7 (1): 7-16 (2023)

Journal of Aviation

https://dergipark.org.tr/en/pub/jav

e-ISSN 2587-1676

A Framework for Airworthiness Certification of Autonomous Systems Within United States Naval Aviation

Donald Costello[1*], Richard Adams[2]

[1*] United States Naval Academy, Annapolis, Maryland, United States (dcostell@usna.edu)
[2] Romeo Alpha, L.L.C., Prince Frederick, Maryland, United States (Richard.adams@romeoalphallc.com)

Article Info

Received: 18 August 2022
Revised: 18 December 2022
Accepted: 26 December 2022
Published Online: 26 February 2023

Keywords:
Airworthiness Certification
Certification of Autonomous Systems
Military Unmanned Aviation

Corresponding Author: *Donald Costello*

RESEARCH ARTICLE

https://doi.org/10.30518/jav.1161725

Abstract

A need has been identified to establish a strategy and framework for airworthiness certification of autonomous systems in Naval Aviation. The purpose of this research was to document a two-day virtual summit that was held in June 2021 to develop an initial strategy and framework for airworthiness certification of autonomous functionality in Naval Aviation air systems. The summit was designed to engage the airworthiness community to start the development of a strategy and framework for airworthiness certification of autonomous systems. The summit was attended by representatives from government, industry and academia. Following two full days of discussions, the group came to a consensus on a strategy and framework for airworthiness certification of airborne autonomous functions based on an example use case: an unmanned aircraft performing a drogue aerial refueling task. This paper summarizes the summit and it's outcomes in an effort to stimulate the certification community to develop methods for certifying autonomous behaviour.

1. Introduction

All modern aircraft have some level of automation. However, there is increasing interest in the implementation of autonomous systems and functions in U.S. military air systems. Current certification standards rely on a human to have overall responsibility for the air vehicle. An autonomous system will not have a human in or on the loop when it operates. These systems present unique airworthiness certification challenges, and a need has been identified to establish an airworthiness certification strategy and framework for these systems.

The various branches of the U.S. armed forces have documented processes for airworthiness certification of military air systems (NAVAIR Instruction 13034.1F: Airworthiness and Cybersecurity Safety Policies for Air Vehicles and Aircraft Systems, 2016; NAVAIR Airworthiness and Cybersafe Process Manual, NAVAIR Manual M-13034.1, 2016; Air Force Instruction 62–601: USAF Airworthiness, 2010; Air Force Policy Directive 62–6: USAF Airworthiness, 2019; Army Regulation 70–62: Airworthiness of Aircraft Systems, 2016; MIL-HDBK-516C: Department of Defense Handbook, Airworthiness Certification Criteria, 2014). These processes primarily utilize a set of airworthiness certification criteria, standards and methods of compliance to establish the airworthiness basis for an air system. While the overall airworthiness processes and procedures can be leveraged for air systems implementing autonomous functions, a strategy and framework for airworthiness certification of air systems with autonomous functions needs to be developed to adapt to this emerging capability. Airworthiness certification processes are designed to ensure that an air system is airworthy and safe for its intended mission in its intended operating environment.

As part of the airworthiness process, an air system is evaluated against a set of airworthiness criteria and standards. Operating limitations, normal and emergency procedures, warnings, cautions and notes are established, and safety risks are assessed and accepted at the appropriate level. In general, current airworthiness criteria and standards used to assess airworthiness and safety of flight of air system functions, hardware and software assume deterministic system behavior.

To stimulate discussion and build consensus on a possible path towards an airworthiness certification strategy and framework for autonomous systems, a two-day summit was held in June 2021 entitled June 2021 Summit on Certification of Autonomous Systems Within Naval Aviation. The summit was held in conjunction with the National Airworthiness Council Artificial Intelligence Working Group (NACAIWG). It was attended by representatives from the United States government airworthiness community (United States Navy (USN), United States Air Force (USAF), United States Army (USA), National Aeronautics and Space Administration (NASA), and the Federal Aviation Administration (FAA)), industry (Boeing, Lockheed Martin, Northrop Grumman, and

Aurora Flight Sciences) and academia (Purdue University, University of Maryland, and the United States Naval Academy). Following two full days of discussions, the group came to a consensus on a strategy and framework for airworthiness certification of airborne autonomous functions based on an example use case (an unmanned aircraft performing a drogue aerial refueling task). This paper summarizes the outcomes of the summit.

The contributions of this paper include documentation of the summit itself, and the framework agreed upon by all participants (all of whom were subject matter experts in safe for flight certification from industry, the United States government, and academia). It allowed an open dialog between the government and industry in an open forum. Most engagement with the industry is only with one company at a time. By using a hypothetical use case, every member of the summit was able to openly contribute to the shared goal of developing a framework for the certification of autonomous systems within naval aviation. Ultimately, the results of the summit were briefed and endorsed by the NACAIWG.

2. Background

This section provides a summary of definitions and concepts, relevant instructions, standards and papers related to airworthiness certification, system safety, development assurance, and certification of autonomous function in aviation.

2.1. Definitions and Concepts

Based on a literature review, ASTM F3060-20 (ASTM International: F3060-20 Standard Terminology for Aircraft, 2020). Standard Terminology for Aircraft, was determined to be the most comprehensive source of relevant definitions addressing autonomous systems and functions. The summit adopted the definitions from ASTM F3060-20 to enable a common understanding of relevant definitions. As an industry consensus standard, ASTM F03060-20 provides a set of relevant standardized definitions applicable to the certification of autonomous systems including definitions for automatic, autonomous, artificial intelligence (AI), adaptive system, complex system, deterministic system, non-deterministic system, intelligent agent, machine learning (ML), run-time assurance (RTA) architecture and safety monitor. It also defines the differences between humans in the loop, humans on the loop and humans out of the loop.

