via the test matrix. The results were evaluated to determine if the system should be recommended for OT

Fig. 3 Pilot’s perspective of two LZs taken from a UH-1 at 200 ft AGL over the turf training area of NAS Patuxent River [95].

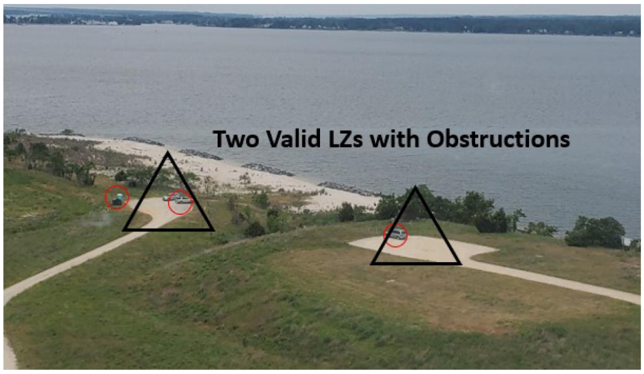

Fig. 4 Pilot's perspective of two LZs that would have been valid if the vehicles were not present, taken from a UH-1 at 200 ft AGL over the turf training area of NAS Patuxent River [95].

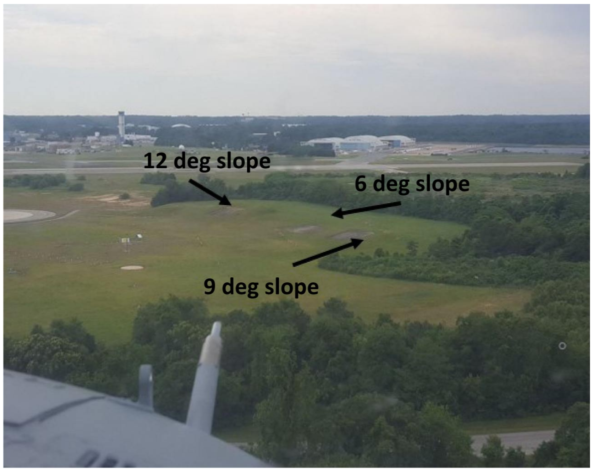

Fig. 5 Pilot's perspective of surveyed LZ used for slope landing evaluation taken from a UH-1 at 200 ft AGL over the turf training area of NAS Patuxent River [96].

because OT requires substantial investment in resources (both time and money). A system that does not receive a positive recommendation for OT from DT typically does not proceed to the next step until mitigation measures are put in place. Ultimately, the test matrix is used to characterize the system under test.

While the test matrix characterizes the system based on its performance in the execution of planned test points, other items are identified during flight-test. Experimental test pilots are trained to find deficiencies in a system. A part 3 deficiency is considered a nuisance, and it is tracked against the system in case there are resources (both time and money) available to fix them in the future. A part 2 deficiency is considered an issue with the system that requires human interaction to overcome (such as pressing extra buttons on a flight management system to accomplish the mission). As with the part 3 deficiency, they are normally tracked for possible correction at a later date. A part 1 deficiency is one that, if not corrected, the system cannot accomplish the mission or may result in a major mishap. Part 1 deficiencies are typically addressed before the system receives an OT transition recommendation.

C. Summary of Developmental Flight-Test Events

The DT of the system under test consisted of six test flights. They were flown as part of the buildup to the AACUS/TALOS final demonstration, the demonstration itself, and follow-on technology maturation assessment by the ONR. All flights took place between 11 December 2017 and 24 January 2018, and they were choreographed by the test team to demonstrate the system's mastery of the requirements levied by the contract. Table 1 summarizes the six test flights in the test matrix.

During DT, the test conductors used both movable and stationary obstructions to force the system to choose individual LZs that met the requirements of the CAL/LZ mission. When evaluating a LZ, the TALOS used LiDAR to build its perception of the LZ. As it approaches a LZ, more data become available to fine-tune its interpretation of the LZ. Figure 6 depicts three images showing the perception model of the LZ building as the test asset approaches. The landing area evaluated was a 50-m-radius seven-sided polygon. Large obstacles were defined as something with a height of 11 in. The system would invalidate an area around the obstacle, although not in a circular shape. The shape is elliptical, with the long axis parallel to the vehicle's approach path. All images were displayed with north up and the distance to the proposed LZ listed to the lower right of the image. The hop number (recorded segment of the test flight) is displayed at the top left corner of the image. The circle in the center of the image is the desired landing spot from the end user. The colors in the image relate to the suitability of the location. Table 2 details the color legend for the TALOS-produced interpretation of the LZ.

Figure 7 depicts the system's interpretation of the LZ for one of the autonomous landings during flight 59F098 and an image of the test UH-1 immediately postlanding [97]. The landing spot was in a field with rolling hills. Figure 8 also depicts images relating to an autonomous landing during flight 59F098; the landing spot was in a simulated forward operating base (FOB) and is considered one of the tougher challenges for the system.

To evaluate the system under test's ability to sense an interloper fouling the LZ, the test team would wait until the system under test approached the LZ; then, one of the test team would drive a golf cart into its path. Upon sensing the fouled LZ, the system would abort the approach and fly an escape route to the hold point. Figure 9 depicts the TALOS's interpretation of a LZ before (left image) and after (right image) a golf cart is driven into it. The golf cart is what creates the orange zone at the bottom of the green zone in the second image. This was done to test the wave-off functionality of the system.

In addition to the test matrix, the safety pilot and test team noted several minor issues during DT. Some of these issues related to the

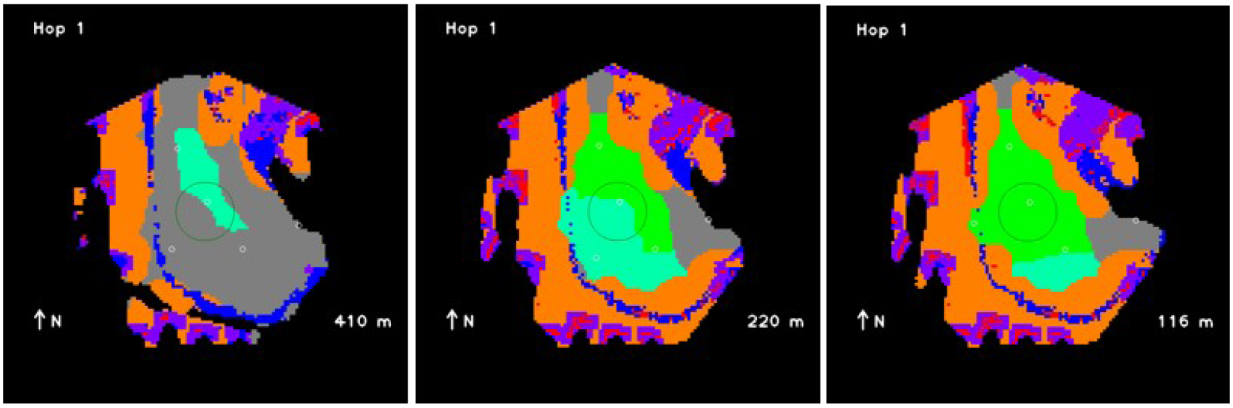

Fig. 6 TALOS LZ Interpretation from 410, 220, and 116 m during flight 59F097. As the vehicle approaches the LZ, its interpretation become clearer [88].