Human in-the-loop systems requires a human to interact with the system to be able to perform its intended functions or control actions. Human on loop systems is characterized by functions where a human can give guidance to an automated system that has the authority to perform functions or control actions without human oversight or actions. Human out-of-the-loop systems are characterized by systems in which a human is not able to intervene or provide guidance to an automatic (or autonomous) system. The system has the authority to perform functions and control actions without human oversight or actions.

For automated functions, a system performs actions without the need for human intervention and may provide the capability for a human to monitor and override the system (to include when the system performs off-nominal or to prevent a mishap). These functions typically operate with a human on the loop as a safety monitor. Current unmanned aerial systems (UAS) such as the MQ-4C and MQ-8 function automatically but with air vehicle operator (AVO) supervision. They follow preplanned routes and utilize a deterministic rule-based architecture. In the absence of human intervention, the system will perform its programmed mission automatically.

For autonomous functions, the system is delegated authority by a human to independently determine a new course of action in the absence of a predefined plan and to accomplish goals based on its knowledge and understanding of its situational observation of the operational environment. The system takes actions that are dependent on sensing and interpreting the external environment. Autonomous systems utilizing ML improve their performance by exposure to data without the need to follow explicitly programmed instructions. These systems have the ability and authority to make decisions independently and self-sufficiently without human intervention. Autonomous functions do not preclude the ability of a human to monitor (on the loop) and override the autonomous function if provisions for human monitoring and intervention are considered in the design. However, for the summit, it was assumed that the autonomous functionality being certified would not have a human in or on the loop.

Deterministic behavior is rule-based. For a given input, a deterministic system will exhibit known behavior based on known input conditions and always produce the same output. There is only one potential output for a defined set of inputs. Automated functions are deterministic.

Non-deterministic systems rely on observation to influence the output. There are multiple potential outputs to a single input. The exact behavior of the system cannot be predicted based on input conditions. Autonomous functions are typically non-deterministic.

DoD Directive 3000.09 (Department of Defense Directive 3009.09: Autonomy in Weapons Systems, 2017) establishes guidelines for autonomy in weapon systems used for lethal, non-lethal, kinetic and non-kinetic use. It provides a set of top-level principles that can be used as a basis for establishing principles for the certification of autonomous functions in air systems. Among these are:

- A human is ultimately responsible for the system and its behavior. A human is responsible for deciding to delegate authority to the system to operate with little or no human intervention.
- Autonomous systems should provide operators with feedback on system status and enable operators to activate and deactivate system functions if needed.
- Autonomous systems should be sufficiently robust to minimize behaviors or failures that lead to safety hazards or unintended consequences.
- Autonomous systems should undergo rigorous hardware and software development and testing, including software validation and verification, lab, ground and developmental flight to ensure the system functions as intended with no unintended or unsafe behaviors. They will function as anticipated in realistic operating environments. Regression tests should be conducted after changes to the system to ensure safety critical systems have not been affected.

2.2. Review of Relevant Airworthiness Instructions and Standards

Relevant Department of Defense (DoD) instructions and handbooks related to airworthiness certification were reviewed. These include:

- NAVAIRINST 13034.1: Airworthiness and Cybersecurity Safety Policies (NAVAIR Instruction 13034.1F: Airworthiness and Cybersecurity Safety Policies for Air Vehicles and Aircraft Systems, 2016).
- NAVAIR M-13034.1: NAVAIR Airworthiness and CYBERSAFE Process Manual (NAVAIR Airworthiness and Cybersafe Process Manual, NAVAIR Manual M-13034.1, 2016).
- AFI 62-601: USAF Airworthiness Instruction (Air Force Instruction 62–601: USAF Airworthiness, 2010).
- AFPD 62-6: USAF Airworthiness Policy (Air Force Policy Directive 62–6: USAF Airworthiness, 2019).
- 70-62: Airworthiness of Aircraft Systems (Army Regulation 70–62: Airworthiness of Aircraft Systems, 2016).
- MIL-HDBK-516C: Airworthiness Certification Criteria Guidance for the DoD (MIL-HDBK-516C: Department of Defense Handbook, Airworthiness Certification Criteria, 2014).

These instructions and policies establish the process and procedures for airworthiness certification for the USN, USAF and USA. They address airworthiness certification processes and procedures but do not specifically address certification criteria and standards for autonomous systems. MIL-HDBK-516C (MIL-HDBK-516C: Department of Defense Handbook, Airworthiness Certification Criteria, 2014) is a tri-service handbook that identifies airworthiness certification criteria, standards and methods of compliance to be considered as part of the airworthiness certification process. MIL-HDBK-516C (MIL-HDBK-516C: Department of Defense Handbook, Airworthiness Certification Criteria, 2014) contains criteria addressing software certification, system safety and vehicle control functions but does not specifically address certification of autonomous systems. These top-level criteria would be generally applicable to autonomous systems. In the future, as the DoD gains experience in the certification of autonomous systems, there may be a need to specifically address autonomous systems within the handbook.

2.3. Review of Relevant System Safety and Development Assurance Standards and Best Practices

Military, civil aviation and industry standards and best practices for system safety and development assurance were reviewed. These include:

- MIL-STD-882E: DoD Standard Practice: System Safety (Department of Defense MIL-STD-882E: Department of Defense Standard Practice, System Safety, 2012).
- Joint Software Systems Safety Engineering Handbook (JSSSEH) (Joint Software Systems Safety Engineering Handbook, 2010).
- FAA Advisory Circular 23.1309E: FAA System Safety Analysis and Assessment for Part 23 Airplanes (Advisory Circular 23.1309.E: System Safety Analysis and Assessment for Part 23 Airplanes, 2011).
- European Union Aviation Safety Agency (EASA) Certification Specification 25 Alternate Means of Compliance AMC 25.1309 (Systems and Equipment) is part of the EASA (Certification Specifications for Large Airplanes CS-25, 2007).
- SAE ARP 4761: Guidelines and Methods for Conducting the Safety Assessment Process on Civil Airborne Systems and Equipment (Guidelines and Methods for Conducting the Safety Assessment Process on Civil Airborne Systems and Equipment, 1996).
- SAE ARP 4754: Guidelines for Development of Civil Aircraft and Systems (Guidelines for Development of Civil Aircraft and Systems, 2010).
- DO-178C, Software Considerations in Airborne Systems and Equipment Certification (Certification Specifications for Large Aeroplanes CS-25 Software Considerations in Airborne Systems and Equipment Certification, 2011).