Table 2 Legend for colors in TALOS LZ interpretation [88]

Color	Meaning
Black	No evaluation performed in the area, or no data available in the area
Gray	No object seen; not enough data to determine if a large-sized object is present
Yellow	No object seen; not enough data to determine if a medium-sized object is present
Teal	No object seen; not enough data to determine if a small-sized object s present
Green	Area is safe for landing; no object seen
Red	Object in this area; not safe for landing
Orange	Too close to an object; not safe for landing
Blue/purple	Terrain is too sloped or too rough for safe landing

software resiliency, which was not evaluated for the autonomous CAL/LZ mission. Yet, other issues noted by the test team directly related the system performance. On flight 59F096 the system selected two landing spots that were not advantageous to the test (one was too close to a road, and one was too close to ground personnel). Although the selected spots met all of the requirements for the system, the safety pilot disengaged the system and selected a more advantageous spot. Also, on flight 59F096 it appeared that the constantly changing cargo load of the vehicle affected the landing performance (both skids did not contact at the same time). On flight 59F097, while performing an escape route, the vehicle tracked outside of the planned route (yet still safely executed the route) due to the fact that the selected route was not planned to properly match the vehicle's maneuverability. On flight 59F101 the local wind conditions were more extreme than

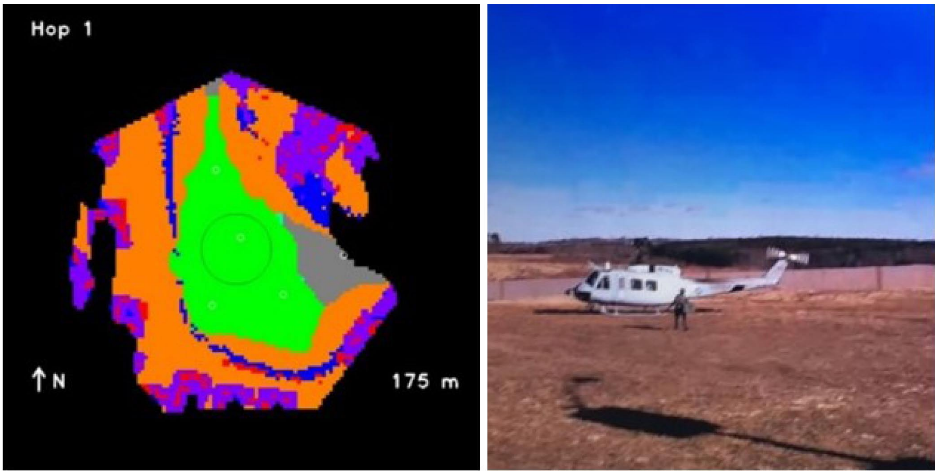

Fig. 7 Two images relating to an autonomous landing in a field during flight 59F098: TALOS interpretation of the LZ (left) [88], and photograph of test vehicle shortly after completing an autonomous landing in LZ pictured on left (right) [97].

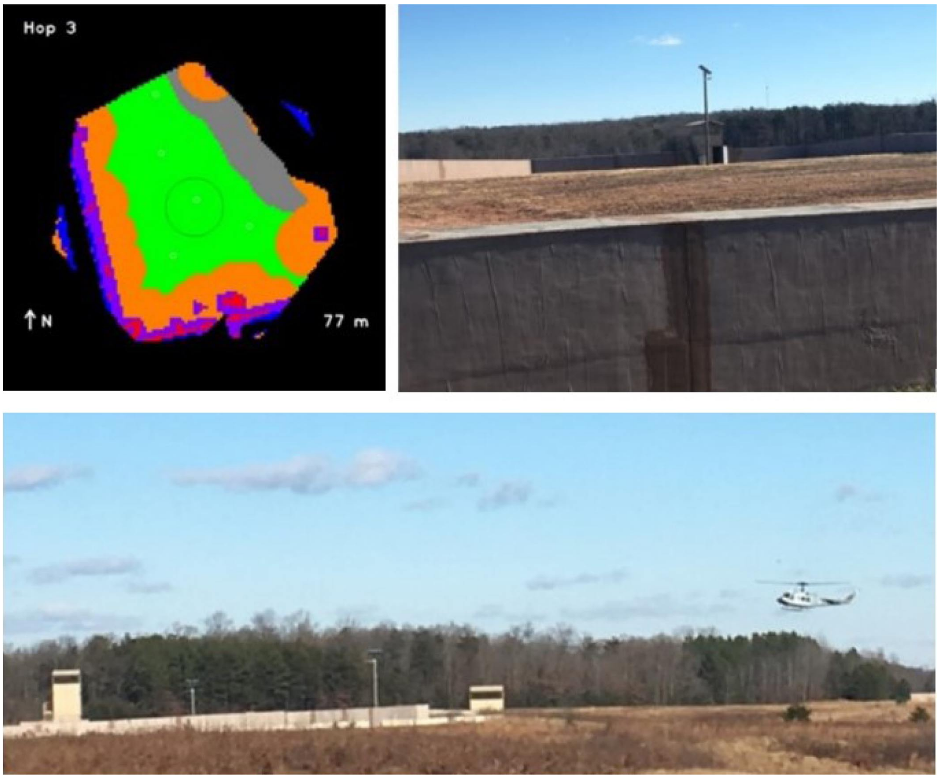

Fig. 8 Three images relating to an autonomous landing in a simulated FOB during flight 59F098: TALOS interpretation of the LZ (top left) [88], photograph of LZ from ground level (top right) [97], and AACUS/TALOS completing an autonomous landing in simulated FOB (bottom) [97].

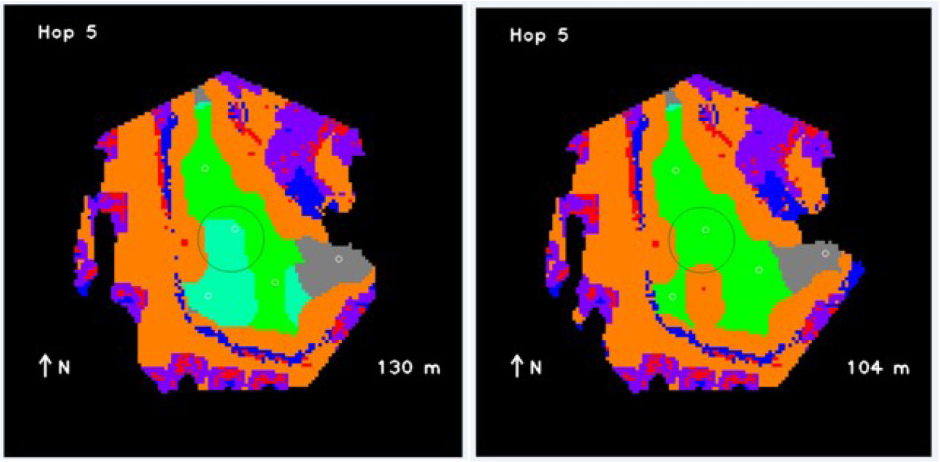

Fig. 9 **TALOS interpretation of a LZ before (left image) and after (right image) a golf cart is driven into it testing the wave-off functionality on flight 59F096 [88].**

seen during past test events (winds were 14 gust 19 kt). Although the winds were well within the limits of the experiment, the vehicle displayed less than optimal performance (still within prescribed limits).

D. DT Results and DT/OT Transition Recommendation

Despite the deficiencies noted, the system was able to perform the mission autonomously under the constraints imposed by the test team. We have determined that the system was able to accurately complete the SWEEP checks under controlled conditions and should proceed to OT.

During six DT events, the system under test performed 33 autonomous landings with zero safety of flight issues (or violations of the requirements placed by the contract). The system also demonstrated the ability to detect if the landing zone was fouled by an interloper, and it executed an escape route to its hold point. However, several deficiencies were identified in the system:

1) The first deficiency is that the system lacks the ability to optimize the landing spot selection (flight 59F096). Once it found a valid point for landing, it ceased looking for a more advantageous spot (part 2 deficiency). We recommend that future software loads have a cost function embedded to help solve this problem.

2) The second deficiency is that the system's actual performance may not be the same as programmed (part 2 deficiency). We recommend that future software loads have an updated model of the performance of the vehicle.

3) The third deficiency is that the system lacks a dynamic c.g. sensing capability, which may lead to an unsteady landing (part 3 deficiency). We recommend that future software loads have an updated c.g. sensing capability.

4) The fourth deficiency is that, during high/gusty wind conditions (yet within the limits of the vehicle/system), the hover and landing performance was safe but not consistent (part 3 deficiency). We recommend that future software loads have improved gust performance.

IV. Operational Flight-Test of AACUS/TALOS

Unlike DT, OT is not carefully scripted. During DT, the test team was tasked with ensuring the system under test can perform to the specifications that were detailed in the contract. All of the DT LZs were designed to test the capabilities of the system under controlled conditions. Unlike DT, OT flight-test is designed to see if the average fleet operator can use the system to perform the mission, as well as determine if the system under test can perform in a mission-representative environment. Operational testers are tasked to determine if the system under test is operationally effective and suitable for the mission [87]. In Sec. IV, we further discuss the aspects of OT (step 6 from Fig. 1). The goals and expectations of the system in OT are covered in Sec. IV.A. The various test points that were tracked during the OT period are outlined in Sec. IV.B. A summary of the OT program is provided in Sec. IV.C. Finally, the AACUS/TALOS system suitability assessment (results from OT) is presented in Sec. IV.D.

Late in 2017, the AACUS/TALOS showed great promise for autonomy. During several technology demonstration flights, the system impressed senior USMC officers. They asked if the system could provide similar results in the field, resupplying actual Marines. The ONR and AFS agreed to allow the system to operate at Twentynine Palms, a USMC base in California, during a major USMC ITX. In the Spring of 2018, the AACUS/TALOS flew 15 flights under operationally relevant conditions.

A. Goals and Expectations of the System in OT

The basic resupply mission is simple: a Marine makes a request for supplies, the request is filled, and a helicopter delivers the supplies to the Marine in the field. The AACUS/TALOS was programmed to fly from one location to the Marine's location, select a LZ near the Marine, land, and allow the Marine to unload the supplies. We evaluated the AACUS/TALOS for the final portion of the resupply mission. We evaluated the system under test for its suitability in the autonomous CAL/LZ mission under mission-representative conditions at Twentynine Palms Marine Corps Base.

As with DT, we used the SWEEP checklist to determine if the system under test can perform the same actions a qualified HAC would under mission-representative conditions. However, during OT, we did not evaluate it against black-and-white requirements. We evaluated it against the safety pilot's (a trained engineering test pilot and fully qualified HAC) opinions to see if the decisions the system under test made would match that of a fully qualified HAC.

B. Operational Flight-Test Matrix

During the ITX at Twentynine Palms Marine Base, the AACUS/TALOS was tasked with resupplying actual Marines. As with DT, we evaluated the AACUS/TALOS for the autonomous CAL/LZ mission (just the landing portion of the resupply mission). However, unlike DT, the LZs the Marines chose were not ideal. The obstacles in them were not pelican cases placed by the test team to determine if the system could distinguish a clear LZ that met the requirements of the system. Instead, the obstacles were whatever was present in the area where the Marine requested resupply.

For the OT evaluation matrix, we once again used the portions of SWEEP that were programmed into the system under test. However, instead of evaluating the performance against the requirements of the system (as we did in DT), we evaluated the system against the expert opinion of the safety pilot (a trained engineering test pilot and fully qualified HAC) while the system performed the autonomous CAL/LZ mission in a mission-representative environment. Table 3 is a flight-test matrix that summarizes the operation flight-test of the

Table 3 Completed OT flight-test matrix of AACUS/TALOS for the autonomous CAL/LZ mission

Flight number: date	Size	Slope	Obstruction	Spot	Wave off	No. of landings	No. aborted
F111: 29 April 2018	Yes	Yes	Yes	Yes	Yes	2	0
59F112: 1 May 2018	Yes	Yes	Yes	Yes	N/A	2	0
59F113: 3 May 2018	Yes	Yes	Yes	Yes	Yes	4	0
59F114: 4 May 2018	Yes	Yes	Yes	Yes	N/A	7	0
59F115: 6 May 2018	Yes	Yes	Yes	Yes	N/A	1	0
59F116: 8 May 2018	Yes	Yes	Yes	Yes	N/A	1	0
59F117: 8 May 2018	Yes	Yes	Yes	Yes	N/A	2	0
59F118: 12 May 2018	Yes	Yes	Yes	Yes	N/A	6	0
59F119: 14 May 2018	Yes	Yes	No	No	N/A	6	0
59F120: 15 May 2018	Yes	Yes	Yes	Yes	N/A	4	0
59F121: 17 May 2018	Yes	Yes	Yes	Yes	N/A	4	0
59F122: 18 May 2018	Yes	Yes	Yes	Yes	N/A	2	0
59F123: 21 May 2018	Yes	Yes	Yes	Yes	N/A	2	0
59F124: 22 May 2018	No	Yes	No	N/A	Yes	2	0
59F126: 23 May 2018	Yes	Yes	Yes	Yes	N/A	2	0

AACUS/TALOS for the autonomous CAL/LZ mission, and the columns can be summarized as follows:

1) "Flight number: date" specifies the flight-test date and flight.

2) "Size" tracks if the safety pilot agreed with the size of the selected LZ.

3) "Slope" tracks if the safety pilot agreed with the slope of the LZ.

4) "Obstruction" tracks if the safety pilot agreed that the LZ was clear of obstructions.

5) "Spot" tracks if the safety pilot agreed with the landing spot chosen by the decision engine.

6) "Wave off" tracks if the safety pilot felt the wave off was executed properly.

7) "No. of landings" tracks the number of autonomous landings during the test flight.

8) "No. aborted" tracks the number of landing aborted by the safety pilot for safety of flight reasons.

To successfully pass OT, and ultimately be given a safety of flight certification and fielded, the system under test will need to demonstrate under operationally relevant conditions that it can complete the autonomous CAL/LZ mission. Unlike DT, where the system merely needed to demonstrate that it met the requirements set in the contract, in OT, the system needed to show that it can perform as a fully qualified HAC to be effective and suitable for the mission (a subjective assessment by the OT team).

C. Summary of Operational Test Events

OT consisted of 15 flights flown between 29 April 2018 and 23 May 2018. They were flown as part of a major field exercise supporting USMC personnel at Twentynine Palms Marine Base. All test flights were flown under mission-representative conditions and not specifically choreographed by the test team to demonstrate the system's mastery of the requirements levied by the contract. The first five flights were system preparation flights to understand the new environment. The final 10 were in direct support of the exercise. Table 3 summarizes the 15 test flights in the OT matrix.

During the first flight in a mission-representative environment, some issues immediately presented themselves. Unlike the LZs of Quantico, those in Twentynine Palms had not been cleared of brush to maximize Marine training. Vegetation in the high desert of California ranges from small shrubs or tumbleweeds to shoulder-high bushes. The test team used the first flight to judge the effect vegetation has on the system. During flight 59F111, the system under test had difficulty (in the opinion of the safety pilot) finding a LZ that met its criteria for obstacle clearance. While evaluating four LZs, only two of them met the requirements for the system under test to perform a landing. The safety pilot noted that the UH-1 could have performed a landing, but it would require extensive crew coordination and pilot judgment (these capabilities were not programmed into the system). Figure 10 is the TALOS interpretation of one of the LZs and a corresponding Google Earth image prepared by the test team from flight 59F111. The safety pilot felt he could land in the LZ, but the TALOS could not find a valid spot based on the extra safety factor programmed into the system.