These standards and best practices document the system safety and software system safety/development assurance processes for military and civil systems. They document a systems-engineering approach for decomposing a system into functions implemented by hardware and software, and establish safety objectives for hardware and software. Both civil and military processes share a common theme of functional decomposition of the system, allocation of functions to hardware and software, and setting safety objectives based on the criticality of the function to be performed in hardware or software.

In general, software functions whose failure results in a catastrophic failure (e.g., loss of aircraft) demands a greater level of development and test rigor (or development assurance) than those that do not. MIL-STD-882E (Department of Defense MIL-STD-882E: Department of Defense Standard Practice, System Safety, 2012) and the JSSSEH (Joint Software Systems Safety Engineering Handbook, 2010) identify Level of Rigor (LOR) Tasks that should be performed based on the criticality of the software function to ensure the software is safe for the intended use. Within the DoD safety framework, software safety risks may be identified and accepted by the appropriate authority for software that does not satisfy the LOR tasks.

Civil airworthiness regulations establish a set of safety objectives and development assurance levels that are assigned at the functional level based on hazard classification. Civil airworthiness certification is compliance based. Unlike the DoD, civil system safety processes do not accommodate safety risk acceptance for software that does not satisfy the required development assurance level. The software must be shown to be compliant with the processes and tasks otherwise the air system may not receive civil certification. The processes defined in ARP 4761 (Guidelines and Methods for Conducting the Safety Assessment Process on Civil Airborne Systems and Equipment, 1996), ARP 4754 (Guidelines for Development of Civil Aircraft and Systems, 2010) and DO-178C (Certification Specifications for Large Aeroplanes CS-25 Software Considerations in Airborne Systems and Equipment Certification, 2011) follow the systems engineering process. Functions are decomposed into hardware and software items. Acceptable failure probabilities are assigned to hardware items, and development assurance levels are assigned to software based on hazard classification (e.g., Catastrophic, Hazardous, Major, etc.) to software items. The development assurance level of the software items establishes the development objectives that must be accomplished. A common theme is that the level of rigor (or development assurance) for safety-critical software is determined by the criticality of the function (i.e., hazard severity and probability of occurrence).

2.4. Other Relevant Standards and Papers on Autonomy

A survey of other relevant standards and papers related to the certification of autonomous systems was performed. Several relevant industry standards and papers were identified. These include:

- Concepts of Design Assurance for Neural Networks (CoDANN) (Concepts of Design Assurance for Neural Networks (CoDANN), 2020), March 2020
- CoDANN II (Concepts of Design Assurance for Neural Networks (CoDANN) II, 2021), May 2021
- EASA Artificial Intelligence Roadmap 1.0 (Artificial Intelligence Roadmap 1.0, 2020), February 2020
- EASA Concept Paper: First Usable Guidance for Level 1 Machine Learning Applications (EASA Concept Paper: First Usable Guidance for Level 1 Machine Learning Applications, 2021), April 2021
- ASTM F3060: Standard Terminology for Aircraft (ASTM International: F3060-20 Standard Terminology for Aircraft, 2020), February 2020
- ASTM TR1-EB: Autonomy Design and Operations in Aviation: Terminology and Requirements Framework, 2019 (Autonomy Design and Operations in Aviation: Terminology and Requirements Framework, 2019)
- ASTM TR2-EB: Developmental Pillars of Increased Autonomy for Aircraft Systems (ASTM International: TR2-EB: Developmental Pillars of Increased Autonomy for Aircraft Systems, 2019)
- ASTM F3269-17: Standard Practice for Methods to Bound Flight Behavior of Unmanned Aircraft Systems Containing Complex Functions (ASTM International: F3269-17 Standard Practice Methods to Safely Bound Flight Behavior of Unmanned Aircraft Systems Containing Complex Functions, 2017)
- ASTM F3269: An Industry Standard on Run-Time Assurance for Aircraft Systems (Nagarajan et al., 2021), January 2021
- Leveraging ASTM Industry Standard F3269-17 for Providing Safe Operations of a Highly Autonomous Aircraft (Skoog et al., 2020), 2020
- An ASTM Standard for Bounding Behavior of Adaptive Algorithms for Unmanned Aircraft Operations (Invited) (Cook, 2017), January 2017

As noted above, ASTM F3060 (ASTM International: F3060-20 Standard Terminology for Aircraft, 2020) provides a set of standardized definitions related to autonomous systems, including the distinction between automated systems and autonomous systems. The standard also defines key terms such as ML, intelligent agent, non-deterministic system, safety monitors and run-time assurance architecture. ASTM F3269-17 (ASTM International: F3269-17 Standard Practice Methods to Safely Bound Flight Behavior of Unmanned Aircraft Systems Containing Complex Functions, 2017) provides information on the implementation of run-time assurance/safety monitor architectures for airborne systems. EASA CoDANN (Concepts of Design Assurance for Neural Networks (CoDANN), 2020), CoDANN II (Concepts of Design Assurance for Neural Networks (CoDANN) II, 2021), and the EASA Concept Paper for First Usable Guidance for Level 1 Machine Learning Applications (EASA Concept Paper: First Usable Guidance for Level 1 Machine Learning Applications, 2021) provide detailed background and certification considerations for autonomous systems. They are particularly focused on the challenges posed by the use of neural networks in aviation and in the broader context of allowing ML and more generally artificial intelligence on-board aircraft for safety-critical applications. The EASA Concept Paper for First Usable Guidance for Level 1 ML Applications (EASA Concept Paper: First Usable Guidance for Level 1 Machine Learning Applications, 2021) also provides guidance for learning assurance of autonomous systems and a set of design objectives/tasks for the development of AI/ML functions.