One of the major concerns from AFS and the ONR was how the system under test would perform under conditions approaching brownout, where the rotor wash picks up dust when landing in a desert LZ that blocks the aircrew view of the ground when approaching touchdown. Several of the LZs chosen during the first five flights at Twentynine Palms were chosen to assess the system's performance in adverse conditions. The fear was that the installed LiDAR could not penetrate dust on landing and could initiate a wave off that was not warranted. No issues were found when operating in near-"brownout" conditions. During flight 59F120, the system was able to complete the CAL/LZ mission despite encountering what the safety pilot considered full brownout (Fig. 11).

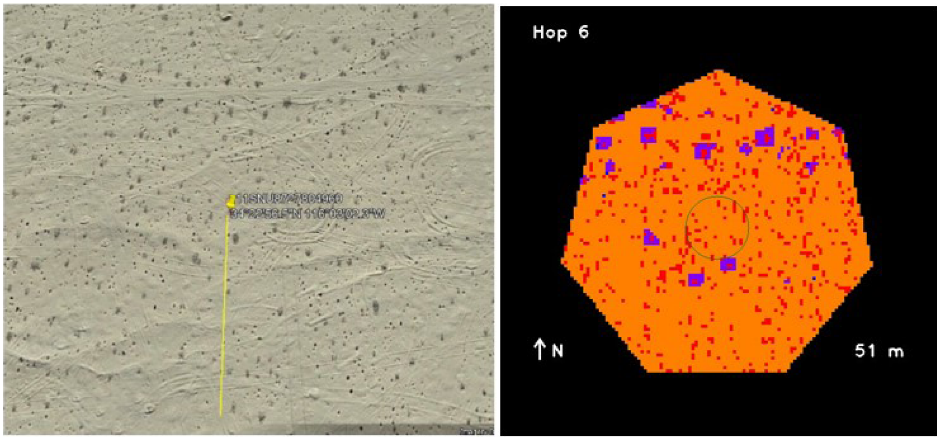

Fig. 10 Two images relating to a LZ during flight 59F111: Google Earth image (left), and TALOS interpretation of the same location (right). TALOS declared location unsuitable; safety pilot disagreed [88].

Fig. 11 System under test performing an autonomous landing during full brownout conditions at Twentynine Palms Marine Base during flight 59F120 [89].

Flight 59F117 was a milestone for the program. It was the first time that the system was used to perform the resupply mission of Marines in the field. The system under test was able to complete the entire mission (to include the autonomous CAL/LZ portion of the flight) autonomously per the requirements of the system.

An issue was found during flight 59F119. The system under test was to fly to a remote location (a dirt runway) for a resupply mission and vehicle refuel. The Marines at the LZ had set up sandbags to indicate to the pilot where to land (a standard operating procedure). However, the system saw the sandbags as an obstruction, and it chose a different landing spot. The safety pilot took control of the aircraft and landed on the runway, in the desired location, to facilitate refueling. The other six landings performed during the flight were all accomplished autonomously at other locations with no issues.

Another issue was noted on flight 59F125. The system under test was directed to resupply Marines in the field with water (mission critical based on the location). Unfortunately, the location the Marines chose for resupply was suboptimal. The foliage in the area made it difficult for the system to select a landing point that met the requirements of its programming. The system under test had a requirement for the size of an obstacle. It was programmed to invalidate the area around detected obstacles. A trained HAC would have evaluated the foliage in the area and dismissed some of the foliage as a nonfactor (yet the system under test identified them as a hazard). Ultimately, the safety pilot had to disengage the system and land manually to accomplish the resupply (because the LZ was compatible with the UH-1, just not the requirements programmed into the AACUS/TALOS).

D. AACUS/TALOS System Suitability

The system under test demonstrated that it could complete the autonomous CAL/LZ mission under favorable conditions (i.e., those that were programmed into the system). During OT, the AACUS/TALOS performed 46 autonomous landings. It also demonstrated extreme promise in controlling a helicopter during brownout conditions. However, under field conditions, the experience and training of the safety pilot was required to complete the landing when the obstacles in the LZ were challenging. The system was programmed with a large safety margin, but that margin negated the ability of the system to perform landings in some of the LZs of Twentynine Palms. In addition, some of the LZs chosen by the system under test were not ideal. The vegetation in the proposed LZs had not been completely cleared as it would have been at an aerodrome or helipad, and the safety pilot had to take control and land at a more advantageous spot (mainly when dealing with an LZ that required interaction with Marines on the ground). The system also had issues when identifying obstacles that could foul a LZ because it was programmed to view an 11 in. obstacle as fouling a LZ. In the field, many of these objects were small shrubs or tumbleweeds. A fully qualified HAC would have identified them as no risk (as the downwash on approach would blow them out of the way). This is also a limitation of the programming in a system that was designed as a technology demonstration, and not a system for operational use. While all 15 flights were flown by the same experimental test pilot, the conclusions in this research were formed by a committee of flight-test experts who had access to the flight-test data.

The results of OT were shared with senior naval officers who currently certify pilots as HACs. They are tasked with certifying the judgment of the pilot to perform critical missions when the conditions are suboptimal. They unanimously agreed that, as evaluated, the AACUS/TALOS did not meet their threshold as being capable of making decisions currently reserved for qualified pilots. When presented with a situation that matches the programming, the system under test was able to complete the mission. However, when presented with a situation that did not fit neatly into the programming, the system could not complete the mission.

We found that the AACUS/TALOS (as programmed and evaluated) was not effective or suitable for the autonomous CAL/LZ mission. Based on these findings, NAVAIR would not grant a safety of flight certification for the system to perform the mission.

V. Analysis of the Test Results as They Relate to Certifying Autonomy

Throughout the 1920s and 1930s, despite meteoric advances in structures, aerodynamics, and propulsion, aircraft handling qualities 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) [98]. The advent of autonomous systems has created a similar daunting task. Currently, certification officials mainly use objective standards to determine if the system can be used by a fully qualified aircrew to complete a mission before granting a flight clearance. However, the commanding officer of a squadron uses a subjective measure to determine if a pilot is ready for full qualification. This creates the same problem aircraft designers have had for improving handling qualities. The designers of autonomous systems will be given a set of performance specifications that are themselves objective ends. However, the behavior described in the specifications, completing a judgment task, requires objective means to an associated subjective end [99]. This research has shown that accomplishing a judgment task (we evaluated the system under test for the CAL/LZ mission) will require new processes, or adjusting current processes to meet the new requirement.

The available flight-test data were evaluated under DT-like conditions (where applicable) to determine if the contractor was able to build a system to a specification of the contract (show that the decision engine would only land in areas that met the conditions of the contract). It was also evaluated under OT-like conditions (where applicable) to determine if the decision engine could execute the task under mission-representative conditions. The flight-test data were also presented to senior officers who currently certified HACs.

AFS developed the decision engine that enabled the system under test to accomplish the task under controlled conditions. During the notional DT phase of this test program, the system under test successfully completed its assigned task 33 times with no issues relating to the landing portion of the test flights. We felt the system under test was able to complete the requirements levied by the contract (objective requirements), and the AACUS/TAOLOS would have passed the DT and transitioned to OT. However, several of the landings were not optimal. In more than one case, the safety pilot took the controls and delivered vehicle to a more favorable location. Once the TALOS found a location that met the minimum requirements it was programmed to execute, it stopped looking for a better solution. The senior naval officers felt that a HAC needs to use their judgment to pick the best available location for landing. While the system can accomplish the CAL/LZ mission by satisfying the SWEEP checklist and executing an autonomous landing, a more ideal landing point offers an extra buffer of safety. One example was flight 59F097. During that flight, the safety pilot disengaged the system and chose a touchdown point to maximize the impending static display following shutdown. The system under test was not aware that a number of high-ranking Marine officers were waiting to see the vehicle. Its only concern was finding a valid landing spot. The safety pilot knew that the closer he could land to the distinguished visitors, the better. This showed the narrow focus of the decision engine. changing the programming for touchdown point is not possible between flights. It was not possible to add judgment in the current build of the software.