3. Example Use Case: Vision Based Receiver Unmanned Aircraft Autonomous Aerial Refueling

An example use case was selected to facilitate the development of an airworthiness certification strategy based on an example application of AI. The use case focused on an unmanned receiver aircraft autonomous aerial refueling task where the receiver aircraft is equipped with a vision-based ML/neural network sensor that provides aerial refueling drogue location in 3-D space to the receiver aircraft's vehicle management system.

The 2006 NASA/Defense Advanced Research Projects Agency (DARPA) study (Schweikhard, 2006) and the 2015 X-47 program (Photo Release -- X-47B Unmanned Aircraft Demonstrates the First Autonomous Aerial Refueling, 2015) both demonstrated the ability of a UAS to receive fuel through the NATO standard method with limited human interaction. However, both programs were flown under intern flight clearances (IFCs). An IFC requires multiple risk mitigation steps from flight certification officials and is not intended for general fleet use. This work focuses on developing a process for obtaining a permanent flight clearance (PFC) for autonomous behavior. A PFC would allow operations with limited risk mitigation steps in place.

Building on prior work (Schweikhard, 2006; Photo Release -- X-47B Unmanned Aircraft Demonstrates the First Autonomous Aerial Refueling, 2015), we assumed that unmanned receiver aircraft would be capable of navigation to pre-contact position (5 to 20 feet directly aft of the drogue) behind the tanker aircraft. Figure 1 illustrates the pre-contact position behind a wing pod of a KC-135. From this point, the UAS would employ a computer vision-based optical sensor with a neural net to identify and track the drogue through contact. The vision sensor provides drogue location from pre-contact (5-20 feet behind the drogue) to contact (probe tip linking with the coupler). The vehicle management system commands the vehicle position to place the aerial refueling probe tip in the drogue based on the drogue position provided by a computer vision system. Following contact, the vehicle management system will station keep based on a position signal (such as a Differential Global Positioning (DGPS) signal as demonstrated in Reference (Schweikhard, 2006; Photo Release -- X-47B Unmanned Aircraft Demonstrates the First Autonomous Aerial Refueling, 2015)) from the tanker aircraft. Consideration of the use case highlighted some of the challenges in applying traditional design and development assurance techniques to autonomous functions. Figure 2 highlights the drogue, coupler and probe tip on an F/A-18F preparing to refuel from a KC-130 wing pod.

Figure 1. EA-18G Growler at the pre-contact position in 2013 behind a KC-135 ("EA-18G at the Pre-Contact Point Behind a KC-10 over Afghanistan," 2013)

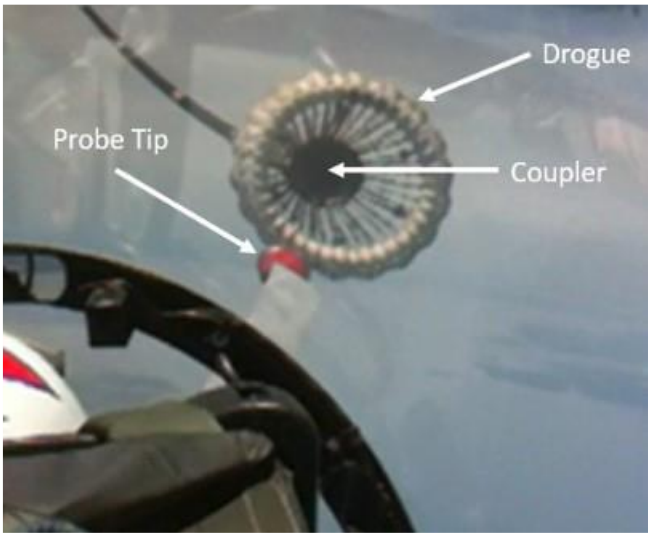

Figure 2. F/A-18F Super Hornet preparing to aerial refuel ("F/A-18F Preparing to Aerial Refuel Over Maryland," 2010)

4. Autonomous Systems Airworthiness Certification Challenges

The emergence of AI/ML functions in airborne systems presents unique airworthiness certification challenges. These challenges include limitations in the application of existing development assurance concepts to AI/ML, system architecture considerations, and the unique challenges associated with machine learning. The design and analysis techniques traditionally applied to deterministic functions may not provide adequate development assurance/safety coverage for AI/ML functions. While the principles of development assurance (utilizing a combination of process assurance and verification coverage criteria, or structured analysis or assessment techniques) may be applied, the non-deterministic nature of AI/ML drives unique design assurance considerations to provide confidence that errors in requirements, design, integration, or interaction effects have been adequately identified and corrected.

Unique challenges associated with certification of AI/ML functions include:

- System Development Assurance: This includes unique considerations for the safety assessment process, definition of requirements for the intended function and architectural considerations.
- Learning Process: This includes the process to train the system, evaluate system performance, and the hardware/software learning environment.
- Data Assurance Process: This includes definition of an end-to-end process to select and manage training data sets throughout the product lifecycle, and considerations regarding training data quality (accuracy, completeness, etc.).
- Training Verification and Data Sets: This includes the development and management of training data and the processes to verify the system functions as intended.