During follow-on testing at Twentynine Palms (considered to be OT data), the system under test was able to complete 46 autonomous landings in mission-representative environments. However, the decision engine displayed issues with distinguishing valid landing zones for the test vehicle. This may have been a byproduct of the demonstration program requiring a large safety buffer (a much larger clear LZ than required for the platform). The software required a large-diameter clear zone for landing. On more than one flight, the safety pilot had to take control of the aircraft and execute a safe landing in an area that the decision engine eliminated as a valid LZ. The judgment that senior naval officers rely upon when granting the HAC qualification on aviators is an intangible that is difficult to quantify or program into a decision engine. Ultimately, we determined (with coordination with military certification officials) the system under test was unsuitable for the CAL/LZ mission and would not be granted a safety of flight clearance as programmed/evaluated.

AFS was able to develop a decision engine and sensor package that could perform the CAL/LZ mission autonomously under controlled conditions. However, when presented with other variables that were not considered, or under field conditions, the decision engine lacked the judgment that a HAC needs to demonstrate to their commanding officer before being fully qualified. This highlights a major issue with certifying autonomous behavior for a safety of flight certification. If requirements are black and white, a simple decision tree can be generated for a decision engine to follow. It is when the decision engine faces off-nominal conditions, or unplanned circumstances present themselves, that its actions do not mirror that of a fully qualified HAC.

Academia and industry have proven that they can build aircraft with autonomous functionality. The AACUS/TALOS was one such example. However, it was a technology demonstration and was never intended for use beyond that. It was given a specific set of requirements to demonstrate, and it was programmed to do so. This research demonstrated that in order to obtain a safety of flight clearance for autonomous functionality, the vehicle must prove that it can perform similar actions to those of a qualified pilot under off-nominal, or mission-representative, conditions.

VI. Conclusions

The existing paradigm for T&E is to define what a system will do given a set of input parameters. Before a safety of flight clearance, certification officials currently need to understand how a system will react when used by a fully qualified pilot/operator when completing a mission (such as the CAL/LZ mission). By removing the pilot/operator (for autonomous systems), it is believed that a flight clearance can be obtained for autonomy based on what the system will not do. To define a box where a system can be allowed to exhibit autonomous behavior, the SWEEP checks performed by H-60 and H-1 aircrew in the USN and USMC were used to complete the CAL/LZ mission. It was possible to evaluate flight-test data of an autonomous system (the AFS AACUS/TALOS UH-1), completing the CAL/LZ mission under controlled conditions (DT) and under mission-representative conditions (OT).

Between the AACUS/TALOS's final demonstration (to include the rehearsals) and the ONR technology maturation assessment, the decision engine under test demonstrated 33 autonomous landings and several wave-off approaches based on a fouled LZ. These flights could be considered DT events because the conditions were controlled to demonstrate the objective requirements of the contract for which the system was acquired under. During these test flights, the decision engine was able to define a safe landing spot that met the constraints of the contract. Therefore, the decision engine would have met the objective requirements of DT. However, several deficiencies were noted with the system. The most troubling was that once the system picked a landing point that satisfied its programming, it did not continue looking for a more advantageous spot. Yet, based on the performance of the system under controlled conditions, it would have passed DT and been recommended for OT (to be evaluated under mission-representative conditions).

During the ITX evaluation period, the AACUS/TALOS was used in a mission-representative environment (OT): Twentynine Palms Marine Base. During 15 flights, the system under test executed 46 autonomous landings in environments similar to those that would be needed to execute the CAL/LZ mission to resupply Marines in the field. However, the OT evaluation is a subjective test: the purpose of which is to determine, to the subjective opinion of the OT organization, if a standard fleet user can use the system under test to complete the desired mission under mission-representative conditions. While the vehicle demonstrated the ability to stay within the clearly defined envelope, several decisions made by the vehicle were in contrast to what a qualified HAC would have made. None of the decisions would have resulted in an unsafe condition. However, the results of an OT report on the data available would have found the system under test unsuitable for the autonomous CAL/LZ mission as programmed.

This paper used legacy test procedures for the evaluation of the system under test. While the procedures provided data on the system and may be a valid method to test an autonomous system, they did not provide a method to correct issues early in the development cycle. Once a system reaches flight-test, it is extremely difficult to fix the system and still meet deadlines. If an autonomous system were certified safe for flight, the most important step would be to ensure the requirements were specified in such a manner that system developers could program in the ability to cope with off-nominal conditions.

Academia and industry have demonstrated that they can develop a system that can exhibit autonomous behavior while completing a mission normally reserved for qualified pilots under controlled conditions; the AACUS/TALOS is one such system. However, when confronted with conditions that were not programmed into the decision engine, the actions of the autonomous system did not match that of a fully qualified pilot. By using the SWEEP checklist as a guarantee of what the system will not do, flight clearance officials can grant a safety of flight clearance for autonomy. However, before authorizing the software package to complete tasks that require a pilot's judgment, the system needs to demonstrate it can accomplish the mission under controlled and off-nominal conditions.

The existing data on the AACUS/TALOS system are promising for the future of unmanned vehicles supporting the CAL/LZ mission. However, the narrowly defined focus of the current AACUS/TALOS architecture is inadequate for the mission need. Future studies are recommended that would expand upon what constitutes a safe landing area. Advanced sensors and an adequate training set of visual images can help support relaxation of what makes a clear LZ (a thin bush may be acceptable, where a small rock may not be acceptable). Until the judgment (subjective means) of a HAC can be translated into an algorithm (objective measure), autonomous behavior will have difficulties being accepted by certification officials.

References

[1] "Teal Group Releases Study, 'World Unmanned Aerial Vehicle Systems, Market Profile and Forecast 2015'," Manufacturing Close-Up, Aug. 2015.

[2] "FAA Creates Commercial Drone Pilot's License (Part 107): American Drone Operators Will Be Able to Apply for a Drone Pilot's License Starting in August 2016," Marketwired, June 2016.

[3] 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.

[4] Li, S., Peng, Y., Lu, Y., Zhang, L., and Liu, Y., "MCAV/IMU Integrated Navigation for the Powered Descent Phase of Mars EDL," *Advances in Space Research*, Vol. 46, No. 5, 2010, pp. 557–570.
https://doi.org/10.1016/j.asr.2010.04.028

[5] "Airborne Software Assurance," U.S. Dept. of Transportation, Federal Aviation Administration Advisory Circular 20-115C, 21 June, 2013.

[6] Webster, M., Cameron, N., Fisher, M., and Jump, M., "Generating Certification Evidence for Autonomous Unmanned Aircraft Using Model Checking and Simulation," *Journal of Aerospace Information Systems*, Vol. 11, No. 5, 2014, pp. 258–279.
https://doi.org/10.2514/1.I010096

[7] Weibel, R. E., and Hansman, R. J., "Safety Considerations for Operation on Unmanned Aerial Vehicles in the National Airspace System," Massachusetts Inst. of Technology International Center for Air Transportation TR-ICAT 2005-1, Cambridge, MA, 2005.

[8] "BCAR Section A: Airworthiness Procedures Where the CAA Has Primary Responsibility for Type Approval of a Product," Civil Aviation Authority Rept. CAP 553, London, Oct. 2011.

[9] 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

[10] Bradley, M., Nealiegh, M., Oh, J. S., Elster, E. A., Rich, N. M., and Rothberg, P., "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

[11] 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

[12] Rodd, S. A., "Government Laboratory Technology Transfer: Process and Impact Assessment," Ph.D. Thesis, Virginia Polytechnic Inst. and State Univ., Blacksburg, VA, 1998.

[13] 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

[14] "NATOPS Flight Manual, MH-60R Helicopter," Chief of Naval Operations Under the Direction of the Commander, Naval Air Systems Command, 2014.