Existing requirement-based hardware and software development and certification processes and standards such as MIL-STD-882E (Department of Defense MIL-STD-882E: Department of Defense Standard Practice, System Safety, 2012), ARP-4761 (Guidelines and Methods for Conducting the Safety Assessment Process on Civil Airborne Systems and Equipment, 1996), ARP-4754 (Guidelines for Development of Civil Aircraft and Systems, 2010), DO-178C (Certification Specifications for Large Aeroplanes CS-25 Software Considerations in Airborne Systems and Equipment Certification, 2011) are well suited to the certification of deterministic systems and functions. These processes and standards utilize formal methods to implement the system engineering process to identify requirements, develop and verify hardware and software functions and items. The process begins with development of functional and performance requirements and, in parallel, decomposition of the system into hardware and software functions. A functional hazard assessment (FHA) is conducted to identify functional failure conditions leading to hazards. Functional failure conditions are assigned a hazard classification that characterizes the probability and severity of the functional failure. As the system is further decomposed into hardware and software items, hazard classifications are assigned to each hardware and software item. For software, a level of rigor or development assurance level is assigned based on criticality of the function/item that establishes the development processes and tasks required to ensure that the function/item performs as intended with an appropriate level of safety. MIL-STD-882E (Department of Defense MIL-STD-882E: Department of Defense Standard Practice, System Safety, 2012), ARP-4761 (Guidelines and Methods for Conducting the Safety Assessment Process on Civil Airborne Systems and Equipment, 1996), ARP-4754 (Guidelines for Development of Civil Aircraft and Systems, 2010), DO-178C (Certification Specifications for Large Aeroplanes CS-25 Software Considerations in Airborne Systems and Equipment Certification, 2011) provide guidelines for the development of software based on the hazard classification/safety criticality of the function. More details on the FHA process can be found in Reference (Guidelines for Development of Civil Aircraft and Systems, 2010).

System architecture is an important consideration in the development of the FHA. Architectural considerations such as redundancy, functional independence, and the degree of human oversight are considered when establishing the criticality and level of rigor/development assurance level for each function/item. MIL-STD-882E (Department of Defense MIL-STD-882E: Department of Defense Standard Practice, System Safety, 2012) implements the concept of software criticality index (SWCI) to establish level rigor tasks for

software development. A SWCI is assigned based on the Software Control Category and severity/consequence of failure. Similarly, DO-178C (Certification Specifications for Large Aeroplanes CS-25 Software Considerations in Airborne Systems and Equipment Certification, 2011) assigns a design assurance level (DAL) to software items based on the functional failure analysis. A set of development objectives are assigned based on the assigned DAL. Use of development assurance methods that utilize a combination of process assurance and verification coverage criteria, or structured analysis or assessment techniques increase confidence that errors in requirements or design, integration, and interaction effects have been adequately identified and mitigated. An overview of the FHA can also be found in Reference (Guidelines for Development of Civil Aircraft and Systems, 2010).

The EASA AI Roadmap (Artificial Intelligence Roadmap 1.0, 2020) notes that traditional development assurance frameworks are not completely adaptable to ML functions. The roadmap identifies several challenges with respect to the trustworthiness of AI/ML functions and the integrity of learning processes.

4.1. Learning Assurance

While the principles of traditional requirement-based development assurance can be applied to autonomous functions using ML, the development assurance process must also consider the unique aspects of the learning function and its implementation during development and fielding. This process is known as a learning assurance. Learning assurance comprises the systematic activities to provide at an adequate level of confidence that the system functions as intended, that errors in the data driven learning process are identified and corrected such that the system satisfies applicable requirements (including safety considerations), and provides sufficient generalization guarantees (EASA Concept Paper (EASA Concept Paper: First Usable Guidance for Level 1 Machine Learning Applications, 2021). The EASA AI Roadmap (Artificial Intelligence Roadmap 1.0, 2020) and the CoDANN phases (Concepts of Design Assurance for Neural Networks (CoDANN), 2020; Concepts of Design Assurance for Neural Networks (CoDANN) II, 2021) identify the challenges associated with learning assurance and provide recommended certification strategies to support certification of autonomous systems implementing ML.

A building block approach was identified that is intended to provide confidence at an appropriate level that an AI/ML function supports the intended functionality safely. The learning assurance 'Process W' developed by EASA addresses the fundamental aspects of this approach. The Learning Assurance Process 'W' is summarized in CoDANN and CoDANN II (Concepts of Design Assurance for Neural Networks (CoDANN), 2020; Concepts of Design Assurance for Neural Networks (CoDANN) II, 2021).

4.2. Architectural Mitigations – Safety Monitors and RTA

Given the challenges with ensuring the trustworthiness of AI/ML functions, architectural considerations such as the implementation of deterministic independent RTA safety monitors can mitigate the inability to fully validate and verify AI/ML functions, ensure behavior is always safe and predictable, and mitigate the risks associated with anomalous/undesired behavior. With this approach, the uncertainty in the AI black box is mitigated by implementing deterministic boundaries and controls around it.

ASTM international committee F38 on Unmanned Aircraft Systems developed ASTM F3269-17 (ASTM International: F3269-17 Standard Practice Methods to Safely Bound Flight Behavior of Unmanned Aircraft Systems Containing Complex Functions, 2017) Standard Practice for Methods to Safety Bound Flight Behavior of Unmanned Aircraft Systems Containing Complex Functions. Acknowledging the challenges in verification of complex functions using conventional software methods, the document was developed to provide industry with a standard practice for certification of UAS containing complex functions. F3269-17 (ASTM International: F3269-17 Standard Practice Methods to Safely Bound Flight Behavior of Unmanned Aircraft Systems Containing Complex Functions, 2017) address a RTA safety monitor architecture concept that implements independent real-time monitoring, prediction, and fail-safe recovery mechanisms that bound the behavior of a complex functions to ensure the safety of a UAS. The RTA architecture implements a deterministic independent safety monitor function that oversees the outputs of the complex function and ensures that the outputs are safe and executable based on a set of predetermined rules that bound acceptable outputs of the complex function. Should the complex function's outputs be determined to be outside of an acceptable range, the safety monitor implements a deterministic response overriding the outputs of the complex system to ensure the air system remains in a safe state.