[15] Koopman, P., and Wagner, M., "Challenges in Autonomous Vehicle Testing and Validation," *SAE International Journal of Transportation Safety*, Vol. 4, No. 1, 2016, pp. 15–24.
https://doi.org/10.4271/2016-01-0128

[16] "NAVAIR Airworthiness and Cybersafe Process Manual, NAVAIR Manual M-13034.1," Naval Air Systems Command, 2016.

[17] 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

[18] "Preliminary NATOPS Flight Manual, MQ-8B/C Unmanned Aircraft System," Chief of Naval Operations Under the Direction of the Commander, Naval Air Systems Command, 2019.

[19] "SAE International Announces Formation of New Standards Committee for Artificial Intelligence in Aviation Systems," Contify Aviation News, June 2019.

[20] 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, 2015.
https://doi.org/10.2514/6.2016-0380

[21] 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, 2016.
https://doi.org/10.2514/6.2017-1744

[22] 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

[23] Graydon, M., "Retrospectively Documenting SAFEGUARD's Possession of the Overarching Properties," *2019 49th Annual IEEE/IFIP International Conference on Dependable Systems and Networks—Supplemental Volume (DSN-S)*, Portland, OR, 2019, pp. 27–28.
https://doi.org/10.1109/DSN-S.2019.00019

[24] 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 (USD(R&E;)), The Pentagon, Washington, D.C., 2015.

[25] Gross, K. H., Fifarek, A. W., and Hoffman, J. A., "Incremental Formal Methods Based Design Approach Demonstrated on a Coupled Tanks Control System," *2016 IEEE 17th International Symposium on High Assurance Systems Engineering (HASE)*, Inst. of Electrical and Electronics Engineers, New York, 2016, pp. 181–188.
https://doi.org/10.1109/HASE.2016.16.

[26] Abraham, J., "Verification and Validation Spanning Models to Code," *AIAA Modeling and Simulation Technologies Conference*, AIAA Paper 2015-1593, 2015.
https://doi.org/10.2514/6.2015-1593

[27] Blanchette, S., "Giant Slayer: Will You Let Software be David to Your Goliath System?" *Journal of Aerospace Information Systems*, Vol. 13, No. 10, 2016, pp. 407–417.
https://doi.org/10.2514/1.I010448

[28] Muli, M., Moudgal, V., and Allen, J., "Best Practices and Recommendations for Model-Based Development Process," SAE International TP 2015-01-2529, Warrendale, PA, 2015.
https://doi.org/10.4271/2015-01-2529

[29] Israelsen, B. W., Ahmed, N. R., Center, K., Green, R., and Bennett, W., "Towards Adaptive Training of Agent-Based Sparring Partners for Fighter Pilots," *AIAA Information Systems-AIAA Infotech @ Aerospace*, AIAA Paper 2017-0343, 2017.
https://doi.org/10.2514/6.2017-0343

[30] Fisher, M., Dennis, L., and Webster, M., "Verifying Autonomous Systems," *Communications of the Association for Computing Machinery*, Vol. 56, No. 9, 2013, pp. 84–93.
https://doi.org/10.1145/2494558

[31] Tobias, E., Tischler, M., Berger, T., and Hagerott, S. G., "Full Flight-Envelope Simulation and Piloted Fidelity Assessment of a Business Jet Using a Model Stitching Architecture," *AIAA Modeling and Simulation Technologies Conference*, AIAA Paper 2015-1594, 2015.
https://doi.org/10.2514/6.2015-1594

[32] Eisemann, U., and Allen, J. L., "New Requirement-Definition and Verification Techniques According to DO-178C, DO-331, and DO-333," *AIAA Infotech @ Aerospace*, AIAA Paper 2016-0223, 2016.
https://doi.org/10.2514/6.2016-0223

[33] Berger, T., Tischler, M., Hagerott, S. G., Cotting, M. C., Gray, W. R., Gresham, J., George, J., Krogh, K., D'Argenio, A., and Howland, J., "Development and Validation of a Flight-Identified Full-Envelope Business Jet Simulation Model Using a Stitching Architecture," *AIAA Modeling and Simulation Technologies Conference*, AIAA Paper 2018-0525, 2017.
https://doi.org/10.2514/6.2018-0525

[34] Walker, S. M., Shan, J., and Liu, L., "DIMMACSS-Stage: A Distributed Intelligence Model for a Multi-Agent Control System Using Simulink and the Stage Robotic Simulator," *AIAA Modeling and Simulation Technologies Conference*, AIAA Paper 2014-0639, 2014.
https://doi.org/10.2514/6.2014-0639

[35] Hangal, S. A., Tak, B., and Arya, H., "Distributed Hardware-in-Loop Simulations for Multiple Autonomous Aerial Vehicles," *AIAA Modeling and Simulation Technologies Conference*, AIAA Paper 2015-0151, 2015.
https://doi.org/10.2514/6.2015-0151.

[36] Mohamedy, A. E.-D. S., Aly, A. M., and Elnashar, A. H., "Modeling and Simulation Hardware-in-the-Loop for Unmanned Aerial Vehicle," *AIAA Modeling and Simulation Technologies Conference*, AIAA Paper 2016-1428, 2016.
https://doi.org/10.2514/6.2016-1428

[37] Baier, C., and Katoen, J., *Principles of Model Checking*, MIT Press, Cambridge, MA, 2008.

[38] Kane, A., "Runtime Monitoring for Safety-Critical Embedded Systems," Ph.D. Dissertation, Electrical and Computer Engineering, Carnegie–Mellon Univ., Pittsburgh, PA, 2015.

[39] Coombes, M., Chen, W.-H., and Render, P., "Site Selection During Unmanned Aerial System Forced Landings Using Decision-Making Bayesian Networks," *Journal of Aerospace Information Systems*, Vol. 13, No. 12, 2016, pp. 491–495.
https://doi.org/10.2514/1.I010432

[40] Gross, K. H., Clark, M., Hoffman, J. A., Fifarek, A., Rattan, K., Swenson, E., Whalen, M., and Wagner, L., "Formally Verified Run Time Assurance Architecture of a 6U CubeSat Attitude Control System," *AIAA Infotech @ Aerospace*, AIAA Paper 2016-0222, 2016.
https://doi.org/10.2514/6.2016-0222

[41] Torens, C., Adolf, F., Faymonville, P., and Schirmer, S., "Towards Intelligent System Health Management Using Runtime Monitoring," *AIAA Information Systems-AIAA Infotech @ Aerospace*, AIAA Paper 2017-0419, 2017.
https://doi.org/10.2514/6.2017-0419

[42] Gross, K. H., Clark, M. A., Hoffman, J. A., Swenson, E. D., and Fifarek, A. W., "Run-Time Assurance and Formal Methods Analysis Nonlinear System Applied to Nonlinear System Control," *Journal of Aerospace Information Systems*, Vol. 14, No. 4, 2017, pp. 232–246.
https://doi.org/10.2514/1.I010471

[43] Schierman, J., Ward, D., Dutoi, B., Aiello, A., Berryman, J., DeVore, M., Storm, W., and Wadley, J., "Run-Time Verification and Validation for Safety-Critical Flight Control Systems," *AIAA Guidance, Navigation and Control Conference and Exhibit*, AIAA Paper 2008-6338, 2008.
https://doi.org/10.2514/6.2008-6338

[44] Lichter, M., Bateman, A., and Balas, G., "Flight-Test Evaluation of a Run-Time Stability Margin Estimation Tool," *AIAA Guidance, Navigation, and Control Conference*, AIAA Paper 2009-6257, 2009.
https://doi.org/10.2514/6.2009-6257

[45] Rabideau, G., Chien, S., and Mclaren, D., "Onboard Run-Time Goal Selection for Autonomous Operations," *SpaceOps 2010 Conference*, AIAA Paper 2010-2203, 2010.
https://doi.org/10.2514/6.2010-2203

[46] Sababha, B., Yang, H. C., and Rawashdeh, O., "An RTOS-Based Run-Time Reconfigurable Avionics System for UAVs," *AIAA Infotech@ Aerospace 2010*, AIAA Paper 2010-3414, 2010.
https://doi.org/10.2514/6.2010-3414

[47] Aiello, M., Berryman, J., Grohs, J., and Schierman, J., "Run-Time Assurance for Advanced Flight-Critical Control Systems," *AIAA Guidance, Navigation, and Control Conference*, AIAA Paper 2010-8041, 2010.
https://doi.org/10.2514/6.2010-8041

[48] Wong, E., Schierman, J. D., Schlapkohl, T., and Chicatelli, A., "Towards Run-Time Assurance of Advanced Propulsion Algorithms," *50th AIAA/ASME/SAE/ASEE Joint Propulsion Conference*, AIAA Paper 2014-3636, 2014.
https://doi.org/10.2514/6.2014-3636

[49] Berezin, S., "Model Checking and Theorem Proving: A Unified Framework," Ph.D. Dissertation, School of Computer Science, Carnegie–Mellon Univ., Pittsburgh, PA, 2002.