Independent safety monitors and RTA architectures provide a deterministic layer of protection around the complex function and mitigates the risk of unpredictable/anomalous behavior of the complex function. In addition, safety monitors provide a layer of failure detection that is independent of the complex system being monitored. As they are deterministic, RTA safety monitor architectures can be developed using traditional hardware and software systems engineering processes (e.g., the systems engineering 'V'). Existing functional, hardware and software safety assessment and development assurance processes and standards (e.g., MIL-STD-882E (MIL-STD-882E: Department of Defense Standard Practice, System Safety, 2012), ARP-4761 (Guidelines and Methods for Conducting the Safety Assessment Process on Civil Airborne Systems and Equipment, 1996), ARP-4754 (Guidelines for Development of Civil Aircraft and Systems, 2010), DO-178C (Certification Specifications for Large Aeroplanes CS-25 Software Considerations in Airborne Systems and Equipment Certification, 2011) can be utilized. System and subsystem functional hazard analyses considering system architecture can be used to establish RTA safety monitor safety objectives and mitigations. This includes allocation of functional, performance and safety requirements to the RTA safety monitor function, and use of system and subsystem functional hazard analyses considering system architecture to establish safety objectives and mitigations. Hardware and software safety objectives should be based on the criticality of function using MIL-STD-882E (Department of Defense MIL-STD-882E: Department of Defense Standard Practice, System Safety, 2012) or civil equivalent (ARP-4761 (Guidelines and Methods for Conducting the Safety Assessment Process on Civil Airborne Systems and Equipment, 1996), ARP 4754 (Guidelines for Development of Civil Aircraft and Systems, 2010). Typically, safety monitors

are considered safety critical functions and developed to the appropriate level of rigor/development assurance.

An example implementation of a RTA safety monitor architecture is provided in Figure 3. In this example, an AI/neural net complex function provides outputs to a vehicle management function. A safety monitor is implemented that monitors the inputs to, and outputs from the complex function. Input monitors assure that system inputs are acceptable for use by the complex function (e.g., valid within an acceptable range of expected values, rates of change, etc.). The output side of the safety monitor checks that the outputs of the complex function are with predefined safety boundaries/range of acceptability before being sent to the vehicle management function. In the event that the output of the complex function falls outside of a pre-defined set of safety boundaries, the safety monitor will override the complex function's output, and provide outputs to the vehicle management function that insures the system remains in a safe state. These outputs can range from simple error declaration that flags an error in the complex function to issuance of safe state commands in lieu of those determined by the complex function.

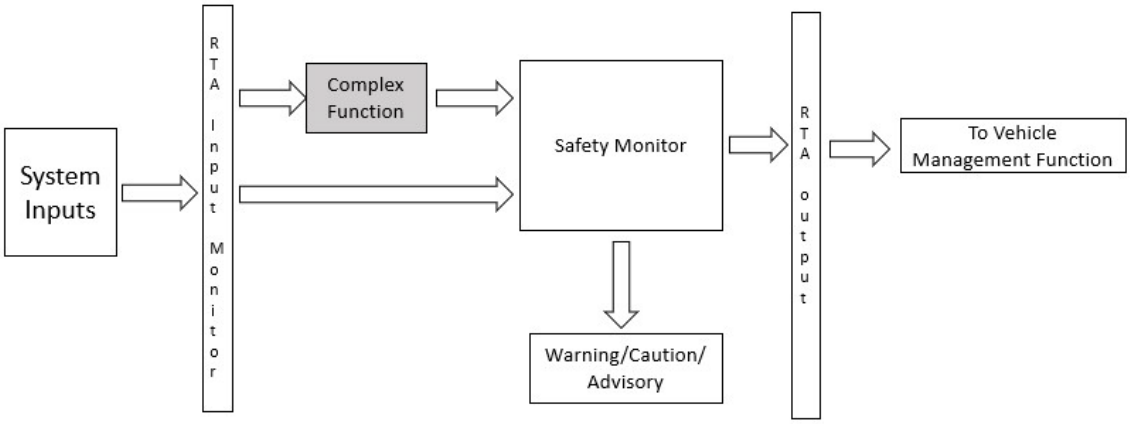

Figure 3. RTA Architecture

5. Materials and Methods

The summit determined that the following concept for aerial refueling could lead to a PFC for a AI/ML function within naval aviation and can be used as a certification strategy and framework:

- Midair Collision Avoidance With the Tanker Aircraft: A 3-D keep-out zone safety boundary shall be established around the tanker aircraft. This keep-out zone serves as a safety boundary to ensure the receiver aircraft does not contact the tanker. Utilizing a DGPS based relative navigation solution, deterministic vehicle management function can be used to position the receiver in the pre-contact position. During the entire tanking evolution, a deterministic safety monitor ensures the receiver aircraft remains in a safe position relative to the tanker.
- Corridor of Autonomy (COA): As part of the RTA architecture the receiver would maintain separation from the tanker and only attempt to make contact with a drogue that is within a defined volume behind the tanker. The summit agreed that the key to obtaining a PFC for autonomous tanking is to establish a COA. The COA would also be considered a flight clearance envelope, where the system will be permitted to exhibit autonomous behavior where a monitor will ensure it will remain within the envelope (Figure 4). The dimensions of the COA are defined relative to the tanker aircraft. It would include the nominal location of the drogue and the nominal location of the optional refueling point (approximately 10 feet forward of the contact position). The COA would include some amount of room for deviations from the nominal positions to allow for perturbations. The UAS would only be allowed to exhibit autonomous behavior if it that behavior would result in the UAS remaining within the COA.
- Transition from Pre-Contact to Contact: Once in the pre-contact position, the vehicle position command transitions from DGPS-based tanker relative position to drogue-relative navigation position using an AI/ML based computer vision sensor. The AI/ML based vision sensor identifies the drogue location and provides relative position commands to the vehicle management function which maneuvers the probe tip into the drogue. This contact point is marked via a relative position to the tanker.
- Fuel Transfer: After the receiver makes contact with the drogue, it will push the drogue in approximately 10 feet (based on DGPS) to allow fuel to transfer. Throughout the tanking evolution safety monitors track the location of the receiver and tanker aircraft to ensure the COA is not breached. This will also continue to mitigate the risk of midair collision.
- Completing the Task: Once the receiver has completed its fuel onload it will return the drogue to the contact point, back out to its pre-contact point, and continue its planned mission.

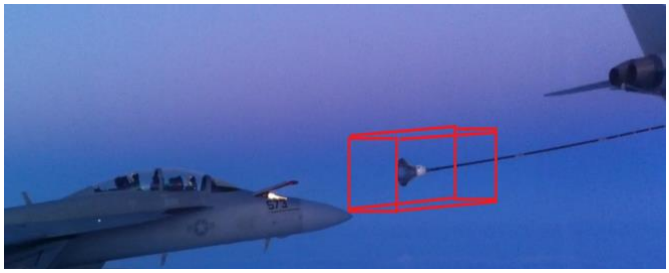

Figure 4. EA-18G Growler at the Pre-Contact Position in 2011 with a Notional COA added to the Image ("EA-18G Completing the Aerial Refueling Task Over Washington State," 2011)

6. Remaining Tasks for Certification

Certification of autonomous systems presents unique challenges. Existing processes and standards form a basic foundation for certification, but are predicated on the use of deterministic systems. Near term practical certification solutions for certification of non-deterministic AI/ML applications are needed. The fundamental work accomplished by ASTM and EASA provides a jumping off point for a development of a tailored development assurance strategy for AI/ML. It is recommended that the NACAIWG develop and document certification, criteria, standards, methods of compliance and processes for AI/ML in DoD air systems. In an effort to help establish a path forward to certification, the NACAIWG is attempting to reach a consensus across the DoD for updates to the relevant guidance documents.

The following key elements should be considered in the development of a tailored certification framework for a specific AI/ML function:

- Implement traditional systems engineering processes (the systems engineering "V"). Utilize existing functional, hardware and software safety assessment and development assurance processes and standards to the maximum extent possible (e.g., MIL-STD-882E (MIL-STD-882E: Department of Defense Standard Practice, System Safety, 2012), ARP-4761 (Guidelines and Methods for Conducting the Safety Assessment Process on Civil Airborne Systems and Equipment, 1996), ARP-4754 (Guidelines for Development of Civil Aircraft and Systems, 2010), DO-178C (Certification Specifications for Large Aeroplanes CS-25 Software Considerations in Airborne Systems and Equipment Certification, 2011)). This includes of system level functional, performance and safety requirements allocation to the AI/ML function. Conduct system and subsystem functional hazard analyses considering system architecture, and establish hardware and software safety requirements and mitigations.
- Leverage the EASA learning process W to develop a development assurance framework for AI/ML functional assurance. Tailor EASA ML development objectives to the proposed system design and architecture. Leverage EASA ML safety assessment objectives, means of compliance and guidance material to establish a framework for AI/ML functional assurance. Implement and adapt the traditional Systems Engineering V with the Learning Assurance W. Guidance and objectives are provided in EASA concept Paper: First Usable Guidance for Level 1 Machine Learning Applications (EASA Concept Paper: First Usable Guidance for Level 1 Machine Learning Applications, 2021).
- Implement an independent, deterministic RTA/safety monitor architecture to mitigate residual risk associated with the uncertainty of the AI/ML function and satisfy system-level safety objectives. This is necessary due to the complexity and nature of AI/ML functions and the challenges associated with ensuring fail-safe behavior of non-deterministic autonomous functions. Because they are deterministic, existing hardware and software systems engineering processes can be used and existing functional, hardware and software safety assessment and development assurance processes and standards (e.g., MIL-STD-882E (MIL-STD-882E: Department of Defense Standard Practice, System Safety, 2012), ARP-4761 (Guidelines and Methods for Conducting the Safety Assessment Process on Civil Airborne Systems and Equipment, 1996), ARP-4754 (Guidelines for Development of Civil Aircraft and Systems, 2010), DO-178C (Certification Specifications for Large Aeroplanes CS-25 Software Considerations in Airborne Systems and Equipment Certification, 2011)) may be utilized. This includes allocation of functional, performance and safety requirements to RTA safety monitor functions, and use of system and subsystem functional hazard analyses considering system architecture to establish safety objectives and mitigations. Hardware and software safety objectives should be based on the criticality of function using MIL-STD-882E (Department of Defense MIL-STD-882E: Department of Defense Standard Practice, System Safety, 2012) or civil equivalent (ARP-4761 (Guidelines and Methods for Conducting the Safety Assessment Process on Civil Airborne Systems and Equipment, 1996), ARP 4754 (Guidelines for Development of Civil Aircraft and Systems, 2010)). Treat safety monitors as safety critical functions. Guidance on the development and implementation of safety monitors and RTA can be found in ASTM F3269-17 (ASTM International: F3269-17 Standard Practice Methods to Safely Bound Flight Behavior of Unmanned Aircraft Systems Containing Complex Functions, 2017).
- Implement provisions for delegation of autonomy to the system as well a human supervision/monitoring/oversight of the autonomous function (human-on-the-loop) as an additional mitigation against unpredictable system behavior or complex failure conditions.
- Utilize an iterative AI/ML development and verification strategy utilizing software simulations, hardware in the loop simulations, and flight test data to train, and verify the AI/ML function.
- Limit ML to the development environment until confidence is gained in the learning function. Once confidence is gained in the integrity of the AI/ML function, consider enabling learning in operational systems.

7. Conclusions and Future Work

The June 2021 summit was a large step in the right direction for certification of autonomous systems in Naval Aviation. By bringing certification officials from the three services and NASA together with industry and academia on a notional use case we were able to have an open and frank dialog on the airworthiness certification issues associated with AI/ML. AI/ML certification guidelines need to be established by the government before we task industry to develop the systems. The results of the summit highlighted the need for standards and guidelines that support the development and certification of autonomous systems.

The results of the summit also provide insight into a potential airworthiness certification framework for AI/ML for a specific use case. We recommend that future summits expand to other use cases. We recommend that a future summit be held with an expanded audience (to include NAVAIR Naval Subject Matter Experts) to further iterate on standards, methods of compliance, validation and verification processes that will be required to certify air systems that incorporate AI/ML functions.