[50] Webster, M., Cameron, N., Jump, M., and Fisher, M., "Towards Certification of Autonomous Unmanned Aircraft Using Formal Model Checking and Simulation," *Infotech@Aerospace 2012*, AIAA Paper 2012-2573, 2012.
https://doi.org/10.2514/6.2012-2573

[51] Good, N., Aboutalib, O., Thai, B., Yamaoka, N., Kim, C., Wilkinson, C., and Findlay, D., "Validation Process of the Physics-Based Modeling of Navigation Sensors for Sea-Based Aviation Automated Landing," *AIAA Modeling and Simulation Technologies Conference*, AIAA Paper 2016-1919, 2016.
https://doi.org/10.2514/6.2016-1919

[52] Bakera, M., Margaria, T., Renner, C. D., and Steffen, B., "Game-Based Model Checking for Reliable Autonomy in Space," *Journal of Aerospace Computing, Information, and Communication*, Vol. 8, No. 4, 2011, pp. 100–114.
https://doi.org/10.2514/1.32013

[53] Sirigineedi, G., Tsourdos, A., White, B., and Zbikowski, R., "Kripke Modelling and Model Checking of a Multiple UAV System Monitoring Road Network," *AIAA Guidance, Navigation, and Control Conference*, AIAA Paper 2010-7582, 2010.
https://doi.org/10.2514/6.2010-7582

[54] Humphrey, L., "Model Checking UAV Mission Plans," *AIAA Modeling and Simulation Technologies Conference*, AIAA Paper 2012-4723, 2012.
https://doi.org/10.2514/6.2012-4723

[55] Verzino, G., "Model Checking Driven Simulation of Sat Procedures," *AIAA SpaceOps 2012 Conference*, AIAA Paper 2012-1275611, 2012.
https://doi.org/10.2514/6.2012-1275611

[56] Humphrey, L., and Patzek, M., "Model Checking Human UAV Mission Plans," *AIAA Guidance, Navigation, and Control (GNC) Conference*, AIAA Paper 2013-5183, 2013.
https://doi.org/10.2514/6.2013-5183

[57] Torens, C., and Adolf, F., "Using Formal Requirements and Model-Checking for Verification and Validation of an Unmanned Rotorcraft," *AIAA Infotech @ Aerospace*, AIAA Paper 2015-1645, 2015.
https://doi.org/10.2514/6.2015-1645

[58] Hansen, J. P., and Wrage, L., "Verification of Real-Time Systems Using Statistical Model Checking," *AIAA Infotech @ Aerospace*, AIAA Paper 2015-1866, 2015.
https://doi.org/10.2514/6.2015-1866

[59] Nandiganahalli, J. S., Lee, S., and Hwang, I., "Flight Deck Mode Confusion Detection Using Intent-Based Probabilistic Model Checking," *AIAA Information Systems-AIAA Infotech @ Aerospace*, AIAA Paper 2017-0345, 2017.
https://doi.org/10.2514/6.2017-0345

[60] Ouimet, M., "Formal Software Verification: Model Checking and Theorem Proving," Embedded Systems Lab., Massachusetts Inst. of Technology TR ESL-TIK-00214, Cambridge, MA, 2008.

[61] Sutcliffe, G., Denney, E., and Fischer, B., "Practical Proof Checking for Program Certification," 2005, https://ti.arc.nasa.gov/m/profile/edenney/papers/escar05.pdf [retrieved 19 Dec. 2017].

[62] Muñoz, C., Dutle, A., Narkawicz, A., and Upchurch, J., "Unmanned Aircraft Aystems in the National Airspace System: A Formal Methods Perspective," *ACM SIGLOG News*, Vol. 3, No. 3, 2016, pp. 67–76.
https://doi.org/10.1145/2984450.2984459

[63] Goodloe, A., Gunter, C., and Stehr, M.-O., "Formal Prototyping in Early Stages of Protocol Design," *Proceedings of the 2005 Workshop on Issues in the Theory of Security*, Jan. 2005, pp. 67–80.
https://doi.org/10.1145/1045405.1045413

[64] Jiang, Y., Liu, J., Dowek, G., and Ji, K., "SCTL: Towards Combining Model Checking and Proof Checking," 2016, https://www.researchgate.net/publication/304547954_SCTL_Towards_Combining_Model_Checking_and_Proof_Checking.

[65] Asokan, S., Kumar, G. S., and Lal, N. J., "Modeling of ALFA Programs Using PVS Theorem Prover," Inst. of Electrical and Electronics Engineers, New York, 2009, pp. 373–375.
https://doi.org/10.1109/ARTCom.2009.134

[66] Muñoz, C. A., "Formal Methods in Air Traffic Management: The Case of Unmanned Aircraft Systems (Invited Lecture), *Proceedings of the 12th International Colloquium on Theoretical Aspects of Computing*, Lecture Notes in Computer Science, Vol. 9399, NASA/Langley Research Center, Hampton, 2015, pp. 58–62.

[67] Jenie, Y. I., Kampen, E.-J. V., Ellerbroek, J., and Hoekstra, J. M., "Safety Assessment of Unmanned Aerial Vehicle Operations in an Integrated Airspace," *AIAA Infotech @ Aerospace*, AIAA Paper 2016-1000, 2016.
https://doi.org/10.2514/6.2016-1000

[68] Guarro, S., Yau, M. K., Ozguner, U., Aldemir, T., Kurt, A., Hejase, M., and Knudson, M., "Formal Framework and Models for Validation and Verification of Software-Intensive Aerospace Systems," *AIAA Information Systems-AIAA Infotech @ Aerospace*, AIAA Paper 2017-0418, 2017.
https://doi.org/10.2514/6.2017-0418

[69] Chien, S., Doubleday, J., Thompson, D. R., Wagstaff, K. L., Bellardo, J., Francis, C., Baumgarten, E., Williams, A., Yee, E., Stanton, E., and Piug-Suari, J., "Onboard Autonomy on the Intelligent Payload Experiment CubeSat Mission," *Journal of Aerospace Information Systems*, Vol. 14, No. 6, 2017, pp. 307–315.
https://doi.org/10.2514/1.I010386

[70] Moses, C., Chipalkatty, R., and Platt, R., "Belief Space Hierarchical Planning in the Now for Unmanned Aerial Vehicles," *AIAA Infotech @ Aerospace*, AIAA Paper 2016-0130, 2016.
https://doi.org/10.2514/6.2016-0130

[71] Carreon-Limones, C., Rashid, A., Chung, P., and Bhandari, S., "3-D Mapping using LiDAR and Autonomous Unmanned Aerial Vehicle," *AIAA Information Systems-AIAA Infotech @ Aerospace*, AIAA Paper 2017-1155, 2017.
https://doi.org/10.2514/6.2017-1155