We recommend a future summit to further decompose the unmanned aerial refueling task, with a focus on establishing guidelines, standards and methods of compliance supporting development and certification of autonomous systems. The example use case may be useful in exploring testing, modeling and simulation methods and strategies to support autonomous system certification.

The summit chose to keep the receiver aircraft as a generic system. Not one that is developed by one of the primes (i.e., Boeing, Lockheed Martin, or Northrop Grumman). This enabled free and open discussion among all attendees of the summit. Prior to full certification of the autonomous system the interactions between the autonomous functionality and the air vehicle would need to be vetted through flight clearance officials. This paper only focused on the application capabilities of the system and not the technical aspects. Once a path towards certification has been approved by airworthiness authorities, examining the technical aspects of the interactions would have been to be accomplished before a truly autonomous system can be given a safety of flight certification within United States Naval Aviation.

Ethical approval

Not applicable.

Conflicts of Interest

The authors declare that there is no conflict of interest regarding the publication of this paper.

Acknowledgements

The authors would like to thank the National Airworthiness Council Artificial Intelligence Working Group who helped publicize the summit in multiple venues. They would also like to thank all of the attendees to the June 2021 summit for their willingness to participate and collaborate in an open frank discussion on finding a path forward for certifying an AI within Naval Aviation. It was only through their efforts that the summit was able to produce the results documented in this paper. In particular, they would like to thank Dr. Xu and the University of Maryland for enabling the summit through a virtual bridge available to all attendees.

References

Advisory Circular 23.1309.E: System Safety Analysis and Assessment for Part 23 Airplanes. (2011). United States Department of Transportation: Federal Aviation Administration.

Air Force Instruction 62–601: USAF Airworthiness. (2010). Secretary of the Air Force.

Air Force Policy Directive 62–6: USAF Airworthiness. (2019). Secretary of the Air Force.

Army Regulation 70–62: Airworthiness of Aircraft Systems. (2016). Department of the Army.

Artificial Intelligence Roadmap 1.0. (2020). European Union Aviation Safety Agency.

ASTM International: F3060-20 Standard Terminology for Aircraft. (2020). ASTM International, Conshohocken, PA.

ASTM International: F3269-17 Standard Practice Methods to Safely Bound Flight Behavior of Unmanned Aircraft Systems Containing Complex Functions. (2017). ASTM International.

ASTM International: TR2-EB: Developmental Pillars of Increased Autonomy for Aircraft Systems. (2019). ASTM International.

Autonomy Design and Operations in Aviation: Terminology and Requirements Framework. (2019). ASTM International.

Certification Specifications for Large Aeroplanes CS-25 Software Considerations in Airborne Systems and Equipment Certification. (2011). RTCA Document DO-178C.

Certification Specifications for Large Airplanes CS-25. (2007). European Union Aviation Safety Agency.

Concepts of Design Assurance for Neural Networks (CoDANN). (2020). European Union Aviation Safety Agency.

Concepts of Design Assurance for Neural Networks (CoDANN) II. (2021). European Union Aviation Safety Agency.

Cook, S. P. (2017). An ASTM Standard for Bounding Behavior of Adaptive Algorithms for Unmanned Aircraft Operations (Invited). AIAA Information Systems-AIAA Infotech @ Aerospace.

Department of Defense Directive 3009.09: Autonomy in Weapons Systems. (2017). United States Department of Defense.

Department of Defense MIL-STD-882E: Department of Defense Standard Practice, System Safety. (2012). United States Department of Defense.

EA-18G at the Pre-Contact Point Behind a KC-10 over Afghanistan. (2013). [Photograph]. In from the Private Collection of CDR Costello.

EA-18G Completing the Aerial Refueling Task Over Washington State. (2011). [Photograph]. In from the Private Collection of CDR Costello.

EASA Concept Paper: First Usable Guidance for Level 1 Machine Learning Applications. (2021). European Union Aviation Safety Agency.

F/A-18F Preparing to Aerial Refuel Over Maryland. (2010). [Photograph]. In from the Private Collection of CDR Costello.

Guidelines and Methods for Conducting the Safety Assessment Process on Civil Airborne Systems and

Equipment. (1996). S-18 Aircraft and Sys Dev and Safety Assessment Committee.

Guidelines for Development of Civil Aircraft and Systems. (2010). S-18 Aircraft and Sys Dev and Safety Assessment Committee.

Joint Software Systems Safety Engineering Handbook. (2010). United States Department of Defense.

MIL-HDBK-516C: Department of Defense Handbook, Airworthiness Certification Criteria. (2014). United States Department of Defense.

Nagarajan, P., Kannan, S. K., Torens, C., Vukas, M. E., & Wilber, G. F. (2021). ASTM F3269 - An Industry Standard on Run Time Assurance for Aircraft Systems. AIAA Scitech 2021 Forum.

NAVAIR Airworthiness and Cybersafe Process Manual, NAVAIR Manual M-13034.1. (2016). Naval Air Systems Command.

NAVAIR Instruction 1304.1F: Airworthiness and Cybersecurity Safety Policies for Air Vehicles and Aircraft Systems. (2016). Naval Air Systems Command.

Photo Release -- X-47B Unmanned Aircraft Demonstrates the First Autonomous Aerial Refueling. (2015). Northrop Grumman Newsroom. https://news.northropgrumman.com/news/releases/photo-release-x-47b-unmanned-aircraft-demonstrates-the-first-autonomous-aerial-refueling

Schweikhard, K. (2006). NASA/DARPA Automatic Probe and Drogue Refueling Flight Test. SAE Guidance and Control Subcommittee Meeting.

Skoog, M. A., Hook, L. R., & Ryan, W. (2020). Leveraging ASTM Industry Standard F3269-17 for Providing Safe Operations of a Highly Autonomous Aircraft. 2020 IEEE Aerospace Conference.

Cite this article: Costello, D., Adams, R. (2023). A Framework for Airworthiness Certification of Autonomous Systems Within United States Naval Aviation. Journal of Aviation, 7(1), 7-16.