[72] Agrawal, P., Ratnoo, A., and Ghose, D., "Image Segmentation-Based Unmanned Aerial Vehicle Safe Navigation," *Journal of Aerospace Information Systems*, Vol. 14, No. 7, 2017, pp. 391–410.
https://doi.org/10.2514/1.I010457

[73] Chao, H., Gu, Y., Gross, J., Rhudy, M., and Napolitano, M., "Flight-Test Evaluation of Navigation Information in Wide-Field Optical Flow," *Journal of Aerospace Information Systems*, Vol. 13, No. 11, 2016, pp. 419–432.
https://doi.org/10.2514/1.I010482

[74] Desai, A., and Lee, D.-J., "Efficient Feature Descriptor for Unmanned Aerial Vehicle Ground Moving Object Tracking," *Journal of Aerospace Information Systems*, Vol. 14, No. 6, 2017, pp. 345–349.
https://doi.org/10.2514/1.I010503

[75] Roback, V. E., Pierrottet, D. F., Amzajerdian, F., Barnes, B. W., Bulyshev, A. E., Hines, G. D., Petway, L. B., Brewster, P. F., and Kempton, K. S., "LiDAR Sensor Performance in Closed-Loop Flight Testing of the Morpheus Rocket-Propelled Lander to a Lunar-Like Hazard Field," *AIAA Guidance, Navigation, and Control Conference*, AIAA Paper 2015-0328, 2015.
https://doi.org/10.2514/6.2015-0328

[76] Cowlagi, R. V., and Sperry, J., "Unifying Artificial Intelligence and Trajectory Optimization for UAV Guidance," *AIAA Guidance, Navigation, and Control Conference*, AIAA Paper 2016-0381, 2016.
https://doi.org/10.2514/6.2016-0381

[77] Muñoz, C., and Narkawicz, A., "Formal Analysis of Extended Well-Clear Boundaries for Unmanned Aircraft," *NASA Center for AeroSpace Information (CASI). Conference Proceedings*, 2016.
https://doi.org/10.1007/978-3-319-40648-0_17

[78] Ghatas, R. W., Jack, D. P., Tsakpinis, D., Vincent, M. J., Sturdy, J. L., Muñoz, C. A., Hoffler, K. D., Dutle, A. M., Myer, R., DeHaven, A. M., Lewis, T., and Arthur, K. E., "Unmanned Aircraft Systems Minimum Operational Performance Standards End-to-End Verification and Validation (E2-V2) Simulation," NASA TM-2017-20780, 2017.

[79] Narkawicz, A., Munoz, C., and Dutle, A., "Coordination Logic for Repulsive Resolution Maneuvers," *16th AIAA Aviation Technology, Integration, and Operations Conference*, AIAA Paper 2016-3156, 2016.
https://doi.org/10.2514/6.2016-3156

[80] Bhandari, S., and Novak, J., "Neural Network Based Control of an Airplane UAV Using Radial Basis Functions," *AIAA Guidance, Navigation, and Control Conference*, AIAA Paper 2015-0076, 2015.
https://doi.org/10.2514/6.2015-0076

[81] Costello, D. H., and Xu, H., "Generating Certification Evidence for Autonomous Aerial Vehicles Decision-Making," *Journal of Aerospace Information Systems*, advance online publication, 22 Oct. 2020.

[82] Thrun, S., Montemerlo, M., Dahlkamp, H., Stavens, D., Aron, A., Diebel, J., Fong, P., Gale, J., Halpenny, M., Hoffmann, G., Lau, K., Oakley, C., Palatucci, M., Pratt, V., Stang, P., Strohband, S., Dupont, C., Jendrossek, L.-E., Koelen, C., Markey, C., Rummel, C., van Niekerk, J., Jensen, E., Alessandrini, P., Bradski, G., Davies, B., Ettinger, S., Kaehler, A., Nefian, A., and Mahoney, P., "Stanley: The Robot that Won the DARPA Grand Challenge," *Journal of Field Robotics*, Vol. 23, No. 9, 2006, pp. 661–692.
https://doi.org/10.1002/rob.20147

[83] "Investigation—PE 16-007, ODI Resume," National Highway Traffic Safety Administration, U.S. Dept. of Transportation, 2016.

[84] Browne, B. A., "Self-Driving Cars: On the Road to a New Regulatory Era," *Journal of Law, Technology and the Internet*, Vol. 8, 2017, p. 1, https://link.gale.com/apps/doc/A558368494/AONE?u=embry&sid;=AONE&xid;=303e7af0 [accessed 4 Nov. 2020].

[85] Roth, M. L., "Regulating the Future: Autonomous Vehicles and the Role of Government," *Iowa Law Review*, Vol. 105, No. 3, 2020, pp. 1411–1446.
https://doi.org/10.4271/2016-01-0128

[86] "Naval Air Training and Operations Procedures Standardization (NATOPS) General Flight and Operating Instructions," Commander Naval Air Forces, 2016.

[87] "Operation of the Defense Acquisition System," U.S. Dept. of Defense Instruction Number 5000.02T, 2020.

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

[89] Jewell, J., "Autonomy Experimentation in an Operationally Relevant Scenario," *Vertical Flight Society's 75th Annual Forum and Technology Display*, Vertical Flight Soc., Fairfax, VA, 2019.

[90] 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.

[91] Whalley, M. S., Takahashi, M. D., Fletcher, J. W., Moralez, E., III, Ott, L. C. R., Olmstead, L. M. G., Savage, J. C., Goerzen, C. L., Schulein, G. J., Burns, H. N., and Conrad, B., "Autonomous Black Hawk in Flight: Obstacle Field Navigation and Landing-Site Selection on the RASCAL JUH-60A," *Journal of Field Robotics*, Vol. 31, No. 4, 2014, pp. 591–616.
https://doi.org/10.1002/rob.21511

[92] Bhandari, S., Farinella, J., and Lay, C., "UAV Collision Avoidance Using a Predictive Rapidly-Exploring Random Tree (RRT)," *AIAA Infotech @ Aerospace*, AIAA Paper 2016-2197, 2016.
https://doi.org/10.2514/6.2016-2197

[93] Lee, H., Lee, D., and Shim, D. H., "Receding Horizon-Based RRT* Algorithm for a UAV Real-Time Path Planner," *AIAA Information Systems-AIAA Infotech @ Aerospace*, AIAA Paper 2017-0676, 2017.
https://doi.org/10.2514/6.2017-0676

[94] Duffie, W. Jr., "Autonomous Flight Technology to Provide Rapid Resupply for Marines," Defense Logistics Agency, Ft. Belvoir, VA, 2017, http://dla.mil/AboutDLA/News/NewsArticleView/Article/1397666/autonomous-flight-technology-to-provide-rapid-resupply-for-marines/ [retrieved 15 March 2020].

[95] Grimes, J., "Images from UH-1 over NAS Patuxent River Taken from the Pilots Perspective at 200 Ft AGL," Private Collection of CDR Costello, 2020.

[96] Weakley, J., "Images from UH-1 over NAS Patuxent River Taken from the Pilots Prespective at 200 Ft AGL," Private Collection of CDR Costello, 2020.

[97] Costello, D., "Images from AACUS Demonstrations," Private Collection of CDR Costello, 2017.

[98] Vincenti, W., *What Engineers Know and How They Know It: Analytical Studies from Aeronautical History*, Johns Hopkins Univ. Press, Baltimore, MD, 1990.

[99] Niewoehner, R. J., O'Conner, J., and Traven, R., "What We're Learning About Learning: Flight-Test Implications," *61st Symposium Society of Experimental Test Pilots*, Soc. of Experimental Test Pilots, Lancaster, CA, 2017.

J. Kuchar
Associate Editor